

# Data analytics to support early warning and early action to prevent mass atrocities: a scoping review

August 2024



HARVARD  
HUMANITARIAN  
INITIATIVE



**Signal**  
Human Security  
+ Technology

# About this report

## Authors

Saira Khan <sup>1</sup>

Fakiha Ahmed

Benjamin Shenouda <sup>2</sup>

Jacqueline Cellini <sup>3</sup>

Katha Ray <sup>4</sup>

Charitra Pabbaraju <sup>5</sup>

Henry Ashworth <sup>6</sup>

Erica Nelson <sup>1,7</sup>

<sup>1</sup> Harvard Humanitarian Initiative, Cambridge, Massachusetts, United States of America

<sup>2</sup> Uppsala University, Uppsala, Sweden

<sup>3</sup> Francis A. Countway Library of Medicine at Harvard Medical School, Boston, Massachusetts, United States of America

<sup>4</sup> Columbia University, New York City, New York, United States of America

<sup>5</sup> Oxford Department of International Development, Oxford, United Kingdom

<sup>6</sup> Department of Emergency Medicine, Highland Hospital, Oakland, California, United States of America

<sup>7</sup> Department of Emergency Medicine, Harvard Medical School, Boston, Massachusetts, United States of America

All listed affiliations reflect the institution with whom the author was affiliated when the research was conducted.

## Keywords

Mass atrocity, conflict prediction, early warning, early action, data analysis

# Table of Contents

<b>Abstract</b>	<b>3</b>
<b>Policy Significance Statement</b>	<b>3</b>
<b>Introduction</b>	<b>4</b>
<b>Methods</b>	<b>5</b>
<b>Search strategy</b>	<b>5</b>
<b>Selection of studies</b>	<b>5</b>
<b>Inclusion and exclusion criteria</b>	<b>6</b>
<b>Data extraction</b>	<b>7</b>
<b>Analysis</b>	<b>7</b>
<b>Results</b>	<b>8</b>
<b>Included studies</b>	<b>8</b>
<b>Study objectives</b>	<b>8</b>
<b>Geographic scope</b>	<b>9</b>
<b>Authorship and data ownership</b>	<b>10</b>
<b>Study methods</b>	<b>10</b>
<b>Study outcomes</b>	<b>13</b>
<b>Study Impact</b>	<b>14</b>
<b>Study Limitations</b>	<b>16</b>
<b>Discussion</b>	<b>17</b>
<b>Limitations</b>	<b>19</b>
<b>Conclusion</b>	<b>20</b>
<b>Data availability statement</b>	<b>21</b>
<b>Competing interests</b>	<b>21</b>
<b>Author contributions</b>	<b>21</b>
<b>References</b>	<b>21</b>
<b>Appendix I. Search terms used for scoping review</b>	<b>26</b>
<b>Appendix II. Summary of reviewed literature</b>	<b>28</b>

## **Abstract**

Globally, there are initiatives to strengthen the efficacy and efficiency of mass atrocity prevention efforts through developing and implementing policies and early warning systems. While substantial progress has been made, the full extent of the analytical tools and techniques used to inform early warning and early action for mass atrocity prevention and the efficacy thereof remains unclear. This scoping review sets out to identify, describe, and understand what data and analytic methods have been developed to support early warning or action for conflict and mass atrocity prevention. This review was based on search terms capturing 1) mass violence, 2) early warning and early action, and 3) analytical methods. The searches resulted in 3,475 articles from 3 databases, which were narrowed to 70 thoroughly analyzed articles. The review found that most literature is concerned with conflict prediction rather than mass atrocity prediction and that there appeared to be substantial variability in the predictive variables and methodologies used for conflict and mass atrocity prediction. Most notable, however, was the overrepresentation of authors affiliated with institutions in high-income countries, while most studies were set in middle- to low-income countries. While a plethora of unique research was reviewed, none of the included studies documented how their methodologies or findings resulted in programmatic or policy change for mass atrocity prevention. In conclusion, while unique and impactful research is being produced on conflict and mass atrocity prediction, more must be done to bring researchers, policymakers, and practitioners together to influence mass atrocity outcomes.

## **Policy Significance Statement**

Conflict and atrocity prevention efforts at the national and international stage are driven by cycles of policymaking and critical data-driven insights. Existing policies have enabled the development of conflict early warning systems and in turn, the scientific and research communities have worked to model the onset, duration, and termination of mass atrocities to better understand its drivers and dynamics. These models hold tremendous potential for changing the ways in which conflict and atrocity prevention efforts are undertaken and have the power to significantly influence policy on the global stage. This paper examines how data and data analytics are being used and if these processes and findings are used for meaningful policy and programmatic change.

# Introduction

Mass atrocities continue to have severe impacts on societies worldwide, resulting in millions of civilian casualties, the destabilization of governance structures and economic systems, and an exacerbation of entrenched socioeconomic disparities. Here, mass atrocities are defined as “large-scale, systematic violence against civilian populations” and include ethnic cleansing, crimes against humanity, war crimes, and genocide (Straus 2016). In the last century alone, an estimated 84 million people were killed by mass atrocities, which is twice the number of people who have died from fighting in wars (Butcher et al. 2020). Mass atrocities gained substantial political interest following the failed international interventions in response to the horrific violence that took place in Bosnia and Rwanda in the 1990s (Anderton and Carter 2015; Krain 2005). Yet, the Early Warning Project estimates that even now 1-2 countries experience a mass killing every year (The Early Warning Project 2024).

Concerning geopolitical dynamics, radicalization, and nationalism witnessed in recent years continue to exacerbate the risk of mass atrocities worldwide (United States Holocaust Memorial Museum 2021). The continued prevalence of mass atrocities such as those perpetrated in Ethiopia, Ukraine, and Myanmar, calls for the improvement of mass atrocity prevention, early warning and early action (EW/EA) efforts. Frameworks and guidelines that are focused on mass atrocity prevention, such as the Framework of Analysis for Atrocity Crimes developed by the United Nations Office on Genocide Prevention (United Nations Office on Genocide Prevention and the Responsibility to Protect 2014) and those from the United States Agency for International Development (USAID) and the United States Holocaust Memorial Museum (USHMM), have established a solid foundation for mass atrocity prevention efforts. Mass atrocity prevention efforts continue to evolve with advancing internet, satellite technologies, geospatial data, media coverage of early warning data/risks. These initiatives and advanced technologies have improved early warning systems in mass atrocities in recent years (Anderton 2016; Butcher et al. 2020), but while significant efforts have been made, the full breadth of the analytical tools and techniques used to inform early warning and early action for mass atrocity prevention and the efficacy thereof, remains unclear.

Thus, the objectives of this scoping review were to identify, describe, and understand what types of data and analytic methods have been developed to support EW/EA for mass atrocity prevention. In addition, this research aims to characterize how these tools have been implemented and their measured impact, and to understand the types of challenges and limitations associated with the implementation of these methods. This review is

considered foundational to the creation of a standardized and evidence-based toolkit to support EW/EA decision-making and ultimately improve the efficiency and efficacy of impact on populations under threat.

## Methods

This scoping review intended to identify and map the body of evidence pertaining to how data analytics have been used to support EW/EA for atrocity prevention. Here, *early warning* is defined as “a process that (a) alerts decision makers to the potential outbreak, escalation and resurgence of violent conflict; and (b) promotes an understanding among decision makers of the nature and impacts of violent conflict,” while *early action* “is a process that utilizes early warning variables to either prevent or intervene upon evolving violence against civilian populations as a means to prevent or mitigate mass atrocities” (OECD 2009, 22).

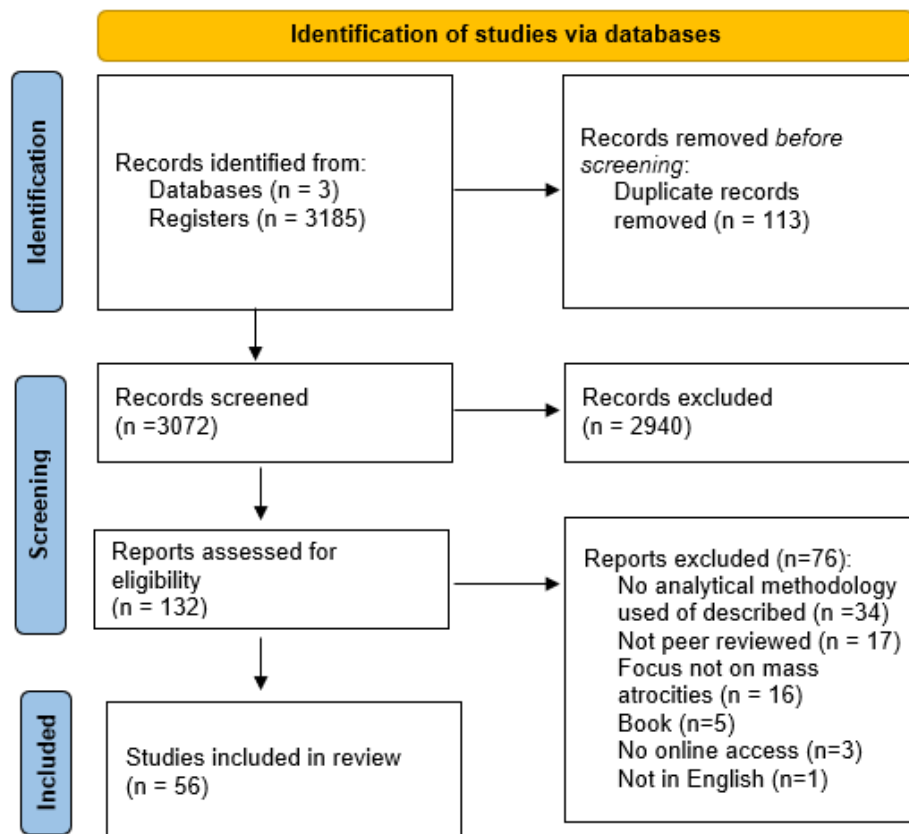
## Search strategy

An a priori (Munn et al. 2018) scoping review protocol was created and followed in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-analysis extension for Scoping Reviews (PRISMA-ScR) checklist. A set of search terms pertaining to 1) mass violence, 2) early warning and early action, and 3) analytical methods were generated based upon prior knowledge, expert consultation, and cross-referencing seminal articles and can be seen in S1 Appendix. These were used to query titles, abstracts, subheadings, key words, and MeSH terms in three distinct databases: Web of Science, Institute of Electrical and Electronics Engineers (IEEE), and Proquest (S1 Appendix). Two literature searches were conducted to capture the most current research trends. The first search was conducted on 19 December 2021 and included studies between January 1995 - December 2021, as much of the literature relevant to EW/EA was developed after the Rwandan genocide in 1994 (OECD 2009). The second literature review was conducted on March 23, 2023, capturing literature published between 19 December 2021 and March 23, 2023, on the same three databases. Due to the lack of available resources for translation services, only English-language studies were considered for this scoping review.

## Selection of studies

The two searches collectively returned 3,475 articles (before de-duplication), from three different databases. After manual de-duplication 3,360 articles were uploaded to Covidence, an online review manager that enables multiple researchers to collaboratively

screen articles for inclusion and exclusion criteria. Two reviewers screened a random sample of five papers by title and abstract to align cognitive models of the inclusion and exclusion criteria. Following this process, two independent reviewers completed the title and abstract screening of the articles with a third-party reviewer utilized to resolve any conflicts. Subsequently, 151 publications were considered for full text review and were blindly assessed by two reviewers with a third, again, resolving conflicts.



**Fig 1. PRISMA flow diagram.** Flow diagram documenting the identification, screening, eligibility, and inclusion stages of the literature considered for this scoping review. Adapted from: Page, MJ, McKenzie, JE, Bossuyt, PM, ... Moher, D (2021) The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, **372**, n71. doi:10.1136/bmj.n71.

## Inclusion and exclusion criteria

This study focused on understanding the data and analytic methods used in mass atrocity prevention efforts in research. As such, the research team defined a set of inclusion and exclusion criteria that would capture conflict-oriented research and literature that was produced after the Rwandan genocide, which is when mass atrocity prevention efforts and accompanying literature increased (OECD 2009). The inclusion criteria specifically included:

1. Peer-reviewed literature
2. Focus on early warning or early action in conflict settings
3. Thoroughly described methodology, which includes descriptions of variables, rationale for the analytical tool used, and a comprehensive description of the methodological steps
4. Published in English during or after 1995, after the Rwandan genocide

This research excluded gray literature, book chapters, dissertations, and other non-peer reviewed work, papers focused on historical analysis conducted exclusively before 1995, or papers that had a focus on medicine, small scale crimes, disasters, or animals.

## **Data extraction**

For each article, two reviewers independently extracted data from each study using the Covidence data extraction form. Extracted data included title, year of publication, first author institutional affiliation, geographic scope of study, funding sources; study characteristics, study objective, research design (quantitative, mixed methods), specific method(s) used, data sources, independent and dependent variable(s), study start and end date, classification of study (conflict prevention, early warning, early action/response, conflict prediction), limitations as defined by the author, reported results, and policy and/or programmatic impact. Conflicts in extraction were reviewed and resolved by an independent reviewer.

## **Analysis**

Data extraction results from Covidence were exported to a comma-separated values (.csv) file for analysis in Microsoft Excel and Notion (software tool for data/thought storage, organization, analysis, and sharing). The findings were analyzed by four researchers, who reviewed and compared article objectives, methodologies, outcome, and reported impact among other indicators. A thematic content analysis approach was conducted to identify common and distinctive themes. Discovered themes were organized, discussed, and theoretically classified by the authors to facilitate comparisons.



# Results

## Included studies

Of the 3,360 studies, 70 were included in the final analysis (Figure 1). Exactly 3,209 studies were excluded at the title and abstract screening stage based on the exclusion criteria described above. An additional 81 studies were excluded at the full text review stage, 41% of which were excluded because they did not provide details about what methodologies were used.

In general, while there was a significant degree of diversity between study objectives, conceptual models, variables, methods, and outcomes, several themes were identified. First, the majority of the reviewed literature is broadly concerned with conflict prediction rather than mass atrocity prediction, specifically, with “conflict” being defined differently across studies. Second, whether the primary objective was focused on conflict or mass atrocity prediction, cross-comparative analysis of the included studies revealed that there was no clear consensus on included predictive variables, the datasets from which variables were derived, or the most effective (set of) methodologies for conflict or mass atrocity prediction. Given the substantial variability and lack of standardization in the studies included in this scoping review, the comparison of the methodology and subsequent results is nuanced.

However, the most important results were thus: almost all reviewed studies were authored by researchers affiliated with institutions in high income countries, while the majority of the studies were set in middle- to low- income countries. The authors acknowledge that the inclusion of English-language manuscripts may be a confounder in this research. Further, none of the included studies documented how their methodologies or findings resulted in programmatic or policy change. There also was no comment on the actual impact of these methods on EWEA efforts.

## Study objectives

Most included studies (89%) were focused on predicting the onset of conflict (68%), genocide (14%), and/or politicide (6%), and other forms of mass atrocity (6%). The second most common objective (35%) was analyzing the changing severity of ongoing mass atrocities, i.e., the escalation, de-escalation, or termination of conflict (22%), genocide (6%), and/or politicide (3%). While the majority of studies analyzed the aforementioned variables in broad terms, approximately 22% of studies specifically analyzed how specific

actors (i.e., peacekeepers, militia, and rebel groups) affected conflict outcomes; 9% sought to quantify the impact of ethnicity on conflict, and 17% analyzed the role of environmental variables on conflict onset, severity, and termination.

It is important to note that the studies predicting the onset, severity, de-escalation, or termination of “conflict” encapsulate different types of conflict including state-sponsored conflict, armed conflict, civil conflict, and ethnic conflict. These terms were not always consistently defined across studies. For example, “armed conflict” was differently defined amongst several studies: Hegre et al (Hegre et al. 2013) use a definition proposed by Themner and Wallensteen (2011) who define armed conflict as “a contested incompatibility between a government and an organized opposition group causing at least 25 battle-related deaths during a calendar year.” Bartusevičius and Gleditsch (Bartusevičius and Gleditsch 2019) on the other hand, defined armed conflict as “an incompatibility involving the systematic use of armed force. ‘Systematic’ implies organized and sustained over an extended period.”

The search terms for this research were largely focused on mass atrocities, specifically including ethnic cleansing, crimes against humanity, war crimes, and genocide in search terms. Nevertheless, an overwhelming majority of returned literature were concerned with conflict rather than specific mass atrocities. This was notable because the only “conflict” specific term included in these searches included “armed conflict”. Despite the fact that there were substantially more mass atrocity terms, studies focused on conflict, broadly speaking; largely outnumbering those considering the prediction of mass atrocities. The set of search terms used to capture the methodological approaches used for conflict and mass atrocity prediction were relatively quantitative in nature and captured concepts relating to modeling and statistical analyses. This resulted in the return of studies that were primarily quantitative, but a few studies (<3) used primarily qualitative analyses.

## **Geographic scope**

This research found that the geographic scale of analysis varied greatly between studies, with 45% of studies conducting country-wide or regional analysis, and 32% of studies conducting analysis on a sub-country scale. Thirty-one percent of studies conduct analysis agnostic of political boundaries, often using ethnic boundaries or specific events (of documented mass killings or genocides, for instance). The geographic scope of cross-country studies typically included upwards of 25 countries across Europe, Africa, Latin America, and Asia. The most studied region (34%) was the African continent, with most publications focusing their analysis on Sub-Saharan Africa or East Africa. Fewer than 15%

of studies focused on their analyses on a singular country, but those that did focused on Nigeria, Iraq, Libya, Rwanda, Bosnia, Sudan, East-Timor, Syria, and England.

## **Authorship and data ownership**

While the geographic focus of most studies was on low- and middle-income countries, almost every author of the reviewed studies were affiliated with 78 unique institutions in high-income countries. Of those institutions, 84% were specifically affiliated with European or American academic institutions, with the remaining studies authored by researchers associated with institutions in Japan, New Zealand, Australia, China, and Israel. The most highly cited institutions were situated in Scandinavian countries.

Similarly, the most commonly used datasets were developed and hosted by European or American institutions with the most commonly cited conflict datasets being the Uppsala Conflict Data Program (UCDP) and the Peace Research Institute Oslo (PRIO). The Political Instability Task Force (PITF) and the Armed Conflict Location & Event Data Project (ACLED) were also frequently cited by reviewed studies and were maintained by institutions in the United States. Other commonly used datasets were produced by the Center for Systemic Peace, Minorities at Risk Project, Correlates of War, all of which are based in America, and ETH Zurich, which is based in Switzerland.

## **Study methods**

Study design, variable selection, and methodological approaches greatly varied among studies, signaling a lack of consensus on conceptual models and methods regarding how to best predict mass atrocities or violent conflict.

### **Variable selection**

Despite the fact that there were clear thematic variable groupings, there was no consistency in the numbers or types of variables used in studies. Explanatory variables captured a wide range of themes pertaining to conflict, economics, climate, governance, ethnicity, military, and socio-economic/demographic/political variables (Table 1). When considering variables pertaining to governance for example, Goldsmith (Goldsmith et al. 2013) used specific indicators such as political assassinations, elections, changes to political institutions, political instability, and election period in a model that forecasts genocide and politicide onset. In contrast, Wayman and Tago (Wayman and Tago 2010) uses only regime type in a model to forecast the onset of mass political killings.

**Table 1. Independent variable groupings.**

<b>Thematic categorization</b>	<b>Examples of variables captured</b>
Conflict	One-sided violence, instances of prior conflict, instability, fatality, deaths
Economics	Gross domestic product, per capita income
Climate	Temperature, precipitation, natural resources, drought
Governance	Regime type, political precipitation, state organization
Ethnicity	Ethnic fractionalization, polarization, ethnic conflict
Military	External military intervention, militia violence
Socio- demographic, economic, political	Infant mortality, education level, Gini index, human capital index
Geographic and Administrative	Boundaries, terrain, population statistics, size of states
Infrastructure	Roads/road networks, night-time lights
Other	Variables that do not fall into aforementioned categories

## **Data sources**

While some diversity persists regarding data sources across studies, approximately 44% of studies utilized three common data streams; including datasets produced by the Uppsala Conflict Data Program and Peace Research Institute Oslo (UCDP and PRIO; 25% of studies), the Political Instability Task Force (PITF; 16% of studies), and the Armed Conflict Location & Event Data Project (ACLED; 16% of studies). Approximately 32% of studies that conducted analysis on a global scale used PRIO and/or UCDP data, while only 16% of research that focused on Africa used the same datasets. Conversely, approximately 47% of the studies that focused on Africa relied on ACLED data, while only one globally scoped study used the dataset. Other commonly used datasets were produced by the Center for Systemic Peace, Minorities at Risk Project, Correlates of War, and ETH Zurich. Despite the

popular use of the handful of datasets listed above, over 60 unique dataset sources were documented across the 70 included studies.

## **Methodological approaches**

Methodologically speaking, regression analysis was most commonly used to predict conflict or mass atrocities. Specifically, 71% of studies used a type of regression analysis as their primary or secondary method of analysis. Approximately 27% of studies exclusively used logistic or probit regression models, while another 33% used regression models in combination with other analytic methodologies such as neural networks (3%; Pinckney and Babak 2022; Schellens and Belyazid 2020) or random forest decision trees (7%; Ettensperger 2022; Hegre et al. 2019; Musumba et al. 2021; Pinckney and Babak 2022; Schellens and Belyazid 2020). It was notable that almost half (49%) of studies predicted genocide, politicide, and/or conflict in a binary manner (i.e., will conflict occur? Yes, or no?), and 24% of these studies used logistic or probit regression models. Bartusevičius and Gleditsch (Bartusevičius and Gleditsch 2019) used a unique approach among the articles reviewed in modeling civil conflict as a two-stage process, where the onset of conflict (stage 2) is conditioned on the presence or absence of pre-existing civil incompatibilities (stage 1).

Just under a quarter of the studies used spatial analytical methods to understand the spatial relationships relevant to conflict and atrocities (21%; Aas Rustad et al. 2011; Buhaug et al. 2011; Cao et al. 2022; Duursma and Read 2017; Ge et al. 2022; Radford 2022; Reeder 2018). The specific spatial methodologies and geographic scope of these articles varied widely. For example, Weidmann (Weidmann and Salehyan 2013) used agent-based modeling to understand the micro-level variations in violence in Baghdad between 2006 and 2009. On the other hand, Hegre et al. (Hegre et al. 2019) take a wider perspective, and develop a model called ViEWS, which systematically monitors all locations at risk of conflict in Africa and the Middle East and produces monthly forecasts at the country and subnational level 36 months into the future. The ViEWS model explicitly incorporates spatial variation into the machine learning models used to develop these forecasts. Duursma and Read (Duursma and Read 2017) took another unique approach by modeling the spatial dispersion of conflict as analogous to disease and examined the effect that deployment of UN Peacekeepers to specific locations had on the spread of violence in Darfur using cluster analysis. Other articles using spatial methods focused on subnational variation of violence based on factors such as geographic distribution of wealth (Buhaug et al. 2011), transfer of state power (Aas Rustad et al. 2011), and the influence of terrain on patterns of violence (Reeder 2018).

While the articles discussed above already used complex analytics, more recently published articles employ even more advanced computational techniques, including machine learning models such as random forests, gradient tree boosting, or neural networks (Ettensperger 2022; Musumba et al. 2021; Pinckney and Babak 2022; Radford 2022; Schellens and Belyazid 2020). Hegre et al. (Hegre et al. 2021) took advantage of increased accessibility of high-performance computing clusters to produce automated monthly forecasts of violence in Africa and the Middle East on a very fine-grained geographical scale.

## **Study outcomes**

### **Reported Results**

Overall, historical violence was found to be one of the most important predictors of conflict (Hegre et al. 2013, 2019; Helman et al. 2020; McDoom 2014; Musumba et al. 2021). Similarly, previous civil wars (Harff 2003; Krain 1997) and previous genocides (Goldsmith et al. 2013) were the most consistent and common predictors of mass killings, genocide, or politicide onset globally. Anderton and Ryan (Anderton and Ryan 2016) also found that low-level violence against civilians was the best predictor for high-level violence against civilians. The studies that came to this conclusion, however, did not exclusively consider historical violence as a singular explanatory variable. These studies also incorporated variables associated with economics (Goldsmith et al. 2013; Musumba et al. 2021), politics (Goldsmith et al. 2013; Harff 2003), ethnicity (Goldsmith et al. 2013), socio-demographic (Harff 2003; Hegre et al. 2013; Helman et al. 2020; Musumba et al. 2021), and climatic variables (Helman et al. 2020) -all of which also contributed to the onset, severity, and termination of conflict and mass atrocities.

When analyzing the relationship between the thematic variables and study outcome more closely, several other notable trends became visible. For instance, studies that analyzed climatic variables such as temperature and precipitation generally found that hotter temperatures were associated with an increased risk of new violence or sustained violence (Anderton and Ryan 2016; O'Loughlin et al. 2012; van Weezel 2020; von Uexkull et al. 2016; Witmer et al. 2017). However, the same studies also found that temperature was generally not the most important predictor of violence compared to other thematic variables, such as sociopolitical and structural conditions.

Fragile and deteriorating political, economic, and ethnic factors were also associated with an increased risk of conflict and genocide. Miodownik and Bhavnani (Miodownik and

Bhavnani 2011) found that poor fiscal policies of a minority group in power increased the risk of conflict onset. Anderton and Carter (Anderton and Carter 2015) and Buhaug et al. (Buhaug et al. 2011) found that low income and local income variation played a role in increasing the risk of conflict: “local income matters more than national income, and we find evidence that pockets of relative wealth within very poor countries are more likely to see conflict outbreaks” (Buhaug et al. 2011). However, although several studies found that economic factors were a notable contributor to the onset of genocide and conflict, they also could not be considered as independent drivers of conflict.

In addition to climatic, political, and economic variables discussed above, ethnic variables were commonly correlated with an increased risk of violence. Several studies found that variables such as the size of an ethnic minority group and the number of excluded ethnic groups play an important role in the prediction of violence (Miodownik and Bhavnani 2011; Tollefsen and Buhaug 2015; Uzonyi and Demir 2020). Similarly, ethnic fractionalization and the mobilization of individuals along ethnic lines increase the potential for war or conflict (Eck 2009; Pinckney and Babak 2022; Wegenast and Basedau 2014). However, the wide range of ethnic variables included in individual studies leaves unclear conclusions about the role of ethnicity in conflict or mass atrocity prediction. However, the relationship between ethnicity and conflict remains debated as explained by Wimmer et al (Wimmer et al. 2009): “the most influential articles argue that ethnicity plays no role in predicting the onset of civil wars...a second group of scholars insists that ethnicity does matter”. Similarly, Hegre et al. (Hegre et al. 2013) and Eck (Eck 2009) cite uncertainty in the ways in which ethnicity and conflict correlate.

Ultimately, the data from these studies shows that while some explanatory variables are stronger predictors than others, there are numerous contributing factors that simultaneously contribute to the onset of, severity, and changing dynamics of violent conflict, genocide, or politicide and the degree of association and importance vary across contexts, timeframes, and datasets.

## **Study Impact**

None of the studies directly reported if their analysis or findings had an impact on programming or policy. However, almost all studies (61%) shared potential impact and implications for policymakers and researchers to improve ongoing and future EW/EA efforts. This research found that the recommendations could be grouped into three categories, including: 1) the incorporation of predictive models, 2) long-term, strategic conflict prevention efforts, and 3) short-term, immediate interventions.

Several authors encouraged policymakers to minimize delayed intervention to prevent the onset and/or escalation of violence by incorporating and relying on predictive forecasting (Ettensperger 2022; Munn et al. 2018; Schellens and Belyazid 2020). Most notably, Harff (Harff 2003) provided a concrete early warning model which provides a framework for a global “watch list” of high-risk countries that is intended to convince policymakers to engage in a timely manner. Goldsmith et al. (Goldsmith et al. 2013) also produced annual lists of at-risk states that could be used by policymakers that seek to prevent mass atrocities. And Hegre et al. (Hegre et al. 2019) emphasized the value of ViEWS, a monthly conflict forecasting dataset for nations across Africa and the Middle East, which aims to be maximally transparent, publicly available, provide uniform coverage and methodological innovations as their four guiding principles.

Several authors pointed to the need for improved conflict prevention and intervention strategies, highlighting that the policy community needs to focus on interventions as a long-term strategy focusing on structural changes to increase effectiveness (Anderton and Carter 2015; Cao et al. 2022; Kathman and Wood 2011). In support, Kathman and Wood (Kathman and Wood 2011) argued that short-term intervention policies have an opposite effect resulting in the escalation of violence in the long run. Thus, the authors proposed that policymakers should not see intervention as a “quick fix” and state that “when deciding to intervene to halt mass murders, third parties should plan to commit for the long haul or refrain from intervention altogether”. Krain (Krain 2017) shared a similar perspective, specifically considering the role of economic sanctions against perpetrators. Krain (Krain 2017) suggested that if policymakers choose to use sanctions as a policy option to halt atrocities, they should use it as part of a long-term strategy, because the use of sanctions as short- to medium-term interventions have shown to be ineffective.

A handful of articles emphasized the importance of immediate intervention in an active, ongoing conflict. This included recommendations regarding the effectiveness of peacekeeping missions and the prioritization of deploying peacekeepers to the region’s most vulnerable to violence in order to stop violence in its earliest stages (Plancikova et al. 2021). According to Duursma and Read (Reeder 2018): “This is in line with recent calls for what has been referred to as ‘rapid reaction forces’ or ‘fire brigades’”, which refers to the idea that rapid intervention may quell ongoing violence and also minimizes the risk of armed violence from spilling over to new areas. According to Eck (Eck 2009), the findings suggest that 60% of conflicts intensify in the first year, thus rapid engagements are necessary, and third parties that aim to prevent an increase in violence should be prepared to act immediately after a conflict breaks out.



In summary, the reviewed studies propose a variety of recommendations to support short- and long- term conflict and mass atrocity prevention efforts. It is evident however, that there appears to be disagreement between the best means of interventions. These recommendations are relevant to researchers, policymakers, and practitioners but, again, no study reported on the incorporation of their methods, findings, or recommendations into real life EWEA efforts.

## Study Limitations

Study limitations cited by the included publications were most frequently concerned with the lack of data/incomplete data, limited generalizability and scaling thereof. Of these, the most frequently cited limitations were the lack of data, incomplete data, or dataset limitations (Aas Rustad et al. 2011; Fox 2000; Goldsmith et al. 2013). For example, Warren and Troy (Warren and Troy 2015) pointed to the limitations of using the Minorities at Risk (MAR) dataset. MAR includes data with a set of inclusion criteria which does not include “the full universe of ethnic groups worldwide”, skewing study results. Wimmer et al. (Wimmer et al. 2009) also cited similar limitations, stating that the MAR dataset depends too much on disadvantaged minorities and is limited in geographic scope. Further, Goldsmith et al., (Goldsmith et al. 2013) have shown how the limited availability of timely data created gaps in their forecasting models.

Another commonly noted limitation was the generalizability of the results, which was mostly mentioned in studies that had a narrow geographic focus - a specific country or region- such as Eastern Africa (Ember et al. 2019), Rwanda (McDoom 2014), or Iraq (Weidmann and Salehyan 2013). Understandably, the authors commented on whether their findings could be applied to cases outside their geographical scope that did not meet the scope conditions (McDoom 2014; Reeder 2018).

Further, some models were based on historical data and did not include the impact of contextual changes or future events. They often only included observations within a “limited time span” (Schellens and Belyazid 2020) and did not consider effects over the longer-term or how conflicts change over time, which means that their analysis could not be extended beyond the set timeframe (Krain 2005; Reeder 2018; Wayman and Tago 2010). As humanitarian settings are subject to sudden contextual changes which are sometimes outside the parameters of the study, future trends can significantly diverge from the current forecasts (Aas Rustad et al. 2011; Ettensperger 2022).

Finally, as noted earlier, much of the literature (49%) took a dichotomous approach. While this was common, authors have also commented on the limitations of taking this approach. For instance, Bartusevičius and Gleditsch (Bartusevičius and Gleditsch 2019) mentioned that this approach does not allow to assess “alternative mechanisms relating specific variables to conflict” that may influence different conflict stages. Such research designs cannot be used to examine many aspects of conflict dynamics such as the intensity of violence or the spread of violence (Eck 2009).

## Discussion

As of late 2023, the Early Warning Project’s Statistical Risk Assessment for Mass Killing indicates that there are currently 14 countries experiencing ongoing mass killings and countries including Afghanistan, Yemen, and Guinea are at risk of new mass killings (“Ranking All Countries - Early Warning Project” n.d.). While the number of mass atrocities has declined over the last few decades, grim predictions remind humanity of the urgency to improve early warning systems to prevent conflict and mass atrocities from occurring. While technological advances have supported the development of more robust EW systems, the full breadth of analytical tools and techniques used to inform EW/EA for mass atrocity prevention is unclear. Given the rapidly evolving nature of predictive analytics and technology, this study set out to characterize the quantitative tools and methodologies used for mass atrocity prevention to create a foundation for the future development and incorporation of analytic methods for EW/EA activities.

This scoping review found that the majority of articles were focused on the prediction of general conflict dynamics while a much smaller portion of the literature analyzed mass atrocities specifically. The majority of studies used a regression-based technique but drew from a large range of explanatory variables. Across studies, variables related to conflict, ethnicity, natural resources, and economic variables, among others, played an important but inconsistent role in conflict and mass atrocity prediction. Along with valuable insights, this study revealed a clear need for the diversification and standardization of methodological approaches to address gaps in mass atrocity prevention research (Straus 2016). For example, utilization of spatial analysis led to significant insights, but was only utilized by 15 of the reviewed articles.

The overwhelming majority of the articles conducted analysis in Africa and Asia, two continents which have been heavily affected by conflict and mass atrocity in the last two decades. Interestingly, the onset, severity, and termination of conflict, genocide or politicide, and other mass atrocities were analyzed on very different geographic scales.

Many studies focused on large-scale/state-based violence, missing sub-national conflict dynamics and their indicators. While this enables country-to-country comparisons, this approach does not capture the sub-national variation of conflict and mass atrocities. Local-level conditions affect the risk of conflict outbreaks as much, if not more, compared to nation-wide dynamics. Despite the high percentage of country-level studies, numerous authors spoke to the importance of more geographically granular analyses, stating that “averaging or aggregating features at the country-level in cross-national studies can obscure a great deal of relevant local variation” (Bulutgil 2015). This sentiment is echoed by Aas Rustad et al. (Aas Rustad et al. 2011): “The aggregation of information at the yearly and/or country level can lead to inaccurate or faulty forecasts. Indeed, most attempts to model the risk of conflict, including efforts at forecasting, suffer from this problem.”

This discrepancy in the use of data sources is likely due to the fact that PRIO and UCDP host a global dataset that goes back to 1946, while ACLED has true global coverage available only after about 2018. ACLED’s first datasets, published in 1997, were focused exclusively on Africa and began covering other regions beginning in 2015. Until 2018, UCDP/PRIO was the only provider of a global, geolocated dataset on conflict events. However, the expanded coverage of ACLED has the potential to augment conflict analysis by providing alternative datasets for the research community to utilize.

As evidenced by the results, many of the datasets have different definitions, scope, and categorizations to study similar phenomena, and this variability resulted in very different conclusions. Wayman and Tago (Wayman and Tago 2010) demonstrated this in their study, where they showed that different combinations of input datasets can have a significant impact on a study’s results. The authors compared two major datasets to see what independent variables best account for the onset of mass killing. They were able to show that the regime type variable had a significant effect when using R.J. Rummel’s (1995) dataset on democides but that this effect disappears/becomes less significant when using the geno-politicide dataset by Barbara Harff (Harff 2003).

Conflict varies greatly across space and time and understanding these complex dimensions can contribute to tangible local action. For instance, it can inform the identification of the most effective locations to deploy peacekeeping missions (Duursma and Read 2017). Granular geographic analysis can also help understand the spatial behavior of rebels on the locations they choose to base operations from (Reeder 2018), expose geographical variations of local income and its relatedness to conflict onsets (Buhaug et al. 2011), and geographical and temporal differences in climate and its impact on conflict (O’Loughlin et al. 2012). With that said, none of the included studies prior to

2010 considered a sub-national focus and were all conducting analysis on a country-level or by ethnic groups.

While the geographic focus of the majority of the studies was in low- and middle- income countries, almost all of the authors were affiliated with institutions in high-income countries. Of the 120 unique co-authors cited in the 70 reviewed papers, 84% of authors were affiliated with institutions in Europe or the United States while the remaining 16% were affiliated with Australian, New Zealand, or Japanese, or Chinese institutions. As previously noted, the most widely used datasets for conflict and mass atrocity research are also developed and hosted in high-income countries. The disproportionate representation of authors from high-income countries in peer-reviewed literature is not a new phenomenon. However, it is rapidly gaining attention across disciplines including global health, public health, and climate science (Garbern et al. 2022; Plancikova et al. 2021; Tandon 2021). It is therefore equally important to highlight the disproportionate representation of authors affiliated with high-income countries in the conflict and mass atrocity prevention literature.

## **Limitations**

There were several critical limitations in conducting this research, many of which stem from the study design. First, the inclusion/exclusion criteria were relatively broad, to avoid the exclusion of critical articles. This resulted in the inclusion of broadly defined terms such as “conflict” and “models”, resulting in noise in the dataset. Further, the search terms yielded mostly quantitative studies, largely filtering out qualitative studies. The research team, however, acknowledges that qualitative data holds valuable insights and depth to research in a different manner compared to quantitative studies. Further, due to the research team’s language restrictions and available resources, only English language papers were included in this review. The reliance on western-centric databases and English language articles may have led to the exclusion of valuable literature developed by non-English speakers throughout the world. Furthermore, the reliance on peer-reviewed articles certainly biased results towards academia, and the future inclusion of grey literature is critical to understand what is being done in programmatic spheres.

Finally, a critical objective of this research included evaluating the impact that methodologies, findings, and recommendations had for other researchers, policymakers, and practitioners. Through an evaluation of the articles, it appears that none of the studies resulted in practical implementation of the proposed methods. The broad nature of analysis may in part be the reason why none of the studies reported the implementation of

efficacy of applied methodologies. This limitation leaves uncertainties as to whether these current methods are usable, replicable, scalable and actionable for mass atrocity prevention. As previously noted, no studies reported that their methodology or their findings were used for policy and programmatic implementation. Due to the nature of this review, we did not consider the long-term impact of these articles and only focused on the reported impact. Additional research is required to understand if and how policy and programmatically aligned organizations are implementing such research to support their mass atrocity prevention efforts.

## Conclusion

The purpose of this paper was to understand and document the various data and analytical methods being used to support EW/EA efforts for conflict and mass atrocity prevention. This scoping review found that most studies' objectives were focused on conflict rather than mass atrocities specifically, revealing that the conflict literature is substantially more robust. There appears to be some consensus among authors on the important role that prior conflict, economic, environmental and geographical variables play in predicting conflict, genocide, and politicide. However, many more studies differed in their understanding of how elements of ethnicity, natural resources, intervening actors, and socio-demographics affect mass atrocities and violent conflict.

An important finding presented in this scoping review is that while there is a clear preference for using standardized regression techniques, there is little agreement on which explanatory variables are most important in these models. Similar relationships explored in different models yielded different results across geographic scales. While spatial relationships are important for conflict analysis and prediction, this study found that less than a quarter of the articles reviewed explicitly consider spatial relationships in their analyses. While there has been an uptick of studies conducted at a more granular level of analysis, the overwhelming majority is still focused on country-wide analyses. While country-wide analyses provide valuable insights about risk, more granular information is needed to inform actionable EWEA efforts.

Finally, and perhaps most critically, it is notable that none of the studies directly reported if their analysis or findings had an impact on programming or policy for conflict and mass atrocity prevention. It is concerning that a wealth of valuable methodologies, findings, and datasets are produced by research institutions in high-income countries, yet none of this research has reportedly been implemented or utilized by practitioners and policymakers navigating conflict and mass atrocities in the low- and middle-income countries in which

these studies are situated. While standalone products such as the ViEWs dataset and the risk assessments developed by the Early Warning Project are open-source and accessible to the public, more efforts need to be made to bridge this wide gap between researchers and academics documenting their work in peer-reviewed papers and EWEA policymakers and practitioners.

## Data availability statement

The data that support the findings of this study are available upon request.

## Competing interests

All authors declare no conflict of interest.

## Author contributions

Conceptualization: S.K, E.N. Methodology: S.K, J.C, E.N, H.A. Data curation: J.C. Investigation: S.K, F.A, B.S, C.P, K.R. Project administration: S.K., E.N. Writing original draft: S.K, F.A., B.S., C.P., E.K. Writing-reviewing & editing: S.K, K.R, B.S, F.A. All authors approved the final submitted draft.

## References

- Aas Rustad, SC, Buhaug, H, Falch, Å, and Gates, S** (2011) All conflict is local: modeling sub-national variation in civil conflict risk. *Conflict Management and Peace Science*, 28(1), 15–40.
- Anderton, CH** (2016) Datasets and trends of genocides, mass killings, and other civilian atrocities. *Economic Aspects of Genocides, Other Mass Atrocities, and Their Prevention*, 52–101.
- Anderton, CH, and Carter, JR** (2015) A New Look at Weak State Conditions and Genocide Risk. *Peace Economics, Peace Science and Public Policy*, 21(1), 1–36. doi:10.1515/peps-2014-0008.
- Anderton, CH, and Ryan, EV** (2016) Habituation to atrocity: low-level violence against civilians as a predictor of high-level attacks. *Journal of Genocide Research*, 18(4), 539–562.

- Bartusevičius, H, and Gleditsch, KS** (2019) A Two-Stage Approach to Civil Conflict: Contested Incompatibilities and Armed Violence. *International Organization*, 73(1), 225–248.
- Buhaug, H, Gleditsch, KS, Holtermann, H, Ostby, G, and Tollefsen, AF** (2011) It's the Local Economy, Stupid! Geographic Wealth Dispersion and Conflict Outbreak Location. *Journal of Conflict Resolution*, 55(5), 814–840.  
doi:10.1177/0022002711408011.
- Bulutgil, HZ** (2015) Social cleavages, wartime experience, and ethnic cleansing in Europe. *Journal of Peace Research*, 52(5), 577.
- Butcher, C, Goldsmith, BE, Nanlohy, S, Sowmya, A, and Muchlinski, D** (2020) Introducing the Targeted Mass Killing data set for the study and forecasting of mass atrocities. *Journal of Conflict Resolution*, 64(7–8), 1524–1547.
- Cao, X, Theodora-Ismena, G, Shortland, A, and Urdal, H** (2022) Drought, Local Public Goods, and Inter-communal Conflicts: Testing the Mediating Effects of Public Service Provisions. *Defence and Peace Economics*, 33(3), 259–279.
- Duursma, A, and Read, R** (2017) Modelling Violence as Disease? Exploring the Possibilities of Epidemiological Analysis for Peacekeeping Data in Darfur. *International Peacekeeping*, 24(5), 733.
- Eck, K** (2009) From Armed Conflict to War: Ethnic Mobilization and Conflict Intensification. *International Studies Quarterly*, 53(2), 369.
- Ember, CR, Jones, EC, Skoggard, I, and Teferi Abate, A** (2019) Warfare, atrocities, and political participation: eastern Africa. *Journal of Aggression, Conflict and Peace Research*, 11(1), 11–23.
- Ettensperger, F** (2022) Forecasting conflict using a diverse machine-learning ensemble: Ensemble averaging with multiple tree-based algorithms and variance promoting data configurations. *International Interactions*, 48(4), 555–578.  
doi:10.1080/03050629.2022.1993209.
- Fox, J** (2000) Religious causes of discrimination against ethno-religious minorities. *International Studies Quarterly*, 44(3), 423–450. doi:10.1111/0020-8833.00166.
- Garbern, SC, Hyuha, G, González Marqués, C, ... Rees, CA** (2022) Authorship representation in global emergency medicine: a bibliometric analysis from 2016 to 2020. *BMJ Global Health*, 7(6), e009538. doi:10.1136/bmjgh-2022-009538.
- Ge, Q, Hao, M, Ding, F, ... Ide, T** (2022) Modelling armed conflict risk under climate change with machine learning and time-series data. *Nature Communications*, 13(1), 2839.  
doi:10.1038/s41467-022-30356-x.
- Goldsmith, BE, Butcher, CR, Semenovich, D, and Sowmya, A** (2013) Forecasting the onset of genocide and politicide: Annual out-of-sample forecasts on a global dataset, 1988–2003. *Journal of Peace Research*, 50(4), 437–452.

- Harff, B** (2003) No Lessons Learned from the Holocaust? Assessing Risks of Genocide and Political Mass Murder since 1955. *The American Political Science Review*, 97(1), 57.
- Hegre, H, Allansson, M, Basedau, M, ... Jansen, R** (2019) ViEWS: a political violence early-warning system. *Journal of Peace Research*, 56(2), 155–174.
- Hegre, H, Bell, C, Colaresi, M, ... Vesco, P** (2021) ViEWS2020: Revising and evaluating the ViEWS political Violence Early-Warning System. *Journal of Peace Research*, 58(3), 599–611. doi:10.1177/0022343320962157.
- Hegre, H, Karlsen, J, Nygard, HM, Strand, H, and Urdal, H** (2013) Predicting Armed Conflict, 2010-2050. *International Studies Quarterly*, 57(2), 250–270.
- Helman, D, Zaitchik, BF, and Funk, C** (2020) Climate has contrasting direct and indirect effects on armed conflicts. *Environmental Research Letters*, 15(10). doi:10.1088/1748-9326/aba97d.
- Kathman, JD, and Wood, RM** (2011) Managing Threat, Cost, and Incentive to Kill: The Short- and Long-Term Effects of Intervention in Mass Killings. *The Journal of Conflict Resolution*, 55(5), 735.
- Krain, M** (1997) State-Sponsored Mass Murder: The Onset and Severity of Genocides and Politicides. *Journal of Conflict Resolution*, 41(3), 331–360.
- Krain, M** (2005) International Intervention and the Severity of Genocides and Politicides. *International Studies Quarterly*, 49(3), 363–387. doi:10.1111/j.1468-2478.2005.00369.x.
- Krain, M** (2017) The effect of economic sanctions on the severity of genocides or politicides. *Journal of Genocide Research*, 19(1), 88.
- McDoom, OS** (2014) Predicting violence within genocide: A model of elite competition and ethnic segregation from Rwanda. *Political Geography*, 42, 34–45. doi:10.1016/j.polgeo.2014.05.006.
- Miodownik, D, and Bhavnani, R** (2011) Ethnic Minority Rule and Civil War Onset How Identity Salience, Fiscal Policy, and Natural Resource Profiles Moderate Outcomes. *Conflict Management and Peace Science*, 28(5), 438–458. doi:10.1177/0738894211418426.
- Munn, Z, Peters, MDJ, Stern, C, Tufanaru, C, McArthur, A, and Aromataris, E** (2018) Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *BMC Medical Research Methodology*, 18(1), 143. doi:10.1186/s12874-018-0611-x.
- 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). doi:10.3390/su13137366.
- OECD** (2009) *Preventing Violence, War and State Collapse: The Future of Conflict Early Warning and Response*, OECD. doi:10.1787/9789264059818-en.



- O’Loughlin, J, Witmer, FDW, Linke, AM, Laing, A, Gettelman, A, and Dudhia, J** (2012) Climate variability and conflict risk in East Africa, 1990-2009. *Proceedings of the National Academy of Sciences of the United States of America*, 109(45), 18344–18349. doi:10.1073/pnas.1205130109.
- Pinckney, J, and Babak, R** (2022) When the levee breaks: A forecasting model of Violent and nonviolent dissent. *International Interactions*, 48(5), 997–1026.
- Plancikova, D, Duric, P, and O’May, F** (2021) High-income countries remain overrepresented in highly ranked public health journals: a descriptive analysis of research settings and authorship affiliations. *Critical Public Health*, 31(4), 487–493. doi:10.1080/09581596.2020.1722313.
- Radford, B** (2022) High resolution conflict forecasting with spatial convolutions and long short-term memory. *International Interactions*, 48(4), 739–758. doi:10.1080/03050629.2022.2031182.
- Ranking All Countries - Early Warning Project** (n.d.). Retrieved February 23, 2024, from <https://earlywarningproject.ushmm.org/ranking-of-all-countries>
- Reeder, BW** (2018) The Political Geography of Rebellion: Using Event Data to Identify Insurgent Territory, Preferences, and Relocation Patterns. *International Studies Quarterly*, 62(3), 696–707. doi:10.1093/isq/sqy016.
- Schellens, MK, and Belyazid, S** (2020) Revisiting the Contested Role of Natural Resources in Violent Conflict Risk through Machine Learning. *Sustainability*, 12(16). doi:10.3390/su12166574.
- Straus, S** (2016) *Fundamentals of genocide and mass atrocity prevention*, Washington, DC: United States Holocaust Memorial Museum.
- Tandon, A** (2021) Analysis: The lack of diversity in climate science research. Retrieved January 20, 2023, from <https://www.carbonbrief.org/analysis-the-lack-of-diversity-in-climate-science-research/>
- The Early Warning Project** (2024) *Early Warning Project Statistical Risk Assessment 2023-24*. Retrieved from <https://earlywarningproject.ushmm.org/storage/resources/2872/Early%20Warning%20Project%20Statistical%20Risk%20Assessment%202023-24.pdf>
- Tollefsen, AF, and Buhaug, H** (2015) Insurgency and Inaccessibility. *International Studies Review*, 17(1), 6.
- United Nations Office on Genocide Prevention and the Responsibility to Protect** (2014) *Framework of Analysis for Atrocity Crimes - A tool for prevention*, New York, USA. Retrieved from <https://www.ohchr.org/sites/default/files/2021-11/Genocide-Framework-of-Analysis-English.pdf>
- United States Holocaust Memorial Museum** (2021) *The Future of Mass Atrocities and Atrocity Prevention*. Retrieved from

[https://www.ushmm.org/m/pdfs/The\\_Future\\_of\\_Mass\\_Atrocities\\_and\\_Atrocidity\\_Prevention.docx.pdf](https://www.ushmm.org/m/pdfs/The_Future_of_Mass_Atrocities_and_Atrocidity_Prevention.docx.pdf)

- Uzonyi, G, and Demir, B** (2020) Excluded Ethnic Groups, Conflict Contagion, and the Onset of Genocide and Politicide during Civil War. *International Studies Quarterly*, 64(4), 857–866. doi:10.1093/isq/sqaa059.
- van Weezel, S** (2020) Local warming and violent armed conflict in Africa. *World Development*, 126, 1.
- von Uexkull, N, Croicu, M, Fjelde, H, and Buhaug, H** (2016) Civil conflict sensitivity to growing-season drought. *Proceedings of the National Academy of Sciences of the United States of America*, 113(44), 12391–12396. doi:10.1073/pnas.1607542113.
- Warren, T, and Troy, K** (2015) Explaining violent intra-ethnic conflict: group fragmentation in the shadow of state power. *Journal of Conflict Resolution*, 59(3), 484–509.
- Wayman, F, and Tago, A** (2010) Explaining the onset of mass killing, 1949-87. *Journal of Peace Research*, 47(1), 3–13.
- Wegenast, T, and Basedau, M** (2014) Ethnic fractionalization, natural resources and armed conflict. *Conflict Management and Peace Science*, 31(4), 432–457.
- Weidmann, NB, and Salehyan, I** (2013) Violence and Ethnic Segregation: A Computational Model Applied to Baghdad. *International Studies Quarterly*, 57(1), 52–64.
- Wimmer, A, Min, B, and Cederman, LE** (2009) Ethnic Politics and Armed Conflict. A Configurational Analysis of a New Global Dataset. *American Sociological Review*, 74(2), 316–337.
- Witmer, FDW, Linke, AM, O’Loughlin, J, Gettelman, A, and Laing, A** (2017) Subnational violent conflict forecasts for sub-Saharan Africa, 2015-65, using climate-sensitive models. *Journal of Peace Research*, 54(2), 175–192. doi:10.1177/0022343316682064.

# Appendix I. Search terms used for scoping review

## Web of Science

(TS=(“War crime” OR “Ethnic violence” OR “ethnic cleansing” OR Warfare OR “Armed conflict” OR battle OR “crimes against humanity” OR Genocide OR “Mass killing” OR “mass murder” OR “mass execution” OR carnage OR slaughter OR atrocity OR atrocities OR Pogrom OR Massacre OR “Arms attack” OR “systematic violence” OR “systemic violence”) AND TS=(Predict OR Forecast OR Estimate OR estimation OR “early warning” OR “early action” decision-making OR decision-support OR Warn OR Prevent OR Mitigate OR mitigation OR Intervene OR intervention OR Model)) AND (ALL=(“Spatial analysis” OR GIS OR “Geographic information system” OR “Events data” OR analys\* OR Mapping OR “Remote Sensor” OR sensing OR “Natural Learning Processing” OR NLP OR Drone OR UAVs OR “Machine Learning” OR ML OR “Artificial Intelligence” OR AI OR Computer) AND PY==(“2022” OR “2021” OR “2020” OR “2019” OR “2018” OR “2017” OR “2016” OR “2015” OR “2014” OR “1995” OR “1996” OR “1997” OR “1998” OR “1999” OR “2000” OR “2001” OR “2002” OR “2003” OR “2004” OR “2005” OR “2006” OR “2007” OR “2008” OR “2009” OR “2010” OR “2011” OR “2012” OR “2013”)) AND DT==(“ARTICLE”) AND TASCA==(“POLITICAL SCIENCE” OR “INTERNATIONAL RELATIONS” OR “PUBLIC ENVIRONMENTAL OCCUPATIONAL HEALTH” OR “MULTIDISCIPLINARY SCIENCES” OR “ENVIRONMENTAL SCIENCES” OR “ENGINEERING ELECTRICAL ELECTRONIC” OR “PSYCHOLOGY SOCIAL” OR “COMPUTER SCIENCE SOFTWARE ENGINEERING” OR “INSTRUMENTS INSTRUMENTATION” OR “ETHICS”))

## IEEE Xplore Digital Library

Filters; Journal articles

1995-2022 results

Predict OR Forecast OR Estimate OR estimation OR “early warning” OR “early action” decision-making OR decision-support OR Warn OR Prevent OR Mitigate OR mitigation OR Intervene OR intervention OR Model

Search within results

“War crime” OR “Ethnic violence” OR “ethnic cleansing” OR Warfare OR “Armed conflict” OR battle OR “crimes against humanity” OR Genocide OR “Mass killing” OR “mass murder” OR “mass execution” OR carnage OR slaughter OR atrocity OR atrocities OR Pogrom OR Massacre OR “Arms attack” OR “systematic violence” OR “systemic violence”

Search within results

“Spatial analysis” OR GIS OR “Geographic information system” OR “Events data” OR analys\* OR Mapping OR “Remote Sensor” OR sensing OR “Natural Learning Processing” OR NLP OR Drone OR UAVs OR “Machine Learning” OR ML OR “Artificial Intelligence” OR AI OR Computer

## ProQuest

Searched for:

ab("War crime" OR "Ethnic violence" OR "ethnic cleansing" OR Warfare OR "Armed conflict" OR battle OR "crimes against humanity" OR Genocide OR "Mass killing" OR "mass murder" OR "mass execution" OR carnage OR slaughter OR atrocity OR atrocities OR Pogrom OR Massacre OR "Arms attack" OR "systematic violence" OR "systemic violence") AND ab(Predict OR Forecast OR Estimate OR estimation OR "early warning" OR "early action" decision-making OR decision-support OR Warn OR Prevent OR Mitigate OR mitigation OR Intervene OR intervention OR Model) AND ab("Spatial analysis" OR GIS OR "Geographic information system" OR "Events data" OR analys\* OR Mapping OR "Remote Sensor" OR sensing OR "Natural Learning Processing" OR NLP OR Drone OR UAVs OR "Machine Learning" OR ML OR "Artificial Intelligence" OR AI OR Computer)

Databases:

6 databases searched

- Applied Social Sciences Index & Abstracts (ASSIA)
- International Bibliography of the Social Sciences (IBSS)
- Policy File Index
- Political Science Database
- Sociological Abstracts
- Worldwide Political Science Abstracts

Filter 1995-2022

## **Appendix II. Summary of reviewed literature**

Author and Date	Geographic Focus (continent, scale)	Objective (as paraphrased from authors)	Objective of study	Independent variables	Data Sources Used	Study Type	Method(s) used	Results (as paraphrased from authors)
Aas Rustad and Binningsbo 2012	Global; event-specific	To examine the relationship between natural resources and post-conflict peace.	The objective is to examine the relationship between natural resources and post-conflict peace.	Climate & Natural Environment	Authors' own created datasets	Conflict prediction+ environment	Piecewise exponential survival model	The authors find that natural resource conflicts in general have a significant effect on peace duration, indicating that peace after natural resource conflicts fails faster than after other conflicts - but the effect becomes insignificant when adding control variables.
Aas Rustad et al 2011	Asia; country-wide and sub-national	Develop a sub-national conflict risk model considering triggers to predict violence	Presentation of a new methodology for estimating sub-national conflict risk. The authors develop a unified conflict prediction model that combines a quantitative assessment of conflict risk at the country level with a series of country-specific sub-national analysis at first-order administrative regions to identify potential triggers (in particular irregular transfer of state power) that might predict the outbreak of violence.	Climate & Natural Environment, Conflict, Ethnic, Geography & Administrative, Socio-demographic/socio-political/socio-economic	Scalar Index of Polities (SIP) dataset, Ethnic Power Relations (EPR) database, CIESIN's Gridded Population of the World, Human Development Reports, Geo-Referencing of Ethnic Groups Data	Conflict prediction	Development of risk index, spatiotemporal modeling, global logit regression model	The authors were able to create a model to predict the onset of (civil) conflict on a subnational level, capturing spatial variability of intrastate conflict and its triggers.

Anderton and Carter 2015	Global; country-wide	Develop an optimization model to understand the risk factors of genocide onset	Develop a constrained mass atrocity optimization model to guide empirical inquiry into risk factors of genocide. The study identifies under which conditions killing civilians would be an "optimal" choice by a regime.	Conflict, Economic, Geography & Administrative, Governance	Political Instability Task Force (PITF), Polity IV Project, PWT 7.1, Minorities at Risk (MAR) Dataset	Genocide prediction	Constrained rational choice model, pooled logit polynomial, random effects logit, rare events logit, Cox proportional hazards	The analysis revealed six significant risk factors for the onset of genocide which include threat posed by rebels against a state, anocracy, new state, low income per capita, discrimination, and Cold War conditions.
Anderton and Ryan 2016	Africa; country-wide	Understand the concept of 'habituation to atrocity' by studying how previous choices to use low level violence on civilians affects an actor's choices to engage in high-level civilian attacks	This study analyses the concept of 'habituation to atrocity', which is characterized as an actor's increased willingness to carry out high-level civilian attacks owing to earlier choices of low-level civilian attacks using a rational choice model and guided an empirical inquiry into the concept of Violence Against Civilians (VAC) studying 49 African countries.	Climate & Natural Environment, Conflict, Ethnic, Geography & Administrative, Governance, Other	ACLED, Polity IV project, UCDP/PRIO Armed Conflict dataset, Stockholm Peace Research Institute, World Bank Data	Actors, Conflict intensity, Conflict prediction	Rational choice model & regressions	The findings suggest that previous periods of low-level violence against civilians increased the number of high-level attacks. Further, the severity of previous low-level violence also has an impact on the number of high-level attacks.
Bartusevičius and Gleditsch 2019	Global; country-wide	To define and demonstrate the value of a two-staged conflict estimation analysis	The objective of the study is to demonstrate how the two conflict stages (incompatibilities - violent conflict) can be conceptually and empirically defined as they relate to the study of civil conflict.	Geography & Administrative, Governance, Secondary, Socio-demographic/socio-political/socio-economic	Conflict Information and Analysis System (CONIAS) data	Conflict prediction	Logit regression	The analysis showed that there is a relationship between horizontal inequalities and civil conflict, but that it primarily is associated with conflict origination. The analysis further found that grievances do not independently facilitate violence, refugee flows can contribute to violence spilling over into

								neighboring countries, and that the influence of media density on decreasing violence exerts itself through militarization.
Beardsley 2011	Global; country-wide	To understand if and how peacekeeping affects the geographic spread of conflict across borders	This article aims to understand if peacekeeping can affect conflict contagion	Conflict, Military, Other	UCDP/PRIO Armed Conflict dataset, International Military Intervention Data, Salehyan 2007 (refugee flows and transnational insurgencies)	Actors, Conflict prediction, Conflict prevention	Event history analysis including spatial dependence (binary time-series cross-sectional data), logistic estimation	The presence of peacekeepers does play a critical role in the occurrence of neighboring conflict. Namely, the risk of conflict increases if a neighboring state experienced conflict without peacekeeping intervention, while the risk of armed conflict does not significantly change when there is a neighboring conflict with peacekeeping.
Besancon 2005	Global; event-specific	To expand the economic inequality-political conflict nexus and test the effects of inequality on intrastate conflicts	The models examine the effects of inequality variables, institutional variables, threshold variables, and a greed variable on the probability of deaths caused by ethnic violence, revolutions, and genocide	Economic, Geography & Administrative, Socio-demographic/socio-political/socio-economic	State Failures Task Force, Gini Index, Human Capital Gini, World Bank Gini	Conflict prediction, Ethnic conflict, Genocide prediction	Ordinal probit model	Economic inequality differently impacts ethnic conflicts, revolutions, and genocides. The results show that economic inequality grievances have an adverse effect in revolutions and that revolutions are more likely in economically unequal societies. In contrast, the effect of economic inequality on ethnic conflict is the opposite and much weaker, meaning higher levels of conflict occur where there is a greater degree of economic equality. Economic inequality appears to have the most ambiguous



								effect on genocide occurrence.
Blair and Sambanis 2020	Global; country-wide	To propose a parsimonious model using event data that incorporates conflict escalation theories to predict civil war onset with higher accuracy and in a shorter timeframe than before.	We use event data to show that a parsimonious model grounded in prominent theories of conflict escalation can forecast civil war onset with high accuracy and over shorter temporal windows than has generally been possible.	Conflict, Economic, Socio-demographic/socio-political/socio-economic, Geography & Administrative	Integrated Crisis Early Warning System (ICEWS) data, Sambanis 2004, Conflict and Mediation Event Observations (CAMEO) codebook	Conflict intensity, Conflict prediction	random forests	The developed procedural model led to accurate conflict predictions within shorter temporal windows in comparison to earlier modeling attempts and that there is a direct connection between theory and active conflict forecasting.
Buhaug et al 2011	Global; sub-national	To understand the relationship between the location of conflict outbreaks and subnational income variation	This article investigates the specific location where conflicts break out and subnational income variation, based on Geographic Information Systems (GIS) data.	Conflict, Economic	PRIO-GRID, G-Econ data set Nordhaus (2006), UCDP/PRIO Armed Conflict dataset	Conflict prediction	GIS analysis, case-control logit design	The results support the hypothesis that location of conflict onset is related to geographical variation in income. The authors find that conflicts are more likely to break out in locations with low income, meaning local income matters more than national income. We also find some evidence that regional inequality makes conflict more likely and that pockets of wealth in very poor states are especially likely to see civil conflict break out.

Bulutgil 2015	Europe; event-specific	To study the factors most relevant to the onset of ethnic cleansing as a distinct phenomenon, using data from 20th century Europe.	This article contributes to the literature on mass violence in general and ethnic cleansing utilizing a definition of ethnic cleansing that is analytically and empirically distinct from related phenomena such as mass killings and/or civilian victimization.	Ethnic, Governance, Military, Socio-demographic/socio-political/socio-economic	Encyclopedia Britannica, other articles and books	mass atrocities (ethnic cleansing)	Logit analysis	The structural conditions that influence ethnic cleansing include more variables than solely ethnicity. Particularly significant are variables that relate to competition along socio-economic divisions such as socio-economic inequality and political competition.
Cederman et al 2009	Global; sub-national	To propose a GIS-based approach as a way to model ethnic center-periphery dyads that confront governments that exclude certain ethno-nationalist groups.	Going beyond an exclusive focus on ethnic minorities, this study employs Geographic Information Systems (GIS) as a way to model ethnic center-periphery dyads that confront governments with excluded groups	Economic, Geography & Administrative, Socio-demographic/socio-economic, Ethnic, Conflict	Atlas Narodov Mira (ANM), Columbia University, UNEP-WCMC 2002, UCDP/PRIO Armed Conflict dataset	Conflict prediction, Ethnicity	GIS, robust and multinomial logit regression	The probability that excluded ethnic groups experience conflict increases relatively compared to the governmental group. When considering ethnic territorial conflict, geographic distance, and the roughness of terrain also play an important role.
Chiba and Gleditsch 2017	Global; country-wide	To examine if additional information on variables considering "plausible motivation and grievances" can improve civil war forecasting models.	The purpose is to examine if dynamic information on mobilization and the behavior of actors from event data can help improve on a model attempting to forecast civil war using measures reflecting plausible motivation and grievances.	Ethnic, Governance, Socio-demographic/socio-political/socio-economic, Conflict	Integrated Conflict Early Warning System (ICEWS), Conflict and Mediation Event Observations (CAMEO), Gleditsch & Ward (1999) list of independent states, Buhaug, Cederman and Gleditsch ('BCG') model	Conflict prediction	Events data analysis, logit models	The authors find that some conflicts are easier to predict than others but that the inclusion of dynamic behavioral information can help in forecasting conflict. Models that include grievance measures tend to perform better than purely event-based models.

Colaresi and Carey 2008	Global; event-specific	To assess under what conditions government security forces actions, result in genocide instead of protecting a population and enforcing national stability.	The objective of the article is to address the question of under what conditions government security force strength results in genocide rather than in increased national stability.	Conflict, Military, Governance, Ethnic	Harff (2003), Correlates of War project, National Material Capabilities v3.02 (Singer 1987), Military Balance series of publications (Military Balance 1963-2004), World Development Indicators (World Bank 2007), Polity IV data set (Marshall and Jaggers 2002).	Genocide prediction	Cross-tabulations and rare event logit model	The impact of security forces on genocide is dependent on executive constraints in a state. Factors related to the dynamics of state failure, democracy, and losing a war significantly affect the inclination of a state to initiate genocide.
Detges 2016	Africa; sub-national	To examine the relationship between drought exposure, limited access to water and transport infrastructure, and high conflict risks are associated.	To assess how far violent and nonviolent outcomes in the wake of drought can be accounted for by regional differences in the provision of key infrastructures that help coping with drought and preventing violence.	Climate & Natural Environment, Geography & Administrative	Global Administrative Unit Layers (GAUL: EC-FAO, 2008), UCDP-GED dataset, standardized precipitation index (SPI), Global Precipitation Climatology Centre, Africa Infrastructure Country Diagnostic (AICD: African Development Bank Group, 2015), Demographic and Health Surveys (DHS: ICF Macro, 2014).	Conflict prediction + environment, Conflict prevention	Logistic regression, GIS analysis	There are important distinctions between types of conflict and its relationship to drought and infrastructure. Overall, however, the calculations suggest that the effect of drought in predicting conflict events is modest.

Dorff et al 2020	Africa; Nigeria; sub-national	To explain the occurrence of violence between armed groups using a network-based approach to understand the inter-dependencies among the actors in in civil conflicts.	The study aims to explain the occurrence of violence between armed groups over time through a network-based approach using the case of Nigeria. The goal is to understand the inter-dependencies among the actors in the network in civil conflicts.	Climate & Natural Environment, Conflict, Governance	ACLED	Actors, Conflict prediction	Network analysis - social relations regression model, latent factor model, additive, and multiplicative effects	In the case of Nigeria, the used network approach was better able to predict which actors would engage in violence at different points in time compared to existing out-of-sample approaches. Significant study results include: 1) Boko Haram has increased violence across Nigeria, whether or not Boko Haram is directly involved, 2) ethnic militias in central Nigeria are generally more violent than elsewhere, 3) armed groups that are active across the entire country are more likely to come into direct conflict with other groups.
Duursma and Read 2017	Africa; Sudan; Darfur	To explore the possibilities of modelling violence using epidemiological methodologies	The objective of this study is to explore the possibilities of modelling violence as analogous to disease, drawing on epidemiological tools	Geography & Administrative, Conflict	UNAMID's Joint Mission Analysis Centres data, Office for the Coordination of Humanitarian Affairs (OCHA) shapefile of Sudan, OCHA on localities in Darfur	Conflict prediction	Spatial analysis of events data along with logit estimates	The analysis finds that predicted conflict events are clustered in localities that experienced prior conflict or are neighbored by localities that experience armed conflict. These findings support the hypothesis that armed violence is spatially inter-dependent.
Eck 2009	Global; event-specific	To address whether mobilization along ethnic lines will lead to intensified conflict compared to conflicts in which there was no mobilization along ethnic lines	The aim of this study is thus to answer the question, given the existence of an ongoing intrastate armed conflict, does ethnic mobilization increase the risk of the conflict intensifying to war?	Ethnic	UCDP/PRIO Armed Conflict dataset	Conflict intensity, Conflict prediction, Ethnicity	Cox proportional hazard regression model	The study finds that conflicts in which rebels mobilize along ethnic lines have a 92% higher chance of intensifying to war

Ember et al 2019	Africa; event-specific	To test whether increased political participation lowers the likelihood of committing atrocities during internal warfare.	The purpose of this paper is to examine when political participation would have similar effects in eastern Africa, and whether more participatory polities commit fewer atrocities against each other.	Climate & Natural Environment, Governance	HRAF Collection of Ethnography (paper version, and online eHRAF World Cultures), Ross (1983), Tuden and Marshall 1980, Murdock et. al. (1980), Ember et. al. (1992, 2013)	Actors, Conflict prevention	Regression, ethnography	In the case of East Africa, the model suggests that greater local political participation predicts less internal warfare and that participatory polities were less likely to commit atrocities during internal warfare.
Esteban et al 2015	Global; group-specific	To present an empirical analysis of how natural resources and other variables may predict mass killings.	This study introduces a model of conflict identifying whether decisions to engage in mass killings can be explained as the result of strategic, rational calculation.	Climate & Natural Environment, Conflict, Economic, Ethnic, Geography & Administrative, Socio-demographic/socio-political/socio-economic	Political Instability Task Force, Correlates of War (CoW), Geo-Referencing of Ethnic Groups (GREG) data set, Minorities at Risk 2009 (MAR) project	mass atrocities (mass killing)	Regression analysis	The empirical analysis confirms that natural resource rents and recent democratization in ethnically polarized nations are a significant predictor of mass killings while on the contrary, high labor productivity is associated with fewer massacres.
Ettensperger 2021	Africa; country-wide	To predict changes of conflict intensity in African countries for 6 months into the future using multi-model ensemble learning techniques.	The goal is to predict changes of conflict intensity in fifty-four African countries for 6 months into the future.	Conflict, Economic, Governance, Infrastructure, Other, Socio-demographic/socio-political/socio-economic	UCDP/PRIO Armed Conflict dataset, REIGN (Bell 2016), WDI (World Bank 2019), V-DEM (V-Dem Institute 2020), and IMF – World Economic Outlook (IMF 2020) data	Conflict intensity, Conflict prediction	Tree-based learning algorithms: a Random Forest (RF) Regressor and the Gradient Tree Boosting model (XGB).	The study found that it may be best to increase training intensity and forest size of the merged dataset. The conflict prediction techniques were tested as case studies in Egypt, Cameroon, and Mozambique. In Egypt, state-based conflict was expected to moderately increase in the forecasted window. Similarly, Cameroon was also predicted as having an increase in conflict and conflict intensity, while Mozambique was predicted to have decreased intensity.

Fox 2000	Global; group-specific	To assess why majority groups, engage in the discriminatory behavior and examine religious causes of ethnic discrimination.	The object of this study is twofold. First, it assesses why majority groups engage in the discriminatory behavior that is believed to be one of major causes of ethnic protest and rebellion. Second, it focuses on the religious causes of ethnic discrimination. However, despite the focus on religion, this study does examine other factors believed to influence the level of ethnic discrimination, and accordingly has wider implications.	Economic, Governance, Military, Socio-demographic/socio-political/socio-economic	Minorities at Risk Phase 3 Dataset	Conflict prediction, Ethnic conflict	Multiple regressions and correlation analyses	Religion and religious variables are significantly associated with multiple forms of discriminations, but the process that results in religious discrimination is typically different compared to other types of discrimination. These general findings among others suggest that religious factors have a distinct influence on conflict.
Fujita et al 2017	Asia; Europe; sub-national	To examine the spatial and temporal correlation of daily deaths and use this information to forecast future deaths.	This article uses information on the daily number of deaths to study temporal and spatial correlations in the data and exploit this information to forecast events of deaths.	Conflict	Violations Documentation Center (VDC) in Syria, UK office for National Statistics	Conflict prediction, Probability of death	Correlations analysis, (vector) auto-regression, and Granger causality analysis	In the comparative analysis of Syria and England, the authors suggest that days with a high number of deaths is typically followed by more conflict and death in a short temporal window. The results further suggests that there is also spatial correlation between cities within Syria.
Goldsmith et al 2013	Global; country-wide	To develop a prototype model to better forecast genocide and politicide 1 year into the future.	In this article we present a forecasting model for genocide and politicide onset which, we believe, improves on existing models in terms of	Conflict, Economic, Ethnic, Governance, Military	Political Instability Task Force (PITF), National Elections across Democracy and Autocracy (NELDA) data,	Genocide prediction, Politicide prediction	Two-stage probit regression models	The model accurately predicted several genocides. The authors found that several of the most significant predictors of genocide and politicide as predicted in their study

			method, scope of the data, and variables used. Specifically, we (1) employ a two-stage approach, (2) use a global dataset including all country-years to forecast one year into the future, and (3) incorporate time-variant factors		Major Episodes of Political Violence data, Ethnic Power Relations (EPR) dataset, Minorities at Risk data, COW IGO Dataset			were state-led discrimination, previous genocide, absence of elections, years passed since previous genocide, ethnic power relations, and more.
Harff 2003	Global; country-wide	To test the effects of a set of variables capturing conflict and political characteristics on the risk of genocide and politicide that began between 1955 and 1997.	The objective of this study was to test the effects of prior conflict, elite characteristics, regime type, and international context on the likelihood of geno-/politicide.	Conflict, Ethnic, Governance, Socio-demographic/socio-political/socio-economic	Genocides and politicides list by the State Failure Task Force	Genocide prediction, Politicide prediction	Case-control procedure and logistic regression	Variables capturing prior genocides, past political upheavals, characteristics of the ruling party, and trade openness were all significant in predicating the likelihood that state failures will lead to geno-/politicide. The analysis resulted in 74% accuracy in distinguishing internal wars and regime collapses that do and do not lead to geno-/politicide.
Hegre et al 2013	Global; country-wide	To predict the conflict state for all countries in the world for the 2010-2050 period	Our goal is to predict the conflict state for all countries in the world.	Conflict, Geography & Administrative, Socio-demographic/socio-political/socio-economic	UCDP GED, UN projections for demographic variables and IIASA projections for education, Fearon and Laitin (2003) for oil dependency, Collier and Hoeffler (2004) for ethnic dominance	Conflict intensity, Conflict prediction	Dynamic multinomial logit model estimation	The authors predict that the global incidence of conflict is likely to decrease from current levels and be reduced to about half of the present number of conflict countries in 2050 (proportionally). Conflict will generally decline in Western Asia and North Africa while conflict will remain in East, Centra, and Southern Africa and East and Southeast Asia.

Hegre et al 2019	Africa; country-wide and sub-national	To develop a system that systematically monitors all locations at risk of conflict and assess the probability of conflict onset, escalation, continuation, and spatial spread.	The objective of VIEWS is to develop a system that systematically monitors all locations at risk of conflict and assesses the probability of conflict onset, escalation, continuation, and geographic diffusion.	Climate & Natural Environment, Conflict, Economic, Geography & Administrative, Socio-demographic/socio-political/socio-economic	UCDP-GED, UCDP-Candidate, PRIO-GRID, World Development Indicators, political excluded ethnic groups, demographic factors, ACLED (protests), V-DEM for data on institutions	Conflict intensity, Conflict prediction	Logistic regression, random forest algorithm, temporal and spatial event data, dynamic simulation model, 'one-step-ahead' modeling	The forecast for conflict between October 2018 and October 2021 is done on a localized basis, identifying areas of continued violence (Burundi, Nigeria, DRC), new state-based conflict (Cameroon and Mozambique). The study finds that the strongest predictor of violence is prior violence but that variables related to natural and social geographics are also important predictors.
Helman et al 2020	Africa; Middle East; sub-national	To understand the relationship between climate and conflict and to test the generalizability of climate-conflict relationships from national to continental scale.	The objective is to test the generalizability of climate-conflict relationships from national to continental scale. It also assesses both the direct effects of climate on violence and indirect effects of climate variability on other variables known to be relevant for conflict research	Climate & Natural Environment, Geography & Administrative, Socio-demographic/socio-political/socio-economic	UCDP/PRIO Armed Conflict dataset, Climate Hazards center InfraRed Temperature with Stations [CHIRTS] dataset, Climate Hazards center Infrared Precipitation with Stations [CHIRPS], Global Subnational Infant Mortality Rates, Version 1 [GSIMR.v1], Climate Change Initiative [CCI] of the European Space Agency [ESA] Land Cover product, ASA's VIPPHEN EVI2 satellite product, NASA Land Information System [LIS]	Conflict prediction + environment	Structural equation modeling	The results suggest that there are numerous complex mechanisms to explain the empirical patterns that underlie climate-conflict relationships across countries and regions. When the model was applied to Africa and the Middle East simultaneously versus individually, the role of temperatures, rainfall, yield, and welfare changed.



Hillebrecht 2016	Africa; Libya; country-wide	To test assumptions about the relationship between international criminal justice and violence by examining the conditions under which the ICC had a deterrent effect on the commission of civilian casualties during 2011 in Libya	This article seeks to empirically test assumptions about the relationship between international criminal justice and violence by evaluating the conditions under which the ICC had a deterrent effect on the commission of civilian casualties during the 2011 crisis in Libya	Conflict, Military, Other	ACLED, CAMEO coding scheme	Actors, Conflict prevention, Probability of death	Regression and event count models	The results suggest that the ICC can deter violence against civilians, particularly during conflict where the government of government-sponsored civilian deaths. The intent to prosecute crimes is also a crime deterrent.
Karreth and Tir 2013	Global; event-specific	To determine to what extent a state's memberships in HSIGOs (highly structured international governmental organizations) reduces the likelihood of conflict in that state's territory	The objective of the study is to determine to what extent a state's memberships in HSIGOs (highly structured international governmental organizations) reduces the likelihood of domestic conflict in that state's territory	Other	Cunningham, Gleditsch, Saleyhan 2009 based on UCDP/PRIO dataset; Boehmer, Gartzke, and Nordstrom (2004)	Conflict intensity, Conflict prevention	Logit models and probit selection models	Membership in HSIGOs reduces the likelihood of low-level conflict to turn into civil war while there was no correlation between conflict escalation and countries' membership in low and medium-structure IGOs.
Kathman and Wood 2011	Global; event-specific	To explore how third-party interventions affect the severity of mass killings.	The objective of the study is to explore the dynamics of third-party interventions in the severity of mass killings.	Conflict, Economic, Governance, Military	Political Instability Task Force (PITF), International Military Interventions (IMI) dataset, National Material Capabilities dataset from the Correlates of War (COW), Polity IV data, ACLED	Actors, Conflict prevention, mass atrocities (mass killing)	Ordered logit model	Short-term intervention appears to worsen genocidal violence while impartial intervention diminishes violence over time. The results also reveal that pro-government interventions can exacerbate hostilities over time.

Crain 1997	Global; country-wide	To test how openings in political structures may predict the onset and severity of genocides and politicides and understand what factors contribute to state-sponsored mass murder.	This study tests how the occurrence of openings in the political opportunity structure predicts onset and differing degrees of severity of genocides and politicides and tests other hypotheses regarding what factors affect state-sponsored mass murder.	Conflict, Ethnic, Governance	Harff and Gurr's (1988) data set, Correlates of War (COW) data set, Bank's (1971) cross-polity data, Polity II data set (Gurr, Jagers and Moore 1989), IMF's Direction of Trade Yearbook	Genocide prediction, Genocide severity, Politicide prediction	Logit models	Civil wars are the most constant predictor of genocide and politicide but openings in political opportunity such as extraconstitutional changes also play an important role in determining the severity of state-sponsored mass murder.
Crain 2005	Global; event-specific	To examine the effectiveness of international intervention to genocide and politicide.	The objective is to better understand the effectiveness of potential responses by the international community to genocides and politicides.	Military	State Failure Task Force, IMI data set by Pearson and Baumann (1993) data set	Actors, Conflict prevention, Genocide prediction, Politicide prediction	Ordered logit models	Interventions that challenge the perpetrator(s) or aid the target group are the only effective type of military response in decreasing the severity of genocide and politicide. Other models of intervention are found to be ineffective.
Crain 2012	Global; event-specific	To examine the effects of "naming and shaming" perpetrators in reducing the severity of ongoing genocide or politicide.	This study examines the effects of network-based naming and shaming approach whether it could reduce the severity of ongoing instances of genocide or politicide.	Conflict	Political Instability Task Force (PITF) Worldwide Atrocities dataset, Amnesty International Cumulative Guide 1962–2000.	Actors, Genocide severity, Politicide severity	Ordered logit models	Overall, the naming and shaming of perpetrators plays an important role in reducing the severity of genocide and politicide. The study finds that as northern media coverage increases, the severity of instances of genocide/politicide decreases and similarly, when UNCHR targets the state committing atrocities, the severity often decreases.
Crain 2017	Global; event-specific	To assess how economic sanctions affect the severity of ongoing instances of	This study examines the effect of economic sanctions on the severity of ongoing instances of genocide or politicide.	Economic	Political Instability Task Force (PITF) for dependent variable, Threat, and Imposition of Economic	Genocide severity, Politicide severity	Ordered logit models	Sanctions typically do not aggravate nor alleviate atrocities in all tested circumstances, contrary to popular belief.

		genocide or politicide.			Sanctions (TIES) dataset			
Marx et al 2019	Asia; Myanmar; sub-national	To demonstrate the utility of very high-resolution satellite imagery in detecting and documenting of human rights violations	The objective of the study was to demonstrate the utilization of small sat imagery for the detection of human rights violations using Myanmar as an example	Other	Human Rights Watch, PlanetScope constellation of smallsats, Google Earth	Conflict monitoring, Conflict prediction	Near infrared value extraction from high-resolution satellite imagery	In the case of Myanmar, satellite imagery analysis resulted in the high accuracy of detecting burning villages in a tight-temporal timeframe. This timely analysis can be used to document committed atrocities.
McDoom 2014	Africa; Rwanda; sub-national	To develop a model identifying areas susceptible and resistant to violence during genocide.	The study presents a theoretical model to help identify areas susceptible and resistant to violence during genocide.	Conflict, Ethnic, Governance	Author-created	Genocide prediction	Duration analysis logit model and within-case comparative analysis of two communes	Overall, communes that experienced more higher levels of segregation, experienced previous and/or neighboring violence, or were ruled by elite as opposed to opposition political parties were more likely to experience violence than communes with the opposite conditions.
Miodownik and Bhavani 2011	Global; group-specific	To identify the conditions under which the marginalization of the ethnic majority from state government may lead to conflict.	This article looks at the conditions under which the marginalization of the ethnic majority from state power triggers conflict.	Climate & Natural Environment, Economic, Ethnic	Wimmer, Cederman, and Min (2009)	Conflict prediction, Ethnicity	Agent-based modeling (ABM)	When considering ethnic minority rule, salience of ethnicity, fiscal policy, and natural resources all affect conflict levels. Namely, conflict increases when: there is a dependence on lootable natural resources, presence of poor fiscal policies, and increased ethnic saliences for all individuals.

Mueller and Rauh 2018	Global; sub-national	To provide a new methodology to predict in-country and between-country armed conflict by using newspaper text.	The purpose of this paper is to add to the forecasting literature in two ways. First, the paper explicitly separates the within-country from the between-country variation before forecasting armed conflict. Second, it provides a new methodology to predict armed conflict by using newspaper text using country fixed effects	Other	New York Times, the Washington Post, and the Economist, UCDP/PRIO Armed Conflict dataset	Conflict prediction	Latent Dirichlet allocation (LDA), panel regression, linear fixed-effects regression	Generally speaking, new reports can be used to create timely predictions of conflict onset both within and between countries. The model accurately predicts the timing of conflict 80% of the time.
Musumba et al 2021	Africa; sub-national	To compare two algorithms in accurately predicting civil conflict in sub-Saharan Africa.	The goal of this study is to compare the logistic regression algorithm to machine learning algorithms in order to predict civil conflict in sub-Saharan Africa.	Conflict, Economic, Geography & Administrative, Socio-demographic/socio-political/socio-economic	ACLED, Spatial Production and Allocation Model (SPAM), 2005 version 1.0	Conflict prediction	Logistic regression, random forests, gradient boosting, multilayer perception, support vector machines, machine learning	The authors find that machine learning algorithms offer the best approach to predicting conflict, given the availability of input data.
Nyseth Brehm 2017	Africa; Rwanda; sub-national	To assess the subnational factors of local violence during the 1994 Rwandan genocide.	The objective of this article is to assess the potential subnational determinants of localized violence during the 1994 Rwandan genocide.	Geography & Administrative, Military, Other	Rwandan Ministry of Local Administration and Community Development, the National University of Rwanda, 1991 Rwandan census, Guichaoua (2010)	Conflict prediction, Genocide severity	Fixed effects regression	A combination of top-down and bottom-up factors influenced local violence during the genocide. Top-down conflict triggers include state-sponsored violence, proximity to the country's capital, presence of Rwandan Armed Forces while bottom-up violence included civilian-driven violence.

O'Loughlin et al 2012	Africa; sub-national	To study the local and temporal differences in climate and conflict in East Africa to further understand the climate-conflict nexus.	This study extends the quantitative approach to the climate-conflict nexus with close attention to local and temporal differences in climate and conflict examining nine countries in East Africa	Climate & Natural Environment	ACLED, SPI6 Standard Precipitation Index, T16 Temperature index	Conflict prediction + environment	Generalized linear modeling, negative binomial modeling, spatial analysis	On their own, the effects of temperature (high temperatures raise risk of violence) and precipitation (high precipitation decreases risk of violence) are significant but when combined with other variables capturing politics and economics, their significance decreases, but remains significant.
Pilster et al 2016	Global; event-specific	To model how a states' military capability and structure of security forces may lead to the onset of genocide.	The article develops a theoretical model looking at states' military capabilities for conducting mass-killings and the structure of their security forces in driving the onset of genocidal violence.	Military	Colaresi and Carey's version of Harff's data on the onset of genocidal violence in state-failure years from 1955 to 2003, Pilster and Bohmelt	Actors, Genocide prediction	Multivariate probit regression models	As the diverse the structure of state security forces increases, the risk of the onset of genocidal violence decreases.
Reeder 2018	Africa; sub-national	To spatially predict unreported conflict based on the location of known rebel activity to study the spatial and temporal patterns of conflict events	The objective of the study is to capture the characteristics of territory occupied by rebels, and how the availability of densely forested terrain influences spatial and temporal patterns of violence in localities proximate to where rebels are located.	Climate & Natural Environment, Geography & Administrative, Other	ACLED, No Peace Without Justice (NPWJ) data	Actors, Conflict prediction	Kernel density estimation (KDE), home range analysis, ecological niche factor analysis (ENFA), Monte Carlo K-function analyses, multivariate regression equation	The results were heterogeneous and revealed that rebel conflict behavior is associated with specific local variables. For example, rebels who relied on natural resources were found to kill civilians more frequently and in higher numbers while rebels who occupied steep terrain were more likely to win combat against government forces.

Schellens and Belyazid 2020	Global; country-wide	To examine the complex relationship between natural resources and socio-economic conditions to understand how they affect the risk of violent conflict.	This paper aims to improve the understanding of natural resources' role in the risk of violent conflicts by accounting for complex interactions with socio-economic conditions	Climate & Natural Environment, Conflict, Economic, Governance, Socio-demographic/socio-political/socio-economic	Global Conflict Risk Index (GCRI), Uppsala Conflict Data Program's (UCDP) One-sided Violence, Non-State Conflict, and Battle-Related Deaths datasets, Environmental Performance Index (EPI)	Conflict prediction+ environment	Logit regression models, neural network models and random forest models	Natural resource variables are found to be important predictors of conflict risk but do not necessarily improve the risk model. When observed in context with political and socio-economic variables several natural resource variables could provide insight into the root-cause of conflict.
Schneider et al 2012	Europe; Bosnia; country-wide	Investigate if events on the battlefield, intervention of international actors, or retaliation of an opponent can explain variability of one-sided violence	Many observers contend that wartime civilian victimization is an instrument of political leaders to achieve a particular goal. This article examines whether retaliation for similar acts by the other side, the developments on the battlefield, or the behavior of international actor's accounts for the ups and downs of this so-called one-sided violence	Conflict	Konstanz One-Sided Violence Event Dataset (KOSVED), ACLED, Penn State Event Data Project	Actors, Conflict intensity, Conflict prediction	Time-series regression; Autoregressive Integrated Moving Average (ARIMA) and Vector Autoregression (VAR) models	Considering the Bosnian civil war, this study found that international intervention is often ineffective, and that one-sided violence is typically either a military strategy or a retaliatory act.
Stanton 2015	Global; event-specific	To study the relationships between militia violence and government violence against civilians during civil war	This article examines pro-government militia violence against civilians in civil wars looking, in particular, at the relationship between militia violence against civilians and government violence	Military	UCDP/PRIO Armed Conflict Dataset, Pro-Government Militias Database (PGMD; Carey, Mitchell, and Lowe 2013)	Actors, Conflict intensity, Conflict prediction	Binary logit regressions, case studies	The research finds that the actions of government and militia are correlated so that when governments target civilians they typically also employ militia forces to target civilians that support the opposing insurgent group. However, the article also

			against civilians. The focus of the article is thus on the consequences of militia involvement in civil war, rather than its causes.					reveals more nuances in this relationship
Tollefsen and Buhaug 2015	Africa; sub-national	To evaluate how civil war risk is impacted by physical inaccessibility and sociocultural barriers.	This paper represents a first comprehensive evaluation of how physical and sociocultural inaccessibility relate to contemporary civil wars - and sheds light on how peripheries can shape space for action and mobilization by shifting focus from the country to the subnational local scale	Climate & Natural Environment, Geography & Administrative	PRIO-GRID, UCDP/PRIO Armed Conflict dataset, UNEP World Conservation Monitoring Center (UNEP WCMC), 2009 GlobCover, Gridded Population of the World v. 3.0 dataset, G-Econ	Conflict prediction	Multivariate logistic regression	Locations that are remote, are surrounded by rough terrain, or are inhabited by minority ethnic groups are more likely experience conflict in Africa, though it is unclear if this would be the same in other parts of the world.
van Weezel 2020	Africa; sub-national	To understand if long-term climatic variations can help explain conflict risk in Africa	This study has aimed to focus on relatively long-term changes, rather than relying on within-year variation in local climatic conditions across Africa and estimate whether these changes can help explain conflict risk	Climate & Natural Environment, Economic, Ethnic	Berkeley Earth Surface Temperature (BEST) dataset, HIRPS (Climate Hazards Group InfraRed Precipitation with Station) dataset, UCDP/PRIO Armed Conflict dataset	Conflict prediction + environment	Bayesian Model Averaging (BMA) and regression analysis	There is a notable positive correlation between temperature and conflict risk but the relationship between precipitation and conflict risk is less clear. Moreover, climatic variability is more likely to impact continuing conflict rather than fuel the onset of new conflict.
von Uexkull et al 2016	Africa; Asia; group-specific	To understand the interplay between drought, sociopolitical context, and risk of conflict	This study presents an actor-oriented analysis of the drought-conflict relationship, focusing specifically on politically relevant ethnic	Climate & Natural Environment, Economic, Governance, Socio-demographic/sociopolitical/socio-economic, Ethnic	UCDP/PRIO Armed Conflict dataset, Spatial Production Allocation Model 2005, Ethnic Power Relations dataset, Defense	Conflict prediction + environment, Ethnicity	Two-level mixed effect logistic regression models	Drought generally plays a small role in conflict risk while factors that more strongly contribute to conflict risk include ethnopolitical exclusion, proximity to conflict, among others. However,

			groups and their sensitivity to growing-season drought under various political and socioeconomic contexts.		Meteorological Satellite Program—Operational Linescan System, World Development Indicators			for populations that are politically excluded in low-income countries or are dependent on agriculture, drought does play an important role in conflict risk.
Warren and Troy 2015	Global; group-specific	To study the causes of violence within ethnic groups to further understand conflict dynamics	The objective is to systematically study the causes of intragroup violence to advance the understanding of conflict dynamics	Ethnic, Governance	Ethnic Power Relations (EPRs) data set (Cederman, Min, and Wimmer 2010), Minorities at Risk (MAR) Project (2009), XPolity index developed by Vreeland (2008)	Actors, Conflict intensity, Ethnic conflict	Parametric and semi-parametric regression techniques.	The likelihood of internal collective violence is related to the relative group size up to a certain point and the nature of this relationship is dependent on sociopolitical circumstances
Wayman and Tago 2010	Global; country-wide	To understand which independent variables account for the onset of mass political killings	The authors aim in this investigation to see what independent variables best account for the onset of mass political killing, with the state-year as the unit of analysis.	Conflict, Economic, Governance	Correlates of War (COW) project (Sarkees, Wayman & Singer, 2003), Polity IV dataset, Arthur Banks's dataset	mass atrocities (mass killing)	Cox proportional hazard model	The authors find that the incorporation of different datasets to represent the same variable in studies leads to very different outcomes. In this study, the author finds that the importance of regime effects appears and disappear according to the dataset used and that regime type generally do have a substantial effect on democide but not on genocide or politicide.



Wegenast and Basedau 2014	Global; group-specific	To study how the relationship between ethnic fractionalization, economics, and natural resources affect conflict	this paper aims to explicitly differentiate between ethnic and non-ethnic violence and strives to reconcile the literatures on the ethnic and economic determinants of armed conflict. It asks why ethnically fractionalized societies may experience armed conflict in the presence of natural resources.	Climate & Natural Environment, Ethnic	UCDP/PRIO Armed Conflict Dataset, Wucherpfennig et al. (2012) dataset, Ethnic Power Relations (EPR) dataset, ELF index of Cederman et al. (2010)	Conflict prediction + environment, Ethnicity	Logit models	The presence and availability of oil increases the risks of ethnic fractionalization and conflict.
Weidmann and Salehyan 2013	Asia; Iraq; sub-national	To understand what factors contributed to instances of increased and decreased violence in Baghdad	The purpose is to find out which variables can explain the variation in violence and segregation we observe in Baghdad, and in particular, if segregation alone is sufficient to explain the decrease in violence	Ethnic	UNs Humanitarian Information Center for Iraq (HIC), Iraq SIGACTS database, M. Izady at the Gulf 2000 project at Columbia University	Actors, Conflict intensity, Conflict prediction	Agent-based modeling	More ethnically homogenous neighborhoods generally experience declining levels of violence. The authors find that violence is more likely to occur where an ethnic minority lives among an ethnic majority, since insurgents may target that minority through ethnic cleansing.
Wimmer et al 2009	Global; country-wide	To understand the relationship between ethnically diverse populations, the state, and conflict	The objective is to examine the linkages between ethnicity in conflict	Ethnic, Geography & Administrative, Governance	Ethnic Power Relations (EPR) data set, UCDP/PRIO Armed Conflict dataset	Conflict intensity, Ethnicity	Multinomial logit models	Ethnically diverse societies are not more likely to result in ethnic conflict. Rather, the study finds that high rates of ethnic exclusion and segmentation from power structures are the several of the main drivers of rebellions.

Witmer et al 2017	Africa; sub-national	To forecast violence in sub-Saharan Africa by incorporating climatic variables and sociopolitical factors	The study uses a climate-sensitive approach to model sub-Saharan African violence in the past and then forecast future violence using sociopolitical factors.	Climate & Natural Environment, Conflict, Geography & Administrative, Socio-demographic/sociopolitical/socio-economic	ACLED, UN World Population Prospects 2012 revision, Freedom House, ESRI, Factiva, Center for International Earth Science Information Network (CIESIN) Gridded Population of the World (GPW) version 3 dataset, Climate Research Unit (CRU) TS3.21 dataset, Community Earth System Model	Conflict prediction+ environment	Regression analysis	Population increases, rising temperatures, and poor governance will likely contribute to higher levels of conflict in the future.
-------------------	----------------------	---	---	--	---	----------------------------------	---------------------	---