International Journal of
*Geo-Information*

MDPI

*Article*

# Identifying Conditioning Factors and Predictors of Conflict Likelihood for Machine Learning Models: A Literature Review

**Timur Obukhov [1] and Maria A. Brovelli [2,]***

[1] Department of Computer, Control and Management Engineering Antonio Ruberti (DIAG), La Sapienza University of Rome, 00185 Roma, RM, Italy; obukhov@diag.uniroma1.it
[2] Department of Civil and Environmental Engineering (DICA), Politecnico di Milano, 20133 Milano, MI, Italy
* Correspondence: maria.brovelli@polimi.it

**Abstract:** In this research, we focused on armed conflicts and related violence. The study reviewed the use of machine learning to predict the likelihood of conflict escalation and the role of conditioning factors. The results showed that machine learning and predictive models could help identify conflict-prone locations and geospatial factors contributing to conflict escalation. The study found 46 relevant papers and emphasized the importance of considering unique predictors and conditioning factors for each conflict. It was found that the conflict susceptibility of a region is influenced principally by its socioeconomic conditions and its political/governance factors. We concluded that machine learning has the potential to be a valuable tool in conflict analysis and, therefore, it can be an asset in conflict mitigation and prevention, but the accuracy of the models depends on data quality and the careful selection of conditioning factors. Future research should aim to refine the methodology for more accurate prediction of the models.

**Keywords:** conflicts; war; conflict susceptibility; conditioning factors; predictors; machine learning

check for updates

**Citation:** Obukhov, T.; Brovelli, M.A. Identifying Conditioning Factors and Predictors of Conflict Likelihood for Machine Learning Models: A Literature Review. *ISPRS Int. J. Geo-Inf.* **2023**, *12*, 322. https://doi.org/10.3390/ijgi12080322

Academic Editors: Godwin Yeboah and Wolfgang Kainz

Received: 8 June 2023
Revised: 25 July 2023
Accepted: 31 July 2023
Published: 2 August 2023



## 1. Introduction

According to the data collected by Uppsala University [1], 294 distinct armed conflicts have occurred in various places around the world since 1946. Considering that there are various types of conflict definitions, including intrastate and internal conflicts, we used the definition described in the United Nations terminology database, UNTERM (https://unterm.un.org/unterm2/en/ (accessed on 30 July 2023)): armed conflict is a situation in which there is resort to the use of armed forces between States or a protracted resort to the use of armed forces between governmental authorities and organized armed groups or between such groups within a State or the use of armed forces between governmental authorities and organized armed groups or between such groups within a State [2].

Sustainable Development Goal 16 calls upon all United Nations Member States to "promote peaceful and inclusive societies for sustainable development" [3]. It has been argued that SDG 16 is the most important goal, without which, none of the other goals can be sustained [4]. The effect of armed conflicts is enormous, for example, there were over 1,590,374 battle-related deaths only for the post-Cold War period from 1989 to 2021 [5] (Figure 1). War is a development issue. War kills, but the consequences extend far beyond these direct deaths. In addition to battlefield casualties, armed conflicts often lead to forced migration, refugee flows, capital flight, and the destruction of societies' infrastructure. Social, political, and economic institutions are indelibly harmed. The consequences of war, and especially civil war, on development are profound [6].

Many complex factors lead to conflicts. Some conditions that increase the likelihood of armed conflicts include the inability of governments to provide good basic governance and the protection of their own populations [7]. Various publications referring to the factors that lead to the conflicts use different terms for such phenomena. Some literature refer to them as drivers of conflicts, others as conditioning factors, and factors that conditioned the evolution of a conflict. Today, armed conflicts are still among the biggest threats to human societies, and identifying the potential drivers is an area of intense scientific research [8]. The potential factors that influence the conflict have been identified in the scientific research literature, including poverty [9], income inequality [10], economic struggle [11], weak governance [12], a preexisting history of conflicts [13], financial assets from natural resource exploitation [14], ethnic fractionalization [15–17], vulnerability to natural disasters [18], and climate change [19]. Referring to the last point, however, after years of research, there has been no consensus among scholars regarding whether or how climate influences the risk of armed conflict [19]. Various organizations have recognized the significance of using conditioning factors in conflict research and analysis. The Joint Research Centre (JRC) of the European Commission launched the Index for Risk Management (INFORM) project in 2012 [20], a significant milestone in the field for assessing risks. INFORM is a global, open-source risk assessment tool for humanitarian crises and disasters on a national level. This project leverages conditioning factors to assess and predict crisis risks. Another project launched in 2017 by Uppsala University—the ViEWS (Violence Early Warning System) project—provides predictions for where armed conflicts are likely to occur on a national level [21]. While organizations like the JRC and Uppsala University focus on applying conditioning factors and machine learning techniques to identify risks at a global or regional level, other scholars emphasize the importance of identifying local dynamics in conflict analysis and predictions [15,22] including personal grudges and local power struggles, which often play a significant role in the onset of a conflict, thereby emphasizing the need for a case-by-case study of conflicts.

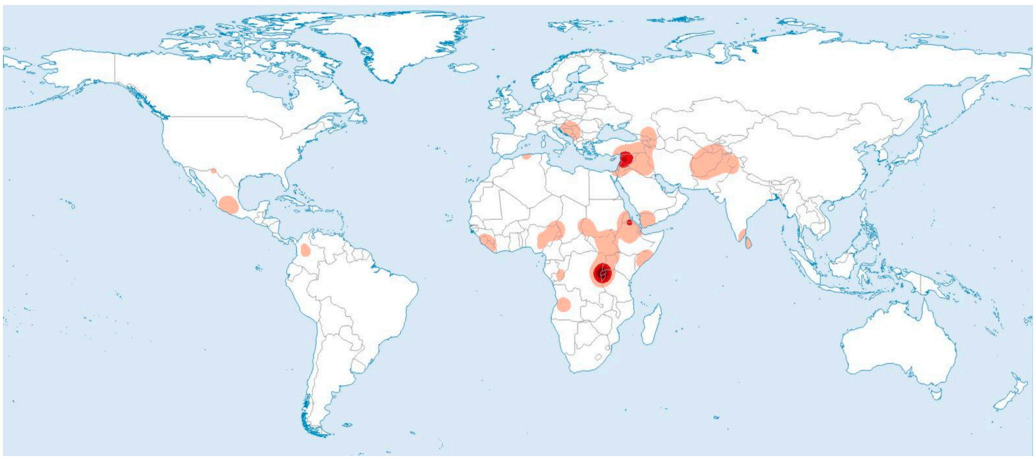

**Figure 1.** Heatmap represents the severity of battle-related deaths for the period 1998–2021 [2].

Over the past decade, machine learning techniques and predictive models have evolved as powerful instruments for predicting conflicts due to their ability to handle vast amounts of data, recognize complex patterns, and generate reliable predictions. Machine learning techniques allow us to capitalize on the increasing availability of data relevant to conflict studies. This includes standard socioeconomic indicators and data on climate change, population movements, social media trends, etc. The capacity of machine learning to integrate and depict insights from this vast and diverse data landscape is another factor propelling its increased use in the field. The aspects of using machine learning and predictive models for identifying the likelihood of conflicts have been introduced previously. However, the specific application of machine learning technology to delineate conditioning factors and predict the onset of new conflicts has gained momentum only in the past few years. This emerging focus reflects the growing recognition of the multifaceted nature of conflicts and the need to approach prediction from a comprehensive, multidimensional perspective.

For this study, the research on the state-of-the-art methodology is crucial to identify the list of conditioning factors that may lead to the escalation of conflicts. Thus, this study evaluated the available literature on the application of machine learning to identify the likelihood of conflict escalation and conditioning factors and their role in conflict escalation. The conditioning factors identified in this study support the application of machine learning technology to conflict susceptibility research for conflict-prone locations worldwide. The conditioning factors identified in this study provide a ground basis for applying machine learning technology in conflict susceptibility research. By leveraging these cutting-edge techniques, we can uncover deeper insights into conflict triggers and offer more nuanced predictions about conflict escalation.

## 2. Methodology

This research focused on identifying a combination of conditioning factors that can be used to determine the likelihood of the eruption of new conflicts and the use in machine learning for susceptibility research. The list of conditioning factors is subject to several preconceived factors, for example, data availability, geographic conditions, and type of conflicts. Considering the above, the list of conditioning factors should be applied specifically to an individual case study.

This paper argues that conditioning factors play a specific role in conflict escalation. The conflict conditioning factors list includes various social-economic aspects, agricultural practices, availability of natural resources, governance practices, climate change, pollution, and different types of terrain and elevation levels.

The methodology for detecting conflicts in the area of interest incorporated data science and machine learning processes. Data science methods have been remarkably effective in flexibly accommodating conditional relationships to uncover differential effects across locations [23]. This requires training and validation datasets and conditioning factors. Then, the machine learning algorithms were applied to produce conflict susceptibility to identify the locations with a high likelihood of conflict eruption. Historical armed conflict datasets were used for the training and validation datasets. The factors that play an essential role in the outbreak of conflicts were used as conditioning factors. The process of employing training and validation datasets, and conditioning factors is described in Figure 2.

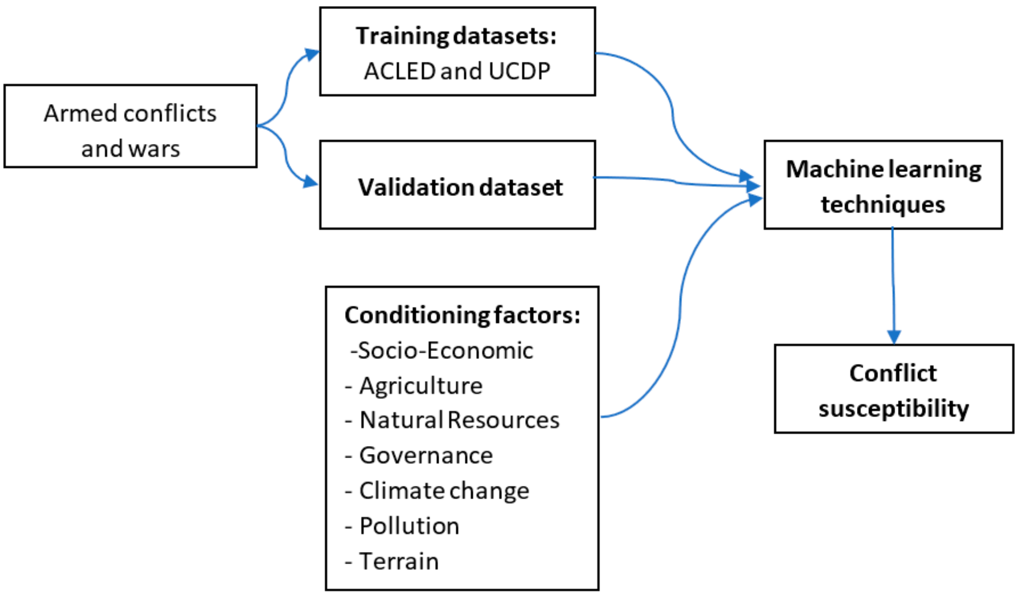


**Figure 2.** The process of predicting conflict susceptibility using machine learning algorithms.

The principle focus of this research was to identify the conditioning factors and locating suitable publicly available data sources to develop a model for predicting the likelihood of conflicts. Figure 2 provides the basic principles and methodology of the concept. Following the identification of training, validation, and conditioning factor datasets, we employed various machine learning techniques to build the models. These techniques included data pre-processing, splitting the conflict data into training and validation datasets using the principle of an 80/20 split, developing models with various machine learning algorithms, validating the models' accuracy, and finally producing conflict susceptibility products depicting the likelihood of conflict eruption.

Data preprocessing is a critical step in applying machine learning applications [24] to conflict susceptibility and developing models for conflict prediction. These models rely on a broad spectrum of publicly available, unstructured data sources that provide data in different models and formats, using different spatial and temporal resolutions and scales. Often, these sources contain inconsistencies, missing values, and noise. The goal of data preprocessing is to transform this raw data into a clean, standardized format that machine learning algorithms can efficiently process. To split the conflict data into training and validation datasets, we used the standard principle of an 80/20 split [25]. This principle provides a balance between training the model to recognize patterns from the conflict datasets (ACLED and UCDP) and evaluating its performance on unseen data. The following step involved the deployment of suitable machine learning algorithms. For example, widely used machine learning algorithms with robust capabilities like Random Forests (RFs), Neural Networks (NNs), and Support Vector Machines (SVMs), were utilized to construct the conflict prediction models. Each algorithm has its strengths and weaknesses. The final stage of applying the machine learning techniques included the validation of the models. Model validation is an important step in our methodology to ensure the models' accuracy, reliability, and generalizability. Several key metrics and techniques can be applied to assess different aspects of model performance. Among these, the Area Under the Receiver Operating Characteristic curve (AUC-ROC) is a crucial metric used in binary classification tasks [26]. Based on the identified accuracy of the models, the appropriate products for the prediction of conflict onset can be selected.

Figure 3 presents the workflow procedure for the literature review of predicting conflict susceptibility using machine learning technologies, which includes indexing databases for the retrieval of articles, screening based on inclusion or exclusion criteria, sorting the articles based on the number of citations, and extraction of the relevant information from each article. The search was conducted using two search engines: Web of Science (https://www.webofscince.com (accessed on 30 July 2023)) and Scopus (https://scopus.com (accessed on 30 July 2023)). The literature search in both search engines was conducted using keywords to optimize the review process. The following queries were prepared for the search criteria in both Web of Science and Scopus.

The initial ("Armed conflicts") AND ("Machine Learning") search provided 80 results from Web of Science and 72 papers from Scopus. Further searches in the returned result using the search words "predictors" or "conditioning factors" provided an even more limited number of publications—46 papers (the list of publications is given in the Supplementary Materials).

The 46 publications were thoroughly reviewed and considered for selection or exclusion based on the following criteria:

1. Articles were selected if:
    (a) Publication was peer-reviewed.
    (b) Publication reported on the use of conditioning factors and/or predictors in machine learning model development for conflict forecasting.
    (c) Publication included a list of factors affecting the onset of wars or conflicts.
2. Articles were excluded if:
    (a) Publication was not related to conflicts and machine learning.
    (b) Full text was not available.
    (c) Publication was not available in English.
    (d) Publication did not have a DOI.

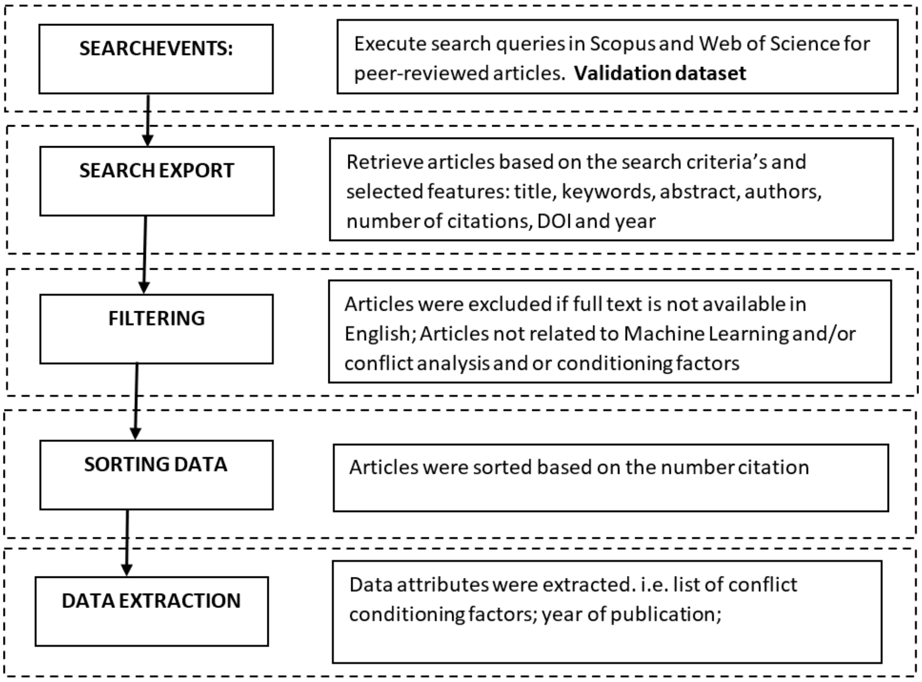


**Figure 3.** Literature review methodology.

After filtering the articles, 18 publications were identified that provide specific conditioning factors or predictors and provide examples of applying them in a machine learning process for conflict onset prediction research.

Figure 4 presents a diagram of the publication categories, such as "Machine Learning and Computer Science", "Political Science", "Environmental Science", "Multidisciplinary Science", etc. and the year when the article was published. The diagram shows that the large majority of selected articles were related to "Machine Learning and Computer Science" and "Political Science" and a large number of articles was published after 2019.

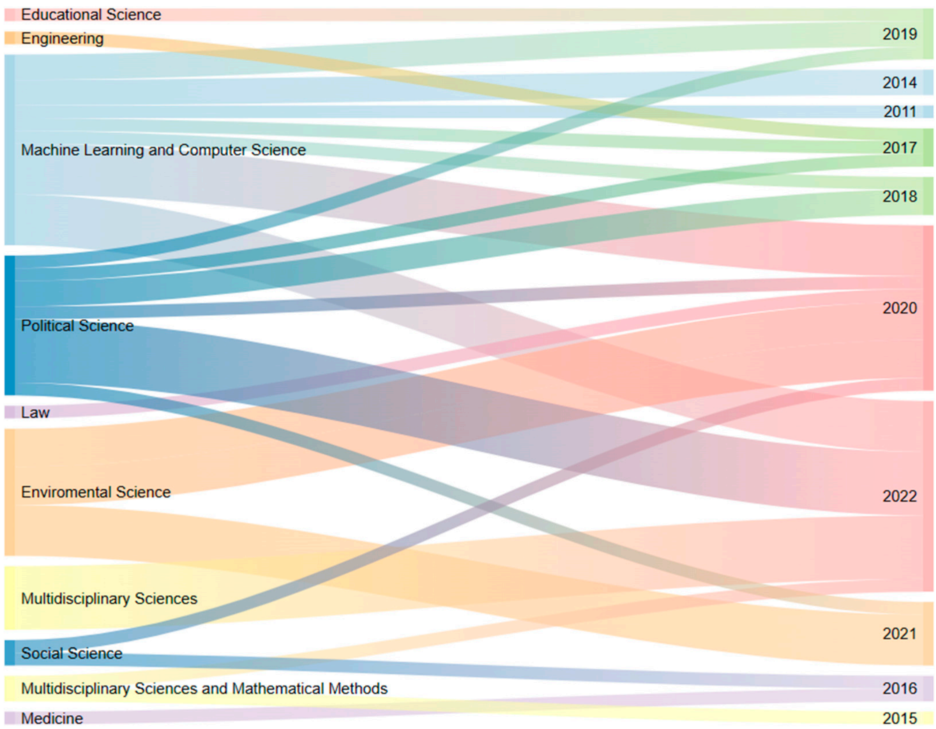


**Figure 4.** Diagram representing publication categories and the year of publication (Web of Science 2023).

It is noteworthy that 15 out of the 18 papers were published after 2017. The majority of the topics and locations covered by these publications focused on global or regional analyses. Specifically, seven papers concentrated on the global scale, four on the African continent, and three on subcontinental African regions. Additionally, four studies focus on country-specific topics of interest, namely India [19], Liberia [27], Kenya [28], and Colombia and Indonesia [19,29]. The Random Forest technique was employed in eight research articles for machine learning processes. In contrast, Boosted Regression Trees (BRT), Logistic Regression, Gradient Boosting, and Neural Network algorithms were utilized in two articles each. Other methods, such as Support Vector Machine (SVM), Dynamic Elastic Net, and Naive Bayesian, were only used once.

The most challenging aspect of the process is identifying the list of conditioning factors to be used. While criteria such as population and GDP are frequently employed, others, such as pollution and poverty, are less commonly used. To address this, we developed a comprehensive list of the conditioning factors cited in the literature and assigned a priority to each.

The significance of a conditioning factor was determined by the number of publications and the frequency of citations in the articles where it is mentioned. This ranking system determined the priority of applying these factors in the machine learning process. However, compared to the number of times a factor was mentioned in an article, the number of publications was substantially lower. Therefore, to normalize both variables, we used the Min-Max method.

Min-Max normalization is a technique that linearly transforms the original data, number of publications, and the times cited into a range from 0 to 1. The Min-Max normalization formula is given by $x - \min(x)/(\max(x) - \min(x))$, where x is the value that needs to be normalized, min(x) is the minimum value in the dataset, and max(x) is the maximum value in the dataset. We conduct the Min-Max normalization process separately for the number of publications (PubN′) and the number of citations (TCit′). To determine the normalized importance of the conditioning factors, which are based on both PubN′ and TCit′, we calculated the mean between both variables. The normalized importance of conditioning factors (NImportanceCF) is equal to the mean of the normalized publication number (PubN′) and the normalized number of citations (TCit′) on the Web of Science platform. To identify the priorities of the conditioning factors, we rescaled the normalized importance of the conditioning factors from a range of [0, 1] to [0, 10] by multiplying the importance of the conditioning factor by 10.

The result of the literature review was a list of conditioning factors or, as they are also called in the literature, "predictors" [27]. In addition, as described above, we assigned priorities for each conditioning factor based on the number of citations and number of times the condition factor was mentioned in the publications. The conditioning factors data were sourced either from international, national, or local open and public sources. The data related to conflicts were sourced from two significant data science projects: The Armed Conflict Location & Event Data Project (ACLED) [29] and Uppsala Conflict Data Program (UCDP) [30]. These projects provide conflict data from around the world. Both datasets provide researchers and policymakers with a vast amount of data on various types of conflicts for analysis and research. Both use a range of data sources, including media reports and other public sources, to create comprehensive data sets of conflict events. Despite their common goal, these two projects have different methodologies and focus on different data granularities.

ACLED is known for its near-real-time reporting of conflict data with high granularity, including a specified taxonomy of event types and event subtypes (see Table 1), including the groups involved, changes in territorial control, and reported fatalities. This classification provides a comprehensive view of conflicts in terms of forms and conflict manifestations. For example, the classifications provide a distinction between events like battles and explosions, which can help understand whether the violence is direct or indirect towards a group of people. Such distinctions can play a critical role while studying the dynamics and impacts of conflicts.

**Table 1.** Taxonomy of event types and subtypes in ACLED [20].

| Event Type | Even Subtype |
| --- | --- |
| Battles | Armed clash |
| | Government regains territory |
| | Non-state actor overtakes territory |
| Explosions/Remote violence | Air/drone strike |
| | Grenade |
| | Remote explosive/landmine/IED |
| | Shelling/artillery/missile attack |
| | Suicide bomb |
| Protests | Excessive force against protesters |
| | Peaceful protest |
| | Protest with intervention |
| Riots | Mob violence |
| | Violent demonstration |
| Strategic developments | Abduction/forced disappearance |
| | Agreement |
| | Arrests |
| | Change to group/activity |
| | Disrupted weapons use |
| | Headquarters or base established |
| | Looting/property destruction |
| | Non-violent transfer of territory |
| | Other |
| Violence against civilians | Abduction/forced disappearance |
| | Attack |
| | Sexual violence |

On the other hand, UCDP provides two critical datasets: a disaggregated UCDP georeferenced event dataset (GED) covering the period from 1989 to 2021; and a yearly UCDP/PRIO armed conflict dataset, covering the period from 1946 to 2021. While both UCDP datasets share common bases in conflict research, they significantly vary in the level of detail spatial and temporal resolution of the data conflicts. UCDP/PRIO data sets are produced in collaboration with the Pierce Research Institute Oslo (PRIO) and provide annual datasets covering global conflicts from 1946 to the present. This dataset includes information on the identity of parties of the conflicts, estimation of the number of battle-related deaths in each year, and the type of conflicts, including interstate and other types of conflicts. The UCDP/PRIO data set is useful for research questions concerning trends and conflicts over time. The GED dataset provides disaggregated and granular records of conflict events. It contains information on individual events of organized violence and the number of fatalities.

Both ACLED and UCDP datasets are used for conflict analysis. While they share the same aim of documenting violent conflicts, they use different methodologies and levels of granularity. ACLED covers political violence and protest events and UCDP GED covers individual incidents of organized violence, both fatal and non-fatal. ACLED and UCDP GED provide highly granular event-based data, and UCDP/PRIO provides aggregate data. In terms of timeline and update aspects, ALCED is near-real time, and UCDP datasets are updated yearly. Overall, all datasets, ACLED, UCDP GED, and UCDP/PRIO, provide invaluable insight into conflict dynamics. The choice between the datasets depends on the specific research question, granularity, and temporal or geographical aspects. However, both ACLED and UCDP datasets are often used complementarily to gain a comprehensive understanding of violent conflicts.

It is understood that the frequency of citations and the number of times a certain conditioning factor has been used in previous research does not automatically reflect the importance of a conditioning factor. This importance is often tied to specific local conditions, with varying factors and combinations differing significantly across geographical regions. First, the impact of these factors largely depends on local circumstances. For example, while access to grazing land may onset a conflict in Darfur, Sudan, in other geographic regions, grazing land might not be a commodity of significant conflict potential.

Second, we broadened our analysis to incorporate the insights of various scholars to provide a more nuanced understanding of feature importance. Different researchers describe varying levels of significance to these factors in contributing to conflicts. Several studies have emphasized the role of political instability, low income, difficult terrain, large populations, and political inequality among ethnic groups as significant predictors of civil conflict [14,31]. These studies have emphasized the role of political inequality among ethnic groups as a significant predictor of civil conflict. Other research has shown how qualitative data on social factors, although challenging to obtain, can offer valuable insights [32]. Some scholars argue for the role of economic shocks, especially those related to weather, in civil conflict, underscoring the predictive potential of environmental data [33]. As suggested by Colin Flint in "The Geography of War and Peace", understanding a conflict requires considering its inherent complexities and unique geographical, social, political, and economic contexts [24]. As these studies reveal, the relative importance of various features remains a contested issue in the field. The conditioning factors should be selected based on such criteria as the context of the study, the availability and quality of data, and the specific theoretical framework adopted. And, as it was already mentioned, the available data can dictate the selection of certain features and conditioning factors.

We encourage researchers to conduct a thorough review of the local conditions specific to their research area, consider the quality and availability of data, and make thoughtful theoretical and methodological decisions on the selection of appropriate conditioning factors for their study. It is important to note that a general ranking of variables might not always offer an applicable solution for predicting conflicts, as the importance and composition of these features depend on the conflict type, location, and specific local circumstances.

We used the list of citations, and the frequency of each conditioning factors as a foundation. This method helps identify which factors have generally been applied in machine learning and conflict prediction research. This approach provides a broader understanding of the features that have drawn attention from the scientific community and the scope of factors considered when applying machine learning for conflict forecasting. The accessibility and relevance of these factors will inevitably vary across different contexts and regions, reflecting the multifaceted and intricate dynamics globally. Therefore, researchers should carefully consider the context, local dynamics, relevance, and applicability of these factors in their individual studies.

## 3. Results

Table 2 provides the result of the analysis of the 18 articles related to the prediction of conflicts with the application of machine learning in combination with the conditioning factors.

**Table 2.** Conditioning factors, number of publications, citations, and priorities [34].

| Conditioning Factor or Variable | Number of Publications | Number of Publications 2017–2022 | Times Cited (WOS Count) | Conditioning Factor Priority |
|---|---|---|---|---|
| Conflict data variables | | | | |
| ACLED | 11 | 8 | 316 | 9 |
| UCDP | 12 | 11 | 70 | 5 |

**Table 2.** *Cont.*

| Conditioning Factor or Variable | Number of Publications | Number of Publications 2017–2022 | Times Cited (WOS Count) | Conditioning Factor Priority |
|---|---|---|---|---|
| Conditioning factor classification: Socioeconomic | | | | |
| GDP | 9 | 8 | 59 | 4 |
| GINI | 1 | 1 | 3 | 0 |
| HDI | 1 | 1 | 3 | 0 |
| Population | 14 | 12 | 309 | 10 |
| Ethnic | 10 | 9 | 38 | 4 |
| Age | 2 | 2 | 2 | 0 |
| Religious | 3 | 3 | 1 | 1 |
| Urban Accessibility | 4 | 2 | 193 | 4 |
| Nightlight Index | 3 | 3 | 3 | 1 |
| Inflation | 1 | 1 | 0 | 0 |
| Social Services | 5 | 5 | 34 | 2 |
| Unemployment | 6 | 5 | 42 | 3 |
| Access to Land | 2 | 2 | 29 | 1 |
| Economic Level | 6 | 6 | 7 | 2 |
| Security | 2 | 2 | 4 | 0 |
| Health | 6 | 4 | 231 | 6 |
| Education | 2 | 2 | 3 | 0 |
| Infrastructure | 3 | 2 | 188 | 4 |
| Poverty | 2 | 1 | 2 | 0 |
| Corruption | 2 | 2 | 6 | 1 |
| Income Inequality | 3 | 2 | 12 | 1 |
| Homicide | 2 | 2 | 6 | 1 |
| Food Security | 1 | 1 | 6 | 0 |
| Conditioning factor classification: Agriculture | | | | |
| Crop | 5 | 4 | 10 | 2 |
| Harvest | 2 | 2 | 1 | 0 |
| Irrigation | 1 | 1 | 1 | 0 |
| Vegetation Index | 2 | 1 | 187 | 3 |
| Conditioning factor classification: Natural Resources | | | | |
| Diamonds | 5 | 3 | 8 | 2 |
| Oil | 7 | 5 | 13 | 3 |
| Metals | 5 | 3 | 7 | 2 |
| Wood | 3 | 1 | 6 | 1 |
| Renewable Energy | 1 | 1 | 0 | 0 |
| Conditioning factor classification: Political and Governance | | | | |
| Allies of Countries | 1 | 1 | 0 | 0 |
| Elections | 3 | 2 | 189 | 4 |
| Democracy | 11 | 9 | 274 | 8 |
| Human Rights | 5 | 4 | 194 | 5 |
| Gov. Effectiveness | 2 | 1 | 0 | 0 |
| Political Type | 2 | 1 | 6 | 1 |
| Conditioning factor classification: Climate Change | | | | |
| Weather Shock | 10 | 9 | 53 | 4 |
| Temperature | 9 | 8 | 232 | 7 |
| Flood | 1 | 0 | 0 | 0 |
| Draught | 1 | 0 | 0 | 0 |
| Precipitation | 8 | 7 | 194 | 6 |
| Conditioning factor classification: Pollution | | | | |
| Air Pollution | 1 | 1 | 0 | 0 |
| Soil Degradation | 1 | 1 | 0 | 0 |

As mentioned earlier, most of the papers were published during the last five years. That proves our observation that applying machine learning, susceptibility, and conditioning factors to analyze and predict the likelihood of eruption of new conflicts is a new subject. We observed little interest in the subject before 2017. However, interest was raised concerning this issue in 2020 when the number of articles increased over time fold. The number of citations for each publication was considered academic value over the paper. The base for citation analysis was selected from the academic publication platform Web of Science.

While conducting the literature review on conflict conditioning factors, we noticed that the conditioning factors could be grouped or classified based on their characteristics and context. These factors were classified into distinct categories such as socioeconomic, agricultural, natural resources, political and governance aspects, and climate change classifications. Analyzing these conditional factors provides a multidimensional aspect through which we can obtain an in-depth insight into potential sources of conflicts and understand which factor plays a more significant role in the onset of a conflict. By examining these factors and their classifications, we can gain a comprehensive understanding of the potential root causes of the conflicts and work towards developing various strategies for conflict prevention and resolution.

One paper, "Climate variability and conflict risk in East Africa, 1990–2009" [35], was published in 2012 and has been cited 185 times, receiving much attention from the academic world. This explains why some conditioning factors have been cited five times more than others. For example, papers using ACLED as a conflict dataset were cited almost five times more (316 citations) than papers using the UCDP dataset.

The results presented in Table 2 include the list of conditioning factors and their grouping, number of publications, number of recent publications (in this case, from 2017 to 2022), number of times that the publication was cited in the Web of Science, and the conditioning factor's priority.

### 3.1. Training and Validation Datasets

The two datasets, ACLED and UCDP, are widely used in various papers analyzing conflict data. These datasets are used for training and validating machine learning models.

The Armed Conflict Location & Event Data Project (ACLED) is a disaggregated data collection, analysis, and crisis mapping project. ACLED collects the dates, actors, locations, fatalities, and types of all reported political violence and protest events around the world. The ACLED team conducts analysis to describe, explore, and test conflict scenarios, and makes both the data and analysis open for free use by the public [29].

The temporal resolution of the dataset is varied between continents. For Africa, the data start from January 1997 [36]. Each data record is referred to as an individual event of a conflict, classified as event type and subtype as presented in Table 1. For example, a type of event is "Battles" with three subtypes: "Armed clash", "Government regain territory", and "Non-state actor overtake territory". ACLED data contain geographical information in the form of the location represented in longitude and latitude and the administrative location. ACLED uses latitude and longitude geo-coordinates based on EPSG:4326 [29].

The data are collected from various sources: traditional media (national and international media outlets), reports from national and international non-government organizations, local partner data (collected by local social organizations through primary and secondary means), and new media such as social media platforms. Conflict data researchers consider an individual process of sourcing data for each conflict as the most reliable [37]. Therefore, ACLED applies data quality verification processes individually to each conflict, location, or country.

Various scholars and academic institutions use the ACLED database, International Organizations such as the United Nations and its Agencies, Funds, and Programmes, ICRC, European Union Agencies, and other government and inter-government institutions, think tanks, as well as media news agencies [36]. In the academic publications, the ACLED dataset was mentioned in 42 publications, mainly related to international relations and political science. The ACLED data are publicly available at the ACLED website www.acleddata.org. ACLED makes its data available to the public under a Non-Commercial Licenses, i.e., ACLED's full dataset is available for use free of charge by noncommercial entities and organizations (e.g., non-profit organizations, government agencies, academic institutions) for non-commercial purposes, subject to the Terms of Use [38].

The UCDP (Uppsala Conflict Data Program) is a project based at Uppsala University in Sweden that provides data and analyses on armed conflicts around the world. It collects new global data on non-state conflicts, or armed conflicts between two groups, neither of which is the state. The dataset includes conflicts between rebel groups and other organized militias, and thus serves as a complement to existing datasets on armed conflicts which have either ignored this kind of violence or aggregated it into civil war. The dataset also includes cases of fighting between supporters of different political parties as well as cases of communal conflict, that is, a conflict between two social groups, usually identified along ethnic or religious lines [39].

The UCDP data includes two main datasets: i. the UCDP/PRIO Armed Conflict Dataset, which provides information about armed conflicts and their characteristics, such as the number of deaths, the type of actors involved, and the location of the conflict covering conflicts in the period from 1946 to 2021 [2]; ii. the Disaggregated UCDP Georeferenced Event Dataset (GED), which provides detailed information about specific events within armed conflicts, such as battles, riots, and protests covering events from 1989 to 2021 [5].

The project's data are used by researchers, policymakers, and journalists to better understand trends in armed conflicts, as well as their causes and consequences. In the academic publications, the UCDP dataset was mentioned in 79 publications, mainly related to international relations, political science, and some public environmental occupational health fields [34]. The UCDP datasets are publicly available at https://ucdp.uu.se/ (accessed on 30 July 2023). Uppsala University provides the UCDP data for free with no restrictions.

Table 2 contains 18 reviewed papers, with 11 referring to the ACLED dataset, and 12 referring to the UCDP dataset. Seven papers referred to both ACLE and UCDP datasets. Most of these articles related to ACLED and UCDP conflict data were published in the last five years, from 2017 to 2022.

It was observed that the article "Climate variability and conflict risk in East Africa, 1990–2009" [35] was cited 185 times. Thus, this article contributed the most to the number of cited articles. In addition, the ACLED database was referred to in this paper as a conflict dataset. Therefore, the number of cited papers for ACLED increased fivefold, with an overall higher priority for ACLED and UCDP. Both datasets were used in the process of training and validating the models.

### *3.2. Conditioning Factors*

#### 3.2.1. Socioeconomic Conditioning Factors

Scholars across the social and political disciplines agree that socioeconomic factors play a significant role susceptibility of a region or community to conflict [33]. These factors include poverty, low per capita income, slow economic growth, high natural resource dependence, low levels of education, and high population [40]. A population may be more prone to engage in or support violent or disruptive behavior to survive or to bring about change when they are struggling to satisfy their fundamental requirements. High levels of inequality may also make a region more prone to war since disenfranchised groups may believe they have no other options for achieving their objectives. An analysis of the socioeconomic situation offers insightful information on the likelihood of conflict and can informs tactics for averting or reducing it.

There are 14 socioeconomic factors identified in Table 2 with priorities from 1 to 10. Priority 10 was given to population; priority 6 was given to health; priority 4 was given to ethnic, GDP, urban accessibility, and infrastructure factors; priority 3 was given to unemployment; priority 2 was given to social services factors; priority 1 was given to religious factors, nighttime light, access to land, corruption, income inequality, and homicide factors.

Population size and growth can be a significant factor in the likelihood of conflict [40–42]. For instance, fast population growth may put a strain on the resources available and heighten the struggle for food, water, and land. This might make conflict more likely, especially in regions where resource-based conflicts have a history or where resources are already scarce. The likelihood of violence can also be influenced by the youth bulge, a demographic phenomenon characterized by an unacceptably high proportion of youth in a population [43,44]. Young people may be more likely to engage in or support violent or disruptive conduct if they perceive that they have little possibility for education and employment.

Table 2 contains 18 reviewed papers with 14 referring to a population dataset. Most of these articles related to population data were published in the last five years, from 2017 to 2022. We identified the Global Human Settlement Layer (GHSL) as the population source. GHSL is a georeferenced layer that provides global coverage information on human settlements and populations. It is produced by elaborating on historical satellite images and data from open sources. The primary datasets consist of gridded layers of built-up areas and the number of inhabitants at a high resolution for four dates: 1975, 1990, 2000, and 2015. In addition, the GHSL allows us to measure the growth of cities and towns over time, including information on population, urbanization rate, and land consumption [36]. The GHS population grid (GHS-POP) is derived from GPW4.1 (multi-temporal: 1975-1990-2000-2015) and R2019A [GHS_POP_MT_GLOBE_R2019A]. This product was distributed as part of the Community pre-release of the GHSL Data Package 2018 (GHS CR2018) [45]; an updated version of the datasets is now available (v2.0).

Health and infant mortality was included in the research on the effects of access to health facilities and infant mortality on the likelihood of violence and conflicts [35,46–48].

Health may be a risk factor for conflicts since it may lead to social and economic instability, resulting in conflicts as during the Ebola outbreak in West Africa in 2014. Conflicts may also result from the spread of infectious diseases since it strains healthcare systems and causes anxiety. Infant mortality can affect conflict in several different ways. High infant mortality rates may indicate a population's poor overall health and well-being, which can feed the dissatisfaction and discontent. High infant mortality rates may also suggest that people lack access to essential resources like healthcare, which increases dissatisfaction and discontent. Finally, high infant mortality rates can impact population growth and demographic change. Although access to health services and infant mortality can contribute to conflict, it is crucial to remember that it is not the only or primary cause of conflict. The risk of conflict in a particular situation is frequently influenced by a variety of factors that interact in an intricate manner.

Table 2 contains 18 reviewed papers. Six referred to health and infant mortality rate data. Most of these articles were published in the last five years, from 2017 to 2022. There were 185 citations for the paper "Climate variability and conflict risk in East Africa, 1990–2009" [35]. The number of cited articles was, therefore, mainly influenced by this article. The priority of this conditioning factor was 6 and was largely influenced by the O'Loughlin's article. Health access data and infant mortality data are available at a country-level resolution and specific cases on a sub-national level. However, the sub-national level data are inconsistent, creating difficulties for analyses on the sub-national and local levels. The World Bank data portal (https://data.worldbank.org/indicator/SH.DYN.MORT (accessed on 30 July 2023)) [49] provides data on child mortality. The World Health Organization [50] provides health-related data at https://www.who.int/data/global-health-estimates (accessed on 30 July 2023). The Demographic and Health Survey portal provides country-specific health-related datasets (https://dhsprogram.com/ (accessed on 30 July 2023)) [51]. The UN Inter-agency Group for Child Mortality Estimates also provides child mortality data (https://childmortality.org/ (accessed on 30 July 2023)) [52].

Gross Domestic Product (GDP) is a measurement of a country's economic output. Conflicts may arise because of economic instability brought on by a low or decreasing GDP value. For instance, societal instability and sometimes even violence might result from poverty and high unemployment rates. Income inequality may rise when GDP is concentrated in the hands of a tiny segment of the population, which can result in social dissatisfaction and perhaps violent outbursts. Low or declining GDP can also contribute to political instability. Political unrest can occasionally result from a drop in GDP because governments may find it difficult to provide the fundamental necessities for their populace. Thus, violence or conflicts may be influenced by GDP, which can be selected as a variable, i.e., predictor for conflict susceptibility [41,42]. Some studies suggest that GDP per capita (PPP) has less importance than other socioeconomic indicators. This may be surprising since low economic development is often mentioned as a major risk factor for conflicts [53]. Thus, the effect of declining GDP will have to be further investigated. Nine papers out of the eighteen have referred to declining or low GDP as a contributing factor to conflicts. The GDP data are presented on a country-level resolution in The World Bank data portal (https://data.worldbank.org/indicator/NY.GDP.MKTP.CD accessed on 30 July 2023) [49]. However, analyses of GDP on the sub-national or local level presents difficulties.

Ethnicity and religious aspects may be one of the factors in violence and conflicts because of their tendency to unite diverse groups of people and give them a sense of common identity and purpose. Ethnicity-related conflicts often happen when two or more groups contest for control of resources or other factors. Such conflicts may also be a result of long-held grudges, bias, or perceived injustices. Grave repercussions, including displacement, violence, and death, may result from these confrontations for the persons involved. Many scholars have used ethnicity factors in their attempt to predict the likelihood of conflicts [8,41,53,54]. Ten papers out of the eighteen have referred to ethnicity factors playing a contributing role in conflicts. The ETH Zurich provides the Ethnic Power Relations (EPR) Dataset Family 2021 at https://icr.ethz.ch/data/epr/ (accessed on 30 July 2023) [55].

Urban accessibility, infrastructure, and transportation hubs can easily become the outbreak site of conflicts since they play a key role for controlling territories and conflict logistics [56]. Road distance has been selected as one of the risk factors by scholars as routes for transporting people and supplies. Roads are often key targets for military activity, although they may also serve as a tool for a central government to secure control over a country's territory [35]. Four papers out of the eighteen have referred to urban accessibility as a risk factor in conflicts. Available road network datasets for the areas of interest can be used in research.

Unemployment. High unemployment rates can result in insecurity and financial instability, which may increase the chance of violence. Six papers out of the eighteen included unemployment as a covariance in their model for the prediction of the likelihood of conflicts and violence. The International Labor Organization provides statistical data on unemployment (ILO statistical data) at https://ilostat.ilo.org/data/ (accessed on 30 July 2023) [57].

Accessibility and social services include an indicator for cell phone coverage, an index of facilities available in each town (e.g., wells, latrines, schools, and clinics), and an estimate of the distance from each town to the nearest usable road. Accessibility increases the likelihood of public goods provision and of third-party intervention to prevent or mitigate violence [27]. Various available accessibility and social services datasets for the areas of interest can be used in research.

Other socioeconomic factors, such as access to land, the nighttime light index, access to land, corruption, income disparity, and homicide, have received less attention.

Land access has a significant role in wars, especially in developing nations. Land grabbing and scarcity can both result in forceful eviction and violence, as can competitiveness and disputes between various groups. Land is a very important resource that frequently causes disputes over ownership since it is necessary for cultivation, habitation, and economic growth. Access to land was used as a factor of conflicts in the paper "Predicting local violence: Evidence from a panel survey in Liberia" [27].

Corruption can contribute to violence by increasing protests and mistrust between different groups and governments and disbelieving in the ability of institutions to provide security and justice. It can lead to a lack of access to resources and increase social and economic inequality following unrest and violence. Corruption can also weaken the ability of the international community to intervene and assist in resolving conflicts. Corruption as a factor leading to a conflict was mentioned in the article "Revisiting the Contested Role of Natural Resources in Violent Conflict Risk through Machine Learning" [58] and The Global Conflict Risk Index (GCRI) by the JRC [47].

Income inequality can contribute to conflicts by creating poverty, marginalizing a population or group, and exacerbating existing divisions. It can leave people feeling unfairness, frustration, anger, and hopelessness, leading to social unrest and violence. Extremist groups can also use income inequality to mobilize support and recruit new members. The same is true for the corruption factor; income inequality was mentioned in "Revisiting the Contested Role of Natural Resources in Violent Conflict Risk through Machine Learning" [58] and "The Global Conflict Risk Index" (GCRI) by the JRC [47].

Homicide may exacerbate conflicts by instilling a sense of dread and distrust between opposing groups. This can result in fights over revenge and a general increase in violence. Homicide rates are another tool for measuring the amount of violence in society, and high rates may indicate a higher likelihood of major conflicts. The homicide index was used as a conflict factor in "The Global Conflict Risk Index" (GCRI) by the JRC [47].

The nighttime light index, which measures the intensity of light emitted from human settlements, collected by satellite imagery, can be used as a covariate to predict potential conflicts. For example, a high level of light emission can indicate high economic activity, population density, and urbanization, which can be associated with high socioeconomic standards. The use of this index as a covariate in conflict prediction models can enhance the accuracy of predictions by adding a spatial and temporal dimension to the conflict susceptibility analysis. The nighttime light index was used in three scientific papers [19,56].

#### 3.2.2. Agriculture

Agricultural resources may be one of the factors that causes conflicts and violence. The scarcity of farming resources, such as land, water, and crops, is mentioned as one of the causes of violence, for instance, in the complex war in Darfur, Sudan [59,60]. Thus, the unequal distribution of resources in agricultural systems can exacerbate social and political unrest and lead to violence. Conflicts can also result from the exploitation of farmers and the deterioration of the land, which can also result in policies that serve the interests of various businesses. Therefore, addressing the agriculture-related issues and promoting sustainable practices is critical to promoting peace and stability and reducing the risk of conflicts.

The agriculture conditioning groupings include four conditioning factors: crops, harvests, irrigation, and vegetation index. Most papers mentioned crops as conditioning factors related to vegetation index and harvests, while irrigation was mentioned in one article. The majority of these publications were released in the last five years, from 2017 to 2022. "Climate variability and conflict risk in East Africa, 1990–2009" received 185 citations [35]. Therefore, this article had the greatest impact on the number of referenced articles and the priority of factors such as "Vegetation Index".

Crops and harvests may become contributing factors in violence and conflicts in developing countries where agriculture is a main source of livelihood. Food security can be used as a tool to manipulate shortages, starvation, and poverty. On the other hand, food security issues, such as harvest failures brought on by pests, drought, or other factors can also spark societal unrest and escalate into violent confrontations. Having control over lucrative crops and rich land may also be a source of strength and influence, making it a target in disputes. Crop and agricultural infrastructure loss can also have long-term economic and social repercussions, resulting in food shortages and impeding post-conflict reconstruction challenges. Crops and harvest were used in several scientific papers using machine learning to predict the likelihood of conflicts [40,58]. Data related to harvest and agriculture can be found on the Statistical data website of the Food and Agriculture Organization of the United Nations [61].

Vegetation Index or Normalized Difference Vegetation Index (NDVI) is a remote sensing indicator that measures the health of vegetation. It has been used in conflict prediction through the indication of changes in vegetation patterns and health. The NDVI offers data on the condition and density of the vegetation, which might serve as evidence of human activity, including conflicts. NDVI values can be used to identify significant changes in vegetation, proving food insecurity and increased likelihood of conflict. For example, the NDVI index has been used to monitor land degradation caused by overgrazing, deforestation, and other human activities. NDVI has also been used to monitor the effects of climate change on agriculture, such as changes in rainfall patterns, temperature, and other environmental factors playing an important role in the decrease in land productivity. Thus, NDVI is a valuable tool in predicting the likelihood of conflicts as it can indicate potential food insecurity and possible competition for resources. Scholars such as O'Loughlin [35,56] have used vegetation index data to identify the areas that are most vulnerable to conflict through machine learning algorithms. Data related to the vegetation index can be produced from Landsat satellite images. Landsat data and NDVI data are publicly available on the United States Geological Survey (USGS) website (https://www.usgs.gov/landsat-missions/data (accessed 30 July 2023)) [62] and https://modis.gsfc.nasa.gov/data/dataprod/mod13.php (accessed on 30 July 2023) [63].

### 3.2.3. Natural Resources

Natural resources, including minerals, oil, and gas, have long been identified as contributing factors to conflict. Competition over access to these resources can create tensions between communities and or conflict parties. For example, the exploitation of minerals, such as diamonds, gold, and coltan, have been linked to conflicts in several countries, including Sierra Leone, the Democratic Republic of Congo, and Angola [64]. In this article, "Natural Resources, Conflict and Security Challenges in Africa", the authors posed the question if natural resources are a curse or blessing and also addressed the complex and multidimensional relationship between natural resources and conflict. Natural resources play a significant part in initiating and sustaining conflicts [65].

Diamonds, a precious mineral sought after for their beauty and rarity, have contributed to conflicts in several countries. This is because diamonds are often found in regions that are home to ethnic or religious groups with long-standing tensions, and the profits from the exploitation of these resources are often used to finance armed groups and perpetuate conflict. The diamond trade in Sierra Leone, which financed the conflict in the 1990s and early 2000s, is one of the examples of the connection between diamonds and conflicts [66]. Armed groups in Sierra Leone exploited the mining of diamonds to fund their operations and keep the conflict going. Several scholars used data on the presence of diamonds as a conditioning factor for conflict prediction using machine learning techniques [19,40,46,67]. Publicly available data related to diamond deposits can be found at the website of the Peace Research Institute Oslo (PRIO; https://www.prio.org/data/10 (accessed on 30 July 2023)) [68].

Oil exploration has played a role in the eruption of conflicts around the world. It is a commodity that is essential for the functioning of modern economies. It is often located in regions with political and social instability. Oil-rich regions often experience internal conflict between local communities and the central government or oil companies [69,70]. In this study, seven articles used oil datasets in predictive models for the likelihood of conflicts. Fuel export data can be found on the World Bank data portal (https://data.worldbank.org/indicator/TX.VAL.FUEL.ZS.UN (accessed on 30 July 2023)) [49].

Precious metal exploration, like oil exploration and other minerals, has contributed to the eruption of conflict. The regions with political and social instability often contain valuable mineral resources, such as gold, diamonds, and cobalt. The extraction and trade of these resources can also add to existing conflicts and provide funding for non-state actors. In regions with weak government structures and limited access to services, exploiting mineral resources can lead to tensions between various powers and local communities. For instance, in the Democratic Republic of Congo, the exploitation of mineral resources, including cobalt and tin, has contributed to the ongoing conflict in this region. Local communities often face exploitation, human rights abuses, and environmental degradation [71–73]. Five papers related to the use of precious metals as a conditioning factor in the prediction models for the likelihood of conflicts. Global mining data are available at the FINEPRINT Project website, funded by European Research Council (ERC) (https://www.fineprint.global/visualisations/viewer/ (accessed on 30 July 2023)) [74].

#### 3.2.4. Political and Governance Aspects

Political and governance aspects in societies play an important role in conflict because they shape the distribution of power and resources. In countries prone to violence, political institutions frequently lack legitimacy, transparency, accountability, and the rule of law. This inadequate governance can contribute to conflict by creating an environment where people and groups feel marginalized and alienated and can create resource disputes. In some cases, local communities or tribe leaders perform the role of the government [75]. Additionally, political and governance systems can lead to conflict by exploiting ethnic, religious, and other identity-based distinctions [76]. For instance, the political elite may utilize these distinctions in other countries as a threat to state stability to preserve their power. Finally, political and governance systems can contribute to conflicts through corruption since the theft of resources and abuse of authority can result in social and economic inequality, generating anger and conflict [77,78]. In this research, we identified four priority contributors to conflicts; they are democracy, human rights, elections, and ethnic leadership.

The political and governance groupings include three conditioning factors: democracy, human rights, and elections. Most of these publications were released in the last five years, from 2017 to 2022. The overall priority of these conditioning factors was eight points for democracy, five points for human rights, and four points for elections.

The lack of democracy can be a prominent contributor to conflicts. When the right to participate in the political process is disrespected and the people's opinions are unheard, frustration and resentment can arise, contributing to social unrest and violent conflict. In non-democratic societies, power is frequently held by an elite group, which can lead to corruption, economic inequality, and human rights violations. The lack of democracy can create a situation when conflicts are more likely to occur, and they will be challenging to resolve. Without democratic institutions ensuring free and fair elections, independent judiciary institutions and free media can also prevent conflicts from being resolved peacefully. Thus, the lack of democracy can be a significant factor in the eruption and persistence of conflicts [79,80]. The expression of various and opposing viewpoints may result from the freedom of speech and expression that is guaranteed in democratic countries [81]. There were 11 publications related to the application of the democracy index in the process of predicting the likelihood of conflicts. Data on democracy can be found on the Center for Systemic Peace POLITY 5 Annual Time-Series 1946–2018 (https://www.systemicpeace.org/inscrdata.html (accessed on 30 July 2023)).

Human rights are a key component of the peaceful resolution of conflicts and maintaining peace and stability in society. The beginning of conflicts can start because of rights violations. Denial of fundamental liberties such as freedom of expression, assembly, and religion can frustrate individuals and lead to violence. The dispute may be addressed by protecting and advancing human rights. When individuals see that their rights are being safeguarded, it can facilitate the development of trust between warring parties or between governments. Thus, protecting human rights or their lack thereof can significantly contribute to prediction models for the likelihood of conflicts [82]. A number of research papers used human-rights-related datasets as a covariant for the prediction of the likelihood of conflicts [34,39]. The CIRI Human Rights Dataset contains standards-based quantitative information on government respect for 15 internationally recognized human rights for 202 countries annually, from 1981 to 2011 [83]. The dataset can be downloaded at https://dataverse.harvard.edu/dataverse/cirihumanrightsdata (accessed 1 August 2023).

Elections. Voting and electing representatives is a basic right in any democratic society. Elections play a crucial role in resolving conflicts and promoting peace. Often, fair and peaceful elections are considered a sign of recovery of a country after protracted conflicts, as it was in 2012 in Timor-Leste [84]. On the other hand, an election that is perceived as unfair can lead to conflicts and undermine peace and stability in society. Thus, following "unfair elections", tensions can arise between different groups and the government. This can erode public trust in the political system. "Unfair elections" can also undermine the legitimacy of elected officials and also the institutions they represent [85,86]. A number of scholars used electoral processes as a conditioning factor in their research [35,87]. The Rulers, Elections, and Irregular Governance (REIGN) dataset describes political conditions [88]. The data are available and can be downloaded at https://oefdatascience.github.io/REIGN.github.io/ (accessed on 30 July 2023).

#### 3.2.5. Climate Change

In the recent publications related to conflicts and the main drivers of conflicts, climate change was recognized as one of the significant elements leading to conflicts [8,9,19,23,56,59]. The livelihood of the people and the accessibility of natural resources like water, food, and land are substantially impacted by changes in precipitation patterns, temperature, and sea levels. Conflict and rivalry over these resources may result from such effects. Climate change may also make natural disasters like floods, droughts, and heat waves more frequent and severe, causing more hardship for populations who are already at risk. Climate change-related population movement can also result in resource rivalry and elevated social tensions in receiving communities. Thus, by escalating already-existing conflicts and igniting new ones, climate change can cause political instability. It is a significant cause of conflict and instability and a thorough response to address its effects and avert other conflicts is needed.

Climate change includes five factors: weather shock, mentioned in ten publications; extreme temperatures, mentioned in nine publications; precipitation, in eight publications; and flood and drought, mentioned in one publication. Most of the reviewed papers were published in the last five years, from 2017 to 2022. For example, "Climate variability and conflict risk in East Africa, 1990–2009" received 185 citations [35]. Therefore, this publication had the greatest impact on the priority of factors such as "Temperature" and "Precipitation". The priorities of the conditioning factors related to climate change are as follows: temperature has seven points, precipitation has six points, and weather shock and natural disasters each have 4 points. There are many data sources available on climate change where specific datasets can be found: the International Monetary Fund (IMF) Climate Change Dashboard (https://climatedata.imf.org/ (accessed on 30 July 2023)) [89]; the National Oceanic and Atmospheric Administration (NOAA) Climate data (https://www.climate.gov/maps-data (accessed on 30 July 2023)) [90]; the Data Distribution Center of the International Panel on Climate Change at https://www.ipcc-data.org/ (accessed on 30 July 2023) [91]; and as part of the United States' Climate Data Initiative (climate.data.gov) which provides access to Federal resources to help America's communities, businesses, and citizens plan and prepare for climate change [92].

## 4. Conclusions and Discussion

Armed conflicts continue to pose a threat to human societies worldwide, and the identification of the underlying causes and drivers is crucial. Various conditioning factors have contributed to conflicts, including poverty, income inequality, weak governance, and ethnic fractionalization. However, there is still no consensus among scholars on the predictors and conditioning factors that influence the exacerbation of conflicts.

This study evaluated the literature on the application of machine learning to predict the likelihood of conflict escalation and the role of conditioning factors. The results showed that machine learning and predictive models could help identify conflict-prone locations and understand the factors that contribute to conflict escalation. We found 46 relevant papers, but the amount of literature on machine learning susceptibility mapping and conditioning factors was limited. Moreover, the application of machine learning technology to predict the likelihood of conflicts gained attention from academics only within the last five years. This study provides research and data sources for each proposed conditioning factor.

We defined a methodology for conflict susceptibility and machine learning, which involves the use of training and validation datasets and conditioning factors. The application of conditioning factors is crucial for accurate predictions. However, it must be noted that no conditioning factor can be applied to any or all environments due to socio-cultural differences. Each geographical location, each country, and each conflict should be studied through the prism of a unique set of conditioning factors. Therefore, to accurately predict the likelihood of conflict escalation, we must identify a unique list of conditioning factors that can be applied to that specific geographic, political, or social scenario.

Academic research, especially in international relations and political science, uses conflict data. The Armed Conflict Location & Event Data Project (ACLED) and Uppsala Conflict Data Program (UCDP) are the most utilized datasets. ACLED tracks political violence and protests around the world, whereas UCDP tracks non-state armed conflicts. Scholars, international organizations, government agencies, and media news outlets freely utilize both databases. "Climate variability and conflict risk in East Africa, 1990–2009", citing the ACLED dataset, is the most referenced conflict study. The training and validating models used both ACLED and UCDP, with ACLED being given preference.

Socioeconomic variables strongly influence conflict susceptibility and strongly link socioeconomic factors and conflict. However, socioeconomic data may be unreliable, making sub-national analysis difficult. In developing countries, agricultural factors are a significant cause of conflicts. Unrest and violent conflicts can result from resource scarcity, unequal distribution, farmer exploitation, and land degradation. Access to natural resources might cause conflicts between communities and parties. Diamonds and valuable metals like gold and cobalt cause rivalries among parties. Mineral resource extraction, especially in unstable regions, can fuel non-state groups and contribute to human rights abuses, environmental damage, and exploitation. Another agriculture-related factor, climate change, also contributes to conflict and instability. It affects the livelihoods and food security of vulnerable communities. In addition, climate change causes natural disasters, migration, and political instability. Political and governance factors can determine the likelihood of conflicts. Democracy, human rights, and elections were the key conditioning elements. The democracy index, CIRI human rights dataset, and REIGN dataset are useful for investigating these conditioning elements in predicting the likelihood of conflicts. Political and governance factors must be considered to understand and prevent conflicts.

It should be noted that conflict dynamics are constantly evolving, and new types and elements may emerge in yet-to-be-studied scenarios. Additionally, certain conditioning factors may have been overlooked due to limitations in data availability during research studies. Therefore, we perceive this research as a living process, adaptable and open to future enrichment with additional conditioning factors as the field of conflict prediction studies continues. As the application of machine learning to conflict prediction gains momentum, we anticipate an increase in the identification and understanding of critical conflict conditioning factors, further enhancing the accuracy and precision of conflict prediction models.

Machine learning can potentially become an indispensable tool for conflict reduction and prevention. Understanding the elements that contribute to conflict escalation and their links to a particular location's society, politics, and geography enables the development of more effective conflict prevention and resolution techniques. However, the availability and quality of data pose obstacles to using machine learning for conflict prediction. In addition, the list of objective conditioning elements must be carefully selected. It is believed that this study presents a viable framework for using machine learning in conflict susceptibility research. Future studies should be focused on refining and enhancing predictive approaches.

**Supplementary Materials:** The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/ijgi12080322/s1, File S1: List of publications used for the extraction of conditioning factors and predictors.

**Author Contributions:** Conceptualization, Timur Obukhov and Maria A. Brovelli; Formal analysis, Timur Obukhov; Investigation, Timur Obukhov; Methodology, Timur Obukhov and Maria A. Brovelli; Supervision, Maria A. Brovelli; Validation, Maria A. Brovelli; Writing—original draft, Timur Obukhov; Writing—review and editing, Timur Obukhov and Maria A. Brovelli. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding. The APC was waived. Author voucher discount code was provided.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** All data used in this research are publicly available.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Gleditsch, P.N.; Wallensteen, P.; Eriksson, M.; Sollenberg, M.; Strand, H. Armed Conflict 1946–2001: A New Dataset. *J. Peace Res.* **2002**, *39*, 615–637. [CrossRef]
2. United Nations General Assembly. Resolution A/RES/57/337 Resolution Adopted by the General Assembly [without Reference to a Main Committee (A/57/L.79)] 57/337. Prevention of Armed Conflict. 2003. Available online: https://documents-dds-ny.un.org/doc/UNDOC/GEN/N02/563/59/PDF/N0256359.pdf (accessed on 30 July 2023).
3. United Nations General Assembly. Transforming Our World: The 2030 Agenda for Sustainable Development. A/RES/70/1. 2015. Available online: https://www.refworld.org/docid/57b6e3e44.html (accessed on 30 July 2023).
4. Kumar, R.; Roy, P. War and peace: Is our world serious about achieving Sustainable Development Goals by 2030? *J. Fam. Med. Prim. Care* **2018**, *7*, 1153–1156. [CrossRef] [PubMed]
5. Davies, S.; Pettersson, T.; Öberg, M. Organized violence 1989–2022 and the return of conflicts between states? *J. Peace Res.* **2023**, *60*. [CrossRef]
6. Gates, S.; Hegre, H.; Nygard, H.M.; Strand, H. Development Consequences of Armed Conflict. *World Dev.* **2012**, *40*, 1713–1722. [CrossRef]
7. Bizău, A.; Stanciulescu, R. Causes of Armed Conflict. *Sciendo* **2022**, *27*, 171–177. [CrossRef]
8. Schleussner, C.; Donges, J.; Donner, R.; Schellnhuber, H. Armed-conflict risks enhanced by climate-related disasters in ethnically fractionalized countries. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 9216–9221. [CrossRef]
9. Slettebak, R.T. Don't blame the weather! Climate-related natural disasters and civil conflict. *J. Peace Res.* **2012**, *51*, 573–589. [CrossRef]
10. Buhaug, H.; Cederman, L.-E.; Gleditsch, K.S. Square Pegs in Round Holes: Inequalities, Grievances, and Civil War. *Int. Stud. Q.* **2014**, *58*, 418–431. [CrossRef]
11. Elbadawi, I.; Hegre, H. Globalization, Economic Shocks, and Internal Armed Conflict. *Def. Peace Econ.* **2008**, *19*, 37–60. [CrossRef]
12. Fearon, J.D. Governance and Civil War Onset. World Bank. 2011. Available online: https://openknowledge.worldbank.org/handle/10986/9123 (accessed on 30 July 2023).
13. Hegre, H.; Karlsen, J.; Nygård, H.M.; Strand, H.; Urdal, H. Predicting Armed Conflict, 2010–2050. *Int. Stud. Q.* **2013**, *57*, 250–270. Available online: http://www.jstor.org/stable/24016137 (accessed on 30 July 2023). [CrossRef]
14. Collier, P.; Hoeffler, A. Greed and Grievance in Civil War. *Oxf. Econ. Pap.* **2004**, *56*, 563–595. [CrossRef]
15. Collier, P.; Hoeffler, A.; Rohner, D. Beyond greed and grievance: Feasibility and civil war. *Oxf. Econ. Pap.* **2009**, *61*, 1–27. [CrossRef]
16. Toft, M.D. Indivisible territory, geographic concentration, and ethnic war. *Secur. Stud.* **2002**, *12*, 82–119. [CrossRef]
17. Denny, E.K.; Walter, B.F. Ethnicity and civil war. *J. Peace Res.* **2014**, *51*, 199–212. [CrossRef]

18. IPCC. Climate Change 2022: Impacts, Adaptation, and Vulnerability. In *Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*; Pörtner, H.-O., Roberts, D.C., Tignor, M., Poloczanska, E.S., Mintenbeck, K., Alegría, A., Craig, M., Langsdorf, S., Löschke, S., Möller, V., et al., Eds.; Cambridge University Press: Cambridge, UK; New York, NY, USA, 2022. [CrossRef]
19. Hao, M.; Ding, F.; Xie, X.; Fu, J.; Qian, Y.; Ide, T.; Maystadt, J.-F.; Chen, S.; Ge, Q.; Jiang, D. Varying climatic-social-geographical patterns shape the conflict risk at regional and global scales. *Humanit. Soc. Sci. Commun.* **2022**, *9*, 276. [CrossRef]
20. De Groeve, T.; Poljansek, K.; Vernaccini, L. *InfoRM Concept and Methodology JRC Report 20140127*; Publication Office of the European Union: Luxembourg, 2014. [CrossRef]
21. Hegre, H.; Buhaug, H.; Calvin, K.; Nordkvelle, J.; Waldhoff, S.; Gilmore, E. Forecasting civil conflict along the shared socioeconomic pathways. *Environ. Res. Lett.* **2016**, *11*, 054002. [CrossRef]
22. Kalyvas, S.N. *The Logic of Violence in Civil War*; Cambridge University Press: Cambridge, UK, 2006. [CrossRef]
23. Mach, K.J.; Adger, W.N.; Buhaug, H.; Burke, M.; Fearon, J.D.; Field, C.B.; Hendrix, C.S.; Kraan, C.M.; Maystadt, J.F.; O'loughlin, J.; et al. Directions for research on climate and conflict. *Earth Future* **2020**, *8*, e2020EF001532. [CrossRef] [PubMed]
24. Garcia, S.; Herrera, F. An Extension on "Statistical Comparisons of Classifiers over Multiple Data Sets" for all Pairwise Comparisons. *J. Mach. Learn. Res.—JMLR* **2008**, *9*, 2677–2694.
25. Dietterich, T.G. Approximate Statistical Tests for Comparing Supervised Classification Learning Algorithms. *Neural Comput.* **1998**, *10*, 1895–1923. [CrossRef]
26. Fawcett, T. An introduction to ROC analysis. *Pattern Recognit. Lett.* **2006**, *27*, 861–874. [CrossRef]
27. Blair, R.A.; Blattman, C.; Hartman, A. Predicting local violence: Evidence from a panel survey in Liberia. *J. Peace Res.* **2017**, *54*, 298–312. [CrossRef]
28. Linke, A.M.; Witmer, F.D.W.; O'Loughlin, J. Weather variability and conflict forecasts: Dynamic human-environment interactions in Kenya. *Political Geogr.* **2022**, *92*, 102489. [CrossRef]
29. Armed Conflict Location & Event Data Project (ACLED). Public Data. 2023. Available online: www.acleddata.com (accessed on 6 January 2023).
30. Uppsala Conflict Data Program (UCDP). Publicly Available Data. 2022. Available online: https://ucdp.uu.se/ (accessed on 6 January 2023).
31. Fearon, J.; Laitin, D. Ethnicity, Insurgency, and Civil War. *Am. Political Sci. Rev.* **2003**, *97*, 75–90. [CrossRef]
32. Cederman, L.; Weidmann, N.; Gleditsch, K. Horizontal Inequalities and Ethnonationalist Civil War: A Global Comparison. *Am. Political Sci. Rev.* **2011**, *105*, 478–495. [CrossRef]
33. Miguel, E.; Satyanath, S.; Sergenti, E. Economic Shocks and Civil Conflict: An Instrumental Variables Approach. *J. Political Econ.* **2004**, *112*, 725–753. [CrossRef]
34. Web of Science. Available online: https://www.webofscince.com (accessed on 6 January 2023).
35. O'Loughlin, J.; Witmer, F.; Linke, A.; Laing, A.; Gettelman, A.; Dudhia, J. Climate variability and conflict risk in East Africa, 1990–2009. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 18344–18349. [CrossRef]
36. Obukhov, T.; Brovelli, M. Defining a methodology for integrating semantic, geospatial, and temporal techniques for conflict analysis. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* **2022**, *XLIII-B4-2022*, 155–161. [CrossRef]
37. Salehyan, I. Best practices in the collection of conflict data. *J. Peace Res.* **2015**, *52*, 105–110. [CrossRef]
38. ACLED Terms of Use and Attribution Policy. Last Update 8 June 2022. Available online: https://acleddata.com/terms-of-use/ (accessed on 30 July 2023).
39. Sundberg, R.; Eck, K.; Kreutz, J. Introducing the UCDP Non-State Conflict Dataset. *J. Peace Res.* **2012**, *49*, 351–362. [CrossRef]
40. Musumba, M.; Fatema, N.; Kibriya, S. Prevention is better than cure: Machine learning approach to conflict prediction in sub-Saharan Africa. *Sustainability* **2021**, *13*, 7366. [CrossRef]
41. Fawaz, M.; Le Quellec, E. Indirect rivalries and civil wars: Empirical evidence. *Def. Peace Econ.* **2022**, 1–28. [CrossRef]
42. Pinckney, J.; Rezaee Daryakenari, B. When the levee breaks: A forecasting model of violent and nonviolent dissent. *Int. Interact.* **2022**, *48*, 997–1026. [CrossRef]
43. Urdal, H. A clash of generations? Youth bulges and political violence. *Int. Stud. Q.* **2006**, *50*, 607–629. [CrossRef]
44. Yair, O.; Miodownik, D. Youth bulge and civil war: Why a country's share of young adults explains only non-ethnic wars. *Confl. Manag. Peace Sci.* **2016**, *33*, 25–44. [CrossRef]
45. Florczyk, A.; Ehrlich, D.; Corbane, C.; Freire, S.; Kemper, T.; Melchiorri, M.; Pesaresi, M.; Politis, P.; Schiavina, M.; Zanchetta, L. *Community Pre-Release of GHS Data Package (GHS CR2018) in Support to the GEO Human Planet Initiative*; Publications Office of the European Union: Luxembourg, 2018; Available online: https://data.jrc.ec.europa.eu/dataset/jrc-ghsl-GHS_pop_gpw4_globe_r2015a (accessed on 30 July 2023).
46. Perry, C. Machine Learning and Conflict Prediction: A Use Case. *Stab. Int. J. Secur. Dev.* **2013**, *2*, 56. [CrossRef]
47. European Commission; Joint Research Centre; Thomakos, D.; Has, S.; Saporiti, F.; Stamatia, K.; Stefano, F. Global Conflict Risk Index (GCRI): New variables in 2018, Publications Office of the European Union, 2018. Available online: https://data.europa.eu/doi/10.2760/258293 (accessed on 30 July 2023).
48. Malembaka, E.B.; Altare, C.; Bigirinama, R.N.; Bisimwa, G.; Banywesize, R.; Tabbal, N.; Boerma, T. The use of health facility data to assess the effects of armed conflicts on maternal and child health: Experience from the Kivu, DR Congo. *BMC Health Serv. Res.* **2021**, *21* (Suppl. 1), 195. [CrossRef]

49. World Bank. Mortality Rate, under-5 (per 1000 Live Births). 2023. Available online: https://data.worldbank.org/indicator/SH.DYN.MORT (accessed on 5 January 2023).
50. WHO. World Health Organization Data. 2023. Available online: https://www.who.int/data (accessed on 25 December 2022).
51. DHS. The Demographic and Health Survey Program Website. Funded by USAID. 2023. Available online: http://www.dhsprogram.com (accessed on 5 January 2023).
52. IGME. UN Inter-Agency Group for Child Mortality Estimates. Child Mortality and Stillbirth Estimates Dataset. 2023. Available online: https://childmortality.org/ (accessed on 5 January 2023).
53. Hoch, J.M.; de Bruin, S.P.; Buhaug, H.; Von Uexkull, N.; van Beek, R.; Wanders, N. Projecting armed conflict risk in Africa towards 2050 along the SSP-RCP scenarios: A machine learning approach. *Environ. Res. Lett.* **2021**. [CrossRef]
54. Sambanis, N.; Shayo, M. Social identification and ethnic conflict. *Am. Political Sci. Rev.* **2013**, *107*, 294–325. [CrossRef]
55. Manuel, V.; Bormann, N.; Rüegger, S.; Cederman, L.; Hunziker, P.; Girardin, L. Integrating data on ethnicity, geography, and conflict: The Ethnic Power Relations data set family. *J. Confl. Resolut.* **2015**, *59*, 1327–1342.
56. Ge, Q.; Hao, M.; Ding, F.; Jiang, D.; Scheffran, J.; Helman, D.; Ide, T. Modelling armed conflict risk under climate change with machine learning and time-series data. *Nat. Commun.* **2022**, *13*, 2839. [CrossRef]
57. International Labor Organization. ILO Statistics Data. 2023. Available online: https://ilostat.ilo.org/data/ (accessed on 30 July 2023).
58. Schellens, M.K.; Belyazid, S. Revisiting the contested role of natural resources in violent conflict risk through machine learning. *Sustainability* **2020**, *12*, 6574. [CrossRef]
59. Hakim, S. The role of climate change in the Darfur crisis. In *The Economic, Social and Political Elements of Climate Change: Climate Change Management*; Leal Filho, W., Ed.; Springer: Berlin/Heidelberg, Germany, 2011; pp. 721–735. [CrossRef]
60. Brown, I.A. Assessing eco-scarcity as a cause of the outbreak of conflict in Darfur: A remote sensing approach. *Int. J. Remote Sens.* **2010**, *31*, 2513–2520. [CrossRef]
61. Food and Agriculture Organization of the United Nations. FAO Statistics Data. 2023. Available online: https://www.fao.org/faostat/en/#home (accessed on 30 July 2023).
62. US Geological Survey. Landsat. Available online: https://www.usgs.gov/landsat-missions/data (accessed on 30 July 2023).
63. US Geological Survey. Moderate Resolution Imaging Spectrodiometer (MODIS) Vegetation Index Products (NDVI and EVI). 2018. Available online: https://modis.gsfc.nasa.gov/data/dataprod/mod13.php (accessed on 30 July 2023).
64. Olanrewaju, F.O.; Joshua, S.; Olanrewaju, A. Natural resources, conflict and security challenges in Africa. *India Q.* **2020**, *76*, 552–568. [CrossRef]
65. Maphosa, S.B. Natural Resources and Conflict: Unlocking the Economic Dimension of Peace-Building in Africa (Policy Brief, Africa Institute of South Africa, Briefing No 74), 2012. Available online: https://studylib.net/doc/8933896/ (accessed on 1 August 2023).
66. Mokuwa, E.; Voors, M.; Bulte, E.; Richards, P. Peasant grievance and insurgency in Sierra Leone: Judicial serfdom as a driver of conflict. *Afr. Aff.* **2011**, *110*, 339–366. [CrossRef]
67. Lujala, P.; Gleditsch, N.P.; Gilmore, E. A diamond curse? Civil war and a lootable resource. *J. Confl. Resolut.* **2005**, *49*, 538–562. [CrossRef]
68. Gilmore, E.; Gleditsch, N.P.; Lujala, P.; Rød, J.K. Conflict diamonds: A new dataset. *Confl. Manag. Peace Sci.* **2005**, *22*, 257–292. [CrossRef]
69. Gutiérrez Rodríguez, J.D. The connection between oil wealth and internal armed conflicts: Exploring the mechanisms of the relationship using a subnational lens. *Extr. Ind. Soc.* **2019**, *6*, 1217–1227. [CrossRef]
70. Bae, J.Y.; Heo, E. Armed conflicts in the Middle East and international oil company returns. *Energy Sources Part B Econ. Plan. Policy* **2018**, *13*, 21–27. [CrossRef]
71. Kelly, J.T.D. "This mine has become our farmland": Critical perspectives on the coevolution of artisanal mining and conflict in the Democratic Republic of the Congo. *Resour. Policy* **2014**, *40*, 167–174. [CrossRef]
72. Kranz, O.; Schoepfer, E.; Tegtmeyer, R.; Lang, S. Earth observation based multi-scale assessment of logging activities in the Democratic Republic of the Congo. *ISPRS J. Photogramm. Remote Sens.* **2018**, *149*, 111–126. [CrossRef]
73. Maystadt, J.F.; De Luca, G.; Sekeris, P.G.; Ulimwengu, J. Mineral resources and conflicts in DRC: A case of ecological fallacy? *Oxf. Econ. Pap.* **2014**, *66*, 721–749. [CrossRef]
74. FINEPRINT 2023. FINEPRINT Global Wirtschaftsuniversität Wien (WU) Institute for Ecological Economics. FINEPRINT Geovisualizations. 2023. Available online: https://www.fineprint.global/visualisations/viewer/ (accessed on 6 January 2023).
75. Taher, A.; Khan, Z.; Alduais, A.; Muthanna, A. Intertribal conflict, educational development, and education crisis in Yemen: A call for saving education. *Rev. Educ.* **2022**, *10*, e3376. [CrossRef]
76. Ashle, S.; Kim, J. Forced migration: Typology and local agency in Southeast Myanmar. *Contemp. Southeast Asia A J. Int. Strateg. Aff.* **2015**, *37*, 211.
77. Orjuela, C. Corruption and identity politics in divided societies. *Third World Q.* **2014**, *35*, 753–769. [CrossRef]
78. Lohaus, M.; Bussmann, M. The politics of survival or business as usual? Exploring the effects of armed conflict on corruption. *J. Int. Relat. Dev.* **2021**, *24*, 149–170. [CrossRef]
79. Hegre, H. Democracy and armed conflict. *J. Peace Res.* **2014**, *51*, 159–172. [CrossRef]
80. Park, J.; James, P. Democracy, territory, and armed conflict, 1919–1995. *Foreign Policy Anal.* **2015**, *11*, 85–107. [CrossRef]

81. Evans, A.L. Fighting words: Targeting speech in armed conflict. *Wash. Int. Law J.* **2021**, *30*, 6. Available online: https://digitalcommons.law.uw.edu/wilj/vol30/iss3/6 (accessed on 30 July 2023).
82. Drange, B. A tug of war: Pursuing justice amid armed conflict. *Nord. J. Hum. Rights* **2022**, *40*, 346–364. [CrossRef]
83. Cingranelli, D.L.; Richards, D.L.; Clay, K.C. The CIRI Human Rights Dataset, Version 2014.04.14. Harvard Dataverse. 2021. Available online: https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/UKCPXT (accessed on 3 January 2023).
84. Leach, M. Timor-Leste in 2012: Beyond International Statebuilding? *Asian Surv.* **2013**, *53*, 156–161. [CrossRef]
85. Cederman, L.-E.; Gleditsch, K.S.; Hug, S. Elections and Ethnic Civil War. *Comp. Political Stud.* **2012**, *46*, 387–417. [CrossRef]
86. Salehyan, I.; Linebarger, C. Elections and Social Conflict in Africa, 1990–2009. *St. Comp. Int. Dev.* **2015**, *50*, 23–49. [CrossRef]
87. Attinà, F.; Carammia, M.; Iacus, S.M. Forecasting change in conflict fatalities with dynamic elastic net. *Int. Interact.* **2022**, *48*, 649–677. [CrossRef]
88. Bell, C.; Besaw, C.; Frank, M. The Rulers, Elections, and Irregular Governance (REIGN) Dataset. One Earth Future. 2021. Available online: https://oefdatascience.github.io/REIGN.github.io/ (accessed on 3 January 2023).
89. International Monetary Fund. Climate Change Indicators Dashboard. 2022. Available online: https://climatedata.imf.org/pages/access-data (accessed on 23 March 2023).
90. National Oceanic and Atmospheric Administration. NOAA Climate Data. 2023. Available online: https://www.climate.gov/maps-data (accessed on 3 February 2023).
91. International Panel on Climate Change. IPCC DDC—Data Distribution Center. 2023. Available online: https://www.ipcc-data.org/ (accessed on 9 April 2023).
92. Climate Data Initiative. Climate Data.Gov. 2023. Available online: https://data.gov/climate/ (accessed on 9 April 2023).