Data Science Program

Capstone Report - Spring 2024

Mapping Conflict Related Sexual Violence

Jakob Aggers

Supervised by Professor Edwin Lo

Abstract:

Conflicts that contain sexual violence make up 15% of all armed conflicts since 1989. In order to reduce the prevalence and prominence of conflict related sexual violence, journalists and policy makers rely on data to influence public opinion and policy. Concerning sexual violence in armed conflict, data on the subject is not easy to access. The goal of this project is to create a resource for journalists and policy makers to view individual instances as well as aggregated data about sexual violence in armed conflict that is both easy to use and accessible.

1. Introduction

In the shadow of armed conflict, atrocities of violence often inflict profound suffering on civilian populations. Among these, conflict-related sexual violence (CRSV) is a particularly grievous violation that can devastate individuals, families, and communities. CRSV often transcends traditional warfare by serving as a deliberate tactic to instill fear, shatter social fabrics, and assert dominance over targeted groups. Despite its pervasiveness and severity, CRSV remains one of the least understood and underreported aspects of armed conflict.

The global response to CRSV has grown significantly in recent years, driven by the recognition that such violence has a broader, more enduring impact than previously acknowledged. Beyond the physical and psychological trauma inflicted on survivors, CRSV contributes to long-term social, economic, and political instability. Consequently, understanding the patterns, perpetrators, and contexts of CRSV is crucial for developing effective strategies for prevention, accountability, and survivor support.

This project seeks to address a critical gap in the research and analysis of CRSV by creating an accessible, data-driven resource that offers insights into the prevalence and patterns of sexual violence in armed conflict. Drawing on authoritative datasets like the Armed Conflict Location & Event Data Project (ACLED) and the Sexual Violence in Armed Conflict (SVAC) dataset, this project aims to empower journalists, policymakers, and researchers with comprehensive, actionable insights.

Through interactive visualizations, the project will provide a unique platform that not only highlights individual incidents but also illustrates broader trends in CRSV. By mapping these incidents across time and space, the project will facilitate a deeper understanding of the geospatial dynamics of CRSV, uncovering trends that are often obscured in broader studies of armed conflict.

The ultimate objective of this project is to enhance the ability of key stakeholders to address CRSV more effectively. By providing a comprehensive view of where, when, and how CRSV occurs, this resource aims to foster greater awareness, inform policy decisions, and ultimately contribute to the global effort to combat sexual violence in conflict settings. Through this project, I hope to bridge the data gap and offer a tool that can aid in the fight against CRSV.

2. Problem Statement

Conflict Related Sexual Violence Data is difficult to access, difficult to visualize, and difficult to understand. The goal of this project is to solve that problem by creating a source of information to be used for journalism and policy-making.

3. Related Work

Over the past few decades, researchers have worked diligently to understand and document conflict-related sexual violence (CRSV). The Armed Conflict Location & Event Data Project (ACLED) has been pivotal in understanding the geographic and temporal patterns of violence globally.

The Sexual Violence in Armed Conflict (SVAC) dataset has further refined the analysis of CRSV by collating data from multiple sources. Studies like Cohen and Nordås's "Sexual Violence in Armed Conflict: Introducing the SVAC Dataset, 1989–2009" (2014) offer structured data that facilitates research on the severity, prevalence, and actors involved in CRSV.

Scholarly research has also enriched the understanding of CRSV. Elisabeth Jean Wood's work, such as "Rape as a Practice of War: Toward a Typology of Political Violence" (2018), delves into the motivations behind sexual violence in conflicts. Kerry Crawford's "Wartime Sexual Violence: From Silence to Condemnation of a Weapon of War" (2017) investigates the global response to CRSV, outlining how international norms have evolved to address these crimes.

Data visualization tools have also advanced the field's understanding. The Uppsala Conflict Data Program (UCDP) offers interactive visualizations of conflict data, although it has a broader focus than just conflict related sexual violence. The Humanitarian Data Exchange (HDX), a vast repository of humanitarian data, including sexual violence, provides valuable visual insights for policy formulation.

In her paper, "In Data We Trust? A Comparison of UCDP GED and ACLED Conflict Events Datasets," Kristine Eck offers an in-depth comparison between UCDP GED and ACLED, identifying the strengths and weaknesses of both datasets. Eck notes that while UCDP GED offers more rigorous classification and actor identifiers, it has a narrower scope, focusing on lethal violence. ACLED, on the other hand, captures a broader range of conflict events, including non-lethal and non-violent activities, but introduces challenges in data quality, particularly in geolocation precision and actor identification. This paper emphasizes the importance of understanding the limitations and potential biases of these datasets before utilizing them in research.

These combined works have contributed significantly to understanding CRSV, but the field still needs specialized tools that focus on visualizing and interpreting data specific to sexual violence in conflict. This project aims to address this gap by building on this existing foundation.

4. Solution and Methodology

The solution I present to this problem is as follows: create a website that aids in the visualization and understanding of sexual violence in armed conflict data to help journalists and policy makers better understand and communicate about conflict related sexual violence.

Open-source data was taken from the following sources:

1.) Armed Conflict Location & Event Data Project (ACLED)
2.) Sexual Violence in Armed Conflict Dataset (SVAC)

The Armed Conflict Location & Event Data Project (ACLED) is a disaggregated dataset that uses the event as its primary unit of analysis. This encompasses both violent and non-violent incidents related to conflicts, with each event geo-referenced and assigned a specific date. The dataset includes a variety of variables: the date and time of the event, location data (with coordinates where possible), actors involved (such as governments, rebel groups, and militias), event type (like battles, protests, riots, and violence against civilians), and the number of fatalities resulting from the event. ACLED captures a wide range of conflict-related events beyond just armed violence, including protests, troop movements, and non-violent activities. Its design allows for a detailed analysis of localized conflict dynamics, providing insights into spatial and temporal trends in conflict-related incidents.

The Sexual Violence in Armed Conflict (SVAC) dataset uses a different unit of analysis: the conflict-actor-year. This unit tracks incidents of sexual violence perpetrated by specific actors in conflicts across each year. Its variables include conflict and actor identifiers (categorizing each incident by the conflict and actors involved), the year of occurrence (spanning from 1989 to 20021), the type and severity of the sexual violence incidents, and sources (such as human rights reports – from Amnesty International, Human Rights Watch, and the U.S. Department of State –, news media, and NGO reports). The SVAC dataset focuses specifically on conflict-related sexual violence, providing comprehensive information on the context, perpetrators, and types of sexual violence. It details incidents involving state forces, rebel groups, and other non-state actors, offering insights into patterns of sexual violence across various conflicts and regions.

I utilized primarily R-Studio to create a Shiny Application website that visualizes this data and information about conflict related sexual violence.

5. Results

Results can be viewed below or at the website. The dashboard is broken up into various parts which are displayed as tabs beneath the header. The first is “Instance Search,” followed by “Heat Map,” Case studies for Ukraine and Myanmar, “SVAC Information,” and “Data Journalism Example”.

The Instance Search tab is intended to visualize the ACLED data on a world map. Each instance of sexual violence in armed conflict populates on a marker that displays the event date, country where the incident occurred, the actor name, the secondary actor, the description of the event, and the source. These markers respond to search criteria in the banner: date range, country, region, actor type, and an overall keyword search. An example can be viewed in Figure 1. The Instance Search tab also shows the frequency and fatalities of the sexual violence data based on the search critera; a user can toggle between graphs of frequency or fatalities or both. An example is provided in Figure 1.2.

In a use-case scenario, a journalist or policy maker could use this tab to find specific instances of sexual violence in armed conflict that are particularly compelling or representative of what is occurring in a certain conflict.

Figure 1.1: An example of the instance search tab

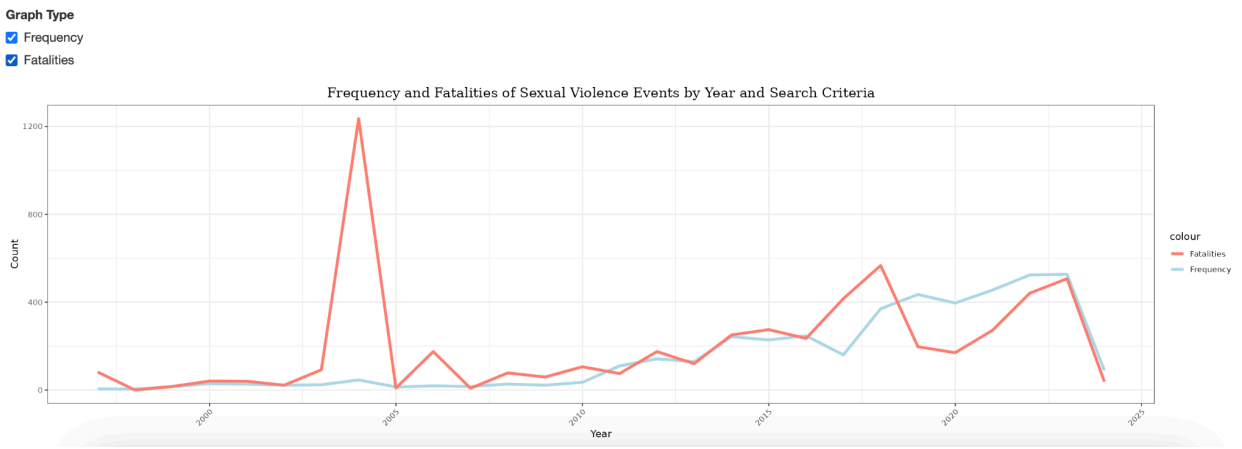

Figure 1.2: An example of the frequency and fatalities graphs which respond to search inputs

The heat map tab is intended to display a heatmap of sexual violence in armed conflict by instance and proximal location. The heatmap also responds to various search criteria. Figure 2 is an example of a heatmap of the hotspot locations for sexual violence in armed conflict in Sudan.

In a use-case scenario, a journalist or policy maker might use this tool to investigate where SVAC instances are congregated within a certain conflict.

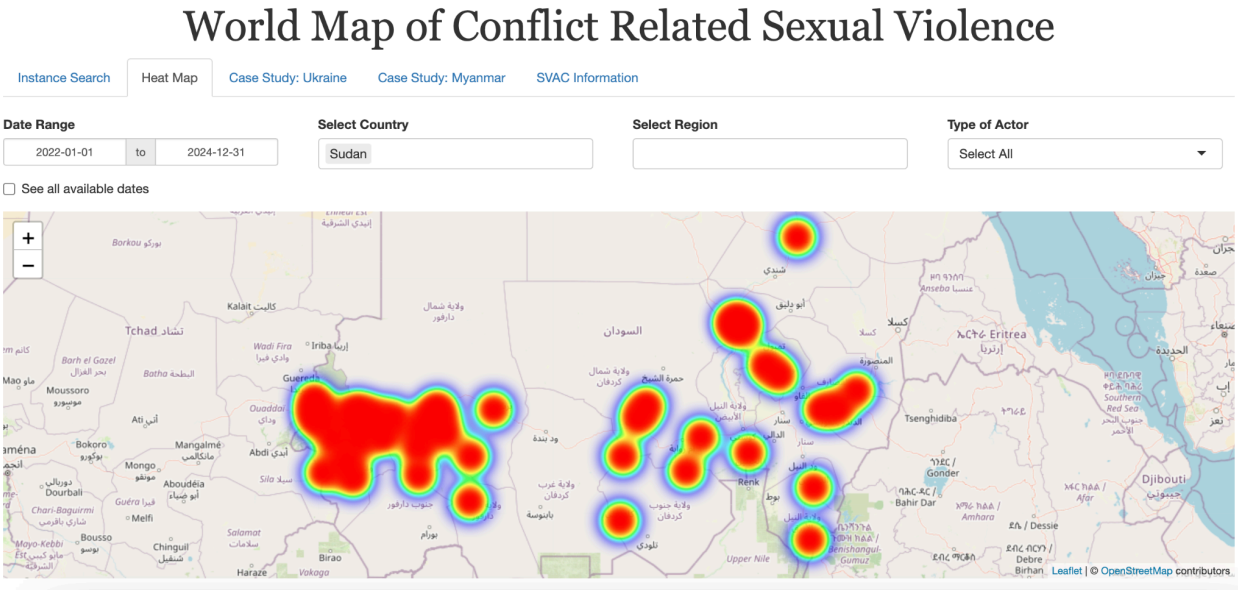

Figure 2: An example of a heat map of CRSV incidents in Sudan since 2022

The following tabs, case studies on Ukraine and Myanmar, are intended to populate the instance markers chronologically, as if in a video. By hitting the "Play" button next to the data slider, the user can watch the instances of CRSV populate with respect to time. In the Ukraine case study, a user can clearly see Russian forces moving West into Ukraine in 2022 as instances of sexual violence populate on the map. In the Myanmar case study, the data populates sporadically, but a user can clearly see a large concentration in the Kachin region of Northern Myanmar, which has had massive amounts of sexual violence in recent years (see Hedström, J., & Olivius, E., 2021).

The SVAC Information tab is intended to provide the user information and resources about conflict related sexual violence. Access to this data is limited, and even more limited is a well-defined understanding about the shortcomings of CRSV data. I intend to provide that here.

The final tab, "Data Journalism Example" is an example of the use-case of this dashboard to create a data journalism story about conflict related sexual violence. The example story is about Russia's invasion of Ukraine. When writing this story, I used only data and information from the dashboard.

The Explosion of Sexual Violence in the Invasion of Ukraine

By: Jakob Aggers

In the pre-dawn hours of February 24, 2022, the world awoke to a new era of conflict in Europe as Russia launched a massive, multi-front invasion into Ukraine. From the historic streets of Kyiv to the strategic port city of Mariupol, Russian forces advanced, unleashing a storm of artillery and airstrikes that shattered cities and marked the start of a devastating and far-reaching war. This brutal aggression not only threatened Ukraine's sovereignty but also precipitated a dire humanitarian crisis, with deep and disturbing impacts including a sharp rise in incidents of sexual violence.

The first documented case of sexual violence during the Russian invasion of Ukraine occurred on February 28, 2022 in Kyiv where a Russian soldier allegedly raped a 28 year old woman. In the month that followed, Russian soldiers committed at least 45 acts of conflict related sexual violence against Ukrainian citizens and soldiers.

In March and April alone, there were horrifying reports across several regions. For instance, on 9 March, near Bucha, Kyiv, "Russian soldiers raped a woman," and in another instance, they "forced other Ukrainian detainees to watch." The brutality escalated by the end of the month, with "Russian soldiers repeatedly raping and killing a woman in front of her son" in Mariupol, highlighting the horrifying personal violations accompanying military advances. April did not see any respite; for example, in Bucha, after the area was liberated, it was reported that "Russian soldiers raped and killed an unspecified number of women... the women's bodies were found burned."

Figure 3: A Data Journalism Example – The Invasion of Ukraine

6. Discussion

6.1 Continued Study:

To improve this dashboard to maximize its intended use, I would add the ability to download datasets based on the search criteria already provided in the dashboard. While my dashboard meets the goal of allowing users to view and visualize CRSV data, it does not allow them to see or use the raw data for themselves. The ACLED datasource is difficult to use (requires an API) and is only available for download at 3 year intervals, so creating a database resource where users could parse and then download their own data would be beneficial.

6.2 Drawbacks:

At the moment, the quality of sexual violence data is poor, mainly due to a large inflation of reported “zeroes” in the data. This made it impossible to accomplish to original goal of this project: creating a model that would predict if a new armed conflict would eventually contain sexual violence based on the presence of risk factors involved. Unfortunately, there is a large inflation of zeroes in the CRSV data, largely due to stigmas around reporting incidents of sexual violence in war. To view (from what I can tell) the only academic resource on the topic, see Changwook Ju’s study: “Determinants of Conflict-Related Sexual Violence: A Meta-Reanalysis Distinguishing Two Classes of Zero Observations.” Further study on the topic is necessary as

well as using such study to improve the collection of data on sexual violence in armed conflict worldwide.

7. Conclusion

This project aims to bridge the gap in understanding conflict-related sexual violence (CRSV) by offering a robust, data-driven resource that empowers journalists, policymakers, and researchers. Leveraging comprehensive datasets like ACLED and SVAC, the project visualizes and analyzes the spatial and temporal trends of CRSV across various conflicts, providing insights that are otherwise challenging to discern. The interactive platform developed here aims to simplify the complex task of accessing and interpreting CRSV data, thereby equipping stakeholders with the necessary tools to make informed decisions and communicate the realities of this violence effectively.

The project's results indicate that the visualization of data is instrumental in uncovering trends and patterns in CRSV, whether through pinpointing specific instances or highlighting broader hotspots. The dashboard, with its diverse range of functionalities, illustrates how spatial and temporal mapping can deepen our understanding of CRSV dynamics, bringing to light the movement of conflict actors and the clusters of violence they leave in their wake. This tool also enables the identification of patterns in CRSV over time, which can be valuable for shaping policies and interventions aimed at reducing such atrocities.

However, the challenges of obtaining quality CRSV data remain a significant barrier. The scarcity of accurate data due to stigmatization and underreporting hampers efforts to create predictive models that could foresee future CRSV incidents. Despite these limitations, the project underlines the potential of data-driven solutions in providing critical insights into conflict-related sexual violence, offering a foundation for future research and improvement in data collection methodologies.

In conclusion, this project contributes to the global effort to address CRSV by making vital data more accessible and interpretable. While there is much work to be done in refining our understanding and data collection of sexual violence in armed conflicts, this initiative marks a step forward in harnessing data science to illuminate the dark realities of CRSV and inspire informed actions toward its prevention and resolution.

8. Bibliography

Balcells, L. (2010). Rivalry and Revenge: Violence against Civilians in Conventional Civil Wars. *International Studies Quarterly*, *54*(2), 291–313. http://www.jstor.org/stable/40664168

Changwook Ju, Determinants of Conflict-Related Sexual Violence: A Meta-Reanalysis Distinguishing Two Classes of Zero Observations, *International Studies Quarterly*, Volume 67, Issue 3, September 2023, sqad044, https://doi.org/10.1093/isq/sqad044

Cohen, Dara Kay. 2016. Rape during Civil War. Ithaca: Cornell University Press. Copy at http://www.tinyurl.com/y3zfy2nb

Cohen, D. K., & Nordås, R. (2014). Sexual violence in armed conflict: Introducing the SVAC dataset, 1989–2009. *Journal of Peace Research, 51* (3), 418–428. DOI: 10.1177/0022343314523028

Crawford, K. F. (2017). *Wartime sexual violence: From silence to condemnation of a weapon of war*. Georgetown University Press. https://doi.org/10.2307/j.ctt1t89jxd

Eck, K. (2012). In data we trust? A comparison of UCDP GED and ACLED conflict events datasets. *Cooperation and Conflict, 47* (1), 124–141. https://doi.org/10.1177/0010836711434463

Hedström, J., & Olivius, E. (2021). The politics of sexual violence in the Kachin conflict in Myanmar. International Feminist Journal of Politics, 23(3), 374–395. https://doi.org/10.1080/14616742.2020.1862690

Wood, E. J. (2018). Rape as a practice of war: Toward a typology of political violence. *Politics & Society, 46* (4), 513–537. https://www.un.org/sexualviolenceinconflict/wp-content/uploads/2019/05/report/rape-as-a-practice-of-war-toward-a-typology-of-political-violence/wood-PS-2018-rape-as-a-practice-of-war.pdf

ChatGPT was used to aid in coding processes and editing of writing.