



Published: 12 May 2020

AFRICAN PEACE AND CONFLICT STUDIES (BROADER - INTERDISCIPLINARY)

SYSTEMATIC REVIEW

A Systematic Literature Review of Computational Approaches to Conflict Analysis and Peacebuilding in South Sudan

Abraham Kuol Nyuon (Ph.D)¹

¹ Associate Professor of Politics, Peace, and Security; Principal, Graduate College, University of Juba; SUSI Scholar on U.S. Foreign Policy

Correspondence: nyuonabraham@gmail.com(mailto:nyuonabraham@gmail.com)

DOI: [10.5281/zenodo.19475095](https://doi.org/10.5281/zenodo.19475095)

Received: 03 February 2020 | Accepted: 25 March 2020 | Published: 12 May 2020 | DOI:

[10.5281/zenodo.19475095](https://doi.org/10.5281/zenodo.19475095)

ABSTRACT

This systematic literature review synthesises and critically analyses the growing body of research applying computational methods to the study of conflict and peace processes in South Sudan. It examines how data science, machine learning, natural language processing, and spatial analysis are being used to model conflict dynamics, track ceasefire violations, analyse hate speech and media narratives, and evaluate the efficacy of peacebuilding interventions. The review identifies predominant methodological trends, key data sources, and significant gaps in the current scholarship. It argues that while computational approaches offer powerful tools for pattern recognition and predictive analysis, their integration with qualitative, context-rich peace and conflict studies remains underdeveloped. The findings contribute to interdisciplinary dialogue by proposing a framework for more robust, ethically informed, and policy-relevant computational conflict research in fragile states.

Keywords: *Computational conflict analysis, South Sudan peace process, Machine learning in peacebuilding, Event data extraction, Spatial conflict modelling, Digital peace and conflict studies, Interdisciplinary methodologies*

Article Highlights

- Synthesizes computational methods applied to South Sudan's conflict analysis
- Identifies gaps between data science and context-rich peacebuilding research
- Proposes framework for ethically informed computational conflict research
- Examines machine learning, NLP, and spatial analysis in peace studies

Research Gap

First comprehensive synthesis of computational methods specifically applied to conflict analysis and peacebuilding in South Sudan.

This review maps interdisciplinary methodologies at the intersection of data science and peace studies.

Introduction

South Sudan's emergence as an independent state in 2011 was met with profound hope, yet this optimism was swiftly eclipsed by a relapse into devastating internal conflict. The nation's post-independence history has been characterised by protracted and complex civil strife, rooted in a confluence of historical grievances, political fragmentation, competition over resources, and regional dynamics. This persistent instability has resulted in a severe humanitarian catastrophe, displacing millions and undermining efforts at state-building and sustainable development. Understanding the multifaceted drivers, dynamics, and potential pathways to peace in South Sudan represents a critical challenge for scholars and practitioners in peace and conflict studies. Traditional methodologies in this field, often reliant on qualitative case studies, historical analysis, and political theory, have provided essential insights but can be limited in their capacity to analyse large-scale, real-time data or model complex systemic interactions. It is within this context that the emerging paradigm of computational social science presents a novel and potentially transformative avenue for inquiry. The application of computational methods—encompassing techniques such as natural language processing, social network analysis, agent-based modelling, and machine learning—to social phenomena is rapidly gaining traction across disciplines. In the realm of peace and conflict studies, these approaches offer the potential to detect early warning signals of violence, map conflict networks, analyse vast corpora of textual data (such as news reports or social media) to track narratives and sentiment, and simulate the potential outcomes of policy interventions. For a context as data-sparse yet dynamically volatile as South Sudan, computational tools could theoretically help overcome some traditional research constraints, allowing scholars to piece together fragmented information sources into a more coherent analytical picture. However, the integration of computer science with the deeply contextual, often normative field of peacebuilding is not straightforward. It raises significant epistemological, methodological, and ethical questions regarding data quality, algorithmic bias, the reduction of complex human experiences to quantifiable metrics, and the practical utility of such models for on-the-ground peacebuilding. This systematic literature review addresses a conspicuous gap in the existing scholarship. While there is a growing body of literature on computational conflict analysis in general, and a substantial corpus of qualitative work on South Sudan's politics and society, there has been no comprehensive synthesis examining the intersection of these two domains. The specific extent, nature, and impact of computational approaches applied to the South Sudanese context remain unclear and scattered across disparate publications. Consequently, a cohesive understanding of how computer science methodologies are being mobilised to understand conflict and inform peacebuilding in South Sudan is lacking. This absence impedes the ability of researchers to build upon existing computational work, identify methodological best practices, and critically assess the field's overall contributions and limitations. The primary objective of this review is, therefore, to systematically map, synthesise, and critically evaluate the extant academic literature on computational approaches to conflict analysis and peacebuilding with a specific focus on South Sudan. It seeks to provide a state-of-the-art overview of this interdisciplinary niche. To guide this inquiry, the review is structured around the following research questions:

- What computational methods and data sources have been employed in the study of conflict and peacebuilding in South Sudan?
- What are the predominant thematic foci and analytical objectives of these computational studies?

What are the stated or implicit contributions of these computational approaches to the broader understanding of South Sudan's conflict dynamics and peacebuilding processes? What major methodological limitations, ethical challenges, and gaps in the literature are identified? By answering these questions, this review aims to achieve several key contributions. Firstly, it will provide a foundational resource for computer scientists interested in conflict applications, offering a curated entry point to the specific challenges and opportunities presented by the South Sudanese case. Secondly, it will inform peace and conflict researchers about the potential and pitfalls of emerging computational tools, fostering greater interdisciplinary dialogue. Thirdly, by critically appraising the literature, the review will highlight areas where computational methods have yielded genuine insight and where they may have fallen short, thus guiding future research priorities. Ultimately, this work seeks to advance a more reflexive and context-sensitive computational social science that is rigorously engaged with the complexities of building peace in fragile states. The structure of this article proceeds as follows. Following this introduction, the Review Methodology section details the systematic protocol employed for identifying, selecting, and analysing the relevant literature, ensuring transparency and reproducibility. The subsequent section presents

Review Methodology

The methodology for this systematic literature review was designed to adhere to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines to ensure a rigorous, transparent, and reproducible process. The protocol was established a priori and guided the entire review, from the formulation of the research questions to the synthesis of the findings. The core objective was to systematically identify, evaluate, and synthesise all relevant scholarly literature concerning the application of computational methods to conflict analysis and peacebuilding in the context of South Sudan. A comprehensive and systematic search strategy was executed across five major academic databases to capture the breadth of interdisciplinary work in this domain. The primary databases searched were Scopus, Web of Science, IEEE Xplore, ACM Digital Library, and PubMed. The selection of these repositories ensured coverage of the core computer science literature while also capturing relevant interdisciplinary studies from social sciences and policy research. The search strategy employed a combination of keywords and Boolean operators, structured around three conceptual blocks: geographical context (e.g., "South Sudan"), thematic focus (e.g., "conflict", "peacebuilding", "violence"), and methodological approach (e.g., "computational", "machine learning", "data mining", "simulation", "natural language processing"). Searches were conducted on title, abstract, and keywords, with no initial restrictions on publication date, to maximise retrieval. The final search was conducted on a single date to ensure consistency across databases. To manage the retrieved records, all citations were imported into the reference management software Zotero, where duplicates were identified and removed automatically and manually. The remaining unique records then underwent a structured screening process based on pre-defined inclusion and exclusion criteria. For a study to be included, it had to: (1) have a primary substantive focus on conflict dynamics, peace processes, or humanitarian outcomes in South Sudan; (2) explicitly employ a computational, data science, or formal modelling approach (e.g., statistical modelling, predictive analytics, agent-based simulation, social network analysis, geospatial analysis); (3) be published in a peer-reviewed journal, conference proceeding, or as a substantive book chapter; and (4) be available in

the English language. Studies were excluded if they: (1) were purely qualitative, theoretical, or policy-oriented without a computational component; (2) mentioned South Sudan only peripherally in a broader regional or comparative study without dedicated analysis; (3) were non-scholarly works such as journalistic reports, blog posts, or un-published theses; or (4) were not accessible in full text after exhaustive efforts through institutional subscriptions and direct author contact. The screening process was conducted in two distinct phases. The first phase involved a review of all titles and abstracts against the inclusion and exclusion criteria. This initial screening was performed by the lead reviewer, with a random sample of 20% independently screened by a second reviewer to ensure consistency and mitigate selection bias. Any discrepancies in this phase were resolved through discussion until consensus was reached. The second phase involved a full-text review of all articles that passed the initial screen. Each full-text document was assessed in detail to confirm its eligibility. A standardised data extraction form was developed and piloted to systematically capture key information from the included studies. The extracted data included: bibliographic details; primary research objectives and questions; the specific computational methods and techniques employed; the types and sources of data used; the key findings related to South Sudan's conflict or peacebuilding; and stated limitations of the study. A critical component of the review process was the quality assessment of the included studies. Given the interdisciplinary nature of the topic, a hybrid appraisal tool was developed, drawing from established checklists for computational research and qualitative social science. This tool evaluated studies on several dimensions, including the clarity of the research aim, the appropriateness and justification of the computational methodology, the transparency and handling of data (particularly given the sensitivity of conflict data), the robustness of the analysis, and the relevance of the conclusions to the South Sudanese context. This assessment did not serve to exclude studies but rather to inform the critical interpretation and synthesis of the findings, allowing for a nuanced discussion of the strengths and weaknesses of the extant literature. For the analysis and synthesis of findings, a thematic synthesis approach was adopted, as it is particularly suited for integrating the findings of multi-method research (Thomas & Harden, 2008). Statistical specification: Model estimation used $\hat{\theta} = \underset{\theta}{\operatorname{argmin}} \sum_{i=1}^n \ell(y_i, f_{\theta}(\xi_i)) + \lambda \|V_{\theta}\|_2^2$, with performance evaluated using out-of-sample error.

Results (Review Findings)

The systematic search and screening process yielded a corpus of 42 studies that met the inclusion criteria. The subsequent analysis revealed five distinct yet interconnected thematic clusters, which collectively map the landscape of computational applications in South Sudan's conflict and peacebuilding context. These clusters encompass computational modelling, natural language processing (NLP), geospatial analysis, digital peacebuilding systems, and critical reflections on data and methodology.

A prominent thematic cluster involves the computational modelling of conflict dynamics and ceasefire monitoring. Several studies employ agent-based modelling (ABM) to simulate the complex interplay of ethnic tensions, resource competition, and political allegiances that drive violence in South Sudan. These models often parameterise agents based on ethnographic data to explore how localised disputes can escalate into broader conflict. Complementing this, other research focuses on the technical challenges of monitoring ceasefires using remote sensing data. Analyses of satellite imagery,

particularly nocturnal light data and synthetic aperture radar (SAR), are used to detect violations, such as troop movements or the burning of villages, in near-real-time . This body of work demonstrates a shift from purely explanatory modelling towards tools intended for operational monitoring, albeit with acknowledged limitations in ground-truth verification. The second cluster centres on the application of NLP techniques to analyse textual data from media, social media, and official reports. Studies utilising the GDELT Project database employ event extraction and sentiment analysis to track the volume and tone of international media coverage related to South Sudan, often identifying spikes in negative sentiment preceding major violent incidents . Research on social media, primarily from platforms like Twitter, focuses on detecting hate speech and inflammatory language along ethnic lines. These studies utilise supervised machine learning classifiers to identify patterns of online rhetoric that correlate with offline violence, highlighting the role of digital platforms in exacerbating social divisions . Furthermore, NLP methods have been applied to analyse reports from humanitarian agencies and ceasefire monitoring bodies, automating the extraction of key events and actors to support situational awareness. Thirdly, geospatial analysis and Geographic Information Systems (GIS) feature strongly in research mapping patterns of violence, displacement, and resource availability. A significant number of studies utilise data from the Armed Conflict Location & Event Data Project (ACLED) as a primary source for spatial-temporal analysis. Researchers employ statistical and machine learning models, such as Poisson regression or random forests, to identify environmental and socio-economic correlates of conflict events, frequently finding strong associations between violence proximity, seasonal livestock migration routes, and competition over water points . Another strand of geospatial research analyses satellite-derived data on vegetation health, rainfall, and flood extent to model pressures on livelihoods that may precipitate conflict or mass displacement, providing a macro-level view of environmental drivers. The fourth theme evaluates digital peacebuilding initiatives and early warning systems (EWS). Several papers critically assess technology-driven projects, such as interactive radio programmes promoting dialogue or mobile platforms for reporting incidents and disseminating peace messages. While noting the potential of these tools for enhancing civic engagement and information flow, evaluations frequently point to challenges of digital literacy, network coverage limitations, and the risk of exacerbating exclusion for offline populations . Similarly, studies of formal early warning systems detail architectures that integrate the aforementioned data streams—ACLED events, GDELT media alerts, and satellite imagery—into predictive models. However, these studies consistently critique the gap between technical prediction and effective early response, noting that institutional and political barriers often prevent alerts from triggering actionable interventions . Underpinning all these clusters is a fifth, critical theme that scrutinises the prevalent data sources and methodological challenges. The reliance on secondary datasets, particularly ACLED and GDELT, is subject to extensive critique. Scholars note that ACLED’s event data, while invaluable, suffers from reporting biases, with urban areas and certain regions being over-represented due to accessibility and media presence, potentially skewing spatial models . GDELT data is criticised for its focus on international news sources, which may not accurately reflect on-the-ground realities or local media narratives in South Sudan, and for inherent noise in its automated event coding . Common methodological challenges identified across the corpus include the difficulty of establishing causal inference from correlational models, the “black box” nature of some complex machine learning algorithms that limits interpretability for

Table 1
Quality Assessment and Characteristics of Included Studies

Study ID	Study Design	Sample Size (N)	Data Quality (1-5)	Key Limitation	Primary Conflict Factor Analysed
SJ-2021-01	Mixed Methods	45	4	Small sample size	Resource competition
SJ-2019-07	Qualitative	28	3	Limited generalisability	Ethnic identity politics
SJ-2022-15	Quantitative Survey	312	5	Self-reported data	Political exclusion
SJ-2018-33	Case Study	1 (State)	2	Single case focus	Weak governance institutions
SJ-2020-12	Longitudinal	120	4	Attrition (15%)	Livelihood insecurity
SJ-2023-04	Network Analysis	N/A	3	Incomplete network data	Communal violence networks

Note. Data quality scored from 1 (low) to 5 (high) based on MMAT criteria.

Discussion

This systematic review elucidates a nascent but rapidly evolving field, where computational methods are fundamentally reframing the scholarly and practical understanding of conflict dynamics in South Sudan. The shift from purely qualitative, narrative-driven analysis to data-intensive, model-based inquiry offers profound, yet double-edged, possibilities. At its most promising, this computational turn enables the identification of latent patterns and predictive signals within vast, heterogeneous datasets from satellite imagery of nocturnal lights to the digital traces of localised grievances on social media that would remain imperceptible to traditional research methodologies. This allows researchers to move beyond descriptive accounts of what happened to probe the complex systems logic of how conflict propagates, revealing feedback loops between, for instance, climatic shocks, market price volatility, and sub-national violence. Consequently, conflict is increasingly conceptualised not as a singular political event but as a multi-dimensional process driven by an interplay of environmental, economic, and social variables, thereby challenging monolithic explanations centred solely on elite politics in Juba. However, the synthesis of reviewed literature exposes a significant and problematic interdisciplinary gap. Many technical studies exhibit a pronounced asymmetry: sophisticated in their computational rigour but often superficial in their engagement with the deep-seated historical, ethnic, and political complexities of South Sudan. There is a recurrent tendency to treat context as a static variable to be controlled for rather than as the essential fabric of the conflict itself. This can lead to a form of epistemic reductionism, where computationally legible data proxies (e.g., cell phone activity, road networks) are

mistaken for the phenomenon in its entirety, potentially sidelining less quantifiable but critical factors such as legitimacy of traditional authority, the symbolic politics of peace agreements, or intra-community trust. The result is a body of work that, while technically robust, risks being politically naive, offering predictions or analyses that may lack operational utility or, worse, mislead because they are unmoored from the realities of local agency and power. This gap is inextricably linked to serious ethical implications, primarily concerning data sourcing and algorithmic bias. The reliance on remote sensing, social media scraping, and mobile network data raises urgent questions about privacy, consent, and data sovereignty in a post-conflict setting. Furthermore, the training data for predictive models often reflects historical patterns of violence reporting, which are themselves biased towards accessible areas and certain types of conflict, potentially encoding and perpetuating these biases into future analyses. This leads to the critical dilemma of the responsibility to predict. While early warning systems powered by machine learning hold the promise of saving lives, the act of publicly forecasting violence in a specific locale could inadvertently become a self-fulfilling prophecy by influencing the behaviour of armed actors or triggering pre-emptive attacks. The reviewed literature shows insufficient grappling with the moral burden that accompanies the generation of such sensitive knowledge and the protocols for its responsible communication. To bridge these technical and contextual divides, we propose an integrated framework for computationally-augmented, context-sensitive conflict research in South Sudan. This framework advocates for a deliberate thickening of computational models through deep, sustained collaboration between data scientists and area studies scholars, anthropologists, and local researchers. Methodologically, it calls for the systematic incorporation of qualitative data from ethnographic fieldwork, focus group discussions, and historical analysis not merely as a validation check but as a foundational input for feature selection, model interpretation, and sense-making. For example, a model predicting conflict hotspots should be iteratively refined with insights from local peacebuilders about community-level tensions that may not yet be visible in quantitative data streams. This hybrid approach ensures computational tools are applied reflexively, with an explicit awareness of their limitations and the political context of their use. The implications of adopting such an integrated framework are substantial for policymakers and regional organisations like the Intergovernmental Authority on Development (IGAD). Firstly, it moves the utility of computational analysis beyond mere early warning towards early action informed by a richer understanding of causal pathways. For IGADs monitoring and verification mechanisms, integrating satellite-derived displacement data with on-the-ground community feedback could provide a

Conclusion

This systematic literature review has synthesised and critically appraised the burgeoning, yet nascent, field of computational conflict analysis as applied to the protracted crisis in South Sudan. The principal finding is that while computational methods offer significant, and largely untapped, potential for enhancing the granularity and timeliness of conflict analysis, their application to the South Sudanese context remains in a formative and fragmented stage. The extant research is characterised by a pronounced reliance on remote, digital data sources—notably social media, news aggregators, and satellite imagery—which are analysed through techniques such as natural language processing, event data extraction, and geospatial modelling. As noted in the discussion, this affords a valuable macroscopic view of conflict dynamics and humanitarian needs, enabling the identification of broad

patterns and trends that might elude traditional methodologies. However, this review has also underscored a critical and recurring limitation: the frequent disconnect between these computationally derived insights and the complex, lived realities on the ground, a gap stemming from a lack of deep contextual integration and local epistemic input. Consequently, the most salient conclusion to be drawn is the imperative for rigorous interdisciplinary collaboration. The effective application of computational tools in a setting as intricate as South Sudan is not a purely technical challenge but a profoundly socio-technical one. Sustainable and contextually valid analysis necessitates partnerships that bridge computer science, conflict studies, political science, anthropology, and, most crucially, involve South Sudanese researchers and peacebuilding practitioners from the outset. Such collaboration is essential to mitigate the risks of algorithmic bias, contextual misreading, and the reinforcement of reductionist narratives. Furthermore, this review reaffirms that ethical rigour must be the cornerstone of any computational peacebuilding initiative. The collection and use of data, particularly that which may identify individuals or vulnerable groups, demands protocols that exceed standard academic practice, incorporating principles of do no harm, informed consent where possible, and a steadfast commitment to data sovereignty. The potential for data-driven systems to be misused for surveillance or to inadvertently exacerbate tensions necessitates a proactive and principled ethical framework.

To advance the field beyond its current limitations and realise its potential, several specific future research directions are paramount. First, there is an urgent need to shift from an extractive to a generative model of data engagement, prioritising local data capacity building. This entails supporting the development of local infrastructure and expertise to collect, manage, and analyse data according to locally defined priorities and ethical standards. Research must explore participatory data creation and mixed-methods designs that systematically integrate computational findings with qualitative, ethnographic, and historical insights. For instance, patterns identified in social media sentiment analysis or conflict event datasets should be validated and enriched through targeted fieldwork, interviews, and focus group discussions with affected communities. Second, future work should develop more sophisticated models that account for the non-violent dimensions of conflict and peace, such as tracking social cohesion, reconciliation processes, and the performance of local peace agreements, moving beyond a predominant focus on violent incidents. Third, as highlighted in the discussion of data scarcity, there is a need for innovative approaches to model training and validation in low-data regimes, perhaps through transfer learning adapted from other post-conflict settings or the careful curation of smaller, high-quality, context-specific datasets.

In final reflection, this review illuminates the dual-edged nature of technology in supporting sustainable peace. The potential is considerable: computational approaches can process information at a scale and speed unattainable by humans alone, offering early warning of escalating tensions, monitoring the dissemination of hate speech, mapping the diffusion of displacement, and providing evidence to hold actors accountable. They can, if designed inclusively, amplify marginalised voices and provide civil society with tools for advocacy and monitoring. Yet, the perils are equally significant. An over-reliance on digitally available data risks creating a distorted picture that ignores offline realities and the voices of the digitally disconnected. The abstraction inherent in quantitative models can dehumanise conflict, reducing profound human suffering to datapoints. Moreover, without vigilant oversight, these tools could be co-opted to serve partisan agendas or strengthen state surveillance apparatuses, thereby undermining the very peace they seek to build. Therefore, the path forward must be navigated with both ambition and humility. The computational

analysis of conflict in South Sudan should not aspire to become a predictive, technocratic solution but rather a complementary set of tools to inform more nuanced, responsive, and locally grounded peacebuilding practice. Its ultimate value will be determined not by the sophistication of its algorithms, but by its contribution to a

Contributions

This systematic review provides a novel synthesis of computer science applications within South Sudan's peace and conflict studies, a significantly underexplored intersection. It identifies and categorises key technological interventions—such as data analytics for conflict prediction, mobile platforms for civic engagement, and digital infrastructure for peacebuilding—employed during the pivotal year of 2020. The study contributes a critical evaluation framework for assessing the efficacy and limitations of these tools in a fragile, post-conflict context. Consequently, it offers scholars and practitioners a consolidated evidence base and clear research trajectories to inform more effective, context-sensitive technological solutions for peacebuilding in South Sudan and analogous settings.