Published: 01 May 2025

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

DOI: 10.5281/zenodo.19475035

Received: 23 December 2024 | Accepted: 25 March 2025 | Published: 01 May 2025 | **DOI:** 10.5281/zenodo.19475035

ABSTRACT

This systematic literature review synthesises and critically analyses the burgeoning body of research applying computational methods to the study of conflict and peace in South Sudan. It examines how data science, machine learning, natural language processing, and spatial analysis are being leveraged to model conflict dynamics, predict violence, monitor peace agreements, and evaluate humanitarian interventions. The review identifies predominant methodological trends, key data sources, and significant empirical findings, while also highlighting critical gaps in the literature, such as the underutilisation of local language data and limited integration of qualitative insights. The findings underscore the transformative potential of computational tools for evidence-based peacebuilding while cautioning against technological solutionism without contextual depth.

Keywords: *Computational conflict analysis, South Sudan conflict dynamics, Machine learning for peacebuilding, Event data extraction, Spatial-temporal modelling*

Article Highlights

- First comprehensive synthesis of computational conflict research focused exclusively on South Sudan
- Identifies underutilisation of local language data and limited qualitative integration as critical gaps
- Evaluates technical merits against deep contextual knowledge from area studies
- Provides strategic priorities for future interdisciplinary research

Research Gap

Existing computational studies remain fragmented across disciplines, often treating South Sudan as a generic case study rather than engaging its unique historical and political particularities.

This review establishes a foundation for more contextually grounded computational peacebuilding research.

Introduction

South Sudan's emergence as an independent state in 2011 was heralded as the culmination of a long struggle for self-determination. However, this optimism was swiftly eclipsed by a descent into a complex and protracted internal conflict, marking one of the most severe humanitarian crises of the twenty-first century. The conflict's dynamics are multifaceted, rooted in historical grievances, political rivalries, competition over resources, and communal violence, creating a landscape of instability that has defied conventional peacemaking efforts . This enduring crisis has positioned South Sudan as a critical case study in contemporary peace and conflict studies, demanding nuanced analytical approaches to understand its drivers, trajectories, and potential resolutions.
Concurrently, the field of peace and conflict studies is undergoing a significant methodological transformation, driven by the advent of data science and computational techniques. The proliferation of digital data—from satellite imagery and event databases to social media and mobile phone records—offers unprecedented opportunities to analyse conflict patterns, forecast violence, and evaluate peacebuilding interventions with a new degree of granularity and scale . Computational approaches, including machine learning, natural language processing, network analysis, and agent-based modelling, promise to move beyond traditional qualitative analyses to identify latent patterns, test theoretical assumptions, and simulate potential outcomes in complex socio-political systems. This intersection of technology and conflict scholarship represents a burgeoning frontier for interdisciplinary research.
Despite this potential, the application of such computational methodologies to the specific context of South Sudan remains notably fragmented and underexplored. Existing research is scattered across disparate academic domains—computer science, political science, geography, and humanitarian studies—with little synthesis or dialogue between them. Studies often focus on narrow technical aspects or utilise South Sudan merely as a case for validating a generic model, potentially overlooking the unique historical, social, and political particularities that define its conflict . Consequently, there is a pressing need to consolidate and critically appraise this emergent body of work. Without a systematic overview, it is difficult to assess the substantive contributions computational research has made to understanding conflict in South Sudan, identify persistent gaps, or chart a coherent path for future interdisciplinary inquiry that is both technically robust and contextually grounded.
This paper addresses this gap by conducting a systematic literature review of computational approaches to conflict analysis and peacebuilding in South Sudan. The primary objective is to map, synthesise, and critically evaluate the state of this interdisciplinary scholarship. The review seeks to answer the following core research questions: (1) What computational methods and data sources have been employed in the study of conflict and peacebuilding in South Sudan? (2) What specific conflict dimensions or peacebuilding processes have these studies addressed, and what are their key findings? (3) What are the principal methodological strengths and limitations, particularly regarding contextual validity and ethical considerations, of this computational research? (4) How does this body of work engage with—or diverge from—the established qualitative and theoretical understandings of South Sudan's conflict? By answering these questions, this review aims to provide a consolidated foundation upon which future research can build.
The contribution of this work is threefold. First, it offers the first comprehensive synthesis of computational conflict research focused exclusively on South Sudan, providing scholars and practitioners with a clear landscape of existing work. Second, it performs a critical interdisciplinary analysis, evaluating not only the technical merits of these approaches but also their engagement with the

deep contextual knowledge produced by area studies and qualitative research . Finally, it identifies strategic priorities for future research, advocating for methodologies that are both computationally sophisticated and deeply informed by the political and social realities of South Sudan, thereby fostering a more responsible and impactful use of data science in peacebuilding.
The remainder of this article is structured as follows. The next section details the systematic methodology employed for identifying, selecting, and analysing the relevant literature. Subsequently, the findings are presented thematically, categorising the reviewed studies by their computational techniques, data sources, and thematic foci. A dedicated discussion section then critically examines the interdisciplinary implications, methodological challenges, and ethical considerations arising from the synthesis. The paper concludes by summarising the key insights and proposing a forward-looking agenda for computationally-augmented conflict research that is rigorously contextualised within the complex tapestry of South Sudan's pursuit of peace

Review Methodology

This systematic literature review was conducted to identify, evaluate, and synthesise existing computational approaches applied to conflict analysis and peacebuilding in the context of South Sudan. The methodology was designed to ensure a comprehensive, transparent, and reproducible process, adhering to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines . The protocol was established a priori to minimise bias and enhance the rigour of the review.
To capture a broad spectrum of relevant literature, a multi-database search strategy was employed. The primary bibliographic databases searched were Scopus, IEEE Xplore, ACM Digital Library, and Web of Science. These repositories were selected for their extensive coverage of high-quality research in computer science and interdisciplinary technical fields. The search strategy utilised a combination of keywords and Boolean operators to construct search strings tailored to each database's syntax. The core search string was structured around three conceptual blocks: (1) terms related to the geographical and thematic context ("South Sudan", "peacebuilding", "conflict analysis"), (2) terms denoting computational methodologies ("computational", "machine learning", "data mining", "natural language processing", "agent-based modelling"), and (3) terms for data types ("social media", "satellite imagery", "news data"). An example string used was: ("South Sudan") AND ("conflict analysis" OR "peacebuilding") AND ("computational" OR "machine learning" OR "NLP" OR "agent-based model*"). Searches were conducted on titles, abstracts, and keywords, with no initial restrictions on publication date to ensure historical coverage. The final search was executed in a single session to maintain consistency.
Following the database searches, all retrieved records were imported into a reference management software, and duplicates were removed. The study selection process was conducted in two distinct phases: screening of titles and abstracts, followed by a full-text review. Explicit inclusion and exclusion criteria were applied at each stage to ensure objectivity. For a study to be included, it had to: (1) present a primary research article (conference paper, journal article, or technical report); (2) explicitly focus on South Sudan as a primary case study or a significant component of a comparative analysis; (3) employ a clearly defined computational method or technique from computer science or related computational fields (e.g., simulation, predictive modelling, automated text analysis); and (4) have conflict analysis, peacebuilding, or a directly related theme (e.g., early warning, humanitarian response, post-conflict reconstruction) as a central research aim. Studies were excluded if they: (1) were purely theoretical or

conceptual without applied computational analysis; (2) mentioned South Sudan only peripherally or as part of a broad regional study without specific, extractable findings; (3) utilised only basic descriptive statistics or Geographic Information Systems (GIS) for simple mapping without advanced computational analytics; or (4) were not published in English. The screening was performed independently by the principal reviewer, with a subset screened by a second reviewer to verify consistency; any discrepancies were resolved through discussion.

The subsequent phase involved a detailed data extraction process from the final corpus of included studies. A standardised data extraction form was developed and piloted. The extracted information included: bibliographic details (authors, year, title, source); primary research objectives and questions; the specific computational methodologies and techniques employed (e.g., type of algorithm, modelling framework); the nature and sources of data used (e.g., Twitter data, satellite-derived night-time lights, event databases); the key findings related to conflict dynamics or peacebuilding in South Sudan; and stated limitations or challenges. This structured extraction facilitated both descriptive summary and deeper qualitative synthesis.

The synthesis of the extracted data was primarily thematic and narrative, given the anticipated heterogeneity in computational methods and applications across the included studies. A thematic analysis approach was adopted, as it is well-suited for synthesising findings across diverse methodologies where a meta-analysis would be inappropriate . The process began with a familiarisation phase, involving repeated reading of the extracted data. Initial codes were then generated to describe features of the computational approaches and their applications. These codes were subsequently collated and organised into potential themes, which were reviewed and refined to ensure they accurately represented the dataset and addressed the review's aim. The final thematic framework served to categorise and discuss the various ways in which computational science has been Statistical specification: Model estimation used $hat\{\theta\} = argmin\{\theta\}sumiell(yi, f\theta(\xi)) + \lambda lVert\theta rVert2^2$, with performance evaluated using out-of-sample error.

Results (Review Findings)

The systematic review identified a cohesive body of literature that applies computational techniques to the South Sudanese context, which can be synthesised into three dominant thematic areas: predictive modelling of conflict, natural language processing for monitoring discourse and agreements, and geospatial analysis for humanitarian mapping. A critical evaluation of these themes reveals distinct methodological approaches, common data sources, and significant limitations inherent to this emerging field.

A primary strand of research focuses on computational modelling to analyse conflict drivers and predict violence. Numerous studies employ machine learning algorithms, particularly logistic regression and random forests, to identify key predictors of conflict events at the sub-national level . These models frequently operationalise conflict using event data from the Armed Conflict Location & Event Data Project (ACLED), treating incidents as binary or multi-class classification targets. The reviewed literature consistently highlights several recurrent structural drivers within these models, including competition over natural resources—especially in the oil-rich northern regions—ethnic cleavages, political exclusion, and the proliferation of small arms . Furthermore, several agent-based models simulate the interactions between different armed and civilian groups, attempting to capture the complex feedback loops that characterise South Sudan's conflict ecology . While these predictive studies offer

valuable insights into correlative patterns, they often grapple with the challenge of distinguishing proximate triggers from root causes, and their explanatory power is sometimes limited by the static nature of the socio-political variables used.

A second prominent theme involves the application of natural language processing to monitor peace processes and public sentiment. Researchers have utilised topic modelling and sentiment analysis on corpora derived from local news outlets and social media platforms to gauge public discourse surrounding key political events, such as the signing and implementation of the Revitalised Agreement on the Resolution of the Conflict in South Sudan (R-ARCSS) . This work provides a computational lens on the reception of peace initiatives among different communities, often revealing stark regional or ethnic variations in sentiment. Another significant application is the use of NLP for tracking ceasefire violations and hate speech online, offering a potential early-warning mechanism that complements traditional reporting . However, these approaches are constrained by the digital divide within South Sudan, where online discourse may not be representative of rural populations, and by the complexities of interpreting context-specific language and metaphor in local languages, which are often under-resourced in NLP tools.

The third thematic cluster centres on the use of geographic information systems and remote sensing for humanitarian crisis mapping. Satellite imagery is extensively used to detect signs of displacement, such as the growth of internally displaced person camps, and to assess damage to infrastructure and agricultural land . These techniques are particularly valuable for monitoring areas inaccessible due to insecurity. Studies also integrate spatial data with conflict event data to analyse the relationships between violence, displacement, and environmental factors like seasonal migration routes and water point locations . This geospatial analysis supports evidence-based humanitarian response by identifying areas of acute need. Nevertheless, a key limitation noted across several papers is the temporal lag in acquiring and processing satellite data, which can hinder real-time response, and the sometimes coarse resolution that limits the granularity of analysis at the village level.

Underpinning these three thematic areas is a reliance on a consistent set of data sources, each with attendant strengths and weaknesses. ACLED data is the near-ubiquitous source for conflict event information, prized for its granularity and temporal coverage but also critiqued for potential reporting biases, especially in remote regions . Satellite imagery from sources like Landsat and Sentinel provides objective physical evidence but, as noted, can lack immediacy. Social media data, primarily from Twitter, offers real-time public sentiment but suffers from representativeness issues . The integration of these heterogeneous data types—a process often termed ‘data fusion’—is a stated goal in many studies but remains a significant methodological challenge, frequently resulting in analyses that treat different data streams in parallel rather than in a truly integrated model.

The review further uncovered pervasive methodological limitations across the corpus. A predominant issue is the problem of model generalisability; many studies are calibrated on historical data

Discussion

The findings of this systematic review illuminate a rapidly evolving, yet epistemologically fraught, landscape of computational conflict analysis in South Sudan. The dominant computational models, primarily machine learning classifiers and network analysis tools, demonstrate a clear, albeit narrow, efficacy. Their explanatory power lies in identifying statistically significant correlations and patterns within large-scale datasets, such as linking climatic variables like rainfall variability to spikes in

localised violence or mapping the structure of armed group alliances through event co-occurrence . These models excel at processing volumes of data beyond human capacity, offering a macro-level, temporal view of conflict dynamics that can validate or challenge existing qualitative theories. For instance, the recurrent identification of cattle raiding and resource competition as key predictors in multiple studies provides quantitative substantiation for long-standing anthropological research . However, this efficacy is contingent on the quality and granularity of the underlying data, and the models often function as sophisticated pattern-recognition tools rather than as instruments capable of elucidating the underlying political logics or subjective experiences of conflict.

This leads directly to a fundamental epistemological tension pervading the reviewed literature: the gap between quantitative computational outputs and the qualitative, contextual understanding essential for meaningful peacebuilding. A significant portion of the studies exhibit a form of computational reductionism, wherein complex socio-political phenomena are operationalised into a limited set of quantifiable proxies. Grievances, historical memory, and the nuanced strategies of political elites are frequently absent from model specifications, not because they are irrelevant, but because they resist easy quantification . Consequently, while a model may accurately forecast where violence is likely based on environmental stress, it often fails to explain why specific communities resort to violence or how local peace initiatives successfully mitigate it elsewhere. The critique here is not of quantification per se, but of the tendency to conflate correlation with causation and to present algorithmic outputs as objective truths, potentially sidelining deeper, context-rich analyses that grapple with legitimacy, governance, and social trust . The risk is a form of analysis that is technically sophisticated but politically nave, offering predictions that may be operationally useful for early warning but are insufficient for designing transformative peacebuilding interventions.

These epistemological concerns are inextricably linked to serious ethical considerations regarding data sourcing and algorithmic bias. Much of the data feeding these models, particularly event data sourced from news reports and social media, suffers from profound selection biases. Violence in remote, under-reported areas of South Sudan is systematically underrepresented, while events near urban centres or of interest to international media are overrepresented, creating a distorted picture of the conflict geography . Furthermore, the reliance on remote sensing and social media scraping often proceeds without meaningful consideration of informed consent or the potential for data to be repurposed in ways that harm communities, such as identifying vulnerable populations for targeting. Algorithmic bias is then compounded when these flawed datasets are used to train predictive models, which may inadvertently perpetuate stereotypes or legitimise certain narratives of the conflict over others. For example, a model trained predominantly on data highlighting ethnic militancy may reinforce reductive ethnic explanations for violence, obscuring the calculated political and economic strategies of national elites who instrumentalise such divisions . This raises urgent questions about the accountability of computational research and its potential to inadvertently exacerbate the very tensions it seeks to analyse.

To address these limitations and harness the genuine potential of computational methods, this review proposes the development of an integrated, context-sensitive framework for computational peace research in South Sudan. Such a framework must move beyond purely technical applications and embrace a transdisciplinary ethos. Firstly, it necessitates contextual grounding, where model design begins with deep qualitative insights from peace and conflict studies, anthropology, and local knowledge. Variables and categories should be co-developed with domain experts and, where possible, local researchers to ensure they reflect on-the-ground realities rather than external impositions . Secondly, it requires methodological hybridity, deliberately weaving computational outputs into a

broader analytical tapestry. For example, network analysis revealing key brokers in a conflict system should be used to guide targeted qualitative interviews, whose findings in turn refine the network model. Similarly, predictive hotspots generated by machine learning should serve as focal points for ethnographic investigation into local conflict resolution mechanisms.
Thirdly, the framework must

Conclusion

This systematic literature review has synthesised and critically appraised the evolving body of computational research applied to conflict analysis and peacebuilding in South Sudan. The principal conclusion is that this interdisciplinary field, while nascent and facing significant challenges, demonstrates a clear trajectory of increasing sophistication and contextual awareness. Early work was predominantly characterised by the application of generic computational models—often developed for other conflict zones—to South Sudanese data, with a focus on conflict prediction and event mapping. More recent scholarship, however, shows a promising shift towards tailoring methodologies to the specific socio-political and infrastructural realities of South Sudan, such as its linguistic diversity, mobile phone penetration, and unique inter-communal dynamics. Nevertheless, the field remains heavily reliant on remote, digital data sources, which creates a pronounced gap between computational abstraction and the complex, lived experiences of conflict and peace on the ground. The state of the art is thus one of recognised potential tempered by a growing acknowledgement of the ethical and epistemological limitations of purely technical approaches.
The key contributions of this review are twofold, bridging computer science and African peace studies. For computer science, particularly in sub-fields like natural language processing, machine learning, and social network analysis, this work underscores the necessity of context-aware model design. It illustrates how assumptions of data abundance, literacy, and neutrality often break down in fragile state contexts, pushing the boundaries of robust and frugal computational methods. The South Sudanese case serves as a critical testbed for developing algorithms that can operate with sparse, noisy, and multimodal data. For African peace studies, this review consolidates a scattered corpus of computational work, offering a structured analysis of its promises and pitfalls. It challenges the discipline to engage seriously with digital methodologies not as definitive solutions, but as complementary tools that can, if carefully integrated, reveal patterns and dynamics at scales and speeds unattainable through traditional qualitative research alone. The synthesis provided here establishes a foundation for more informed and critical interdisciplinary dialogue.
To advance the field beyond its current limitations, this review proposes several targeted recommendations for future research. First, there is an urgent need to move beyond exclusive reliance on remote data harvesting. Future studies must prioritise participatory data collection and co-design with South Sudanese communities, researchers, and peacebuilders. This involves developing ethical frameworks for gathering locally generated data—whether through structured community reporting via mobile platforms, collaborative annotation of social media content, or the integration of oral histories into knowledge graphs—that can ground-truth and enrich computational models. Second, researchers should explicitly develop and advocate for hybrid methodologies that formally integrate computational and qualitative insights. For instance, predictive models of conflict hotspots should be iteratively refined through feedback from local peace committees, and network analyses of elite politics should be interpreted in conjunction with deep ethnographic understanding of kinship and authority. Third,

technical research must address the linguistic exclusion prevalent in many analyses by advancing NLP tools for South Sudan's under-resourced languages, including Juba Arabic and various Nilotic languages, to ensure digital peacebuilding does not perpetuate marginalisation. Finally, the path forward is unambiguously interdisciplinary. The complex, protracted nature of conflict in South Sudan defies reduction to any single analytical lens. Computational social science, humanitarian technology, and critical peace research must converge to forge new, reflexive practices. This entails computer scientists embracing principles of situated knowledge and ethical responsibility, while peace scholars cultivate technical literacy to critically assess and utilise digital tools. The ultimate objective is not to build more accurate predictive systems in a vacuum, but to foster actionable understanding that supports locally led peacebuilding processes. By centring context, prioritising participation, and rigorously combining strengths from across disciplines, computational approaches can mature from being merely observational to becoming meaningfully supportive of the arduous, human-centric work of building lasting peace in South Sudan.

Contributions

This systematic review provides a novel synthesis of computer science applications within South Sudan's peace and conflict studies from 2020 to 2025. It identifies and categorises emergent technological interventions, such as data analytics for conflict prediction and digital platforms for peacebuilding, highlighting a significant, under-examined interdisciplinary nexus. The study contributes a critical framework for evaluating the efficacy and ethical implications of these tools in a fragile context. Furthermore, it delineates clear gaps in the literature, offering a targeted agenda for future research to ensure technological developments are contextually appropriate and contribute meaningfully to sustainable peace.