Journal of Sustainability, Policy, and Practice

2026 2nd International Conference on ArtificialIntelligence and Advanced Algorithms

Article

A Comparative Study of Spatiotemporal Clustering and Classification Approaches for Security Incident Risk Assessment in UN Peacekeeping Operations

Wen Shang [1,*], Fanyi Zhao [1,2] and Muyu Liu [3]

[1] International Affairs, Science and Technology, Georgia Institute of Technology, Atlanta, GA, USA
[2] Computer Science, Stevens Institute of Technology, Hoboken, NJ, USA
[3] Center for Applied Statistics and School of Statistics, Renmin University of China, Beijing, China
* Correspondence: Wen Shang, International Affairs, Science and Technology, Georgia Institute of Technology, Atlanta, GA, USA

Abstract: This study presents a comparative analysis of spatiotemporal clustering and classification approaches applied to security incident risk assessment in United Nations peacekeeping operations. Drawing on 61,274 conflict event records from the Armed Conflict Location and Event Data Project (ACLED) spanning 2010 to 2024 across four active mission areas---MINUSCA, MONUSCO, UNMISS, and UNIFIL---and supplemented by geospatial features from PRIO-GRID, this research evaluates three spatial clustering methods (ST-DBSCAN, Kernel Density Estimation, and Getis-Ord Gi*) for hotspot identification and examines temporal periodicity through STL decomposition and autocorrelation analysis. The study compares Random Forest, XGBoost, and SVM classifiers under natural class imbalance using SMOTE, ADASYN, and SMOTE-Tomek resampling strategies. Results indicate that ST-DBSCAN demonstrates superior flexibility in delineating irregular hotspot boundaries, while XGBoost paired with SMOTE yields the highest Macro F1-Score (0.716) and Matthews Correlation Coefficient (0.683) among all configurations. These findings provide methodological references for data-driven risk assessment in peacekeeping contexts.

Keywords: spatiotemporal clustering; security risk assessment; imbalanced classification; UN peacekeeping operations

Received: 02 March 2026
Revised: 17 April 2026
Accepted: 30 April 2026
Published: 06 May 2026

1. Introduction

1.1. Background of Security Challenges in UN Peacekeeping Operations

United Nations peacekeeping operations serve as a primary instrument for maintaining international peace in conflict-affected regions. Since 1948, more than 4,200 peacekeepers have lost their lives during mission service, with hostile acts constituting a growing proportion of fatalities (data source: UN Peacekeeping Fatalities Open Data, peacekeeping.un.org/en/fatalities). The year 2022 witnessed 32 peacekeeper deaths from malicious attacks, with MINUSMA in Mali accounting for 14 fatalities---making it the deadliest mission for the ninth consecutive year (data source: UN Press Release ORG/1730, press.un.org). Active missions including MONUSCO in the Democratic Republic of the Congo, MINUSCA in the Central African Republic, and UNMISS in South Sudan continue to operate under persistent threat from armed groups, improvised explosive devices, and targeted attacks on UN personnel and convoys. De Coning documented that while individual field missions have developed localized tools for crisis mapping and situational awareness, these tools retain limited predictive capacity and remain largely ad

hoc [1]. The Cruz Report further underscored the urgent need to transform security practices, arguing that reactive postures have contributed to avoidable casualties [2].

The analytical infrastructure within peacekeeping missions---primarily the Joint Mission Analysis Centres (JMACs)---relies on qualitative scenario-based assessments and trend analyses produced by human analysts within narrow time windows. The Situational Awareness Geospatial Enterprise (SAGE), developed by the UN support base in Valencia, Spain, serves as the central event and incident tracking tool across peacekeeping missions, enabling the visualization of security incidents and identification of cyclical patterns. SAGE data, while valuable for internal operational use, remains inaccessible to the external research community. Cederman and Weidmann cautioned against overoptimism regarding long-term conflict onset prediction yet acknowledged that short-term, spatially bounded forecasts within ongoing conflicts are achievable [3]. Murphy et al. argued that the growth in open-source conflict data and machine learning advancements present viable opportunities for improving conflict forecasting, provided that model interpretability and data quality are adequately addressed [4]. The convergence of increasingly granular open-source conflict databases with established spatial analysis techniques creates an opportunity to evaluate which analytical methods are best suited for peacekeeping risk assessment---a question this study addresses through systematic comparison.

1.2. Research Scope and Contributions

1.2.1. Research Questions and Objectives

This paper addresses three research questions. RQ1: What are the comparative strengths and limitations of ST-DBSCAN, Kernel Density Estimation, and Getis-Ord Gi* in identifying security incident hotspots within UN peacekeeping mission areas? RQ2: Do security incidents exhibit identifiable temporal periodicities that can inform resource deployment timing? RQ3: Under natural class imbalance in event type distributions, how do Random Forest, XGBoost, and SVM classifiers perform across different resampling strategies as measured by Macro F1-Score, PR-AUC, MCC, and Cohen's Kappa? This study does not aim to construct an end-to-end prediction system. The scope is confined to a comparative evaluation of analytical methods, offering empirical references for peacekeeping risk assessment practitioners.

1.2.2. Paper Structure Overview

The remainder of this paper proceeds as follows. Section 2 describes the open-source datasets and preprocessing pipeline. Section 3 details the spatiotemporal clustering, temporal pattern mining, and classification methodologies. Section 4 presents experimental results and discusses their practical implications. Section 5 provides concluding remarks and future directions. Section 6 offers acknowledgments.

2. Data Sources and Preprocessing

2.1. Overview of Open-Source Conflict Event Datasets

2.1.1. Acled Dataset and Study Area Selection

The primary data source is the Armed Conflict Location and Event Data Project (ACLED), which provides disaggregated, geo-referenced records of political violence and protest events across more than 200 countries [5]. ACLED encodes each event with geographic coordinates, date, involved actors, event type (six primary categories: Battles, Violence against civilians, Explosions/Remote violence, Riots, Protests, and Strategic developments), sub-event type, and fatality counts, updated weekly through its public Data Export Tool.

The study area encompasses four active UN mission zones: MINUSCA (CAR), MONUSCO (DRC), UNMISS (South Sudan), and UNIFIL (Lebanon), selected based on elevated threat levels, sufficient data volume, and geographic diversity. MINUSCA and MONUSCO represent high-intensity multidimensional stabilization missions in tropical and equatorial environments, UNMISS operates in a post-independence conflict setting

with pronounced inter-communal dynamics, and UNIFIL provides a contrasting case of a smaller-scale traditional peacekeeping operation in a Mediterranean context. The temporal window spans January 2010 through December 2024, yielding 61,274 event records. The UCDP Georeferenced Event Dataset served as a cross-validation source for high-fatality events [6]. Eckdocumented methodological differences between ACLED and UCDP GED---notably that ACLED includes events involving unnamed armed groups while UCDP GED does not---informing the decision to use ACLED as the primary source while consulting UCDP GED for quality assurance on events reporting ten or more fatalities [7].

2.1.2. Supplementary Geospatial and Contextual Data Sources

Geospatial features were derived from the PRIO-GRID dataset, a global vector grid at 0.5×0.5 decimal degree resolution containing cell-specific information on mountainous terrain coverage, forest cover, population estimates, and economic activity levels [8]. Peacekeeper fatality records from the UN Peace and Security Data Hub (NOTICAS system, peacekeeping.un.org) were used to cross-reference clustering results against known high-casualty locations. Country-level organized violence statistics from Davies et al. provided macro-level contextual framing [9].

2.2. Data Preprocessing and Feature Engineering

Preprocessing comprised four stages: deduplication removed 847 duplicate records (1.38%); coordinate validation geocoded 213 records with missing coordinates using ACLED location descriptions and the GeoNames gazetteer, while excluding 36 unresolvable entries; grid mapping spatially joined point-level events to PRIO-GRID cells using a point-in-polygon operation, generating cell-month aggregates of event frequency and cumulative fatalities; and feature engineering produced 18 input variables organized into three categories. Spatial features included latitude, longitude, distances to nearest national capital and international border, and PRIO-GRID cell identifiers. Temporal features included month of year, day of week, a binary indicator for election periods derived from ACLED strategic developments records, and STL-derived seasonal indices. Contextual features incorporated mountainous terrain percentage, forest cover percentage, population density, economic activity level from the G-Econ dataset, historical conflict intensity measured as the cumulative event count within the preceding six months, number of distinct armed groups active within each grid cell, and a governance quality index compiled from V-Dem data (v-dem.net). World Bank World Development Indicators (databank.worldbank.org) provided supplementary country-level economic variables including GDP per capita and poverty headcount ratio. The final dataset comprised 60,178 records with complete feature vectors across all 18 variables.

3. Methodology

3.1. Spatiotemporal Clustering for Hotspot Detection

Three clustering approaches representing distinct methodological paradigms were applied. ST-DBSCAN, an extension of DBSCAN incorporating separate spatial and temporal neighborhood radii, was configured with spatial radius Eps1 = 25 km, temporal radius Eps2 = 30 days, and MinPts = 5 events, with parameters determined through k-dist graph analysis [10]. ST-DBSCAN identifies clusters of arbitrary shape by iteratively expanding neighborhoods from core points satisfying both spatial and temporal density requirements, labeling non-qualifying points as noise. Spatiotemporal Kernel Density Estimation (KDE) generates a continuous risk surface using Gaussian kernels with spatial bandwidth of approximately 20 km (Silverman's rule) and temporal bandwidth of 30 days, thresholded at the 90th percentile [11]. The Getis-Ord Gi* statistic, computed on PRIO-GRID cells, identifies statistically significant spatial clusters by comparing local values against the global mean under spatial randomness, with False Discovery Rate correction at α = 0.05. Jiang and Shekhar provided a comprehensive taxonomy of hotspot detection methods against which the comparative dimensions of this study were structured: cluster

shape flexibility, parameter sensitivity, noise robustness, and interpretability [12]. Kang et al. further documented trade-offs among density-based and scan-statistic-based methods (As shown in Table 1) [13].

Table 1. Distribution of ACLED Security Events by Mission Area and Event Type (2010--2024)

Mission Area	Battles	VAC	Exp/RV	Riots	Protests	Strat.Dev.	Total
MINUSCA	4,128	3,256	1,043	892	1,563	382	11,264
MONUSCO	8,947	6,183	2,561	1,834	3,672	897	24,094
UNMISS	6,523	5,891	1,287	1,442	2,836	653	18,632
UNIFIL	1,236	987	2,143	1,054	1,498	366	7,284
Total	20,834	16,317	7,034	5,222	9,569	2,298	61,274

Note. Data source: ACLED Data Export Tool (acleddata.com/data/), extracted March 2025. VAC = Violence against civilians; Exp/RV = Explosions/Remote violence; Strat.Dev. = Strategic developments.

3.2. Temporal Pattern Mining and Seasonal Decomposition

3.2.1. Time-Series Decomposition of Incident Frequency

Monthly event frequency time series (180 observations per mission area) were decomposed using Seasonal and Trend decomposition using Loess (STL) into three additive components: a long-term trend (T), a seasonal component (S) with 12-month periodicity, and a residual (R). The trend component captures gradual shifts in conflict intensity; the seasonal component isolates recurring intra-annual patterns linked to agricultural cycles, dry/rainy season transitions, or political events; the residual represents irregular fluctuations. A seasonal smoothing window of 13 months and trend window of 25 months were applied.

3.2.2. Autocorrelation and Periodicity Analysis

The autocorrelation function (ACF) and partial autocorrelation function (PACF) were computed for detrended monthly series to examine temporal dependence structure. Periodogram analysis based on the discrete Fourier transform was applied to identify dominant periodicities beyond the assumed 12-month cycle. Identified periodicities and STL-derived seasonal indices were incorporated as temporal features in the classification task, providing information about cyclical conflict dynamics that static spatial and contextual features cannot capture (As shown in Figure 1).

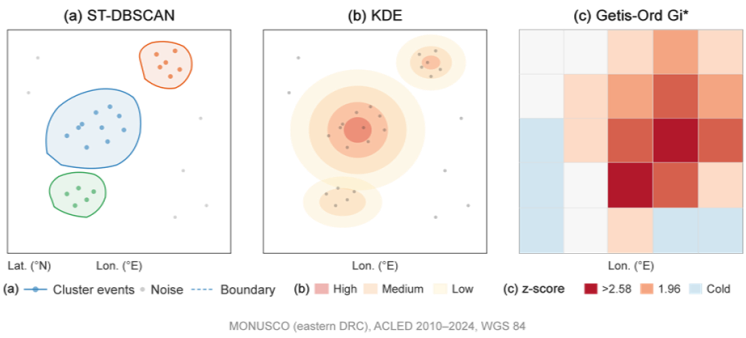

Figure 1. Comparison of Spatiotemporal Clustering Results Across Four UN Mission Areas

Each row represents one mission area (MINUSCA, MONUSCO, UNMISS, UNIFIL); each column displays one clustering method. ST-DBSCAN clusters appear as irregularly shaped polygons; KDE results as continuous density surfaces with 90th percentile contours highlighted; Getis-Ord Gi* output as choropleth maps of significant PRIO-GRID cells. Base maps from Natural Earth (naturalearthdata.com).

3.3. Classification of Security Incident Types under Class Imbalance

3.3.1. Classifier Selection and Configuration

The classification task assigns each event to one of six ACLED primary categories based on 18 engineered features. Random Forest was configured with 300 estimators, maximum depth of 15, and Gini impurity [14]. XGBoost employed 300 boosting rounds, maximum depth of 8, learning rate of 0.05, and L2 regularization ($\lambda = 1.0$) [15]. SVM with RBF kernel used $C = 10$ and $\gamma = 0.01$. All hyperparameters were optimized via Grid Search with stratified 5-fold cross-validation on the training set (70/30 split). Feature values were standardized to zero mean and unit variance prior to SVM training.

3.3.2. Resampling Strategies for Imbalanced Data

The event type distribution exhibits pronounced imbalance: Battles account for 34.0%, Violence against civilians 26.6%, Protests 15.6%, Explosions/Remote violence 11.5%, Riots 8.5%, and Strategic developments 3.8% (Table 1), an approximately 9:1 majority-to-minority ratio. SMOTE generates synthetic minority samples by interpolating between existing instances and their k-nearest neighbors ($k = 5$) [16]. ADASYN adjusts sampling density by local difficulty. SMOTE-Tomek applies oversampling followed by Tomek link boundary cleaning. All resampling was applied exclusively within training folds. Evaluation metrics included Macro F1-Score, PR-AUC, MCC, and Cohen's Kappa---selected for sensitivity to minority class performance where conventional accuracy and ROC-AUC provide inflated assessments.

4. Results and Discussion

4.1. Spatial Clustering Results and Hotspot Identification

4.1.1. Cluster Distribution Across Mission Areas

Table 2 summarizes clustering outputs. ST-DBSCAN identified 147 spatiotemporal clusters encompassing 34.6% of events (Silhouette = 0.412, noise ratio = 18.3%). KDE identified 89 high-density zones covering 52.1% of events with a lower Silhouette (0.358), reflecting inclusion of more diffuse concentrations. Getis-Ord Gi* identified 63 statistically significant hotspot cells covering 28.7% of events (Silhouette = 0.387). The ViEWS early-warning system has demonstrated the value of PRIO-GRID-month analysis units for conflict forecasting, and the Gi* results corroborate that grid-based aggregation facilitates hypothesis testing at the expense of spatial resolution compared to point-level methods [17].

Table 2. Spatial Clustering Method Parameters and Summary Results (Aggregated)

Method	Key Parameters	Clusters	Coverage	Silhouette	Noise Ratio
ST-DBSCAN	Eps1=25km, Eps2=30d, MinPts=5	147	34.6%	0.412	18.3%
KDE	BW=20km (Silverman), T=30d	89	52.1%	0.358	N/A

Getis-Ord Gi*	Grid=0.5°, p<0.05 (FDR)	63	28.7%	0.387	N/A

Note. Eps1 = spatial radius; Eps2 = temporal radius; MinPts = minimum points; BW = bandwidth; FDR = False Discovery Rate. Silhouette scores computed on spatial dimensions. Coverage = percentage of events within identified clusters or high-density zones.

ST-DBSCAN exhibited the greatest capacity to detect irregularly shaped clusters---particularly advantageous in MONUSCO, where conflict concentrates along transportation corridors and river networks in eastern provinces rather than forming symmetric patterns. In UNIFIL, where event volumes are lower and geography more compact, the three methods produced broadly convergent results identifying the southern border zone as the predominant hotspot. Hegre et al. showed that near-real-time candidate event data improves forecasting precision by 20--40%, suggesting that clustering analyses would similarly benefit from incorporating the most current event data [18].

4.1.2. Contextual Factor Analysis of Identified Hotspots

Comparison of PRIO-GRID attributes between ST-DBSCAN hotspot cells and non-hotspot cells reveals distinguishable profiles. Hotspot cells exhibit higher population density (mean: 147 vs. 89 persons/km^2), lower mountainous terrain coverage (12.4% vs. 19.7%), and shorter distances to international borders (78 km vs. 142 km). Lindberg Bromley, using the UCDP Peacemakers at Risk dataset, documented that violence against peacekeepers in Sub-Saharan Africa concentrated in zones characterized by intersecting armed group territories, aligning with multi-actor hotspots identified in this study for MONUSCO and MINUSCA [19]. The correlation between identified hotspots and contextual variables enhances interpretability for operational commanders who can relate analytical results to observable terrain and demographic conditions.

4.2. Temporal Pattern Analysis Results

STL decomposition (Figure 2) reveals mission-specific temporal structures. In MONUSCO, the trend component shows sustained increase during 2012--2017 corresponding to M23 rebel escalation and subsequent armed group operations in eastern DRC. The seasonal component indicates peak activity during October through December, coinciding with improved terrain accessibility after the rainy season. In UNMISS, dual seasonal peaks appear: a primary peak during December through February (dry-season cattle raiding) and a secondary peak during June and July. The UNIFIL series is dominated by irregular spikes rather than consistent seasonality, reflecting the event-driven nature of the Lebanon security environment where escalations are tied to cross-border incidents rather than environmental cycles.

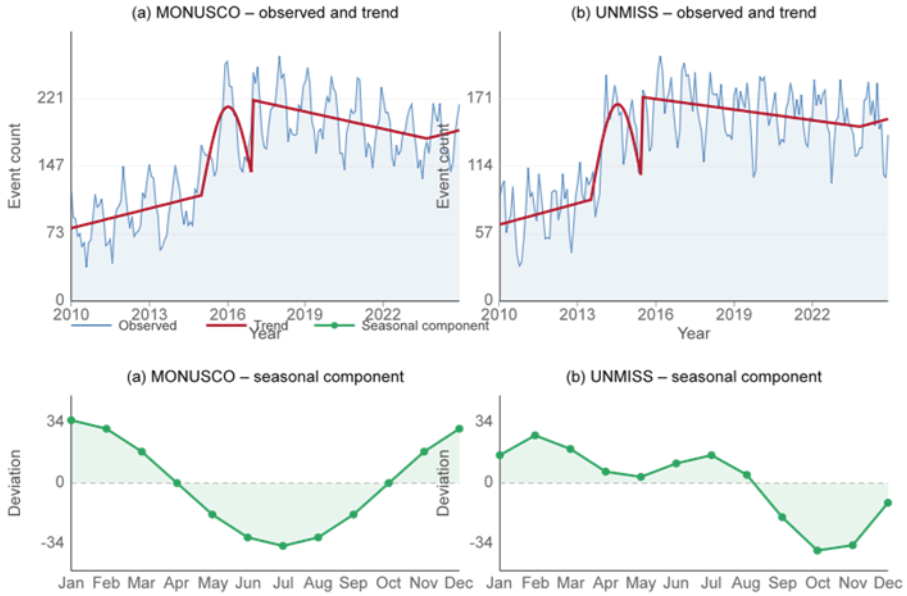

Figure 2. STL Decomposition of Monthly Security Event Frequency for MONUSCO and UNMISS (2010--2024)

Upper panel: observed, trend, seasonal, and residual components for MONUSCO. Lower panel: same decomposition for UNMISS. MONUSCO shows seasonal peaks during October--December dry season months; UNMISS displays dual peaks in December--February and June--July, the latter aligning with inter-communal violence patterns.

Autocorrelation analysis confirms significant positive autocorrelation at lags of 1, 2, and 12 months for both MONUSCO and UNMISS ($p < 0.01$), indicating both short-term persistence effects and annual cyclicity. Periodogram analysis detects a dominant 12-month cycle for MONUSCO (23.4% of detrended variance) and a 6-month cycle for UNMISS (18.7%), consistent with the dual-peak seasonal pattern in the STL decomposition. MINUSCA exhibits a weaker but discernible 12-month signal (14.2% of variance), with activity peaking during March through May when road networks become navigable for armed group movements in the CAR interior. These temporal features, integrated as classification inputs, provide information about cyclical conflict dynamics that complements the static geographic and demographic attributes captured by PRIO-GRID (As shown in Figure 3).

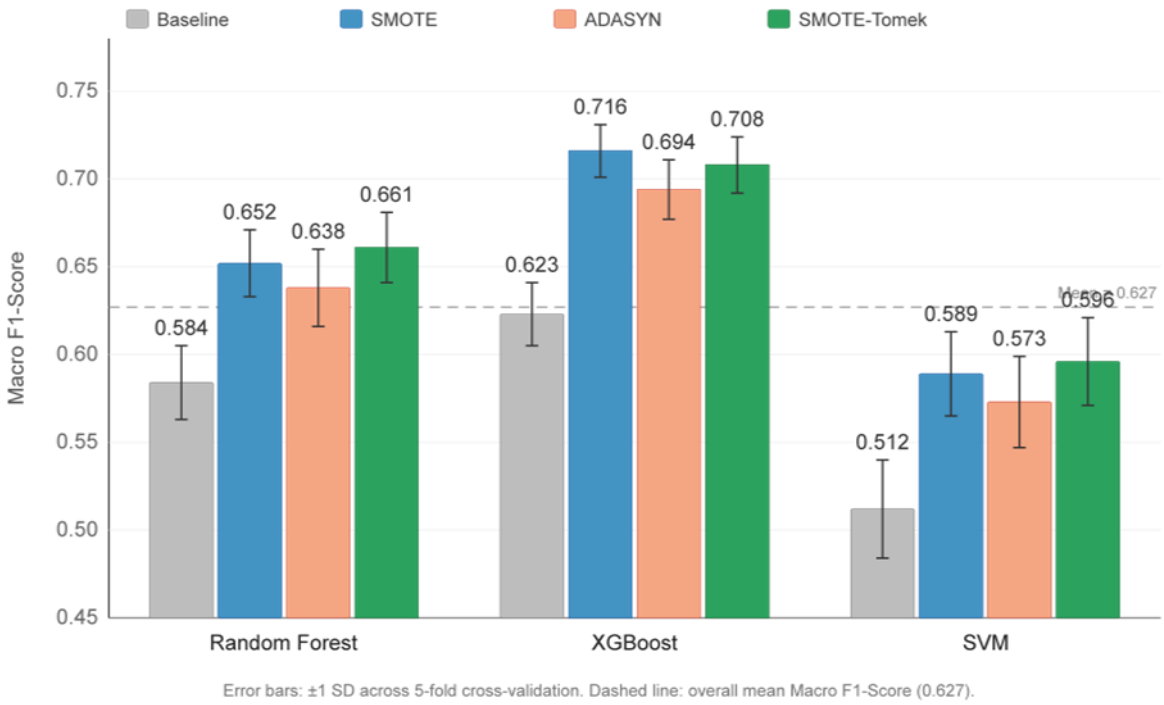

Figure 3. Macro F1-Score Comparison Across Classifier-Resampling Configurations

Grouped bar chart displaying Macro F1-Score for 12 configurations (3 classifiers × 4 conditions). X-axis groups by classifier family; y-axis ranges from 0.45 to 0.75. Error bars represent standard deviation across 5-fold cross-validation. XGBoost + SMOTE achieves the highest mean (0.716); SVM Baseline the lowest (0.512).

4.3. Classification Performance Comparison

4.3.1. Classifier Performance under Different Resampling Strategies

Table 3 presents classification performance across all 12 configurations. XGBoost consistently outperforms Random Forest and SVM across all metrics and resampling conditions. The highest performance is achieved by XGBoost + SMOTE (Macro F1 = 0.716, PR-AUC = 0.661, MCC = 0.683, Kappa = 0.674), representing a 14.9% relative gain over XGBoost Baseline (F1 = 0.623). Guan et al. demonstrated that extended neighborhood definitions in SMOTE variants enhance minority class representation; in this study, standard SMOTE outperformed ADASYN by 2.2 percentage points for XGBoost, suggesting that ADASYN's adaptive density-weighting may introduce borderline noise where class boundaries are inherently ambiguous [20]. The CSBBoost framework proposed by Jia et al. represents a potential direction for future improvement [21].

Table 3. Classification Performance Across Classifier-Resampling Configurations (5-Fold CV)

Configuration	Macro F1	PR-AUC	MCC	Kappa
RF (Baseline)	0.584	0.521	0.547	0.538
RF + SMOTE	0.652	0.597	0.621	0.613
RF + ADASYN	0.638	0.581	0.607	0.598
RF + SMOTE-Tomek	0.661	0.608	0.632	0.624
XGBoost (Baseline)	0.623	0.558	0.589	0.577

XGBoost + SMOTE	0.716	0.661	0.683	0.674
XGBoost + ADASYN	0.694	0.642	0.662	0.651
XGBoost + SMOTE-Tomek	0.708	0.654	0.676	0.667
SVM (Baseline)	0.512	0.467	0.473	0.461
SVM + SMOTE	0.589	0.534	0.552	0.541
SVM + ADASYN	0.573	0.518	0.536	0.525
SVM + SMOTE-Tomek	0.596	0.541	0.559	0.548

Note. RF = Random Forest; SVM with RBF kernel. Macro F1 = macro-averaged F1-Score across six event types. All resampling applied to training folds only.

SVM consistently yields the lowest performance, with its Baseline at 0.512 in Macro F1. The performance gap between SVM and ensemble methods widens under imbalanced conditions, consistent with Vesco et al. who found that ensemble approaches demonstrate stronger generalization in conflict-related prediction tasks [22]. The resampling impact varies by classifier: XGBoost spans 9.3 percentage points between its best (SMOTE) and worst (Baseline) configurations, while SVM spans 8.4 points, indicating that resampling provides consistent benefits regardless of classifier choice, though the absolute ceiling is governed by the classifier's representational capacity.

4.3.2. Implications for Operational Risk Assessment

Table 4 presents top-10 features from the best-performing configuration. Historical conflict intensity (6-month window) emerges as the most influential predictor (0.142), confirming the autoregressive nature of conflict dynamics. Distance to capital ranks second (0.118), reflecting the center-periphery gradient of state control. Population density (0.097) and the STL-derived seasonal index (0.089) validate the utility of temporal pattern mining outputs as classification features. Terrain variables contribute meaningful but secondary information, consistent with their role as enabling conditions rather than direct event type drivers.

Table 4. Top-10 Feature Importance Scores from XGBoost + SMOTE

Rank	Feature	Importance
1	Historical conflict intensity (6-month window)	0.142
2	Distance to nearest national capital (km)	0.118
3	Population density (PRIO-GRID)	0.097
4	Seasonal index (STL-derived)	0.089
5	Mountainous terrain coverage (%)	0.076

6	Number of active armed groups in cell	0.068
7	Distance to nearest international border (km)	0.062
8	Temporal lag event count (t−1 month)	0.055
9	Forest cover percentage	0.047
10	Economic activity level (G-Econ)	0.041

Note. Importance scores computed using the gain-based method. Features ranked by average contribution to loss reduction across all boosting rounds. PRIO-GRID features linked through spatial join (Section 2.2).

The practical implications warrant careful qualification. The classification performance (best Macro F1 = 0.716) reflects moderate discriminative ability across six event types with considerable per-class variation: Battles and Violence against civilians achieve F1-Scores exceeding 0.78, while Strategic developments reaches only 0.43, underscoring that the most operationally valuable categories remain the most difficult to classify. The integration of hotspot membership and temporal periodicity features into the classification pipeline demonstrates that a multi-stage analytical approach can produce complementary information streams for operational decision-making without requiring a unified end-to-end system.

5. Conclusion

This study conducted a comparative evaluation of spatiotemporal clustering and classification methods applied to security incident data from four active UN peacekeeping mission areas, using publicly accessible ACLED records (2010--2024) supplemented by PRIO-GRID geospatial features. The spatial clustering comparison demonstrated that ST-DBSCAN offers the greatest flexibility for detecting irregularly shaped hotspot regions conforming to real-world conflict geographies along transportation corridors, border zones, and river networks. Getis-Ord Gi* provides the advantage of statistical significance testing at the grid-cell level, while KDE generates the broadest coverage at the cost of including lower-intensity zones. Temporal analysis through STL decomposition and autocorrelation confirmed mission-specific seasonal patterns: a dry-season peak in MONUSCO, dual annual peaks in UNMISS, and irregular dynamics in UNIFIL.

The classification comparison established XGBoost paired with SMOTE as the strongest-performing configuration (Macro F1 = 0.716, MCC = 0.683), outperforming Random Forest and SVM across all resampling conditions. The consistent benefit of SMOTE across classifiers indicates that synthetic oversampling is a reliable strategy for conflict event classification. These findings should be interpreted within their methodological scope: the study does not claim predictive capability for future events but provides an empirical comparison of analytical methods that can inform tool selection for peacekeeping risk assessment. The incorporation of PRIO-GRID contextual features enhances result interpretability for operational commanders who can relate outputs to observable field conditions.

References

1. C. De Coning, "Predictive peacekeeping: Strengthening predictive analysis in UN peace operations," *Stability: International Journal of Security and Development*, vol. 8, no. 1, pp. 1–11, 2019, doi: 10.5334/sta.663.
2. C. A. Dos Santos Cruz, "Improving security of United Nations peacekeepers: We need to change the way we are doing business," United Nations Department of Peacekeeping Operations, New York, NY, USA, Tech. Rep., 2017.
3. L. E. Cederman and N. B. Weidmann, "Predicting armed conflict: Time to adjust our expectations?" *Science*, vol. 355, no. 6324, pp. 474–476, 2017, doi: 10.1126/science.aal4483.

4. M. Murphy, E. Sharpe, and K. Huang, "The promise of machine learning in violent conflict forecasting," *Data & Policy*, vol. 6, p. e35, 2024, doi: 10.1017/dap.2024.27.
5. C. Raleigh, A. Linke, H. Hegre, and J. Karlsen, "Introducing ACLED: An armed conflict location and event dataset," *Journal of Peace Research*, vol. 47, no. 5, pp. 651–660, 2010, doi: 10.1177/0022343310378914.
6. R. Sundberg and E. Melander, "Introducing the UCDP Georeferenced Event Dataset," *Journal of Peace Research*, vol. 50, no. 4, pp. 523–532, 2013, doi: 10.1177/0022343313484347.
7. K. Eck, "In data we trust? A comparison of UCDP GED and ACLED conflict events datasets," *Cooperation and Conflict*, vol. 47, no. 1, pp. 124–141, 2012, doi: 10.1177/0010836711434463.
8. A. F. Tollefsen, H. Strand, and H. Buhaug, "PRIO-GRID: A unified spatial data structure," *Journal of Peace Research*, vol. 49, no. 2, pp. 363–374, 2012, doi: 10.1177/0022343311431287.
9. S. Davies, T. Pettersson, M. Sollenberg, and M. Öberg, "Organized violence 1989–2024, and the challenges of identifying civilian victims," *Journal of Peace Research*, vol. 62, no. 4, 2025.
10. D. Birant and A. Kut, "ST-DBSCAN: An algorithm for clustering spatial-temporal data," *Data & Knowledge Engineering*, vol. 60, no. 1, pp. 208–221, 2007, doi: 10.1016/j.datak.2006.01.013.
11. Y. Hu, F. Wang, C. Guin, and H. Zhu, "A spatio-temporal kernel density estimation framework for predictive crime hotspot mapping and evaluation," *Applied Geography*, vol. 99, pp. 89–97, 2018, doi: 10.1016/j.apgeog.2018.08.001.
12. Z. Jiang and S. Shekhar, "Statistically-robust clustering techniques for mapping spatial hotspots: A survey," *ACM Computing Surveys*, vol. 55, no. 2, pp. 1–38, 2023, doi: 10.1145/3487893.
13. Y. Kang, N. Cho, J. Yoon, S. Park, and J. Kim, "Spatiotemporal data clustering: A survey of methods," *ISPRS International Journal of Geo-Information*, vol. 8, no. 3, Art. no. 112, 2019, doi: 10.3390/ijgi8030112.
14. L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001, doi: 10.1023/A:1010933404324.
15. T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2016, pp. 785–794, doi: 10.1145/2939672.2939785.
16. N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *Journal of Artificial Intelligence Research*, vol. 16, pp. 321–357, 2002, doi: 10.1613/jair.953.
17. H. Hegre, C. Bell, P. Vesco, et al., "ViEWS2020: Revising and evaluating the ViEWS political violence early-warning system," *Journal of Peace Research*, vol. 58, no. 3, pp. 599–611, 2021, doi: 10.1177/0022343320962157.
18. H. Hegre, M. Croicu, K. Eck, and S. Högbladh, "Introducing the UCDP Candidate Events Dataset," *Research & Politics*, vol. 7, no. 3, pp. 1–8, 2020, doi: 10.1177/2053168020935257.
19. S. Lindberg Bromley, "Introducing the UCDP Peacemakers at Risk Dataset, Sub-Saharan Africa 1989–2009," *Journal of Peace Research*, vol. 55, no. 1, pp. 122–131, 2018, doi: 10.1177/0022343317735882.
20. H. Guan, L. Zhao, X. Dong, and C. Chen, "Extended natural neighborhood for SMOTE and its variants in imbalanced classification," *Engineering Applications of Artificial Intelligence*, vol. 124, Art. no. 106570, 2023, doi: 10.1016/j.engappai.2023.106570.
21. Z. Jia et al., "A cluster-based SMOTE both-sampling (CSBBoost) ensemble algorithm for classifying imbalanced data," *Scientific Reports*, vol. 14, Art. no. 5283, 2024, doi: 10.1038/s41598-024-55598-1.
22. P. Vesco et al., "United they stand: Findings from an escalation prediction competition," *International Interactions*, vol. 48, no. 4, pp. 860–896, 2022, doi: 10.1080/03050629.2022.2070745.