CREATIVE SCIENCE

Journal home page: csj.snru.ac.th



A geospatial clustering and visualization system for gaur encroachment monitoring in the Khao Phaeng Ma Non-Hunting Area, Thailand

Ekachai Thammasat*, Natthaphon Ardhan, Saichon Rittikraithirathorn and Phanisorn Soisaengthong

Expert Center of Innovative Industrial Robotics and Automation, Thailand Institute of Scientific and Technological Research, Pathum Thani, 12120, Thailand

Article history:

Received: 06-05-2026
Revised : 16-10-2026
Accepted: 20-11-2026
Published: 23-11-2026

Keywords:

Human-wildlife conflict;
Gaur encroachment;
DBSCAN;
Geospatial dashboard.

Abstract

In Thailand, traditional manual patrolling is increasingly inadequate for mitigating human-wildlife conflicts in high-value ecological zones such as the Khao Phaeng Ma Non-Hunting Area. This study applies a geospatial data analytics framework to characterize the spatio-temporal distribution of gaur (*Bos gaurus*) encroachment, adapting the Cross-Industry Standard Process for Data Mining (CRISP-DM) methodology to a longitudinal dataset of ranger patrol observations (N = 474 incidents) spanning October 2023 to April 2026. Exploratory data analysis revealed a temporal pattern in which incident frequency peaked mid-week and declined markedly on weekends a pattern consistent with variation in human activity levels, though potentially confounded by patrol scheduling. Applying the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm (ε = 500 m, MinPts = 2) across three ecologically defined seasons Summer/Dry Peak (February–May, n = 240), Rainy/Wet Season (June–October, n = 101), and Winter/Dry Onset (November–January, n = 133) the framework identified high-density encroachment zones comprising 10 dispersed spatial clusters during summer, 4 concentration zones during the rainy season, and 9 encroachment hotspot areas during winter. Of the 474 records, 41 (8.6%) were classified as noise points (Summer: 14, Rainy: 12, Winter: 15) and excluded from cluster characterization. Cluster analysis identified critical high-risk interfaces at the intersections of the western forest boundary, local road networks, and seasonal cash-crop plantations. These outputs were integrated into an interactive Geospatial Dashboard to support ranger patrol planning and early-warning deterrent placement. Findings are subject to observation bias from variable patrol effort and the absence of individual animal identification, acknowledged as limitations and directions for future research.

Check for updates

* **Corresponding author** : ekachai@tistr.or.th

DOI: https://doi.org/10.55674/cs.v19i1.268385

© [2026] The Author(s).

1. Introduction

As the global expansion of human-dominated landscapes increasingly overlaps with wildlife habitats, human-wildlife conflict has emerged as a critical challenge in modern conservation biology [1]. In tropical regions, interactions between local agricultural communities and large megafauna often result in severe socio-economic consequences, including crop depredation, livestock predation, and fatal encounters [2, 3]. In Thailand, wildlife conservation efforts through the establishment of protected areas have successfully recovered several threatened species. However, this recovery has paradoxically intensified human-wildlife conflict at the forest-agriculture interface, particularly concerning the Asian elephant (*Elephas maximus*) and the globally vulnerable gaur (*Bos gaurus*) [4, 5]. This ecological dilemma is critically evident in the Khao Phaeng Ma Non-Hunting Area within the Dong Phayayen, Khao Yai Forest Complex, a designated UNESCO World Heritage site [6]. Following successful reforestation initiatives initiated in 1994, Khao Phaeng Ma Non-Hunting Area transformed from a degraded landscape into a robust secondary forest, catalyzing a remarkable recovery of the local gaur population [7]. Recent ecological surveys indicate that the gaur population has surged to approximately 300 individuals, a density that the park hosts substantial gaur populations that regularly range into adjacent agricultural areas of the isolated 8-square-kilometer reserve [8, 9]. Consequently, seasonal resource scarcity drives these large bovines beyond the protected boundaries, leading to chronic crop raiding in surrounding

maize, cassava, and fruit orchards [10]. Throughout this paper, a gaur encroachment incident (hereafter 'record' or 'sighting') refers to a single confirmed ranger observation of one or more gaur individuals outside the designated forest boundary of the Khao Phaeng Ma Non-Hunting Area., recorded at a specific location and time regardless of herd size. Beyond agricultural losses, the encroachment has precipitated a surge in wildlife-vehicle collisions on peripheral road networks, consistent with documented patterns of wildlife-vehicle collisions in adjacent protected areas of the Dong Phaya Yen-Khao Yai Forest Complex [11].

Traditional mitigation strategies in Khao Phaeng Ma Non-Hunting Area, such as manual ranger patrolling and static electric fencing, have proven increasingly inadequate against the dynamic spatial ecology of gaurs [12]. The highly contiguous and irregular movement patterns of gaurs necessitate a paradigm shift from reactive intervention to proactive, evidence-based management [13]. Recently, the integration of intelligent innovation, the Internet of Things (IoT), and machine learning (ML) has substantially advanced wildlife monitoring, offering scalable solutions for early warning systems (EWS). Advanced data analytics algorithms can now process vast amounts of geospatial sighting records to identify spatial concentration patterns of gaur encroachment [14, 15].

This study presents a retrospective geospatial data analytics framework for characterizing gaur (*Bos gaurus*) encroachment patterns in the Khao Phaeng Ma Non-Hunting Area. The primary objective is to apply density-based spatial clustering to identify historical hotspot distributions and the spatio-temporal factors associated with recorded wildlife occurrences. By transforming raw ranger patrol logs into an interactive geospatial visualization tool, this system provides evidence-based insights to support park authorities in patrol planning and resource allocation, contributing to the broader goal of sustainable human–wildlife coexistence. To guide this analysis, three specific research questions were formulated: (RQ1) to identify the primary spatial hotspots of recorded gaur encroachment within the Khao Phaeng Ma Non-Hunting Area and to examine how cluster configurations vary across ecological seasons; (RQ2) to determine whether encroachment hotspot distributions differ significantly across the three ecological seasons (Summer/Dry Peak, Rainy/Wet Season, and Winter/Dry Onset); and (RQ3) to characterize the temporal patterns including day-of-week and time-of-day associated with elevated sighting frequencies in the ranger patrol dataset.

2. Materials and Methods

To systematically analyze the spatial distribution of recorded gaur encroachment, this study adapts the CRoss-Industry Standard Process for Data Mining (CRISP-DM) framework to transform retrospective geospatial datasets into decision-support visualizations [16, 17]. Within this framework, evaluating point-based observational data presents analytical challenges due to inherent spatial noise, isolated sightings, and varying observation densities [18, 19]. To address these spatial complexities and execute a structured data science lifecycle, the methodology is partitioned into six sequential phases:

Domain Understanding

The Khao Phaeng Ma Non-Hunting Area, situated in Wang Nam Khiao District, Nakhon Ratchasima, Thailand, encompasses a mountainous terrain strategically connected to the northern boundary of the Dong Phayayen-Khao Yai Forest Complex [6]. Fig. 1 illustrates the geographic location of the study area, covering approximately 8 square kilometers. As a high-value ecological zone, it serves as a high-density core for an estimated population of 300 gaurs. This figure has the park supports substantial gaur populations that regularly range into adjacent agricultural areas, creating a classic threshold response where resource depletion forces wildlife into anthropogenic landscapes [8, 20]. The habitat transitions from interior core and buffer zones into highly vulnerable forest edges and encroachment zones. These borders intersect directly with local road networks and peripheral commercial maize plantations, which act as dominant spatio-temporal drivers for wildlife encroachment [10].

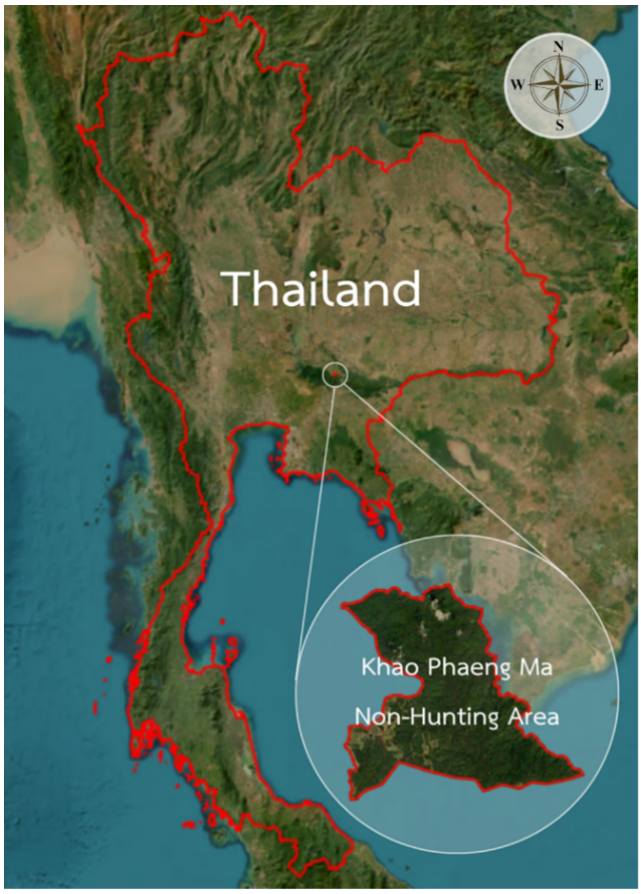

Fig. 1 Geographic location and landscape characteristics of the Khao Phaeng Ma Non-Hunting Area.

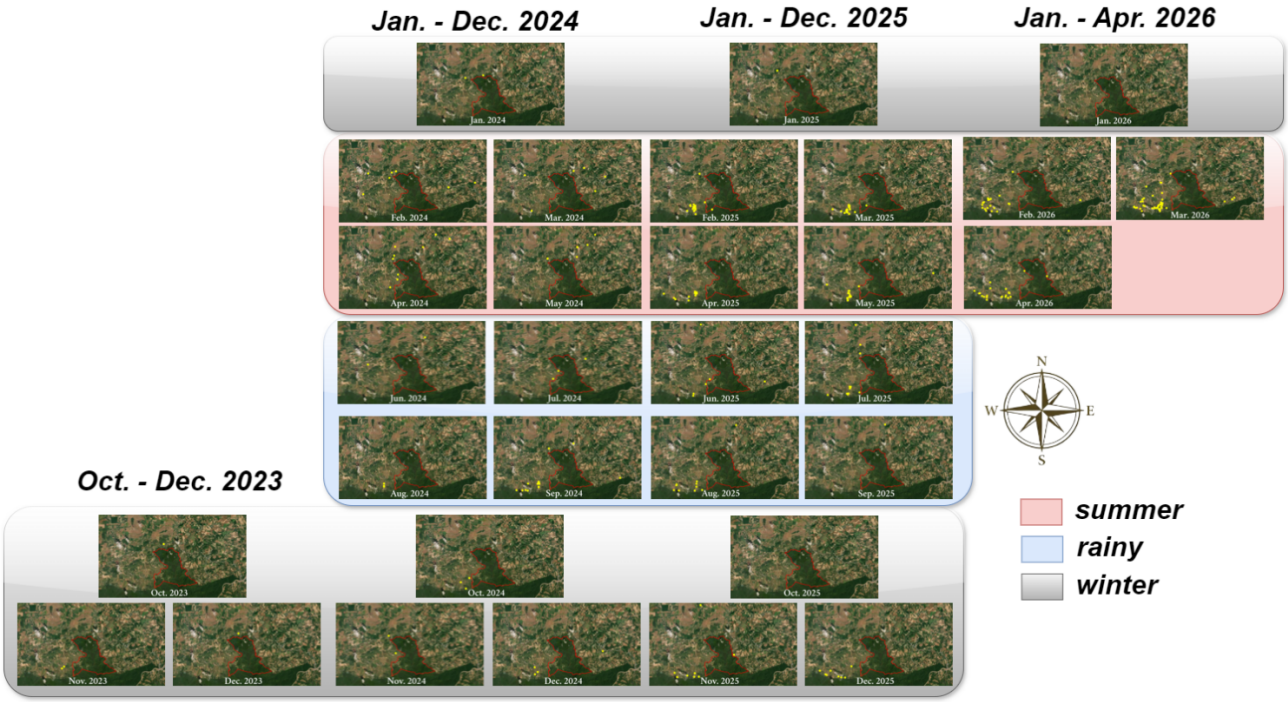

Fig. 2 Monthly temporal distribution of recorded gaur encroachment (Oct. 2023 – Apr. 2026).

Data analysis, supported by recent behavioral studies, reveals that movement patterns are heavily dictated by these surrounding crop cycles, with though field literature associate's gaur encroachment with post-16:00 and nocturnal movements [9], recorded sightings in this study peak at midday on Wednesdays, a pattern consistent with ranger patrol scheduling. To address chronic crop damage and rising traffic accidents, this study synthesized stakeholder requirements to prioritize key influencing factors, such as seasonal water availability and harvest cycles. Implementing an integrated data analytics framework successfully identified high-risk spatial zones, which align with zones of documented wildlife-vehicle conflicts along adjacent highway corridors in the Dong Phaya Yen-Khao Yai Forest Complex [11]. These findings are further substantiated by a systematic longitudinal dataset initiated in 2023, providing a robust empirical foundation to enhance conservation measures.

Data Collection

The data collection phase was initiated alongside the project proposal in 2023, establishing a continuous monitoring system for wildlife encroachment. While traditional ranger-based monitoring often suffers from observation bias, systematically standardizing patrol logs into structured geospatial formats significantly enhances their empirical utility for conservation modeling [21]. Under this framework, park rangers from Khao Phaeng Ma Non-Hunting Area have been systematically recording field observations on a daily basis, including species identification, precise dates, timestamps, and geographic coordinates. Patrol routes systematically cover the Khao Phaeng Ma Non-Hunting Area boundary and its adjacent community interface zones, encompassing the western agricultural edge, peripheral road networks, and village settlements in Wang Nam Khiao District. Patrol effort is maintained on both weekdays and weekends; however, recorded sighting frequencies are notably lower on weekends; this pattern may reflect variation in patrol scheduling rather than intrinsic gaur behavior, and is discussed further in Section 4, which acts as an anthropogenic deterrent that temporarily suppresses gaur activity in monitored zones rather than reflecting any reduction in patrol effort. Additional patrols are initiated upon receipt of community-reported encroachment incidents from surrounding villages.

Specifically, the longitudinal dataset analyzed in this study encompasses the period from October 2023 to April 2026. The temporal and spatial distributions of these recorded gaur encroachments are compiled and visualized at monthly intervals, represented as the yellow dots illustrated in Fig. 2. Preliminary observations indicate that the majority of these records are heavily concentrated toward the western region, particularly during the onset of the summer season. These records are maintained within the organization's internal database, serving as a centralized repository for operational data used in both daily management and long-term conservation efforts. This longitudinal dataset provides the foundational evidence required to analyze spatio-temporal patterns and serves as the primary ground-truth data for training and validating the spatial pattern analysis.

Field observations were recorded by park rangers using mobile smartphones equipped with the Google Maps GPS application (horizontal accuracy: ±3–15 meters depending

on canopy conditions). Each record represents a single gaur encroachment incident (hereafter interchangeably referred to as a 'sighting' or 'record'): one confirmed ranger encounter logged at the time and location of observation, regardless of herd size. Coordinate data were manually entered into structured Microsoft Excel spreadsheets at the end of each patrol shift and subsequently uploaded to the project's web-based Geospatial Dashboard for centralized data management and visualization. Two periods of incomplete data coverage were identified: October 2025 and January 2026, attributable to manual logging constraints and human recording error at the field level. Records from these periods were retained in the dataset to preserve temporal continuity; analysts should note that sighting counts for these months may be underrepresented relative to periods with consistent patrol coverage.

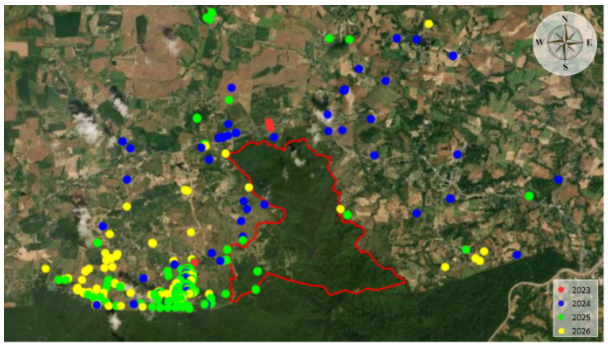

Fig. 3 Yearly spatial distribution of recorded gaur encroachment.

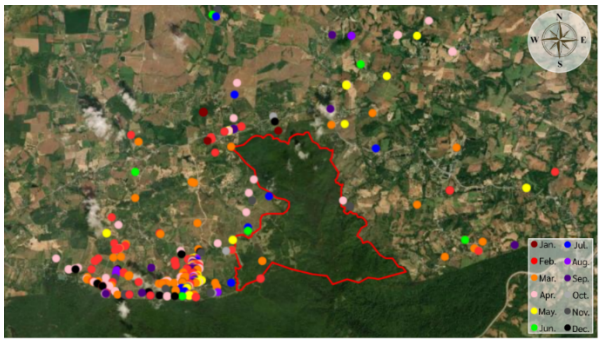

Fig. 4 Monthly spatial distribution of recorded gaur encroachment.

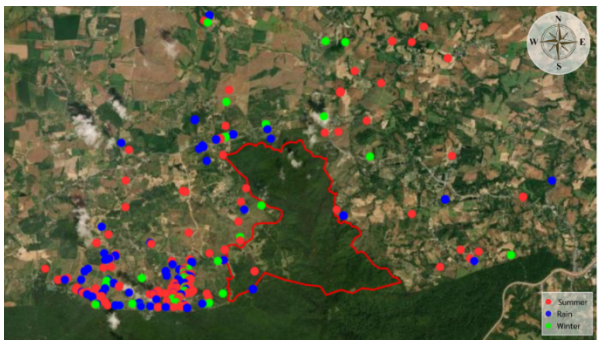

Fig. 5 Seasonal spatial distribution of recorded gaur encroachment.

Data Pre-processing

To transform raw ranger field logs into an analytically viable format, a multi-stage pre-processing pipeline was engineered, yielding the structured yearly, monthly, and seasonal spatial distributions in Figs. 3, 4, and 5. Managing wildlife tracking data requires rigorous validation to mitigate localization errors inherent in manual GPS logging [22]. First, data cleaning algorithms were applied to eliminate anomalous records with missing timestamps or invalid GPS coordinates falling outside the 8-square-kilometer study boundary. Second, spatial coordinates were projected onto the Universal Transverse Mercator Zone 47N coordinate reference system (EPSG: 32647), enabling Euclidean distance calculations in meters as required for the DBSCAN proximity analysis and aggregated across distinct temporal horizons.

The dataset was stratified into three ecological seasons aligned with the Thai meteorological calendar: the Summer season (Dry Peak: February–May), the Rainy season (Wet Season: June–October), and the Winter season (Dry Onset: November–January). This seasonal classification reflects both the national meteorological standards of the Thai Meteorological Department and the agricultural crop cycles of surrounding communities, which directly influence gaur movement and encroachment behavior.

This spatial aggregation is crucial for harmonizing disparate data points and revealing historical encroachment patterns without the distortion of spatio-temporal autocorrelation [23]. Finally, continuous timestamps and dates were decoded into discrete categorical features, explicitly capturing day-and-night cycles, the post-16:00 peak activity periods, and regional wet or dry seasons [24]. This systematic refinement successfully compiled the unstructured database into a tabular master feature matrix optimized for downstream spatial clustering and spatial pattern analysis.

Exploratory Data Analysis

Following the pre-processing phase, an Exploratory Data Analysis (EDA) was conducted to identify baseline behavioral patterns for conservation management. Temporal analysis of the dataset reveals that the peak in recorded gaur sightings occurs during midday on Wednesdays, suggesting a specific window of elevated land usage mid-week when human disturbance is minimal. Conversely, the analysis of repeat sightings across monthly, weekly, and hourly scales highlights a significant anomaly: on Saturdays, recorded sightings drop to near-zero, with virtually no recurring presence in the monitored zones. These temporal findings are consistent with variation in human activity levels across the week; however, interpreting the lower weekend sighting frequency as gaur behavioral avoidance requires patrol effort data to distinguish between reduced gaur activity and reduced ranger coverage. This limitation is discussed further in Section 5. Consequently, to account for these distinct shifts in movement, a seasonal spatial clustering approach was determined to be essential for the spatial analysis framework.

Data Modeling

To extract robust spatial features, the processed observation records were analyzed using the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm. This unsupervised learning technique is specifically chosen for its superior ability to identify core encroachment zones of arbitrary shapes and effectively isolate spatial outliers within geospatial datasets [18, 25]. Unlike centroid-based partitioning methods such as K-Means, DBSCAN does not require a prior specification of the number of clusters, making it uniquely suited for ecological monitoring where the true number of wildlife aggregation zones is inherently unknown and highly variable [25]. Compared to kernel density estimation (KDE), DBSCAN provides explicit, discrete cluster boundaries and formally designates noise points, enabling direct seasonal comparison without arbitrary bandwidth selection. Compared to spatial autocorrelation statistics such as Getis-Ord Gi*, DBSCAN operates on raw incident coordinates without rasterization or a pre-defined analysis scale, preserving the point-process nature of the sighting data. Hierarchical DBSCAN (HDBSCAN) was also considered but rejected in favor of fixed-ε DBSCAN to maintain an ecologically interpretable and seasonally consistent neighborhood radius (ε = 500 m) aligned with the estimated home-range movement scale of gaur groups in this landscape. Furthermore, because spatial patterns of gaur encroachment typically follow irregular geographic features rather than uniform geometric layouts, DBSCAN's density-connectivity principle is ideal for discovering highly contiguous, non-spherical clusters [26].

Most importantly, the algorithm explicitly designates low-density peripheral data as noise, preventing anomalous, non-recurring sightings or inherent GPS telemetry errors from skewing the core distribution analysis [27]. Building directly upon the insights derived from the exploratory data analysis specifically the temporal anomalies on weekends and the post-16:00 shifts the dataset was stratified by season and operational windows before clustering. This stratification ensures that the algorithm's hyperparameters are calibrated to ecological realities rather than global averages. The algorithm operates by defining an ε -neighborhood for every point p, denoted as $N_\epsilon(p)$, which is mathematically expressed as:

$$N_\epsilon(p) = \{q \in D \mid d(p,q) \leq \epsilon\} \quad (1)$$

where D represents the dataset of gaur encroachment coordinates, and $d(p,q)$ is the Euclidean distance between points. By specifying a minimum point threshold ($MinPts$), the algorithm distinguishes between core points, which satisfy the density condition $|N_\epsilon(p)| \geq MinPts$, border points, and noise points (isolated records that do not belong to any dense cluster) [18]. This clustering process filters out erratic spatial noise caused by GPS inaccuracies or temporary manual logging errors, yielding a dataset that accurately represents core encroachment zones and high-risk areas. Ultimately, this spatial baseline ensures that the spatial analysis is based on statistically significant density patterns rather than accidental spatial fluctuations.

Data Visualization

The final phase of the framework involves transforming the spatial clustering outputs into an actionable Geospatial Dashboard, as shown in Fig. 6. This web-based system utilizes Geographic Information System (GIS) tools to overlay the identified high-density encroachment zones directly onto local topographic and community maps. Rather than serving as a static visualization, this dynamic dashboard functions as a critical decision-support tool for park authorities [28, 29]. It integrates the temporal rhythms identified during the exploratory data analysis with the spatial boundaries computed by the DBSCAN model to automatically highlight conflict zones. By interacting with the user interface, conservation managers can review current-period encroachment indicators, prioritize ranger patrol routes toward historically dense encroachment zones, and allocate manpower to sectors identified as high-frequency in the seasonal cluster analysis [15].

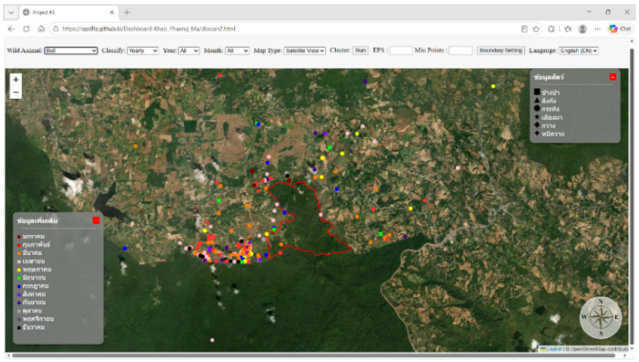

Fig. 6 Actionable Geospatial Dashboard.

3. Results and Discussion

Empirical evaluation of the longitudinal dataset indicates that gaur encroachment outside the Khao Phaeng Ma Non-Hunting Area is substantially influenced by seasonal resource depletion and anthropogenic activity cycles [8, 20]. Furthermore, the spatio-temporal distributions (illustrated across Figs. 2, 3, 4, 5) demonstrate a pronounced escalation in encroachment events from January to June. During this dry-season temporal window, spatial coordinates are densely aggregated along the western boundary, specifically at the critical ecological interface where the protected forest intersects with peripheral commercial maize and jackfruit plantations [10]. Exploratory data analysis further unmasked distinct behavioral adaptations to human presence: gaur activity peaks mid-week on Wednesdays when ambient community noise is stable but drops to near-zero on Saturdays, a pattern consistent with reduced patrol activity on weekends, as documented in the Data Collection section. Whether this reflects intrinsic gaur behavior or patrol scheduling artifacts cannot be determined from presence-only records alone and remains a limitation of this study.

Table 1 Seasonal clusters.

Season	Cluster ID	Map Indicator Color	Central Latitude	Central Longitude	Associated Geographic Core / Interface	Number of Points
Summer	Cluster 1	Red	14.36209°N	101.76569°E	Western Boundary Interface	188
(Dry Peak)	Cluster 2	Orange	14.38986°N	101.77325°E	North-Western Border Zone	12
	Cluster 3	Pink	14.36233°N	101.74760°E	Western Agricultural Edge	4
	Cluster 4	Yellow	14.37968°N	101.76576°E	Central-Western Buffer Zone	3
	Cluster 5	Green	14.36767°N	101.82314°E	Eastern Peripheral Forest Edge	4
	Cluster 6	Blue	14.40956°N	101.81012°E	Northern Road Intersection	3
	Cluster 7	Purple	14.36422°N	101.77991°E	South-Central Fringe	4
	Cluster 8	Dark Purple	14.37863°N	101.83381°E	Eastern Commercial Orchard Interface	2
	Cluster 9	Grey	14.39901°N	101.79709°E	North-Central Transit Interface	3
	Cluster 10	Black	14.39119°N	101.79538°E	North-Central Buffer Zone	3
	Noise	-	-	-	-	14
Rainy	Cluster 1	Red	14.36169°N	101.76621°E	Core Southwest Encroachment Zone	68
(Wet Season)	Cluster 2	Orange	14.35862°N	101.75129°E	South-Western Low-Disturbance Edge	11
	Cluster 3	Pink	14.39025°N	101.77401°E	North-Western Forest Edge	6
	Cluster 4	Yellow	14.40888°N	101.79613°E	Northern Sector Fringe	4
	Noise	-	-	-	-	12
Winter	Cluster 1	Red	14.39368°N	101.76784°E	North-Western Concentration Zone	72
(Dry Onset)	Cluster 2	Orange	14.36634°N	101.77377°E	South-Western Boundary Buffer	14
	Cluster 3	Pink	14.36170°N	101.75974°E	South-Western Agricultural Border	10
	Cluster 4	Green	14.36064°N	101.77955°E	Southern Border Zone	6
	Cluster 5	Blue	14.38890°N	101.77145°E	North-Western Edge Interface	5
	Cluster 6	Grey	14.41385°N	101.77076°E	Far-Northern Highway Intersection	3
	Cluster 7	Dark Grey	14.37502°N	101.79784°E	South-Central Concentration Area	3
	Cluster 8	Black	14.37821°N	101.81827°E	South-Eastern Sector Edge	3
	Cluster 9	Uncolored	14.39091°N	101.78287°E	North-Central Interior Fringe	2
	Noise	-	-	-	-	15

By applying the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm, the pre-processed coordinate data was successfully segmented into high-density spatial zones, effectively isolating core encroachment zones from stochastic logging errors or temporary wandering (classified as noise via Equation 1) [22, 27]. The granular geographic centroids and color-mapped profiles of these detected seasonal clusters are systematically documented in Table 1. Because the algorithm does not assume rigid spherical shapes, it successfully traced highly irregular and contiguous paths.

The spatial clustering of recorded gaur encroachment exhibits distinct variations across different ecological periods, showing widespread dispersal across 10 clusters along the western boundary and peripheral zones during the summer season (Fig. 7), spatial consolidation into 4 primary clusters around the southwest encroachment zone during the rainy season (Fig. 8), and a shift into 9 clusters expanding toward the northern and southern sectors during the winter season (Fig. 9). By filtering out erratic spatial noise through the calibrated neighborhood distance ($\varepsilon = 500$ meters) and minimum points thresholds ($MinPts = 2$), The ε value of 500 meters corresponds to the estimated neighborhood radius for gaur groups in this landscape. These parameter values (ε = 500 m, $MinPts$ = 2) were applied consistently across all three seasonal subsets (Summer, Rainy, and Winter) to ensure direct comparability of cluster structures across ecological periods. The $MinPts$ value of 2 was intentionally set at the minimum threshold to ensure that even repeated but spatially isolated sightings which represent high-priority, low-frequency encroachment corridors are captured as clusters rather than discarded as noise. This conservative parameterization is appropriate for the limited dataset size characteristic of manual ranger-log systems in small protected areas. Spatial clustering was implemented using the DBSCAN package (version 1.1.12)

in R statistical software (version 4.3.2). Noise points (cluster label = 0) were excluded from geographic characterization but are reported to assess algorithm sensitivity (Summer: 14/240, 5.8%; Rainy: 12/101, 11.9%; Winter: 15/133, 11.3%). To assess the robustness of the clustering results to parameter choices, a sensitivity analysis was conducted using alternative hyperparameter combinations: ε = 300 m with $MinPts = 2$; ε = 700 m with $MinPts = 2$; and ε = 500 m with $MinPts = 3$. The analysis revealed that across all tested configurations, the number of seasonal clusters ranged from 4 to 11 in summer, 2 to 5 in the rainy season, and 2 to 9 in winter (totaling 8 to 23 clusters across parameter combinations). Crucially, the primary cluster boundaries and the dominant Cluster 1 encroachment zone (Western Boundary Interface) were robustly identified under all parameter settings without spatial drift. These findings confirm that while finer or broader sub-clusters merge or fractionate depending on spatial scale, the core spatial encroachment topology remains stable under reasonable parameter variations. The selected baseline values (ε = 500 m, $MinPts = 2$) were retained as the optimal configuration based on their superior ecological interpretability for ranger patrol operations and strong alignment with the inflection point of the *k*-distance plot.

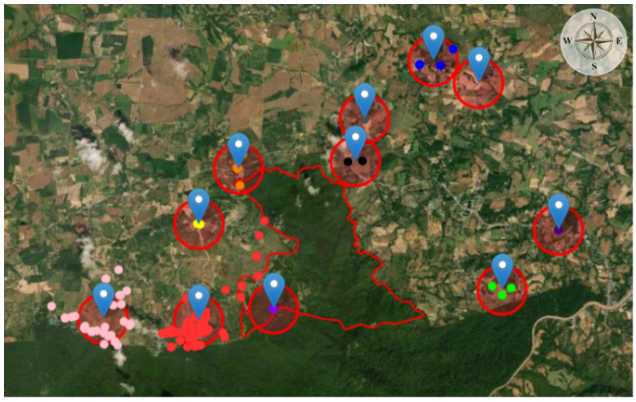

Fig. 7 Spatial clustering of recorded gaur encroachment during the summer season.

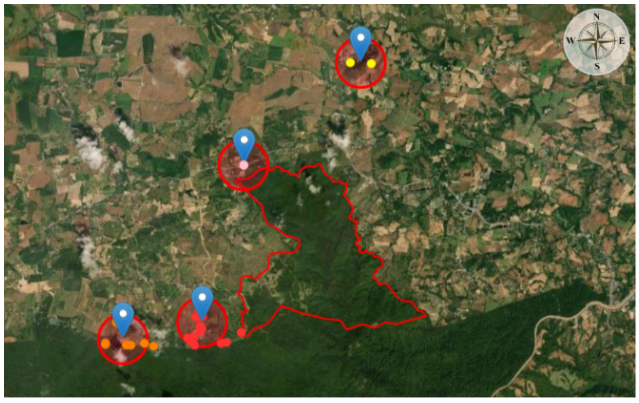

Fig. 8 Spatial clustering of recorded gaur encroachment during the rainy season.

The framework confirmed that sighting density is geographically concentrated, with Cluster 1 (Western Boundary Interface) consistently accounting for the largest proportion of incidents across all seasons. Specifically, during the Summer (Dry Peak), the model detected a high degree of herd dispersal, yielding 10 distinct clusters where intense encroachment tightly hugs the Western Boundary Interface (Cluster 1: 14.36209°N, 101.76569°E) and extends to critical peripheral areas such as the Northern Road Intersection (Cluster 6) and the Eastern Commercial Orchard Interface (Cluster 8). Conversely, during the Rainy (Wet Season), the abundance of internal forest resources leads to spatial consolidation, reducing the output to 4 primary clusters concentrated heavily around the Core Southwest Encroachment Zone (Cluster 1: 14.36169°N, 101.76621°E). As the landscape transitions into Winter (Dry Onset), the gaurs shift their spatial focus toward the northern and southern sectors, establishing a major North-Western Concentration Zone (Cluster 1: 14.39368°N, 101.76784°E) and expanding into the Far-Northern Highway Intersection (Cluster 6) and South-Central Concentration Area (Cluster 7). The identified high-risk clusters spatially coincide with agricultural zones reported by rangers as frequent crop-raiding sites, suggesting that encroachment patterns are substantially influenced by agricultural land placement along the western forest boundary.

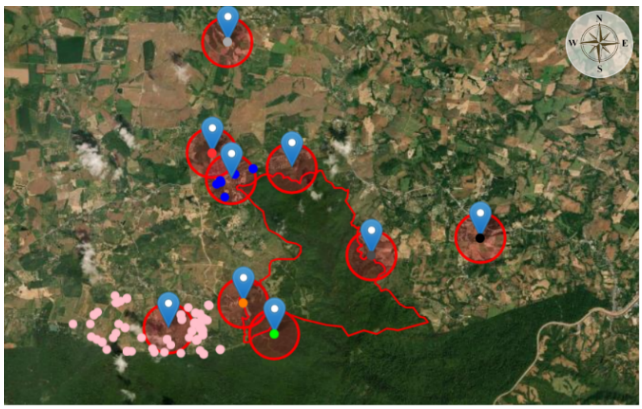

Fig. 9 Spatial clustering of recorded gaur encroachment during the winter season.

The culmination of this framework, the web-based Geospatial Dashboard (Fig. 6), successfully translates complex unsupervised learning outputs into current-period encroachment intelligence accessible for field operations [28]. Prior to the deployment of this framework, park rangers relied entirely on manual, resource-intensive patrolling, which left surrounding communities highly vulnerable to sudden nocturnal crop raiding, particularly during daylight patrol hours (particularly the midday window of 10:00–14:00, as identified in the EDA), consistent with operational patrol scheduling rather than necessarily reflecting intrinsic crepuscular behavior. By overlaying the dynamic DBSCAN cluster outputs onto community topographic maps, the dashboard provides an

intuitive decision-support interface that enables evidence-based manpower allocation [29]. This allows managers to prioritize patrol coverage toward the identified western encroachment zones, particularly during the summer season (February–May) when cluster density is highest. Decisions regarding deterrent placement and patrol scheduling remain at the discretion of park authorities and are informed, not automated, by the dashboard outputs. Ultimately, these results demonstrate that transitioning from reactive manual monitoring to a data-driven evidence-based framework supports better-informed decision-making by conservation authorities while safeguarding the ecological corridor under the Thailand 4.0 initiative. It is important to note that the identified encroachment hotspots may partially reflect areas of higher patrol intensity rather than absolute gaur preference, as the dataset is inherently limited to ranger-observed incidents.

4. Conclusion

This study presents a retrospective geospatial data analytics framework for characterizing the spatio-temporal distribution of gaur (*Bos gaurus*) encroachment incidents in the Khao Phaeng Ma Non-Hunting Area, Nakhon Ratchasima Province, Thailand. Applying the CRISP-DM methodology and the DBSCAN algorithm (ε = 500 m, $MinPts$ = 2; R 4.3.2, DBSCAN v1.1.12) to 474 ranger patrol observations spanning October 2023 to April 2026, the framework identified 23 spatial encroachment clusters across three ecological seasons: 10 clusters during the Summer/Dry Peak (February–May), 4 clusters during the Rainy/Wet Season (June–October), and 9 clusters during the Winter/Dry Onset (November–January). The highest cluster concentration was consistently observed along the western forest boundary interface, a zone of persistent gaur–agricultural community overlap. The spatio-temporal patterns identified in this study provide an empirical foundation for evidence-based patrol planning. Seasonal variation in cluster geography indicates that patrol priorities should be adjusted across ecological seasons rather than maintained as static year-round allocations.

The accompanying Geospatial Dashboard translates these historical cluster outputs into an accessible, web-based interface enabling park rangers to visualize encroachment hotspots overlaid on operational maps, supporting tactical patrol scheduling. This study is subject to several important limitations. The dataset consists of ranger-observed incidents, and sighting frequency therefore reflects patrol coverage intensity as well as gaur activity; patrol effort data were unavailable for formal bias correction. The retrospective, descriptive nature of the analysis precludes causal inference regarding behavioral drivers of encroachment. The absence of individual animal identification data prevents movement trajectory analysis. Cluster patterns may partially reflect areas of higher patrol intensity rather than areas of absolute gaur preference. Future research should incorporate systematic patrol effort recording, environmental covariates (vegetation phenology, water availability, crop calendars), and camera trap or GPS-collar data to enable predictive spatial risk modeling and behavioral interpretation.

5. Suggestions

Based on the operational challenges encountered during the longitudinal monitoring phase (specifically the data gaps in October 2025 and January 2026 due to manual logging constraints and human error), future research should prioritize the integration of automated data acquisition systems. Transitioning from manual ranger logs to a systematic deployment of network-connected camera traps (camera trapping systems) equipped with computer vision algorithms would significantly mitigate human-induced reporting errors. This hybrid data-collection approach will ensure an uninterrupted, high-fidelity stream of spatio-temporal data, thereby improving the continuous calibration and predictive accuracy of density-based clustering models (DBSCAN) within the smart conservation framework.

6. Acknowledgement

This research was supported by the Fundamental Fund for the Fiscal Year 2026, allocated through the Science, Research and Innovation Promotion Fund, under the project titled “Development of an Intelligent Detection and Warning System for Conserving Forest Resources and Preventing Wildlife from Leaving Forest Reserves.” The authors would also like to express their sincere gratitude to the Khao Phaeng Ma Non-Hunting Area for providing operational facilities and invaluable data support throughout this study. Special thanks are extended to the park rangers for their dedication in maintaining the daily field logs.

7. Declaration of Generative AI in Scientific Writing

The authors declare that generative artificial intelligence (AI) tools were utilized solely to enhance the readability, language clarity, grammar, and academic style of the manuscript. The underlying scientific content, experimental design, data analysis, interpretation of results, and conclusions were entirely conceived and validated by the authors. Consequently, the authors maintain full responsibility for the accuracy, originality, and integrity of the submitted work.

8. CRediT Author Statement

Ekachai Thammasat: Methodology, Investigation, Data Curation, and Original Draft. **Natthaphon Ardhan:** Conceptualization, Methodology, and Data Collection. **Saichon Rittikraithirathorn:** Data Processing. **Phanisorn Soisaengthong:** Data Visualization.

9. Research Involving Human and Animals Rights

Not applicable.

10. Ethics Approval and Consent to Participate

This study was conducted using secondary observational data collected as part of routine operational wildlife

monitoring by Khao Phaeng Ma Non-Hunting Area rangers. No direct animal capture, handling, or experimental intervention was performed. The data was provided to the research team under an institutional collaboration between TISTR and the Khao Phaeng Ma Non-Hunting Area. A formal research permit application to the Department of National Parks, Wildlife and Plant Conservation (DNP) is currently under review. Coordinate data is restricted to authorized personnel in accordance with wildlife protection protocols.

11. Declaration of Competing Interest

The authors declare no competing financial interests or personal relationships that could have influenced the work reported in this paper.

12. Data Availability Statement

Data Availability Statement: The GPS coordinate data used in this study are restricted to authorized personnel of the Khao Phaeng Ma Non-Hunting Area due to wildlife protection and conservation security concerns. Aggregated cluster statistics are available from the corresponding author upon reasonable request.

13. References

[1] Dickman, A. J. (2010). Complexities of conflict: the importance of considering social factors for effectively resolving human–wildlife conflict. *Animal Conservation*, *13*(5), 458-466. https://doi.org/10.1111/j.1469-1795.2010.00368.x

[2] Yng, T. K., & Hasan, N. N. N. (2022). Human–wildlife conflicts and the impact on local communities' support for Khao Yai National Park, Thailand. *Parks*, *28*(1), 61–70. https://doi.org/10.2305/iucn.ch.2022.parks-28-1tky.en

[3] Suttidate, N., Radeloff, V. C., & Sukmasuang, R. (2025). Agricultural land use change drives crop-raiding of Asian elephants in a human-elephant conflict zone in Thailand. *Global Ecology and Conservation*, *64*, e03945. https://doi.org/10.1016/j.gecco.2025.e03945

[4] Chockcharoen, R., Pharejaem, T., Saisamorn, A., & Pattanaviboor, A. (2020). Gaur recovery and management in Khao Phang Ma Non-Hunting Area, northeastern Thailand. *BULLetin*, *3*, 15–22.

[5] Supanta, J., Poolkhet, C., Yindee, M., Manatchaiworakul, W., & Wongtawan, T. (2025). Human-elephant conflict in Thailand over the past decade (2014-2023): Occurrence, geographical distribution, and temporal trends. *Animals*, *15*(9), 1304. https://doi.org/10.3390/ani15091304

[6] IUCN. (2025, October 11). *Dong Phayayen-Khao Yai forest complex: 2025 conservation outlook assessment*. IUCN World Heritage Outlook. https://whc.unesco.org/en/list/590/documents/

[7] Prayong, N., & Srikosamatara, S. (2017). Cutting trees in a secondary forest to increase gaur *Bos gaurus* numbers in Khao Phaeng Ma Reforestation area, Nakhon Ratchasima Province, Thailand. *Conservation Evidence*, *14*, 5–9.

[8] Duengkae, P., Ariyaraphong, N., Tipkantha, W., Jairak, W., Baicharoen, S., Nguyen, D. H. M., Korboon, O., Singchat, W., Panthum, T., Ahmad, S. F., Kaewkhunjob, E., Chaisonkhram, C., Maikaew, U., Muangmai, N., Ieamsaard, G., Sripiboon, S., Paansri, P., Suksavate, W., Chaiyes, A., … Srikulnath, K. (2022). Coincidence of low genetic diversity and increasing population size in wild gaur populations in the Khao Phaeng Ma Non-Hunting Area, Thailand: A challenge for conservation management under human-wildlife conflict. *PLoS ONE*, *17*(8), e0273731. https://doi.org/10.1371/journal.pone.0273731

[9] Prayoon, U., Suksavate, W., Chaiyes, A., Paansri, P., Siriaroonrat, B., Utara, Y., Tipkantha, W., Baicharoen, S., Jairak, W., Kaewkhunjob, E., Chaisonkhram, C., Maikaew, U., Ieamsaard, G., Sripiboon, S., Vinitpornsawan, S., Tunhikorn, S., Wanna, N., Sriracha, S., Sornsa, T., … Duengkae, P. (2024). Home range and habitat utilization of gaur (*Bos gaurus*) in transition zone between protected forest and human-dominated landscape, Eastern Thailand. *Global Ecology and Conservation*, *50*, e02811. https://doi.org/10.1016/j.gecco.2024.e02811

[10] Prayoon, U., Suksavate, W., Chaiyes, A., Winitpornsawan, S., Tunhikorn, S., Angkaew, C., Wanna, N., Sriracha, S., Pattanakiat, S., Pattanavibool, A., & Duengkae, P. (2022). Importance of protected areas in conservation and recovery of gaur (*Bos gaurus*) in Thailand. *BULLetin*, (7), 25–33.

[11] Rattanawanawong, N., Bhumpakphan, N., Kutintara, U., & Sukmasuang, R. (2022). Wildlife-vehicle collisions in Khao Yai National Park, Thailand: Impact on native species for some conservation management. *Biodiversitas*, *23*(6), 3050–3061. https://doi.org/10.13057/biodiv/d230632

[12] Chaiyarat, R., Thongkrathok, P., Maisuwan, W., Chantra, A., Chimplee, J., Cheachean, N., Assawaklang, S., & Youngpoy, N. (2023). *Variation in water use by mammals in Khao Phaeng Ma Non-hunting Area, Thailand* [Preprint]. Research Square. https://doi.org/10.21203/rs.3.rs-3227096/v1

[13] Pimm, S., Alibhai, S., Bergl, R., Dehgan, A., Jewell, Z., Joppa, L., Kays, R., & Loarie, S. (2015). Emerging technologies to conserve biodiversity. *Trends in Ecology & Evolution*, *30*(11), 685-696. https://doi.org/10.1016/j.tree.2015.08.008

[14] Jebet, I. K., Chege, I. M., & Otuko, M. S. (2024, June 6). *The role of machine learning in mitigating human wildlife conflict in Kenya*. In *Proceedings of the 12th Data Science Africa Workshop (DSA 2024)*, Nyeri, Kenya, 1–5.

[15] Ronoh, E., Mirau, S., & Dida, M. (2022). Human-wildlife conflict early warning system using the Internet of Things and short message service. *Engineering, Technology & Applied Science Research*, *12*(2), 8273–8277. https://doi.org/10.48084/etasr.4662.

[16] Wirth, R., & Hipp, J. (2000, April 11–13). CRISP-DM: Towards a standard process model for data mining. In *Proceedings of the 4th International Conference on the Practical Applications of Knowledge Discovery and Data Mining*, Manchester, UK, 29-39.

[17] Bemthuis, R., Govers, R., & Asadi, A. (2025). A CRISP-DM-based methodology for assessing agent-based simulation models using process mining. *Journal of Simulation, 20*(5), 519–540. https://doi.org/10.1080/17477778.2025.2508245

[18] Ester, M., Kriegel, H. P., Sander, J., & Xu, X. (1996, August 2–4). A density-based algorithm for discovering clusters in large spatial databases with noise. In E. Simoudis, J. Han, U. Fayyad (Eds.), *KDD'96: Portland, OR,* USA*, 226–231.*

[19] Rahmani, A., Davoodi, M., & Calabrese, J. M. (2025). Stable Trajectory Clustering: An Efficient Split and Merge Algorithm. *arXiv preprint arXiv:2504.21808*. https://doi.org/10.48550/arXiv.2504.21808

[20] Nyhus, P. J. (2016). Human–wildlife conflict and coexistence. *Annual Review of Environment and Resources*, *41*, 143–171. https://doi.org/10.1146/annurev-environ-110615-085634

[21] Ahumada, J., Fegraus, E., Birch, T., Flores, N., Kays, R., O'Brien, T., Palmer, J., Schuttler, S., Zhao, J., Jetz, W., Kinnaird, M., Kulkarni, S., Lyet, A., Thau, D., Duong, M., Oliver, R., & Dancer, A. (2019). Wildlife Insights: A platform to maximize the potential of camera trap and other passive sensor wildlife data for the planet. *Environmental Conservation, 47*(1), 1–6. https://doi.org/10.1017/S0376892919000298

[22] Urbano, F., Cagnacci, F., Calenge, C., Dettki, H., Cameron, A., & Neteler, M. (2010). Wildlife tracking data management: A new vision. *Philosophical Transactions of the Royal Society B: Biological Sciences, 365*(1550), 2177–2185. https://doi.org/10.1098/rstb.2010.0081

[23] Fletcher, R. J., Hefley, T. J., Robertson, E. P., Zuckerberg, B., McCleery, R. A., & Dorazio, R. M. (2019). A practical guide for combining data to model species distributions. *Ecology, 100*(6), e02710. https://doi.org/10.1002/ecy.2710

[24] Calabrese, J., Fleming, C. H., & Gurarie, E. (2016). ctmm: An R package for analyzing animal relocation data as a continuous-time stochastic process. *Methods in Ecology and Evolution, 7*(9), 1124–1132. https://doi.org/10.1111/2041-210X.12559

[25] Hahsler, M., Piekenbrock, M., & Doran, D. (2019). dbscan: Fast density-based clustering with R. *Journal of Statistical Software*, *91*(1), 1–30. https://doi.org/10.18637/jss.v091.i01

[26] Edelhoff, H., Signer, J., & Balkenhol, N. (2016). Path segmentation for beginners: an overview of current methods for detecting changes in animal movement patterns. *Movement Ecology*, *4*(1), 21. https://doi.org/10.1186/s40462-016-0086-5

[27] Tomkiewicz, S. M., Fuller, M. R., Kie, J. G., & Bates, K. K. (2010). Global positioning system and associated technologies in animal behaviour and ecological research. *Philosophical Transactions of the Royal Society B: Biological Sciences, 365*(1550), 2163–2176. https://doi.org/10.1098/rstb.2010.0090

[28] Slingsby, A., & van Loon, E. (2016). Exploratory visual analysis for animal movement ecology. *Computer Graphics Forum*, *35*(3), 471–480. https://doi.org/10.1111/cgf.12923

[29] Aljibawi, M., Algabri, H., & Ibrahim Rasool, Z. (2025). Adaptive clustering using enhanced DBSCAN: A dynamic approach to optimizing density-based clustering. *Statistics, Optimization & Information Computing, 14*(4), 1980–1991. https://doi.org/10.19139/soic-2310-5070-2484