



Sea Turtle Trajectories Prediction via Long Short-Term Memory and Kalman-Filter

Carlos A. Rovetto R.^a, Eric E. Flores^b, Kexy Rodriguez^a,
Ivonne Nuñez^c, Andrzej Smolarz^d, Dimas Concepción^e,
and Elia E. Cano^a

^a Facultad de Ingeniería en Sistemas Computacionales, Universidad Tecnológica de Panamá, Panamá

^b Coiba Scientific Station (COIBA AIP), Panamá, Panamá

^c CDS Research Group, Institute of Computer Science, University of Bern, 3012 Bern, Switzerland

^d Lublin University of Technology, Department of Computer and Electrical Engineering, Poland

^e Escuela de Innovación Digital, Instituto Técnico Superior Especializado, Panamá, Panamá

✉ carlos.rovetto@utp.ac.pa, kexy.rodriguez@utp.ac.pa, elia.cano@utp.ac.pa; eflores@coiba.org.pa; ivonne.nunez@unibe.ch; a.smolarz@pollub.pl; dconcepcion@itse.ac.pa

KEYWORDS ABSTRACT

sea turtle
tracking;
*Chelonia
mydas*;
Time series
prediction;
Kalman filter;
geospatial
analysis;
biotelemetry;
machine
learning

Tracking sea turtle migration is hindered by noisy and incomplete geolocation data, as well as irregular sensor transmission. These limitations make it challenging to model trajectories and accurately interpret ecological patterns. This study presents a predictive framework for modelling the trajectories of green turtles (*Chelonia mydas*) using satellite telemetry and artificial intelligence techniques. Georeferenced data from SPOT-375B tags were pre-processed to address noise, data gaps, and spatial anomalies. A Long Short-Term Memory (LSTM) neural network was trained with normalized time series data to forecast future positions, capturing the temporal dependencies of turtle movement. A Kalman filter was applied post-prediction to enhance trajectory continuity and reduce uncertainty through recursive state estimation. Experimental results show that the approach yields an average MAE of 0.0986, MSE of 0.0307, and RMSE of 0.1288, and reduces mean prediction error by 43.75 % relative to a recurrent neural network (RNN) baseline while requiring ~36 % of its CPU time. This integrated pipeline enhances the reliability of wildlife trajectory forecasting and provides a scalable solution for ecological tracking under uncertain detection conditions, facilitating a deeper understanding of species behavior and more effective conservation strategies.

Carlos A. Rovetto R.; Kexy Rodriguez; Eric E. Flores;
Ivonne Nuñez; Andrzej Smolarz; Dimas Concepción;
Elia E. Cano

Sea Turtle Trajectories Prediction via Long Short-Term
Memory and Kalman-Filter



ADCAIJ: Advances in Distributed Computing
and Artificial Intelligence Journal
Regular Issue, Vol. 14 (2025), e33385
eISSN: 2255-2863 - <https://adcaij.usal.es>
Ediciones Universidad de Salamanca - CC BY-NC-ND

1. Introduction

The conservation of threatened marine species, such as the green turtle (*Chelonia mydas*), demands scientific tools capable of accurately understanding their movement patterns over time (Abang Shakawi et al., 2024). This species undertakes extensive migrations between feeding and nesting grounds, crossing diverse habitats under dynamic environmental conditions (Tariq et al., 2025). Understanding these displacements is essential for identifying critical protection zones, designing safe marine corridors, and establishing public policies to mitigate the effects of climate change, pollution, and other anthropogenic pressures (Perneel et al., 2025). To achieve this, it is crucial to have detailed spatiotemporal data that reflect the mobility of individuals in open ocean areas, where direct observation is virtually impossible (Chambault, Gaspar, et al., 2021; Noguchi et al., 2025).

In this context, satellite biotelemetry has transformed marine wildlife research by enabling remote and continuous monitoring of individuals in the wild (Alsaïdi et al., 2024). Through transmitters attached to the turtles' carapaces, georeferenced coordinates are collected and transmitted via systems such as Argos, allowing researchers to track migratory routes with precision (Ning et al., 2025). However, the data obtained come with several inherent challenges: intermittent transmissions due to animal submersion, signal loss under atmospheric interference, location errors, and irregular time intervals between records (Steenacker et al., 2023). These factors result in a heterogeneous and partially noisy dataset that requires rigorous preprocessing and advanced analytical techniques (Bokani et al., 2025; Chambault, Fossette, et al., 2021).

Over the past five years, the use of machine learning (ML) algorithms has shown steady growth in ecological and conservation applications (Santos et al., 2023). In particular, LSTM neural networks have proven effective in modeling nonlinear temporal sequences, with multiple studies applying them to analyze the movements of marine animals (Restrepo et al., 2023). For example, (Zakry et al., 2024) developed a framework for modeling and predicting animal movement using LSTM networks applied to behavioral and velocity data, demonstrating their utility for long-term simulations. Similarly, (Shakawi et al., 2024) proposed an architecture based on autoencoders and recurrent neural networks to improve trajectory imputation in time series with missing values, achieving notable improvements in accuracy over conventional methods. However, most of these studies do not comprehensively address data quality issues nor incorporate statistical filters to reduce positional uncertainty (Li et al., 2023; Wijeyakulasuriya et al., 2020).

Additionally, recent works have explored integrating predictive models with Bayesian estimation techniques, such as the Kalman filter, to smooth trajectories and enhance spatial coherence. (Rider et al., 2024) introduced a hybrid LSTM-H model that combines LSTM networks with Kalman filters to predict livestock movement in sensor-based monitoring environments, achieving up to 80 % error reduction compared to traditional approaches. Similarly, (Pasanisi et al., 2024) presented the KeySORT algorithm, which integrates an adaptive Kalman filter to estimate animal poses with improved temporal consistency. Nevertheless, these implementations remain scarce in marine studies, where data tend to be more variable, subject to immersion-induced errors, and characterized by extended periods without transmission (Jeantet et al., 2020; Kipnis et al., 2023).

These thematic connections are clearly illustrated in the keyword co-occurrence analysis derived from the scientific literature reviewed in this study, as illustrated in Figure 1. The conceptual map illustrates the clustering structure of key research topics, showing the interdisciplinary ecosystem surrounding sea turtle studies. The central node, «sea turtle», is directly connected to terms such as

hawksbill turtle, marine turtle, satellite tracking, and endangered species, underscoring the pivotal role of spatiotemporal tracking and conservation. Several thematic clusters emerge:

- Brown/Central Node (sea turtle): The term sea turtle appears as the largest and most connected node, serving as the central hub linking tracking technologies (satellite/telemetry), ecological themes (foraging, animal movement, endangered species), and computational approaches (machine/deep learning).
- Orange/Blue cluster (AI and deep learning): Includes deep learning, machine learning, neural networks, and artificial intelligence. This group connects with abundance estimation and aerial video footage, reflecting the growing use of computational models for video detection, pattern recognition, and population estimation.
- Turquoise cluster (telemetry and data reliability): Comprises telemetry, satellites, correlative model, and accuracy. It emphasizes data acquisition and validation methods in satellite tracking, highlighting methodological reliability in movement studies.
- Green cluster (movement ecology and conservation): Groups terms such as foraging, animal movement, marine vegetation, and endangered species (along with related nodes like scute pattern and Bayesian hierarchical). This cluster underscores the ecological and conservation-oriented context of long-term behavioral monitoring.
- Red cluster (marine turtles and population monitoring): Encompasses marine turtle, aerial video footage, and abundance estimation, stressing the importance of observation and counting techniques for population assessment.
- Pink cluster (complementary tracking technologies): Includes atlas tracking and accelerometer, linked to satellite tracking. It points to the integration of multiple tracking devices to enhance behavioral and movement inference.
- Purple cluster (satellite tracking and ecological-taxon linkages): Centered on satellite tracking, this cluster connects with tropical ecology and hawksbill turtle, functioning as a bridge between the central sea turtle node and other thematic clusters.

This semantic analysis reinforces the methodological approach adopted in this research, which integrates satellite sensors, predictive algorithms, and statistical inference to generate robust and interpretable models of animal behavior in marine environments. The convergence of these domains reflects an emerging scientific agenda that prioritizes data precision, temporal modeling, and applied utility for conservation.

Accordingly, the present study proposes a methodology for predicting and correcting migratory trajectories of green turtles, based on the integration of satellite sensors, machine learning, and geospatial processing. First, SPOT-375B transmitters were deployed on eleven green turtles monitored along the Pacific coast of Panama, generating over 22,000 georeferenced records. A rigorous data cleaning process was followed, including outlier detection, removal of land-based points using marine shapefiles, and temporal normalization. Subsequently, an LSTM network was trained to forecast future coordinates based on the time series of latitude and longitude. These predictions were then refined using a Kalman filter to reduce positional variance and improve trajectory continuity. Unlike prior studies, this approach not only focuses on prediction but also includes an interactive visualization phase through a georeferenced mobile application prototype. Furthermore, the study reports quantitative model

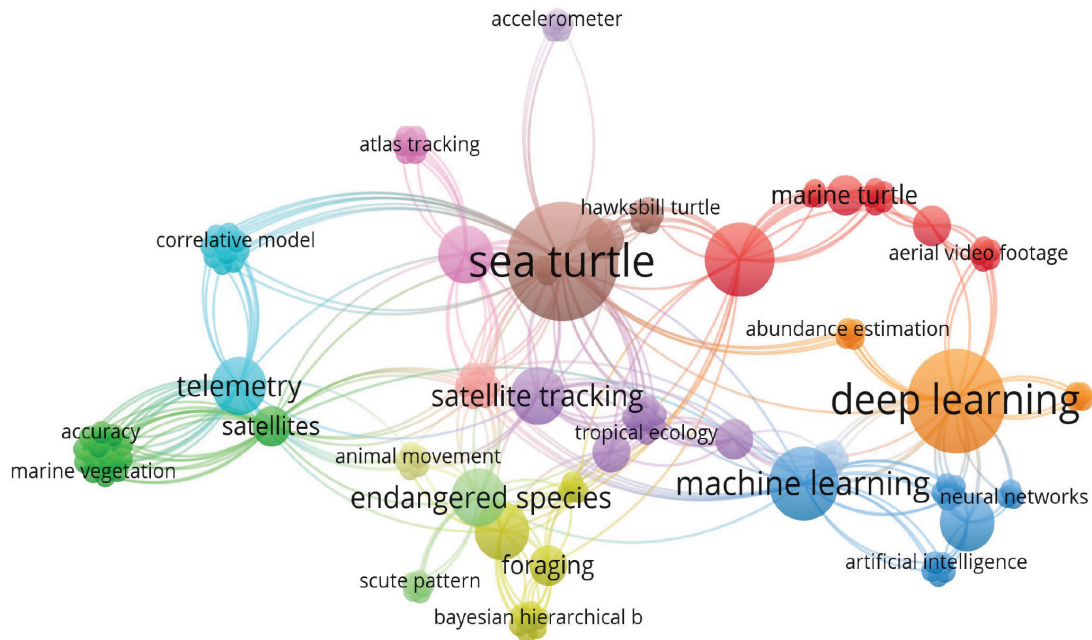


Figure 1. Map of co-occurrence of key terms in the study of satellite monitoring of sea turtles

performance metrics (MSE, MAE, RMSE) for each sensor, along with a comparative analysis between raw and smoothed data. This methodology enables not only accurate predictions but also coherent trajectories that can be interpreted by biologists, conservation managers, and policymakers. Lastly, the study outlines the potential applicability of the model to other marine species and its adaptability to data-limited scenarios.

The remainder of the paper is structured as follows: Section 2, Materials and Methods, details the procedures used in data collection, including the instruments, techniques employed, and methods of analysis. Section 3, Theory/Calculation, outlines the theoretical foundations and any relevant calculations or models applied to support the analysis. Section 4, Results, presents the main findings derived from the research. Section 5, Discussion, interprets these findings, highlighting their implications and relevance. Section 6, Limitations of this study, and finally, Section 7 Conclusions and Future Research, conclude by summarizing the key contributions and future steps for this work.

2. Materials and Methods

This section describes the instruments, data, and procedures used to predict and smooth the trajectories of *Chelonia mydas*. It details the biotelemetry system with SPOT-375B transmitters and the Argos link, as well as the collection of data from eleven turtles on Panama's Pacific coast. Subsequently, the preprocessing applied to clean records, prioritize reliable location classes, eliminate land points using marine masks, and normalize time series is explained. For modeling, the LSTM architecture trained on georeferenced sequences is presented, along with the application of a Kalman filter that

smooths trajectories, interpolates missing data, and dynamically adjusts noise according to the quality of the observations. Finally, the evaluation metrics (MSE, MAE, RMSE), the comparison with other reference models, and the development of a mobile visualization application are indicated.

2.1. Methodological Pipeline

As shown in Figure 2, the methodology follows a structured six-step pipeline. This integrated workflow enhances the accuracy and interpretability of sea turtle tracking, offering a robust tool for ecological monitoring and conservation efforts:

- 1) **Data Collection:** Eleven green turtles were equipped with SPOT-375B satellite tags, which transmitted location data via the Argos satellite system. These georeferenced records captured movement patterns along the Pacific coasts of Panama.
- 2) **Pre-processing:** The raw data were cleaned to remove duplicate entries, null values, and inaccurate positions caused by environmental noise or signal loss. This step also involved normalizing the time series inputs.
- 3) **Modeling with LSTM:** An LSTM neural network was trained on the pre-processed time series to learn the temporal dependencies in turtle movement. The model was validated using a split-training approach and optimized using early stopping criteria.
- 4) **Trajectory Smoothing:** To improve positional coherence and reduce geolocation uncertainty, a Kalman filter was applied to the model outputs, providing recursive correction and interpolation of the predicted states.
- 5) **Trajectory Prediction:** The smoothed output generated continuous trajectories for each turtle, enabling interpretation of their spatiotemporal behavior.
- 6) **Visualization:** A mobile application prototype was developed to visualize the predicted trajectories and movement history interactively on a geospatial interface.

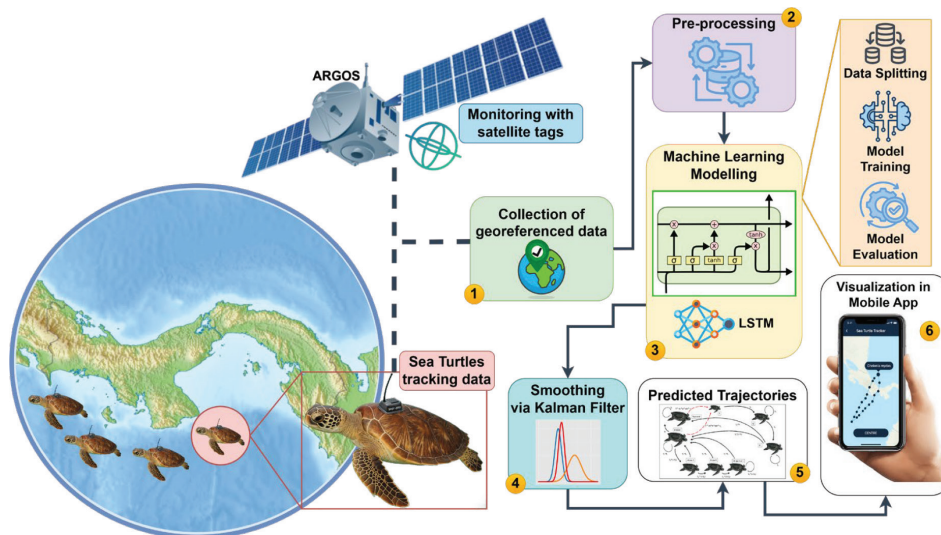


Figure 2. Methodological pipeline for sea turtle trajectory prediction and smoothing

2.2. Biotelemetry System and SPOT Sensor Specifications

Biotelemetry is a fundamental approach in modern wildlife research, enabling continuous and non-invasive monitoring of animal movements, physiology, and behavior through remote sensor systems. It bridges the gap between ecological field observation and quantitative modeling by providing high-resolution spatiotemporal data that reveals migration routes, habitat preferences, and behavioral patterns in dynamic environments. Among the various biotelemetry modalities, satellite telemetry offers the most extensive spatial coverage, making it particularly suitable for marine species that traverse large oceanic areas where direct observation is impossible (Azizan et al., 2023). In this study, SPOT-375B transmitters were attached to the shells of green turtles to record spatial coordinates during their migratory movements in the Pacific Ocean (Miao et al., 2024). Each device transmits signals via the Argos system at a frequency of 401.678 MHz (± 30 kHz), with individual identification using a PTT (Platform Transmitter Terminal) code (Benscoter et al., 2022). Transmissions depend on the immersion status of the sensor, which operates under the condition $\sigma \in [0.1, 5] \text{ S/m} \Leftrightarrow \text{antenna exposed to air}$, where σ represents the electrical conductivity of the medium surrounding the antenna. Transmission only occurs when the antenna is exposed to air, generating a set of observations that are not uniformly spaced in time (Mestre et al., 2024).

Let $T = \{t_1, t_2, \dots, t_n\} \subset \mathbb{R}^+$ denote the sequence of valid observation timestamps. The inter-observation intervals are bound by: $\Delta t_i = t_{(i+1)} - t_i, \Delta t_i \in [90, 600]$ seconds. Each valid georeferenced observation is represented as a triplet: $z_i = (\phi_i, \lambda_i, q_i), \forall i \in \{1, \dots, n\}$, where ϕ_i and λ_i are latitude and longitude (in WGS84 coordinates), and $q_i \in \{3, 2, 1, 0, A, B\}$, is the Argos location class indicating positional accuracy. The temporal density and quality of received data are influenced by biological and environmental factors, including dive duration, satellite pass availability, antenna orientation, and radio interference from nearby devices. These characteristics result in a non-uniform time series, which must be addressed during preprocessing and model design. The SPOT-375 B Argos tag, developed by Wildlife Computers, is a satellite-linked transmitter optimized for marine animals that frequently surface, such as sea turtles, sharks, and marine mammals. At the surface, the tag emits brief Argos transmissions, whose Doppler-shifted frequencies enable precise geolocation, with accuracies reaching approximately 250 meters under optimal conditions. Its pressure-sealed epoxy housing and integrated sensors ensure reliable operation in harsh marine environments and depths of up to 2000 m. The device supports fully programmable transmission schedules and provides key data products, including Argos locations, Time-at-Temperature (TAT) histograms, and Percent-Dry timelines. The SPOT-375B has an estimated operational lifespan of ~ 915 days and is engineered for minimal hydrodynamic impact, measuring $99 \times 55 \times 21$ mm and weighing 152 g. Table 1 summarizes the complete technical specifications of the SPOT-375 B Argos tag.

The main components of the SPOT-375B satellite tag are shown in Figure 3. The device is attached to the turtle's carapace following standard biotelemetry protocols to minimize stress and hydrodynamic drag. Each label indicates a functional element: (a) Argos transmission antenna for satellite communication; (b) communication port used for configuration and data retrieval; (c) LED status indicator that provides visual feedback during activation and transmission; (d) magnetic activation switch for manual start or shutdown; and (e) wet/dry immersion sensors that detect surfacing events and control transmission timing. These components directly affect the temporal resolution and positional accuracy of the transmitted data, shaping both the frequency and precision of geospatial observations used throughout this study (Rodríguez-Martínez et al., 2024; Rovetto et al., 2023).

Table 1. SPOT-375B (Argos) – Technical Specifications

Feature	Specification / Detail
Model	SPOT-375B (Hard-Shell Turtle / Argos)
Localization system	Argos Doppler (when antenna is surfaced)
Operating frequency	401.678 MHz
Integrated sensors	Temperature; Wet/Dry sensor
Temperature sensor range	−40 °C to +60 °C
Temperature sensor resolution	0.05 °C
Pressure rating / Maximum depth	~2000 m for most shapes/configurations
Dimensions (L × W × H)	~99 mm × 55 mm × 21 mm for the «Hard-Shell Turtle» version
Weight	~152 g for that configuration
Optimal storage temperature range	0 °C to 5 °C
Conductivity (Wet/Dry sensor) limits	0.1 to 5 S/m*
Estimated deployment/battery life	~915 days for that configuration (but variable, depending on programming, duty cycle, and environmental conditions)
Data products transmitted	Argos locations; Time-At-Temperature histogram (TAT); Percent-Dry timeline
Attachment / Mounting options	Over 50 different shapes/configurations available (fin mount, flipper mount, float mount, towed, anchored)
Recovery pinger (optional)	UHF pinger for active tracking and recovery when at surface
Customizable transmission schedule	User-configurable uplink schedule to optimize battery life and data resolution
Localization accuracy (Argos)	Up to ~250 m under good conditions (Argos LC3 class)
Transmission power	Nominal 0.5 W (varies by duty cycle).
Note on shape/config variability	Specifications (size, weight, battery life) vary depending on shape, mount, and deployment scenario.



Figure 3. SPOT-375B sensor specifications: (a) Argos transmission antenna, (b) communication port, (c) LED status indicator, (d) magnetic activation switch, and (e) wet/dry immersion sensors

2.3. Data Collection and Preprocessing

2.3.1. Data Collection

SPOT transmitters generate multiple data files in comma-separated values (CSV) format, each containing up to 32 columns with a mix of telemetry and sensor readings. These outputs are structured to support both location tracking and behavioral inference and require preprocessing prior to modeling. The primary files include:

- All.csv – Raw, undecoded transmissions, often used for device diagnostics.
- Argos.csv – Timestamped location data derived from satellite passes.
- Behavior.csv – Categorical records of inferred behaviors (e.g., surface activity).
- Corrupt.csv – Contains rejected messages, often flagged due to signal degradation.
- DDN.csv – Classifies hourly periods as dry, deep, or undefined immersion states.
- DivePDT.csv – Records depth-temperature profiles when available.
- Histos.csv – Aggregated histograms of activity, such as time spent at depth or temperature bins.

Additionally, the tags integrate Fastloc GPS technology, which enables location estimation even when the antenna is briefly exposed. Fastloc operates using a time-difference-of-arrival approach and a proprietary LocSolve algorithm, solving pseudo-range equations via least squares minimization. This method yields a high-accuracy geolocation point, even with short surfacing windows.

The resulting location data are assigned to Argos location classes based on estimated positional error, with Class 3 indicating the highest precision. These classifications guide subsequent filtering and model training stages. A summary of the location class distribution is presented in Table 2.

Table 2. Types of messages

Class	Estimated error	Messages received per satellite pass
3	< 250m	4 messages or more
2	250m << 500m	4 messages or more
1	500m << 1500m	4 messages or more
0	> 1500m	4 messages or more
A	No accurate estimation	2 messages
B	No accurate estimation	2 messages
Z	Invalid location (available only for Service Plus/Auxiliary Location)	

2.3.2. Data Preprocessing

Raw data transmitted from SPOT-375B sensors are subject to various sources of distortion during satellite communication, including antenna immersion, signal overlap from neighboring devices, and atmospheric interference (Wang et al., 2022). These factors introduce noise, message loss, and duplication, necessitating a robust preprocessing stage to prepare the data for subsequent modeling (Anselin & Rey, 2022). The collected data were parsed and evaluated according to their Argos location class, which reflects the estimated accuracy of each georeferenced point. As shown in Table 3, most records fall

into low-precision categories (Classes 1 and «Others»), yet a significant portion (19 %) corresponds to Classes 2 and 3, which offer sufficient spatial reliability for trajectory analysis. During preprocessing:

- Corrupted or duplicated entries were removed.
- Records classified as Class 2 or 3 were prioritized for model training and evaluation.
- Lower-accuracy data were retained only when their displacement magnitudes were within biologically plausible thresholds.

Additionally, georeferenced points falsely located on land were identified using a marine mask shapefile and excluded from the dataset. This ensured that all retained data points correspond to realistic sea-based movements. Table 3 summarizes the distribution of data quality per SPOT ID. On average, only 8.45 % of records were classified as Class 3, and 10.03 % were classified as Class 2, with over 60 % belonging to unclassified or low-confidence categories.

Table 3. Records by SPOT and percentage by category

SPOT ID	Total Data	Percent category messages Classes			
		3	2	1	Others
241136	3474	6.56	8.69	16.12	68.62
241137	1328	10.47	12.50	18.30	58.73
241138	1064	5.17	8.08	21.33	65.41
241139	2284	6.09	8.89	17.78	67.25
241140	1451	5.31	11.58	25.16	57.96
241141	1544	7.51	11.98	27.53	52.98
241142	1201	9.08	5.50	18.82	66.61
241143	1354	8.79	11.45	19.42	60.34
262534	1978	16.38	6.37	20.98	56.27
262536	3239	7.41	14.26	28.00	50.32
262537	3853	10.17	11.06	22.50	56.27
Averages		8.45	10.03	21.45	60.07

Table 4 reports the active tracking periods per transmitter. The average monitoring duration across the eleven SPOT devices was approximately 50 days, with some tags exceeding 90 days of transmission. These time ranges provide sufficient temporal resolution to analyze mid-scale migratory behavior and evaluate predictive model performance under varied tracking lengths. The selection of Class 2 and Class 3 data also aligns with statistical consistency:

- Standard deviation for Class 3: 106.71
- Standard deviation for Class 2: 130.03

These are significantly lower than the deviations for Class 1 (241.53) and unclassified categories (603.56), justifying their use for most calculations.

Table 4. Time of measurement

SPOT ID	Initial Date / Hour	Final Date / Hour	Days
241136	02/26/23 13:24	05/28/23 3:04	90.57
241137	02/26/23 13:25	03/30/23 1:55	31.52
241138	02/26/23 13:24	03/25/23 0:37	26.47
241139	02/26/23 13:25	04/26/23 11:19	58.91
241140	04/18/23 1:02	05/26/23 23:19	38.93
241141	04/18/23 1:01	05/24/23 4:32	36.15
241142	04/18/23 1:00	05/21/23 15:35	33.61
241143	04/18/23 1:01	05/24/23 15:46	36.61
262534	03/31/24 11:23	05/18/24 15:59	48.19
262536	03/31/24 11:27	05/31/24 23:32	61.50
262537	03/31/24 11:29	06/24/24 2:37	84.63
Averages (± 50 days)			49.74

3. Theory/Calculation

3.1. Geometric Foundations for Trajectory Estimation

To analyze the migratory trajectories of *Chelonia mydas*, two complementary distance metrics were employed: Euclidean distance and the Haversine formula, each suitable for different spatial modeling contexts (Goodwin et al., 2022). This dual-metric framework enhances the interpretability of movement patterns under different spatial and temporal resolutions, reinforcing the methodological rigor of the predictive modeling pipeline (Alia et al., 2022).

- **Euclidean Distance:** Let $P_i=(x_i, y_i)$ and $P_{i+1}=(x_{i+1}, y_{i+1})$ be two consecutive projected GPS points in a local planar coordinate system. The Euclidean distance between these points is computed as: $d_E(P_i, P_{i+1}) = \sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2}$. This measure assumes flat Earth approximation and is computationally efficient. It was applied primarily for short-range displacement estimations within localized regions after coordinate projection (e.g., UTM).
- **Haversine Formula for Geodesic Distance:** Given the global nature of sea turtle migration, the Earth's curvature must be considered. For this, the Haversine formula computes the great-circle distance d_H between two points $A=(\phi_1, \lambda_1)$ and $B=(\phi_2, \lambda_2)$ defined by their latitude and longitude in radians:

$$\Delta\phi = \phi_2 - \phi_1, \Delta\lambda = \lambda_2 - \lambda_1,$$

$$a = \sin^2\left(\frac{\Delta\phi}{2}\right) + \cos\phi_1 \cos\phi_2 \sin^2\left(\frac{\Delta\lambda}{2}\right), \quad c = 2 \arctan2(\sqrt{a}, \sqrt{1-a}),$$

$$d_H = R * c$$

where: $R=6371$ km is the Earth's mean radius, ϕ denotes latitude, and λ denotes longitude, and d_h yields the distance in kilometers. The Haversine method was applied to all raw GPS coordinates prior to any spatial filtering, enabling robust inter-point distance measurement over curved surface domains. It was particularly suitable for: Cross-region movement tracking, Error-tolerant interpolation between distant measurements, Validation of displacement magnitudes against known migratory bounds.

- Daily Displacement Calculation: Let, $T_i = \{t_j, \phi_j, \lambda_j\}_{j=1}^{n_i}$, where n_i be the time-ordered position dataset of the turtle i . The total displacement over the tracking period is:

$$D_i = \sum_{j=1}^{n_i-1} \cdot d_H((\phi_j, \lambda_j), (\phi_{j+1}, \lambda_{j+1}))$$

- And the average displacement per day: $\bar{d}_i = \frac{D_i}{\Delta t_i}$, $\Delta t_i = \frac{t_{n_i} - t_1}{1 \text{ day}}$, i.e., total great-circle distance divided by the number of elapsed days. These metrics were tabulated for each SPOT-tagged individual to characterize individual migratory behavior.

3.2. Predictive Model

To model and predict the migratory behavior of *Chelonia mydas*, we implemented a supervised Deep Learning model based on LSTM networks, a variant of recurrent neural networks (RNNs) designed to learn from temporal sequences with long-range dependencies (Labrada-Martagón et al., 2024). LSTMs address the problem of gradient vanishing due to their internal architecture, which features gates that regulate the flow of information (Hardin et al., 2024). Specifically, each LSTM cell contains three main gates: the input gate, which decides what new information is incorporated into the cell's state; the forget gate, which removes irrelevant information from the past; and the output gate, which determines what part of the internal state is used to generate the output. This structure allows LSTMs to capture complex temporal patterns, making them widely used in tasks such as language modeling and time series prediction (Nathan et al., 2022). Their ability to retain and update information in a controlled manner makes them a powerful tool for problems where context and data sequence are critical (Cullen et al., 2024).

Problem Setup: Given a time-ordered sequence of positions $S = \{(\phi_t, \lambda_t)\}_{t=1}^T$, (WGS84) where ϕ_t and λ_t are the latitude and longitude at time t , the goal is to learn a function f_θ parameterized by θ , such that: $(\hat{\phi}_{t+1}, \hat{\lambda}_{t+1}) = f_\theta((\phi_t, \lambda_t), \dots, (\phi_{t-k+1}, \lambda_{t-k+1}), u_{t-k+1:t})$, where k is the length of the input sequence (lookback window), $(\hat{\phi}_{t-k+1}, \hat{\lambda}_{t-k+1})$ denotes the predicted coordinates for the next time step, and denotes optional covariates (e.g., time-gap Δt , Argos class indicators). We report one-step-ahead errors; longer horizons are obtained by recursive rollout.

3.2.1. LSTM Architecture

We use length-10 input windows of GPS pairs (ϕ, λ) , z-score standardized (zero mean, unit variance) to stabilize training (Kot et al., 2022). The network consists of two LSTM layers (150 and 100 units), followed by a fully connected layer with two outputs (latitude and longitude). Standard LSTM nonlinearities are used inside the cells (sigmoid gates and tanh activation); no ReLU is applied within

LSTM cells. To mitigate overfitting, a dropout rate of $p = 0.30$ is applied after each LSTM layer. An LSTM cell at time step t maintains an internal state $c_t \in \mathbb{R}^h$ and a hidden state $h_t \in \mathbb{R}^h$, updated with input x_t and previous hidden/state (h_{t-1}, c_{t-1}):

$$\begin{aligned} f_t &= \sigma(W_f x_t + U_f h_{t-1} + b_f), \text{ (forget gate)}, & i_t &= \sigma(W_i x_t + U_i h_{t-1} + b_i), \text{ (input gate)} \\ o_t &= \sigma(W_o x_t + U_o h_{t-1} + b_o), \text{ (output gate)}, & \tilde{c}_t &= \tanh(W_c x_t + U_c h_{t-1} + b_c), \text{ (cell input)} \\ c_t &= f_t \odot c_{t-1} + i_t \odot \tilde{c}_t, & h_t &= o_t \odot \tanh(c_t) \end{aligned}$$

where $x_t \in \mathbb{R}^d$ is the input at time t , $\odot$ element-wise multiplication, and σ sigmoid activation function. The LSTM maps the full sequence to a predicted position $\hat{y}_t \in \mathbb{R}^2$ (latitude, longitude) $\hat{y}_t = W_y * h_t + b_y$.

Figure 4 illustrates the architecture used in this study. Left: schematic of a single LSTM cell with forget (F), input (I), and output (O) gates; σ denotes the sigmoid activation and $\tanh$ the state nonlinearity. The cell receives the previous hidden state h_{t-1} and cell state c_{t-1} and produces h_t and c_t . Right: stacked LSTM layers followed by a dense layer mapping the input sequence $\{x_1, \dots, x_n\}$ to predictions $\{y_1, \dots, y_n\}$. Panel labels and icons (scaling/adding) indicate gate operations; arrows show data flow.

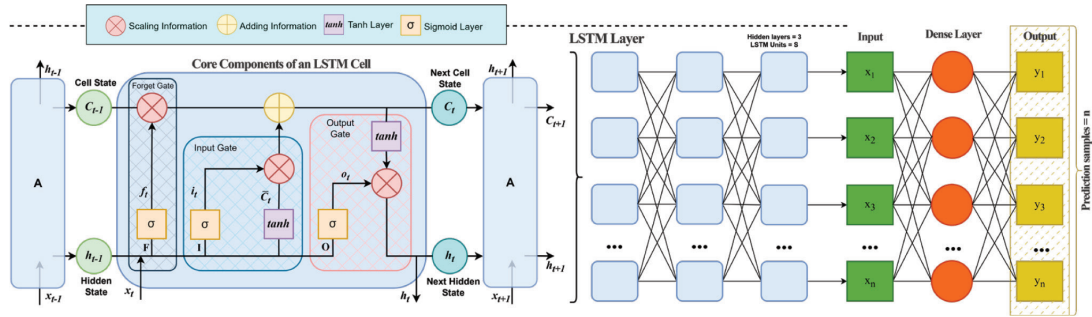


Figure 4. LSTM Architecture Model Overview

3.2.2. Model Training and Optimization

Let the input sequence of georeferenced positions be defined as a time series: $x = \{x_t\}_{t=1}^T$, $x_t \in \mathbb{R}^d$, where each vector x_t encodes the multivariate features at time t , such as: latitude and longitude, velocity or displacement, sensor quality indicators, and environmental covariates. All sequences are normalized using the z-score method: $x' = \frac{x - \mu}{\sigma}$, with μ and σ calculated per feature (latitude, longitude). The preprocessed dataset is split into training and testing sets while preserving temporal order. Given a ground-truth target sequence $\{y_t\}_{t=1}^{T'}$, the model is trained by minimizing the empirical loss over the prediction window:

$$L_{MSE} = \frac{1}{n} \sum_{i=1}^n \left[(\phi_i - \hat{\phi}_i)^2 + (\lambda_i - \hat{\lambda}_i)^2 \right]$$

where the optimizer is Adam with a learning rate η , and to prevent overfitting, regularization (Early stopping) based on validation loss with a patience of 10 epochs is applied. The trained model is applied to the full sequence to predict future points recursively. The model was trained using mean squared error (MSE) as the loss function (on scaled latitude and longitude), optimized with Adam at a learning rate of 1×10^{-3} , a batch size of 32, and a look-back window of $k=10$ steps. Training was performed for up to 9,500 epochs with Early Stopping (patience = 20, restoring the best weights), applying L2 weight decay of 1×10^{-3} , on recurrent and dense kernels, global-norm gradient clipping at 1.0, and model checkpointing based on validation MSE.

We apply z-score standardization fit exclusively on the training split and reuse the same transform for validation/test. Model inputs include (ϕ, λ) (latitude, longitude) plus optional covariates: the normalized inter-observation interval Δt and a one-hot encoding of Argos class $\{3, 2, 1, 0, A, B\}$. Sequences are constructed at the observed timestamps; irregular sampling is handled explicitly by providing Δt as a feature. Targets are de-standardized after prediction to compute downstream geodesic errors in kilometers.

We use a time-based, walk-forward split per tag: for each SPOT transmitter, the earliest 80 % of its timeline forms the training set and the latest 20 % the test set, strictly preserving temporal order to prevent leakage. From the training period, a temporally stratified 10 % slice is held out for validation. To assess sensitivity to temporal cut points, we perform 5× rolling-origin resamples within each tag and report the mean across resamples. Additionally, we run five random seeds for weight initialization and batching, and average the results across seeds.

We mitigate overfitting with dropout 0.30 after each LSTM layer (applied to inputs of the subsequent layer), L2 regularization of 1×10^{-3} on recurrent and dense weights, and global-norm gradient clipping at 1.0 to stabilize training under heterogeneity. Evaluation is group-aware: splits are strictly time-ordered within each tag to prevent identity leakage and optimistic errors. The network uses ReLU activations in the hidden layers and the architecture LSTM(150) → Dropout(0.30) → LSTM(100) → Dropout(0.30) → Dense(2), where the final dense layer outputs the predicted latitude and longitude.

3.2.3. Evaluation Metrics

To assess predictive performance, the following metrics were computed, where $y_i = (\phi_i, \lambda_i)$ are true positions and $\hat{y}_i$ the predicted ones. This formulation enables robust modeling of non-linear spatial-temporal dynamics inherent in sea turtle migration patterns (Gupte et al., 2022). The LSTM effectively captures dependencies in the input sequence, while the structured loss minimization guides learning toward accurate geospatial predictions (Sharmila & Sabarish, 2021). Metrics are reported per tag and averaged across tags/seeds; this avoids planar distortions and reflects true geodesic error.

$$MAE = \frac{1}{n} \sum_{i=1}^n \|y_i - \hat{y}_i\|, \quad MSE = \frac{1}{n} \sum_{i=1}^n \|y_i - \hat{y}_i\|^2, \quad RSME = \sqrt{MSE}$$

3.3. Kalman Filter for Geospatial Smoothing and Interpolation

The Kalman filter was employed not merely as a denoising tool, but as a recursive Bayesian estimator over sequential geospatial observations, enabling prediction and correction under sensor uncertainty and missing measurements (Christiaan et al., 2024). This section introduces an extended vector-state formulation adapted to two-dimensional geographical positions with temporal evolution (Li et al., 2021). Let $x_k \in \mathbb{R}^n$ be the latent state vector at time step k , representing the true (but unobserved) geospatial position of a turtle (e.g., latitude, longitude, velocity components) (Hays et al., 2021).

The filter is initialized from the first valid observation. The starting state is set to $x_0 = [\phi_1, \lambda_1, v_{\phi,0}, v_{\lambda,0}]^T$ (with velocities initialized from nominal priors), and the starting state covariance is a diagonal matrix: $P_0 = \text{diag}(\sigma_{\phi,0}^2, \sigma_{\lambda,0}^2, \sigma_{v_{\phi,0}}^2, \sigma_{v_{\lambda,0}}^2)$, where $\sigma_{\phi,0}^2, \sigma_{\lambda,0}^2$ are derived from the Argos location class of the first fix (position uncertainty), and $\sigma_{v_{\phi,0}}^2, \sigma_{v_{\lambda,0}}^2$ reflect the initial uncertainty in velocities. This setting encodes the initial positional accuracy while allowing the filter to learn velocities from subsequent observations.

In this implementation, the Kalman Gain Adaptation with Dynamic Noise Calibration, where to account for time-varying measurement quality, the observation noise matrix R_t can be adapted dynamically (DiMatteo et al., 2021). For each message class $c \in \{3, 2, 1, 0, A, B, Z\}$, empirical variances σ_c^2 are computed, and R_t is assigned. This strategy allows the Kalman filter to weight GPS data according to confidence category, reducing the influence of low-quality fixes while maximizing the reliability of high-quality observations. The Kalman filter is a recursive estimator that operates in two phases: prediction and update, adapting to noisy satellite telemetry data (Cai et al., 2022).

- State-Space Model for Position Estimation

Let the hidden state vector at time t be: $x_k = \begin{bmatrix} \phi_k \\ \lambda_k \\ v_{\phi,k} \\ v_{\lambda,k} \end{bmatrix}$, $z_k = \begin{bmatrix} \phi_k^{GPS} \\ \lambda_k^{GPS} \end{bmatrix}$, where (ϕ_t, λ_t) denote the latitude

and longitude, and (ϕ'_t, λ'_t) their respective temporal derivatives (velocity components). The system

evolves linearly under the dynamic model:

$$x_{t+1} = F x_t + w_t, \quad \text{with } F = \begin{bmatrix} 1 & 0 & \Delta_t & 0 \\ 0 & 1 & 0 & \Delta_t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

and $w_t \sim N(0, Q)$ representing process noise, capturing environmental and behavioral irregularities (e.g., drift, tide response). The transition uses the observed time step Δt . During long intervals without transmissions ($\Delta t \geq 600s$), the filter runs in prediction-only mode and inflates process noise to reflect growing uncertainty; operationally, Q is scaled with Δt (i.e., $Q_t \propto \Delta t$) so that covariance growth matches the gap length. This down-weights subsequent extrapolations and prevents overconfident tracks until the next valid update arrives, at which point the class-conditioned R_t is used in the measurement update.

- Geospatial Filtering of Terrestrial Points

Given that SPOT telemetry may erroneously report points on land, a spatial filtering procedure was applied using high-resolution coastline shapefiles. Let:

$D = \{d_1, \dots, d_n\}$, where $d_i = (\phi_i, \lambda_i)$ is a GPS location (latitude and longitude).

$L = \{P_1, \dots, P_n\}$, a set of polygonal landmass representations from Natural Earth.

Contains (P_j, d_i) is a boolean function that returns true if $d_i \in P_j$.

We define the on-land indicator function: $OnLand(d_i) = \bigvee_{j=1}^m \text{Contains}(P_j, d_i)$.

The filtered dataset is then: $D_{sea} = \{d_i \in D \mid OnLand(d_i)\}$.

This ensures that all retained data points are strictly oceanic, aligning the dataset with realistic migration behavior and eliminating spatial anomalies before model evaluation.

- Observation Model with Missing Data Handling

The measurement model is defined as: $z_t = Hx_t + \nu_t$ with $H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$, $\nu_t \sim N(0, R)$.

However, due to submerged intervals or satellite occlusion, we define a masked observation model: $z_t^{(M)} = M_t z_t$, $M_t \in \{0, 1\}^{2 \times 2}$, where M_t is a diagonal masking matrix indicating the availability of latitude/longitude at time t . During missing observations, the filter switches to a prediction-only mode.

- Trajectory Confidence Quantification via Covariance Propagation

Rather than relying on the filtered state alone, we propagate the full state covariance matrix P_t to quantify uncertainty over time. For each predicted position, we construct an elliptical confidence region via the Mahalanobis distance: $\xi_\alpha = \{x \in \mathbb{R}^2 : (x - \hat{x}_t)^T S_t^{-1} (x - \hat{x}_t) \leq x_{2,\alpha}^2\}$ with: $\hat{x}_t = H\hat{x}_t^{posterior}$, projected position estimate, $S_t = HP_tH^T + R$ innovation covariance, $x_{2,\alpha}^2$ quantile of the chi-squared distribution at confidence level α . These ellipses were rendered per step to visualize geospatial uncertainty, adding interpretability to the model beyond pointwise predictions.

- Kalman Gain Adaptation with Dynamic Noise Calibration

To account for time-varying measurement quality (as determined by message class in SPOT telemetry), we introduce a data-driven adjustment to the observation noise matrix R . For each class $c \in \{3, 2, 1, 0, A, B, Z\}$, we compute empirical variances σ_c^2 , and assign: $R_t = \sigma_c^2 * I_2$. This strategy allows the Kalman filter to dynamically weight GPS data by confidence category, reducing the impact of class B or Z positions while maximizing the utility of class 3 readings.

4. Results

4.1. Analysis of Distances Traveled

Based on the corrected and smoothed geolocation data, the total distances traveled by each turtle during its respective tracking period were calculated. For each route, the Haversine formula was used with latitude and longitude coordinates expressed in radians, which allowed for accurate estimation of the geodesic distances between successive points on the Earth's surface. The results reveal substantial variability in migratory behavior. Turtle SPOT ID 262537 recorded the longest total distance, traveling 320.87 km in 84.63 days, representing a daily average of 3.79 km/day. In contrast, turtle 241138 exhibited the least mobility, with a distance of 13.80 km traveled in 26.47 days, equivalent to 0.52 km/day, as shown in Table 5. The overall average daily displacement among all active turtles was 1.92 km/day, noting that those with longer tracking periods tended to show greater cumulative distances, although not necessarily at a constant rate of displacement. These findings reinforce the importance of adequate temporal resolution to accurately capture interindividual migration patterns. Table 5 summarizes these results, enabling a comparison of both the total displacement and the daily pace for all individuals monitored.

Table 5. Kms traveled

SPOT ID	Total Kms	Kms x Day
241136	185.13	2.04
241137	72.81	2.31
241138	13.80	0.52
241139	69.71	1.18
241140	60.46	1.55
241141	61.59	1.70
241142	72.12	2.15
241143	47.68	1.30
262534	70.43	1.46
262536	144.57	2.35
262537	320.87	3.79

Figure 5 presents a comparative visualization of the average daily displacement for each SPOT ID. There is marked variability between individuals, with turtle 262537 standing out with a displacement of more than 3.7 km/day, in contrast to 241138, whose daily average was less than 0.6 km/day. This graph allows for the immediate identification of differences in migratory behavior, possibly associated with individual, (inter-nesting or post-nesting movements), environmental, or tracking quality factors.

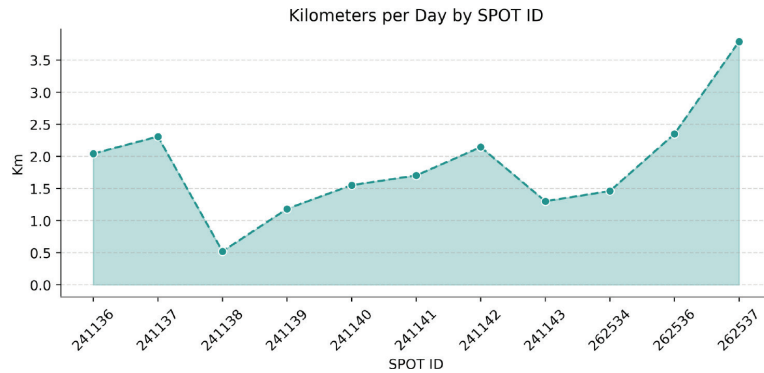


Figure 5. Average daily displacement (km/day) per SPOT ID

4.2. Performance of the LSTM Model in Trajectory Prediction

The predictive model based on LSTM neural networks was evaluated using time series of geo-referenced coordinates obtained from SPOT sensors. For each turtle, the model was trained with historical latitude and longitude sequences, generating predictions of future positions that were compared against actual trajectories. Accuracy was assessed using three standard metrics: Mean Squared Error

(MSE), Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE). Table 6 reports the errors obtained for each transmission. The average error across all devices was 0.0307 (MSE), 0.0986 (MAE), and 0.1288 (RMSE). The best performance was observed in SPOT ID 262534 (RMSE = 0.0130), while the highest error corresponded to SPOT ID 241141 (RMSE = 0.3754), reflecting the impact of noisy or incomplete records. These results demonstrate that the LSTM model effectively captures nonlinear temporal dependencies in animal telemetry data, providing reliable estimates of turtle locations even during periods without signal reception.

Table 6. Error measurement parameters

No	SPOT ID	MSE	MAE	RMSE
1	241136	0.0842	0.2516	0.2903
2	241137	0.0017	0.0269	0.0417
3	241138	0.0034	0.0523	0.0587
4	241139	0.0458	0.1438	0.2141
5	241140	0.0014	0.0280	0.0384
6	241141	0.1409	0.3017	0.3754
7	241142	0.0008	0.0218	0.0284
8	241143	0.0002	0.0114	0.0153
9	262534	0.0001	0.0100	0.0130
10	262536	0.0394	0.1460	0.1985
11	262537	0.0203	0.0914	0.1426
	Minimum	0.0001	0.0100	0.0130
	Maximum	0.1409	0.3017	0.3754
	Average	0.0307	0.0986	0.1288

To contextualize these findings, the LSTM model was compared with other representative neural architectures: conventional Recurrent Neural Networks (RNN), feed-forward Deep Neural Networks (DNN), and temporal Convolutional Neural Networks (CNN). The results, summarized in Table 7, reveal a clear performance gradient across the models. RNNs exhibited limitations in handling irregular and long temporal sequences, resulting in lower predictive accuracy and higher computational demands.

Table 7. Average performance comparison across models

Model	MSE	MAE	RMSE	CPU Time (%)
RNN	0.0545	0.1747	0.2290	100
CNN	0.0412	0.1391	0.1903	78
DNN	0.0368	0.1237	0.1724	65
LSTM	0.0307	0.0986	0.1288	36

CNNs achieved moderate improvements by effectively capturing local temporal features, while DNNs further enhanced performance by leveraging deeper nonlinear transformations, albeit without explicit temporal memory. LSTMs, in contrast, demonstrated a distinct advantage by combining sequence memory with computational efficiency, enabling more accurate and robust trajectory predictions under the challenging conditions of noisy biotelemetry data.

Figure 6 illustrates the average error values for MSE, MAE, and RMSE across the four models. The grouped bar plot clearly shows that LSTM consistently achieves the lowest error in all three metrics, followed by DNN and CNN, while RNN exhibits the highest error.

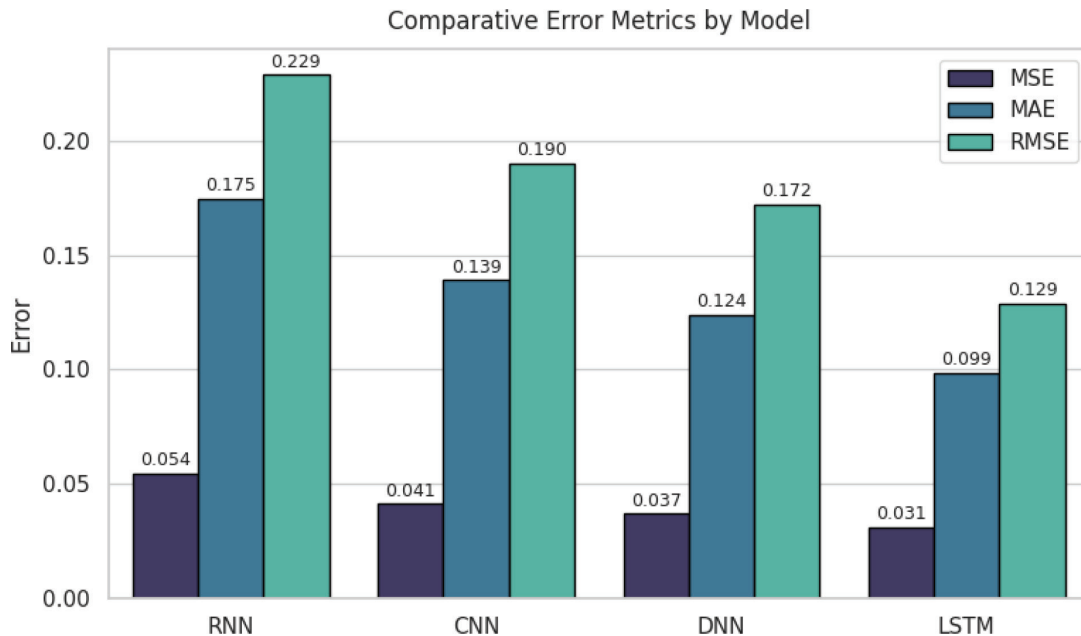


Figure 6. Comparative error metrics (MSE, MAE, RMSE) for RNN, CNN, DNN, and LSTM

Beyond predictive accuracy, efficiency is a critical factor for practical deployment. Figure 7 presents the relative comparison of prediction error, CPU time, and computational resources, with RNN normalized to 100 %. The results show that while CNN and DNN reduced prediction error by 24.4 % and 32.5 % relative to RNN, LSTM achieved the largest improvement with a 43.75 % reduction. In terms of computational cost, CNN and DNN required 78 % and 65 % of RNN runtime, respectively, whereas LSTM required only 10.43 %. Similarly, LSTM's computational resource usage was 36.46 % of the RNN baseline.

Overall, these results confirm that although CNN and DNN provide efficiency gains compared to traditional RNNs, the LSTM architecture offers the best balance of predictive accuracy and computational efficiency. This makes it particularly suitable for near-real-time applications in sea turtle migration monitoring, where both precision and efficiency are essential for operational feasibility.

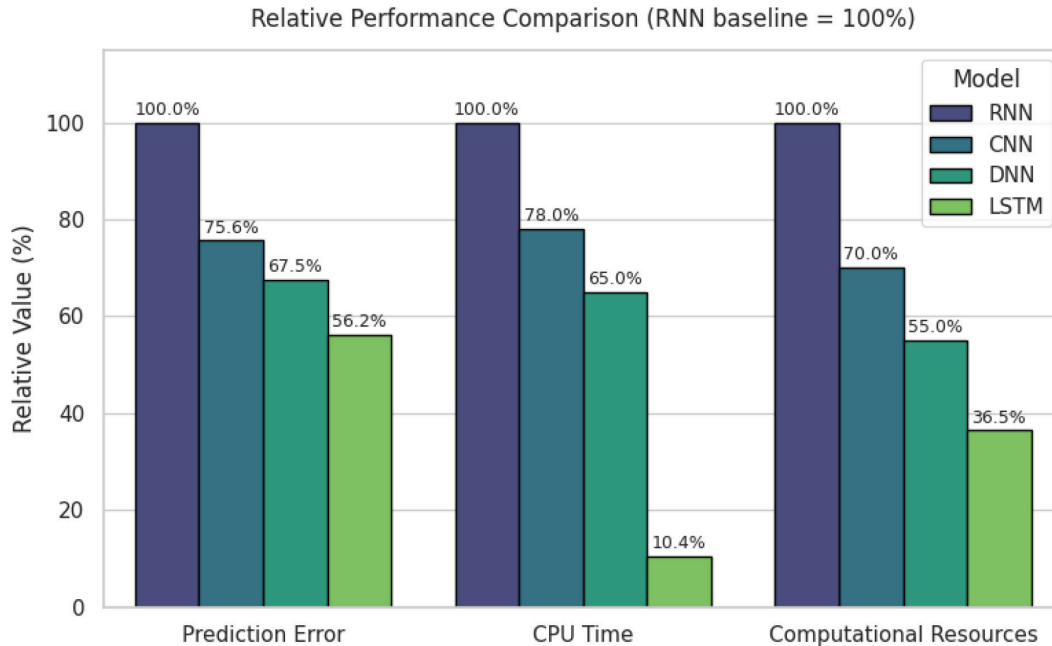


Figure 7. Relative performance comparison (RNN baseline = 100 %)

4.3. Results of Geospatial Smoothing and Filtering

The application of the Kalman filter to the predicted trajectories markedly enhanced the spatial coherence of the movement data. By integrating statistical smoothing with error correction, the method not only interpolated missing positions during periods without transmissions but also reduced abrupt displacements associated with noise and signal instability. As shown in Figure 8, the contrast between the raw trajectory (panel A) and the filtered version (panel B) highlights how the process reconstructs continuous and biologically plausible routes, generating intermediate points that bridge gaps in areas of limited satellite coverage. This refinement provides a more accurate basis for ecological interpretation, ensuring that the reconstructed tracks reflect migratory behavior rather than artifacts of data acquisition.

In Figure 9, for each location sequence, the model estimated new smoothed coordinates that corrected abrupt variations generated by noise or signal loss. Panel A displays all the original points reported by the SPOT ID 241139 transmitter, categorized by age (green = old; red = recent). This representation includes duplicate values, null positions, and erroneous locations on land. In contrast, Panel B shows the result after applying quality filtering and geographic exclusion, retaining only valid and accurate coordinates in the open sea, which drastically reduces the volume of data but improves its reliability.

In addition to smoothing, a geospatial filtering process was implemented that eliminated coordinates located on land. These erroneous positions, which represent a considerable fraction of the

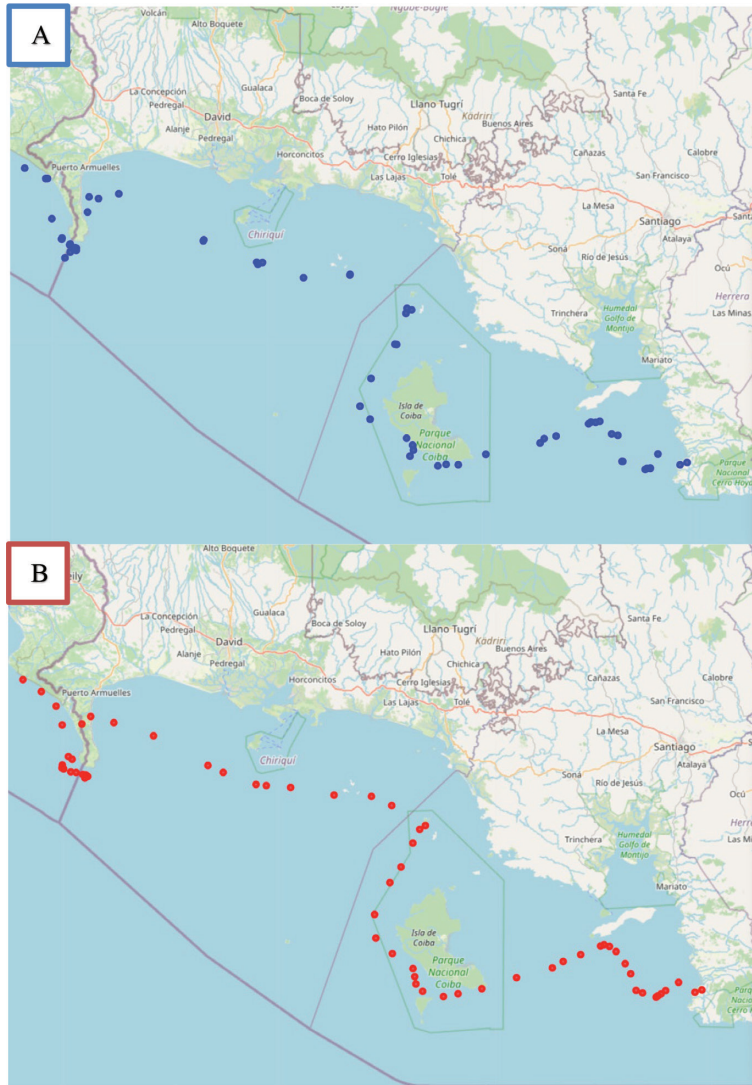


Figure 8. Kalman Filter effect: (A) Raw SPOT-375B positions (blue points). (B) Kalman-filtered trajectory (red points)

original records in some transmitters, were detected through a geometric intersection operation with coastal polygons from the Natural Earth dataset. As a result, only the subset of ocean points was retained, ensuring that the analysis accurately reflected marine mobility. The combination of both processes, Kalman smoothing and geospatial inspection, resulted in a significant reduction in the number of points considered valid. This adjustment contributed to a more accurate and reliable representation of actual movement patterns, strengthening the quality of predictions and their usefulness for marine conservation decision-making.

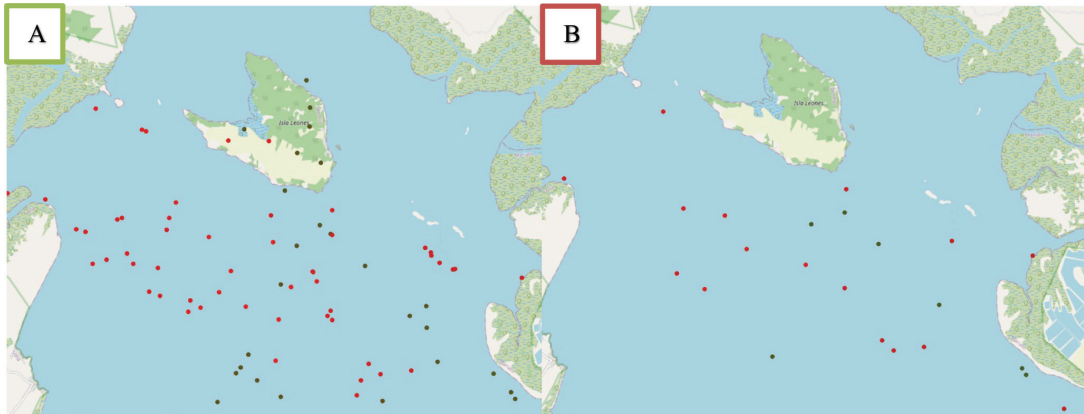


Figure 9. Geospatial cleaning: (A) Raw positions reported by SPOT ID 241139 (green = older, red = recent), including duplicates, nulls, and on-land fixes. (B) Positions after filtering (removal of duplicates/nulls and exclusion of on-land points), retaining only valid open-sea locations

4.4. Prototype Interface Application for Trajectory Visualization

As part of the practical validation of the proposed model, a prototype mobile interface called Turtle Tracker was developed for the interactive tracking of sea turtles equipped with SPOT-375B sensors. In its current version, the application enables users to visualize estimated and smoothed trajectories generated by the LSTM model and Kalman filter, focusing on reconstructed paths from the turtles monitored in this study. The system integrates basic features for historical consultation, user authentication, and visualization of the current location on georeferenced maps.

The prototype was built using AppGyver Community Edition, a cross-platform low-code framework that enables the deployment of mobile and web applications from a unified component-based model. The app connects to a Firebase backend responsible for user authentication (email/password), data storage (Cloud Firestore), and file management (Cloud Storage) for captures and metadata. Cloud Functions orchestrate periodic tasks, including the ingestion and normalization of new Argos transmissions, the deserialization and publication of LSTM and Kalman predictions, and expose verified endpoints for secure app communication. Geospatial visualization is implemented using LeafletJS and OpenStreetMap layers, rendering historical trajectories (as lines and temporal markers) and the most recent state for everyone. To enhance robustness and minimize latency, the app maintains local caching on the device and utilizes near-real-time Firestore subscriptions when network connectivity is available. In offline mode, it switches to deferred reads and automatic synchronization once the connection is restored. All trajectory data are synchronized with the Argos repository through an integration pipeline (ingestion → validation → terrestrial filtering → Kalman smoothing → geographic tiling for maps), ensuring consistency between the displayed tracks and model outputs. This modular architecture facilitates future integration of cloud-based inference services and webhooks for operational alerts and management dashboards.

Figure 10 illustrates the four screens of the Turtle Tracker prototype and clarifies their role in the workflow: the splash screen presents the app identity; the authentication screen provides email/password fields with «Log in», «Forgot password?» and «Sign up» options; the welcome screen introduces

the purpose of the app and offers a single call-to-action («Start Tracking») to enter the tracking module; the exploration panel shows a feed of map thumbnails with reconstructed routes along the Panamanian coast; each card corresponds to one tracked turtle and is rendered from SPOT-375B positions filtered to remove on-land records and smoothed (LSTM + Kalman) to ensure spatial continuity. A top bar with the app logo and menu icon, and a bottom navigation (Explore / Saved / Updates) indicate the current section and available views.

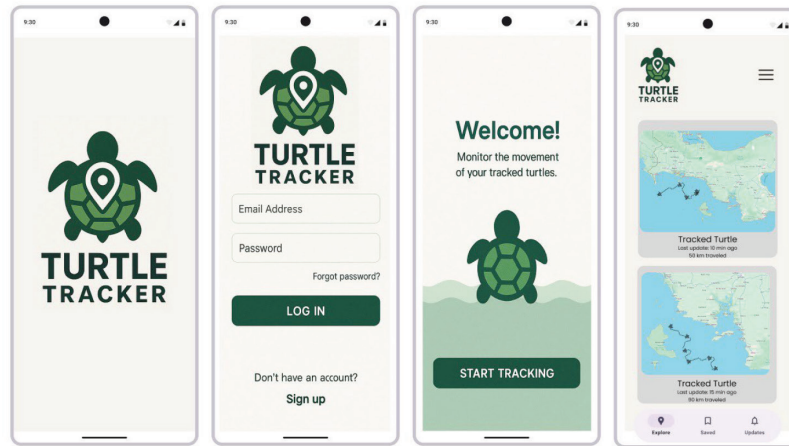


Figure 10. Mobile interface of the Turtle Tracker prototype: (left-right). (a) Splash/logo screen. (b) User authentication. (c) Welcome screen. (d) Exploration

It is essential to note that Turtle Tracker is presented as a prototype, currently limited to visualizing turtle trajectories. However, its modular design anticipates further development. Future updates are planned to incorporate additional functionalities such as the ability to select individual turtles, overlay multiple trajectories for comparison, filter data by date or transmission quality, and access metadata linked to each position. Advanced features envisioned for conservation managers include automated alerts for anomalous behavior, dashboards summarizing traveled distances and time spent in key habitats, overlap analysis with high-risk zones such as fishing corridors, and annotation tools for integrating field observations. This phased development approach ensures that while the present version serves as a proof-of-concept for translating model outputs into a mobile interface, the roadmap aims to transform it into a comprehensive decision-support platform. By doing so, the system will evolve from a visualization prototype into a practical management tool that can enhance ecological monitoring and support marine conservation initiatives.

5. Discussion

This study shows that combining LSTM-based trajectory prediction with Kalman filtering substantially improves the reliability of satellite telemetry analysis for *Chelonia mydas*. The two-stage architecture effectively addressed the common challenges of marine tracking data, including positional

noise, temporal irregularity, and the occurrence of outliers. Empirical evaluation, as measured by MSE, MAE, and RMSE, confirmed that the smoothed outputs achieved higher spatial coherence, yielding trajectories that are more consistent with ecological expectations.

A key contribution lies in the selective use of data with varying accuracy levels. While over 60 % of the raw location records fell into low-confidence categories, the model retained useful patterns by prioritizing high-quality Argos Classes 2 and 3 and applying geospatial filtering through shapefiles. This strategy preserved biologically meaningful sequences without overfitting or discarding large segments of data, a limitation present in many existing approaches.

The use of SPOT-375B tags under real-world conditions validated their capacity to deliver temporally dense, georeferenced inputs across different immersion profiles. The Fastloc GPS system proved particularly effective during short surfacing intervals, enhancing spatial resolution and enriching the training data with high-precision points. An average tracking duration of nearly 50 days per turtle enabled the model to learn both short-range movement and broader migratory behavior, essential for meaningful ecological interpretation.

Kalman filtering further enhanced the robustness of predictions by interpolating across missing intervals and reducing artificial fluctuations. Unlike linear interpolation, this statistical estimator preserved movement continuity without distorting the underlying biological signal. This is particularly valuable in marine environments, where long transmission gaps are common due to submersion or satellite occlusion.

The model also benefited from ecological constraints integrated into preprocessing, such as the removal of terrestrial points using high-resolution coastline shapefiles and the enforcement of biologically plausible displacement thresholds. These safeguards ensured that the output remained interpretable and ecologically valid. Furthermore, the mobile application prototype successfully translated predictive data into an accessible visualization platform, bridging computational results with field-based conservation needs.

Compared to prior studies, this work presents a unified pipeline that encompasses telemetry acquisition, ML-based prediction, error correction, and user-oriented visualization. Although similar hybrid models have been applied to terrestrial species, their adaptation to oceanic tracking remains limited. This research confirms that applying such methods to marine contexts is both feasible and advantageous, given the specific noise characteristics and transmission constraints of aquatic telemetry.

6. Study Limitations

- Temporal coverage and sampling: The model's effective resolution is bounded by Argos duty cycles and surfacing behavior; long dives create irregular gaps that may miss fine-scale movements. Walk-forward evaluation reduces leakage, but short monitoring windows (≈approximately 50 days on average) limit the ability to make seasonal inferences.
- Sample size and selection: Results are based on 11 individuals and predominantly post-nesting females from specific rookeries; this introduces selection and site biases and constrains generalization to other populations or life stages. The study generated ~22,000 recordings, which is a reasonable but still limited number for broad inference.
- Telemetry/measurement quality: Location accuracy varies by Argos class and satellite geometry. Filtering (e.g., removing on-land fixes) and class prioritization improve quality but may

introduce bias toward favorable conditions. Tag loss or transmission failure causes informative censoring (two tagged turtles failed to transmit their tags).

- Preprocessing dependence: Outcomes depend on thresholds (e.g., plausibility bounds, land-mask resolution) and coordinate choices; different settings can slightly shift errors and path continuity.
- Modeling assumptions: The LSTM is trained on (ϕ, λ) sequences with limited covariates; the Kalman filter uses a simple kinematic model and class-conditioned noise. These choices may oversmooth tortuous segments or underreact to abrupt reorientations.
- Uncertainty and evaluation: Errors are reported on de-standardized coordinates, but uncertainty is summarized mainly via point metrics (MAE/MSE/RMSE); interval calibration and per-segment confidence are not yet reported.
- External validity and domain shift: Trained on Panama's Pacific coast, the model has not been externally validated on other regions/species; covariate shift (bathymetry, currents, fisheries pressure) may degrade performance.
- Environmental drivers omitted: Oceanographic covariates (e.g., sea-surface temperature, surface currents) are not included here by design. The goal of this study was to develop a predictive/corrective pipeline directly from raw satellite positions (LSTM + Kalman). Additionally, obtaining consistent, gap-free oceanographic fields aligned in space/time with all tracks for the full monitoring period was not feasible within the scope. We therefore rely on indirect proxies already present in the data (e.g., Δt between fixes and Argos class indicators) and explicitly acknowledge that excluding environmental drivers can restrict ecological realism and reduce predictive stability under anomalous conditions. Future work will incorporate a small set of harmonized environmental layers to assess incremental benefit relative to the baseline model.
- Scale and feasibility constraints: Deployments were carried out under contract No. 191-2022 (SENACYT, Panama). Tagging free-ranging turtles requires substantial field effort and strict attachment protocols to avoid harm/stress; tag loss is common in green turtles due to the oily carapace. Costs are a limiting factor ($\approx$ US\$2,000 per sensor plus $\approx$ US\$600/tag for ARGOS service), which constrains sample size and tracking duration. Despite these constraints, the study scale is comparable to prior work with similar numbers of satellite tags in Costa Rica.
- System constraints: The mobile app is a visualization prototype; it does not perform on-device inference, and its outputs depend on the offline pipeline.

6.1. Ethical and Operational Considerations

- Animal welfare and permitting: Tagging follows national permits and welfare protocols to minimize handling time, stress, and hydrodynamic drag; nevertheless, attachment and adhesives can fail, and cumulative impacts must be monitored. Only animals that meet the health/size criteria are tagged, which can introduce selection bias.
- Sensitive location data: High-resolution tracks of a protected species entail disclosure risks. Data sharing should follow governance policies (including temporal/spatial blurring where appropriate) and be controlled to ensure access to sensitive sites.
- Field logistics and costs. Hardware, satellite service, and vessel time constraints limit sample size and tracking duration, thereby affecting statistical power and temporal coverage.

- Computing scalability and real-time use: Real-time operation at scale would require: (i) cloud ingestion with bandwidth-aware telemetry compression, (ii) low-latency inference services, (iii) fault tolerance and queueing for satellite pass bursts, (iv) automated drift monitoring/re-training, and (v) cost/battery budgeting aligned with duty cycles.
- Reproducibility and security: Software versions, model checkpoints, and preprocessing scripts must be versioned; the prototype app should enforce authentication, least-privilege access, and secure storage of telemetry and metadata.

7. Conclusions and Future Research

This study proposed and validated a hybrid architecture that combines LSTM neural networks with Kalman filters to predict and correct the migratory trajectories of the green turtle (*Chelonia mydas*) based on satellite biotelemetry data collected via SPOT-375B sensors. The results reveal that this methodology not only enhances the continuity and spatial accuracy of the trajectories but also enables a more realistic representation of migratory behavior under real oceanic conditions. The model's ability to mitigate the inherent limitations of satellite telemetry, such as positional errors, irregular temporal intervals, and signal loss due to submersion, constitutes a substantial contribution to the ecological monitoring of marine species.

One of the main strengths of the proposed approach lies in its capacity to integrate data of varying quality without compromising the ecological validity of the predictions. While a significant fraction of the data was classified into low-accuracy categories, the application of geospatial filters, the selective use of reliable Argos classes, and the statistical analysis of positional deviation enabled the construction of robust trajectories consistent with the known behavior of the species. This ability to preserve biologically meaningful information, even in the presence of noise or uncertainty, represents an advancement over previous methodologies that typically discard such records or fail to quantify the impact of their modeling.

Moreover, the inclusion of an interactive mobile visualization application provides an effective bridge between computational models and their practical application in management and conservation. The ability to explore smoothed trajectories on georeferenced maps in near real-time adds value not only to scientific research but also to decision-making processes of environmental institutions, NGOs, and government authorities.

Based on these findings, multiple research avenues emerge that could further strengthen and expand the scope of this methodology. One approach involves enriching the current models by incorporating environmental variables such as ocean currents, sea surface temperature, or salinity gradients, which directly influence turtle navigation. Although GPS position information provides insight into ocean dynamics that include these variables, this model's information is limited. Including this information would enable the model to operate in a deeper ecological context, allowing for more responsive predictions that are better suited to the physical environment. Ocean currents play a crucial role in helping turtles navigate, while temperature influences their food resources and habitats. Finally, salinity contributes to currents that indirectly shape turtle routes and food availability.

Additionally, there is potential to evolve towards multimodal deep learning architectures that combine geospatial data with behavioral signals extracted from the sensors themselves (e.g., dive patterns or surface activity). This approach could capture more complex relationships between movement and behavior, supporting the simultaneous prediction of location and physiological state.

From an algorithmic perspective, integrating attention mechanisms or more advanced architectures, such as Transformers adapted for time series, could enhance the model's sensitivity to abrupt events or long-term patterns, including route changes or reproductive behaviors. These mechanisms could complement the Kalman filter, enhancing the system's adaptability to nonlinear or interrupted trajectories.

Another promising direction is adapting the model to other marine species with similar conservation needs. Although the methodology was calibrated for *Chelonia mydas* along the Pacific coast of Panama, it can be applied to different contexts, such as hawksbill turtles, sharks, or migratory pelagic fish, through adjustments or retraining processes, especially when transfer learning strategies are employed. The adjustment processes for hawksbill turtles will be less complex due to their similarity to the green turtle, as both belong to the same family, *Chelonia*, and inhabit the Pacific coast of Panama. The methodology would work similarly for other species, such as sharks or migratory pelagic fish, where only the data obtained by GPS sensors would need to be verified. This would also apply in other geographic contexts. This versatility would expand the model's impact in comparative studies or in ecosystems that are less intensively monitored. In parallel, the model's predictive capabilities can be used to simulate trajectories under different environmental or management scenarios, such as in response to climate change, industrial fishing zones, or commercial shipping routes. These simulations would provide critical insights for planning marine protected areas and implementing safe migratory corridors, anticipating potential conflicts between marine fauna and human activities.

Finally, the integration of this approach into real-time monitoring platforms is anticipated, which would enable continuous and automated tracking of individuals, with early warnings in the event of unusual deviations or potential risks. This operational application, combined with accessible visualization capabilities, would consolidate a decision-support system based on ecological intelligence, well-suited to addressing the challenges of marine conservation in the 21st century.

In future work, expanding collaboration across international telemetry networks and integrating larger multi-species datasets will be essential for enhancing both the robustness and generalizability of the proposed methodology. Similarly, the adoption of standardized data pipelines and open-access platforms would accelerate reproducibility, ensure methodological transparency, and strengthen the translation of predictive models into actionable policies for biodiversity management.

Conflict of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper. No financial, personal, or professional relationships influenced the design, execution, interpretation, or reporting of this research. All data collection, analysis, and conclusions were carried out independently and with scientific integrity.

Author Contributions

Conceptualization, C.R., E.E.F., and K.R.; methodology, I.N., and D.C.; validation, I.N., D.C., and A.S.; formal analysis, I.N., D.C., and C.R.; investigation, K.R., and E.E.F.; data curation, C.R., A.S., and E.E.C.; writing original draft preparation, C.R.; writing review and editing, I.N., and D.C.; supervision, C.R., and E.E.C. All authors have read and agreed to the published version of the manuscript.

Acknowledgements

We would like to thank the Moreno family in Arenas, Veraguas Province, for granting us access to El Gato Beach, and Joelbin De la Cruz, Daniel Rivas, Bernardo Peña, Wilfredo Poveda, and Adrian Rivas for their assistance during the fieldwork and satellite tag installation. Additionally, we thank Jeffrey Seminoff from NOAA for his advice and suggestions regarding materials and the satellite tag installation procedure. We also express our gratitude to the National Research System of Panama (SNI) and to the Research Group in Engineering Technologies Applied to Society (GIITAS) for their support and collaboration in the development of this research project. The satellite tagging and monitoring program was part of a project funded by the Secretariat of Science, Technology, and Innovation of Panama (SENACYT) under Contract No. 191-2022. This work was conducted under Scientific Permit No. ARB-077-2022 was issued by the Ministry of Environment of Panama.

References

- Abang Shakawi, A. M. H., Shabri, A., & Hassan, R. (2024, October 1). *Forecasting Green Sea Turtle (Chelonia mydas) Landing in Sarawak Using Grey Model*. | EBSCOhost. <https://doi.org/10.55230/mabjournal.v53i4.3050>
- Alia, A., Maree, M., & Chraibi, M. (2022). On the exploitation of GPS-based data for real-time visualisation of pedestrian dynamics in open environments. *Behaviour & Information Technology*, 41(8), 1709–1723. <https://doi.org/10.1080/0144929X.2021.1896781>
- Alsaidi, M., Al-Jassani, M. G., Bang, C., O’Corry-Crowe, G., Watt, C., Ghazal, M., & Zhuang, H. (2024). Localization and tracking of beluga whales in aerial video using deep learning. *Frontiers in Marine Science*, 11. <https://doi.org/10.3389/fmars.2024.1445698>
- Anselin, L., & Rey, S. J. (2022). Open Source Software for Spatial Data Science. *Geographical Analysis*, 54(3), 429–438. <https://doi.org/10.1111/gean.12339>
- Azizan, N. H., Naharudin, N., Hashim, N., & Rusli, M. U. (2023). Site Suitability Analysis for Sea Turtle Nesting Area by using AHP and GIS. *IOP Conference Series: Earth and Environmental Science*, 1217(1), 012031. <https://doi.org/10.1088/1755-1315/1217/1/012031>
- Benscoter, A. M., Smith, B. J., & Hart, K. M. (2022). Loggerhead marine turtles (*Caretta caretta*) nesting at smaller sizes than expected in the Gulf of Mexico: Implications for turtle behavior, population dynamics, and conservation. *Conservation Science and Practice*, 4(1), e581. <https://doi.org/10.1111/csp2.581>
- Bokani, A., Yadegaridehkordi, E., & Kanhere, S. S. (2025). LSTM-H: A Hybrid Deep Learning Model for Accurate Livestock Movement Prediction in UAV-Based Monitoring Systems. *Drones*, 9(5), 346. <https://doi.org/10.3390/drones9050346>
- Cai, L., Aikio, A., Kullen, A., Deng, Y., Zhang, Y., Zhang, S.-R., Virtanen, I., & Vanhamäki, H. (2022). GeospaceLAB: Python package for managing and visualizing data in space physics. *Frontiers in Astronomy and Space Sciences*, 9. <https://doi.org/10.3389/fspas.2022.1023163>
- Chambault, P., Fossette, S., Heide-Jørgensen, M. P., Jouannet, D., & Vély, M. (2021). Predicting seasonal movements and distribution of the sperm whale using machine learning algorithms. *Ecology and Evolution*, 11(3), 1432–1445. <https://doi.org/10.1002/ece3.7154>

- Chambault, P., Gaspar, P., & Dell’Amico, F. (2021). Ecological Trap or Favorable Habitat? First Evidence That Immature Sea Turtles May Survive at Their Range-Limits in the North-East Atlantic. *Frontiers in Marine Science*, 8. <https://doi.org/10.3389/fmars.2021.736604>
- Christiaanse, J. C., Antolínez, J. A. A., Luijendijk, A. P., Athanasiou, P., Duarte, C. M., & Aarninkhof, S. (2024). Distribution of global sea turtle nesting explained from regional-scale coastal characteristics. *Scientific Reports*, 14(1), 752. <https://doi.org/10.1038/s41598-023-50239-5>
- Cullen, J. A., Domit, C., Lamont, M. M., Marshall, C. D., Santos, A. J. B., Sasso, C. R., Al Ansi, M., Hart, K. M., & Fuentes, M. M. P. B. (2024). A comparative framework to develop transferable species distribution models for animal telemetry data. *Ecosphere*, 15(12), e70136. <https://doi.org/10.1002/ecs2.70136>
- DiMatteo, A., Lockhart, G., & Barco, S. (2021). Normalizing home ranges of immature Kemp’s ridley turtles (*Lepidochelys kempii*) in an important estuarine foraging area to better assess the spatial distribution. *Marine Biology Research*, 17(1), 57–71. <https://doi.org/10.1080/17451000.2021.1896004>
- Goodwin, M., Halvorsen, K., Jiao, L., Knausgård, K., Martin, A., Moyano, M., Oomen, R., Rasmussen, J. H., Sørvald, T., & Thorbjørnsen, S. (2022). Unlocking the potential of deep learning for marine ecology: Overview, applications, and outlook. *ICES Journal of Marine Science*, 79. <https://doi.org/10.1093/icesjms/fsab255>
- Gupte, P. R., Beardsworth, C. E., Spiegel, O., Lourie, E., Toledo, S., Nathan, R., & Bijleveld, A. I. (2022). A guide to pre-processing high-throughput animal tracking data. *Journal of Animal Ecology*, 91(2), 287–307. <https://doi.org/10.1111/1365-2656.13610>
- Hardin, E. E., Cullen, J. A., & Fuentes, M. M. P. B. (2024). Comparing acoustic and satellite telemetry: An analysis quantifying the space use of *Chelonia mydas* in Bimini, Bahamas. *Royal Society Open Science*, 11(1), 231152. <https://doi.org/10.1098/rsos.231152>
- Hays, G. C., Laloë, J.-O., Rattray, A., & Esteban, N. (2021). Why do Argos satellite tags stop relaying data? *Ecology and Evolution*, 11(11), 7093–7101. <https://doi.org/10.1002/ece3.7558>
- Jeantet, L., Planas-Bielsa, V., Benhamou, S., Geiger, S., Martin, J., Siegwalt, F., Lelong, P., Gresser, J., Etienne, D., Hiélaud, G., Arque, A., Regis, S., Lecerf, N., Frouin, C., Benhalilou, A., Murgale, C., Maillet, T., Andreani, L., Campistron, G., ... Chevallier, D. (2020). Behavioural inference from signal processing using animal-borne multi-sensor loggers: A novel solution to extend the knowledge of sea turtle ecology. *Royal Society Open Science*, 7(5), 200139. <https://doi.org/10.1098/rsos.200139>
- Kipnis, D., Levy, Y., & Diamant, R. (2023). Sonar Point Cloud Processing to Identify Sea Turtles by Pattern Analysis. *IEEE Journal of Oceanic Engineering*, 48(2), 431–442. <https://doi.org/10.1109/JOE.2022.3214274>
- Kot, C. Y., Åkesson, S., Alfaro-Shigueto, J., Amorocho Llanos, D. F., Antonopoulou, M., Balazs, G. H., Baverstock, W. R., Blumenthal, J. M., Broderick, A. C., Bruno, I., Canbolat, A. F., Casale, P., Cejudo, D., Coyne, M. S., Curtice, C., DeLand, S., DiMatteo, A., Dodge, K., Dunn, D. C., ... Halpin, P. N. (2022). Network analysis of sea turtle movements and connectivity: A tool for conservation prioritization. *Diversity and Distributions*, 28(4), 810–829. <https://doi.org/10.1111/ddi.13485>
- Labrada-Martagón, V., Islas Madrid, N. L., Yáñez-Estrada, L., Muñoz-Tenería, F. A., Solé, M., & Zenteno-Savín, T. (2024). Evidence of oxidative stress responses of green turtles (*Chelonia*

- mydas*) to differential habitat conditions in the Mexican Caribbean. *Science of The Total Environment*, 946, 174151. <https://doi.org/10.1016/j.scitotenv.2024.174151>
- Li, J., Xu, W., Deng, L., Xiao, Y., Han, Z., & Zheng, H. (2023). Deep learning for visual recognition and detection of aquatic animals: A review. *Reviews in Aquaculture*, 15(2), 409–433. <https://doi.org/10.1111/raq.12726>
- Li, X., Sindihebura, T. T., Zhou, L., Duarte, C. M., Costa, D. P., Hindell, M. A., McMahon, C., Muelbert, M. M. C., Zhang, X., & Peng, C. (2021). A prediction and imputation method for marine animal movement data. *PeerJ Computer Science*, 7, e656. <https://doi.org/10.7717/peerj-cs.656>
- Mestre, J., Patrício, A. R., Sidina, E., Senhoury, C., El'bar, N., Beal, M., Regalla, A., & Catry, P. (2024). Movement patterns of green turtles at a key foraging site: The Banc d'Arguin, Mauritania. *Marine Biology*, 172(1), 1. <https://doi.org/10.1007/s00227-024-04558-4>
- Miao, Y., Li, B., & Guo, X. (2024). Research on multi-submersible positioning prediction based on Kalman filter and LSTM neural network. *Journal of Physics: Conference Series*, 2898(1), 012026. <https://doi.org/10.1088/1742-6596/2898/1/012026>
- Nathan, R., Monk, C. T., Arlinghaus, R., Adam, T., Alós, J., Assaf, M., Baktoft, H., Beardsworth, C. E., Bertram, M. G., Bijleveld, A. I., Brodin, T., Brooks, J. L., Campos-Candela, A., Cooke, S. J., Gjelland, K. Ø., Gupte, P. R., Harel, R., Hellström, G., Jeltsch, F., ... Jarić, I. (2022). Big-data approaches lead to an increased understanding of the ecology of animal movement. *Science*, 375(6582), eabg1780. <https://doi.org/10.1126/science.abg1780>
- Ning, H., Li, Z., Akinboyewa, T., & Lessani, M. N. (2025). An autonomous GIS agent framework for geospatial data retrieval. *International Journal of Digital Earth*, 18(1), 2458688. <https://doi.org/10.1080/17538947.2025.2458688>
- Noguchi, N., Nishizawa, H., Shimizu, T., Okuyama, J., Kobayashi, S., Tokuda, K., Tanaka, H., & Kondo, S. (2025). Efficient wildlife monitoring: Deep learning-based detection and counting of green turtles in coastal areas. *Ecological Informatics*, 86, 103009. <https://doi.org/10.1016/j.ecoinf.2025.103009>
- Pasanisi, E., Pace, D. S., Orasi, A., Vitale, M., & Arcangeli, A. (2024). A global systematic review of species distribution modelling approaches for cetaceans and sea turtles. *Ecological Informatics*, 82, 102700. <https://doi.org/10.1016/j.ecoinf.2024.102700>
- Perneel, M., Adriaens, I., Aernouts, B., & Verwaeren, J. (2025). Consistent multi-animal pose estimation in cattle using dynamic Kalman filter based tracking. *Smart Agricultural Technology*, 11, 101014. <https://doi.org/10.1016/j.atech.2025.101014>
- Restrepo, J., Webster, E. G., Ramos, I., & Valverde, R. A. (2023). Recent decline of green turtle *Chelonia mydas* nesting trend at Tortuguero, Costa Rica. *Endangered Species Research*, 51, 59–72. <https://doi.org/10.3354/esr01237>
- Rider, M. J., Avens, L., Haas, H. L., Hatch, J. M., Patel, S. H., & Sasso, C. R. (2024). Where the leatherbacks roam: Movement behavior analyses reveal novel foraging locations along the Northwest Atlantic shelf. *Frontiers in Marine Science*, 11. <https://doi.org/10.3389/fmars.2024.1325139>
- Rodríguez-Martínez, K., Rovetto, C., Cano, E., & Flores, E. E. (2024). Optimization of satellite biotelemetry data in seaturtles through outlier removal techniques. *2024 9th International Engineering, Sciences and Technology Conference (IESTEC)*, 188–193. <https://doi.org/10.1109/IESTEC62784.2024.10820233>

- Rovetto, C., Cruz, E., Flores, E., Nuñez, I., Rodriguez, K., & Cano, E. (2023). Behavioral data analysis of sea turtles from the Pacific coast of Panama, using biotelemetry. *2023 VI Congreso Internacional En Inteligencia Ambiental, Ingeniería de Software y Salud Electrónica y Móvil (AmITIC)*, 1–7. <https://doi.org/10.1109/AmITIC60194.2023.10366354>
- Santos, A. J. B., Cullen, J., Vieira, D. H. G., Lima, E. H. S. M., Quennessen, V., Santos, E. A. P. dos, Bellini, C., Ramos, R., & Fuentes, M. M. P. B. (2023). Decoding the interesting movements of marine turtles using a fine-scale behavioral state approach. *Frontiers in Ecology and Evolution*, 11. <https://doi.org/10.3389/fevo.2023.1229144>
- Shakawi, A. M. H. A., Shabri, A., & Hassan, R. (2024). Forecasting Green Sea Turtle (*Chelonia mydas*) Landing in Sarawak Using Grey Model. *Malaysian Applied Biology*, 53(4), 115–124. <https://doi.org/10.55230/mabjournal.v53i4.3050>
- Sharmila, S., & Sabarish, B. A. (2021). Analysis of distance measures in spatial trajectory data clustering. *IOP Conference Series: Materials Science and Engineering*, 1085(1), 012021. <https://doi.org/10.1088/1757-899X/1085/1/012021>
- Steenacker, M., Tanabe, L. K., Rusli, M. U., & Fournier, D. (2023). The influence of incubation duration and clutch relocation on hatchling morphology and locomotor performances of green turtle (*Chelonia mydas*). *Journal of Experimental Marine Biology and Ecology*, 569, 151954. <https://doi.org/10.1016/j.jembe.2023.151954>
- Tariq, U., Ahmed, I., Khan, M. A., & Bashir, A. K. (2025). Bridging biosciences and deep learning for revolutionary discoveries: A comprehensive review. *IAES International Journal of Artificial Intelligence (IJ-AI)*, 14(2), 867–883. <https://doi.org/10.11591/ijai.v14.i2.pp867-883>
- Wang, M., Xu, C., Zhou, C., Gong, Y., & Qiu, B. (2022). Study on Underwater Target Tracking Technology Based on an LSTM–Kalman Filtering Method. *Applied Sciences*, 12(10), 5233. <https://doi.org/10.3390/app12105233>
- Wijeyakulasuriya, D. A., Eisenhauer, E. W., Shaby, B. A., & Hanks, E. M. (2020). Machine learning for modeling animal movement. *PLOS ONE*, 15(7), e0235750. <https://doi.org/10.1371/journal.pone.0235750>
- Zakry, K. A., Soria, M. S., Hipiny, I., Ujir, H., Hassan, R., & Hardi, R. (2024). *Chelonia mydas* detection and image extraction from field recordings. *IAES International Journal of Artificial Intelligence (IJ-AI)*, 13(2), 2354–2363. <https://doi.org/10.11591/ijai.v13.i2.pp2354-2363>