Deep Learning Comparative Analysis of Predictions for Wolves' Movements

Robin Michel
HEC Lausanne
MSc. in Business Analytics
Email: robin.michel@unil.ch

***Abstract*—This study undertakes a comprehensive analysis of wolves' movements and their interactions with elks in Alberta, Canada, aiming to compare the performance of deep learning models in predicting wolf movements. Employing GPS tracking data from 2001 to 2011, we explore the efficacy of multiple predictive models using wolves' historical data, specifically ARIMA, FNN, RNN, LSTM, GRU, and an enhanced LSTM model that incorporates a proximity metric between wolves and elks. Each model's performance is assessed using a combination of quantitative metrics and visual comparisons of predicted versus actual movement paths on geographical maps. The findings reveal that neural network models, particularly LSTM, outperform the ARIMA baseline. This study also suggests that incorporating elk proximity metrics into the LSTM model may to some extent enhance predictive accuracy, compared to using historical data alone. Despite these results, all models struggle to precisely predict wolf movements, often failing to represent credible real-world movement patterns. This analysis highlights the complexity of animal movement prediction and suggests that further refinement and integration of ecological data are necessary to enhance model performance.**

I. Introduction

The analysis of animal movement patterns, particularly those of large predators such as wolves, plays a pivotal role in ecological research and wildlife management [1]. Wolves, with their complex behaviors and significant impact on ecosystems, present a unique challenge for ecologists seeking to understand and predict their movements [2]. In Alberta, Canada, where wolves and elks coexist within shared territories, predicting wolf movements is not only critical for conservation efforts but also for maintaining the balance of predator-prey dynamics [3]. As elks are one of the main preys for wolves [4][5], this study will also analyze elk movements to provide a comprehensive understanding of this ecological interaction. However, despite the ecological importance, few studies have employed advanced machine learning techniques to forecast these movements.

This gap in the literature is particularly pronounced in the context of using deep learning models for comparative analysis. While traditional studies have focused on observing and analyzing movement patterns, there is a significant opportunity to leverage the predictive power of machine learning to forecast future movements. This research aims to fill this gap by conducting a comparative analysis of several deep learning models to determine which is the most effective at predicting wolf movements based on historical data.

The importance of such predictive modeling extends beyond academic interest. Understanding wolf movements can inform wildlife management strategies, aid in the preservation of endangered species, and help mitigate potential conflicts between wolves and human activities. Furthermore, the outcomes of this research could provide insights into the broader implications of predator movements on ecological communities.

To address these environmental challenges effectively, the study will first analyze the movements of both species, utilizing data from 16 wolves [6] and 42 elks [7], to establish an understanding of their interactions. Subsequently, it will proceed with the modeling phase, evaluating various deep learning models for their ability to accurately predict future movements of wolves. This comprehensive approach aims not only to assess the predictive capabilities of each model but also to explore how enhancements, such as the integration of elk proximity data, might improve their accuracy and applicability to real-world conservation efforts.

II. Research Question and Relevant Literature

In the field of environmental research, the analysis of animal movements has been extensively explored through various modeling techniques. However, these studies often focus on descriptive rather than predictive analytics. Lewis et al. (1997) [8] provided a seminal contribution using spatially explicit mathematical models to analyze territorial patterns in wolves, focusing on the role of scent marking in shaping territories. While this study offers critical insights into territorial dynamics, it largely overlooks the potential of predictive modeling to anticipate future movement patterns.

Later, Hebblewhite et al. (2005) [9] explored the spatial aspects of wolf and elk interactions, providing a method to decompose predation risk into stages of encounter and attack, modeled via resource selection functions. This approach underscores the potential for predictive models to integrate detailed spatial analyses of predator-prey dynamics, moving beyond simple locational data to embrace the complexity of ecological interactions.

Building on the need for more predictive approaches, Franke et al. (2006) [10] employed Hidden Markov Models to predict wolf kill-sites in Alberta, demonstrating the utility of complex statistical models in decoding spatial behavior from GPS tracking data. This study marked a significant advance by integrating spatial data into predictive models, yet it did not address the temporal complexities that deep learning architectures like LSTM and GRU are well-equipped to handle.

Regarding more recent studies, Ditmer et al. (2023) [11] explore the potential for wolf recolonization in Colorado using a Spatial Absorbing Markov Chain (SAMC) model to predict dispersal patterns and conflict risks. This innovative model integrates dynamic elements such as movement resistance from terrain and human structures, various mortality risks, and site fidelity based on habitat quality. Despite its advancements in blending ecological, geographical, and social factors, the study notably ignores deep learning techniques, which could potentially enhance predictive accuracy by capturing complex nonlinear relationships and temporal dynamics.

These studies collectively highlight a critical gap in the current ecological research landscape—the limited use of advanced machine learning techniques, particularly deep learning, that might seamlessly integrate and process nonlinear and temporally complex data for predictive analysis. The literature underscores the need to not only analyze but also predict animal movements by using complex machine learning models.

In response to the identified gaps in the literature, this study will conduct a comparative analysis of deep learning models to assess their effectiveness in predicting the movements of wolves in Alberta, Canada, based on historical data. This research will employ a variety of deep learning models, assessing their ability to accurately predict future locations. We will also explore whether incorporating proximity metrics between wolves and elks can enhance model performance, thereby refining the predictive process.

III. Methodology

This section details the methodology employed in the study, outlining the key processes and techniques used to develop and evaluate the predictive models. It includes discussions on pre-processing steps, model description, and the criteria used for assessing model performance.

A. Pre-Processing

The initial phase of our analysis involved the alignment and integration of two distinct datasets: one detailing wolf geolocations and the other recording elk's in Alberta, Canada. To ensure the coherence and relevance of our analysis, it was required to align these datasets spatially and temporally.

The datasets were filtered to exclude any wolves that did not share the same geographical territory as elks. This alignment was done to maintain the focus of our study on areas where interactions between these two species were most likely. Following this spatial alignment, we performed a temporal synchronization. We adjusted the datasets to retain only those elk time series that coincided with the time frame of at least one wolf, thereby eliminating any discrepancies in observation periods between the two species.

The merged dataset, thus refined, consisted of wolf and elk observations that were both temporally aligned and geographically co-located, providing a robust foundation for further analysis of their movements and interactions.

B. Proximity Metrics Computation

To effectively integrate ecological interactions into the predictive modeling of wolf movements, a proximity metric between wolves and elk was computed. Given the large volume of observations, direct computation of distances between every wolf and elk was computationally impractical. Instead, a dimension reduction approach was adopted through clustering to manage the data more feasibly.

K-means clustering was utilized to identify regions of high elk density. Despite the elbow method suggesting fewer clusters, an arbitrary decision was made to set the number of clusters at 15. This decision was informed by a heat map analysis of elk presence (see Figure 1), which identified more than five distinct regions with significant elk activity. This approach ensured that the clustering captured all key elk hotspots effectively. Comparisons of clustering algorithms, including Gaussian Mixture Models (GMM) and assessments using the Calinski-Harabasz and Davies-Bouldin indices, confirmed that K-means provided the most suitable clustering for this task (see Table 1).

TABLE I
Clustering Performance

	Calinski Higher the better	**Davies** Lower the better
K-means	687884	0.581
GMM	501084	0.761

The centroids of these 15 elk clusters, represented by the red circles, were computed and treated as static throughout the study, based on the assumption that elk movement patterns remain consistent over time. This simplification allowed for the consistent application of proximity metrics without the need for updating centroids dynamically.

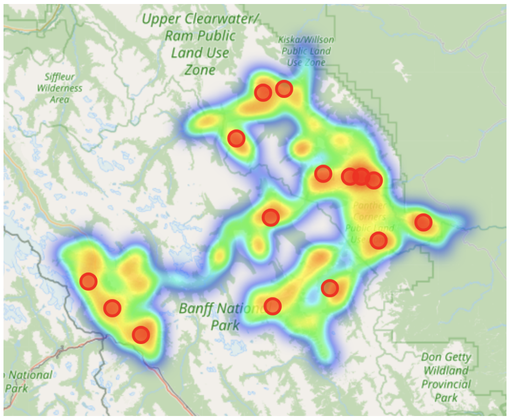

Fig. 1. Elk Clustering using K-means

These centroids then served as reference points for computing the proximity metrics. The metric used was defined as $\frac{1}{\text{wolf\_nearest\_distance}+1}$, measuring the inverse of the distance from a wolf observation to the nearest elk cluster centroid. This proximity metric was later incorporated into the loss function of an LSTM model, aiming not only to minimize the prediction error but also to reduce the geographical distance between predicted wolf locations and areas of high elk density.

Analysis of the proximity map (see Figure 2), which illustrates the locations of wolves relative to the elk cluster centroids, clearly indicates that wolves frequently approach these high-density elk areas. Warmer circles on the map represent wolf observations where proximity to elk is higher, visually demonstrating closer interactions. This observation supports the hypothesis that interactions between wolves and elk are probable, reinforcing the relevance of including these proximity metrics in predictive models. The proximity map thus serves as a critical tool for confirming the spatial dynamics observed between these two species.

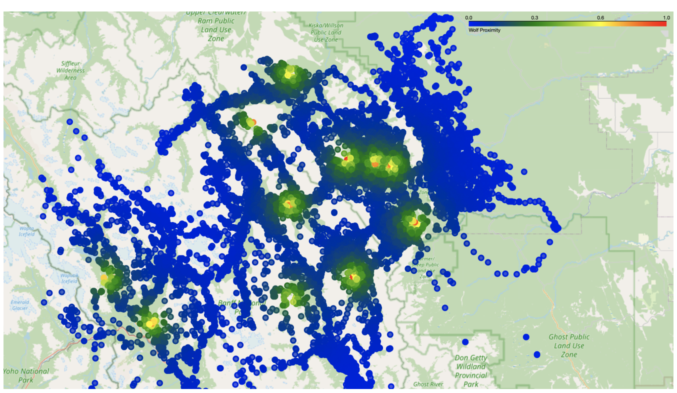

Fig. 2. Proximity Map between Wolves and Elks

C. Model Application and Description

The modeling of wolf movements is based on over 54,000 observations across 16 distinct wolf time series. For each wolf, both latitude and longitude data series were handled simultaneously, requiring models to predict these two interdependent variables concurrently. Each model was trained independently on the 16 time series, using the historical latitude and longitude observations as inputs, meaning that each model had to be trained 16 times, once for each time series. Given the extensive dataset and the necessity to manage multiple time series, strategic compromises were implemented to accommodate computational limitations. To enable the application of advanced deep learning models such as LSTM and GRU, training was restricted to 10 epochs with a batch size of 32. Although these parameters were less than ideal, they were necessary to balance computational feasibility with the potential for meaningful model insights. The data was divided using a train, validation, and test split to ensure robust model evaluation and to mitigate overfitting. As we are working with time series, overlapping windows were employed to improve the models' ability to learn temporal dependencies, which is essential for predicting movement sequences accurately. We experimented with a range from 5 to 40 steps for overlapping windows; however, windows spanning 20 to 30 steps yielded the best results.

Additionally, hyperparameter tuning was carried out through grid search to optimize each model's performance, tailoring parameters specifically to the dynamics of the time series data. Details on these optimizations will be provided in subsequent sections. This setup introduces the discussion on the machine learning models employed, including ARIMA, FNN, RNN, LSTM, and GRU. Detailed descriptions of the chosen models and explanations of their suitability for our specific study context will be provided below.

1) ARIMA: The ARIMA (Autoregressive Integrated Moving Average) model is a popular statistical approach used for analyzing and forecasting time series data, particularly well-suited for data that shows evidence of non-stationarity. ARIMA models work by describing the autocorrelations in the data. They are composed of three main parts: the autoregressive (AR) part, which indicates that the evolving variable of interest is regressed on its own prior values; the integrated (I) part, which indicates that the data values have been replaced with the difference between their values and the previous values (this differencing process helps in making the data stationary); and the moving average (MA) part, which incorporates the dependency between an observation and a residual error from a moving average model applied to lagged observations.

In the context of this study, ARIMA serves as the baseline model due to its simplicity and linear nature. While ARIMA models are powerful for many forecasting scenarios, they typically do not capture complex nonlinear patterns as effectively as some advanced neural network models. Therefore, using ARIMA allows us to establish a benchmark for the performance of more sophisticated models. By setting ARIMA as the baseline, the study aims to illustrate the potential enhancements in predictive accuracy that might be achieved with neural network approaches, especially in handling the dynamic and complex time series data of wolf movements. This comparative analysis helps to validate whether the additional complexity and computational expense of neural network models are justified by a significant improvement over the baseline ARIMA model.

2) Feedforward Neural Network: Feedforward Neural Networks (FNNs), also known simply as multilayer perceptrons, consist of layers of nodes in which the signal direction moves strictly forward—from input nodes, through hidden layers (each node connected to every node in the previous layer), to output nodes. FNNs are generally used for pattern recognition and classification due to their structure, which excels in approximating any function or correlation between inputs and outputs. However, they do not inherently possess memory components to store any information about previous inputs, which is essential for processing time series data where past values influence future ones.

To adapt FNNs for time series analysis, we employed overlapping windows—a technique that segments the data into overlapping sequences that serve as independent observations. Each segment provides the network with a snapshot of recent temporal data, allowing the model to infer patterns that extend over the sequence's span. This method introduces a form of "pseudo-memory", enabling FNNs to capture temporal dependencies somewhat akin to more sophisticated time series models. While this adaptation does not endow FNNs with the genuine sequential data processing capabilities of LSTMs, or GRUs, it enables FNNs to serve as a preliminary tool to evaluate how well neural networks might improve upon baseline models like ARIMA in predicting complex time series data.

3) Recurrent Neural Network: Recurrent Neural Networks (RNNs) are particularly well-suited for handling time series data, as a result of their architecture that processes sequences of information iteratively. Unlike feedforward neural networks, RNNs maintain a form of internal state or memory, making them capable of capturing temporal relationships within data. This memory is achieved by looping the output of a layer back to its input, which allows the network to carry forward information across the steps of the sequence. However, RNNs are not without limitations. One significant challenge they face is the vanishing gradient problem—during training, gradients, which are used to update the weights, can shrink exponentially as they propagate back through each time step, making it difficult for the model to learn correlations from data points that are far apart in time.

This issue restricts the effectiveness of standard RNNs in modeling long-term dependencies within time series data, as they struggle to retain earlier information in longer sequences. While RNNs can model short-term dependencies with reasonable success, their inability to capture longer-range temporal relationships can hamper their performance in more complex scenarios. This shortcoming sets the context for the introduction of Long Short-Term Memory (LSTM) networks, which include additional mechanisms to manage information over extended periods more effectively. Despite these challenges, RNNs still represent a significant advancement over simpler models like FNNs, providing a deeper understanding of time-dependent data dynamics.

4) Long Short-Term Memory: Long Short-Term Memory (LSTM) networks are a refined type of Recurrent Neural Network (RNN) specifically designed to address and overcome the limitations of standard RNNs, including the vanishing gradient problem that can impede learning in sequences of long durations. The key to LSTM's effectiveness lies in its sophisticated architecture, which includes several gates that manage the flow of information: the input gate, forget gate, and output gate. These gates function as regulators that selectively control which information should be retained or discarded as the network processes data through time.

The input gate decides which values from the input should be used to modify the memory state. The forget gate can remove information that is no longer relevant to the network's task, helping to prevent issues related to over-cluttering of the memory state, which is essential for maintaining the network's performance. The output gate then determines what the next hidden state should be, which contains information on previous inputs and helps predict future ones. By manipulating these gates, LSTMs can maintain a balance between the memory of old information (which is necessary for understanding context and dependencies over time) and new input data (which keeps the model updated with the latest observations).

In the specific context of modeling wolf movements, LSTMs are advantageous due to their ability to manage long-term dependencies effectively. These dependencies are critical in animal movement data, where past behaviors and locations could influence future movements over extended periods. LSTMs utilize their gated architecture to retain important past information for lengthy sequences, which potentially enhances their ability to forecast future positions. This capability allows LSTMs to capture more complex temporal patterns, suggesting that they might be better suited for predicting wolf locations by utilizing historical data to infer future movements.

5) Gated Recurrent Unit: Gated Recurrent Units (GRUs) [12] are a streamlined variant of Long Short-Term Memory (LSTM) networks, both of which are designed to effectively handle long-term dependencies in sequence data. GRUs achieve similar outcomes to LSTMs by using a more simplified internal structure, which incorporates only two gates—reset and update gates—compared to the three gates (input, output, and forget) used in LSTMs. The reset gate determines how to combine new input with the previous memory, and the update gate defines how much of the past information to keep versus the new information to add.

This reduction in complexity generally results in GRUs being less computationally intensive, allowing for faster training times without a substantial decrease in performance. This makes GRUs particularly advantageous in scenarios where computational resources are limited or when rapid model training is required. In the context of modeling wolf movements, where the data involves complex temporal patterns and potentially large datasets, GRUs can offer a good balance between computational efficiency and modeling effectiveness. While GRUs may not always match the capability of LSTMs, their simplified architecture could still handle the sequence dependencies involved in tracking animal movements over time, potentially with fewer resources and faster execution than more complex models like LSTMs.

6) Enhanced Long Short-Term Memory: While the standard LSTM model uses only historical data to predict future movements, the enhanced LSTM model builds upon this approach by incorporating a proximity metric into its learning process. This metric quantifies the closeness of wolf locations to regions of high elk density, integrating a spatial dimension into the model's predictive capabilities. By embedding this proximity metric into the loss function, the enhanced LSTM assigns a higher penalty to prediction errors that occur closer to these critical regions. This adjustment not only encourages more accurate forecasts in areas where elk are concentrated but also ensures that spatial relationships influencing wolf behaviors are factored into the model's calculations.

This approach could potentially enhance the model's ability to generalize across different scenarios and provide more accurate predictions where they are most needed. Incorporating both spatial and temporal data considerations offers a more comprehensive modeling strategy that could improve performance, particularly in ecological studies where understanding the dynamics between predator locations and prey densities is essential.

D. Parameter Choice

While the ARIMA model utilized an auto.arima() function to automatically determine the optimal order of p, d, and q parameters, the selection of hyperparameters for the neural network models was systematically managed through a grid search.

The parameter tuning was conducted on a standard CPU, constraining the scope of our experiments but necessitating efficiency in our choices. Aware of the constraints posed by the available computational resources, a deliberate trade-off was made regarding the number of epochs and batch size used during training. While limiting the epochs to 10 and setting the batch size to 32 are far from optimal and may not fully exploit the potential of the models, these parameters were necessary compromises. This approach allowed the incorporation and testing of more complex models such as LSTM and GRU, which are computationally demanding but offer significant advantages in modeling time series data.

The grid search was designed to explore combinations of three key hyperparameters: the number of layers, the number of nodes per layer, and the learning rate. For each neural network model, the grid search was performed separately on each of the 16 time series, with each series representing a distinct pattern of wolf movements. The layers were varied at three levels, $l = \{3, 5, 7\}$, and the nodes were tested at four, $n = \{32, 64, 128, 256\}$, to assess how depth and width impacts model capacity to capture complex temporal dependencies without incurring significant computational costs or overfitting. Learning rates were explored at three scales, $lr = \{1e\text{-}5, 1e\text{-}4, 1e\text{-}3\}$, to determine the most effective rate for convergence that would not compromise the stability of the model updates. It was observed that most models performed better with a higher number of nodes, ranging from 64 to 256, but fewer layers, suggesting that wider neural network structures tend to be more effective than deeper ones for this particular application.

Unfortunately, due to computational resource limitations, the grid search for LSTM and GRU models was restricted. Hyperparameter tuning for these models was conducted with only two levels of layers and nodes, specifically $n = (64, 128)$ and $l = (3, 5)$. This adjustment represented a balance between model complexity and computational feasibility. Additionally, only learning rates of 1e-4 and 1e-5 were tested because higher rates led to instability issues such as exploding gradients, which resulted in infinite or NaN values during training.

The primary metric used to gauge the performance of each configuration was the Mean Squared Error (MSE), which measures the average of the squares of the errors—that is, the average squared difference between the estimated values and what is estimated. MSE was chosen for its sensitivity to large errors, making it suitable for emphasizing accuracy in predictions of wolf locations. The grid search validated each combination using a straightforward train-validation split, simplifying the validation process while still providing robust insights into each model's predictive accuracy.

Given the computational limitations of running these intensive models on a CPU without GPU acceleration, considerations such as training duration and risk of kernel crashes also influenced the final selection of hyperparameters. The chosen parameters for each time series model were those that offered the best balance between computational efficiency and predictive performance, as evidenced by the lowest MSE scores.

E. Movements Prediction

Predictions of wolf movements were made by forecasting the future positions based on the last observations from the training dataset, specifically targeting the last 20% of the observations designated as the test set. This approach allows for a direct evaluation of the model's accuracy against actual data. The models utilized the final input data from the training segment, which consists solely of latitude and longitude values.

The prediction process is iterative, with each model outputting a forecast for one step ahead and then feeding this predicted output back as the new input. This method implements a rolling window technique on the input data, where the current input array is adjusted to discard the oldest input and incorporate the latest prediction at the end. Such a strategy ensures the continuity of data through the predictive model, thus simulating real-time data processing and allowing the model to generate a series of predictions that extend through the length of the test set.

As explained, the models were trained on each of the 16 individual wolf time series using a grid search approach to optimize the hyperparameters. Once the best parameters were established, the best models for each time series were used to make predictions, each tuned to optimally forecast based on its learned patterns from the training phase.

F. Model Evaluation

The effectiveness of the six models tested was compared using three key numerical metrics: Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and the Haversine Average Distance, each calculated separately for latitude and longitude predictions. RMSE and MAE provide measures of the average magnitude of the errors in the predictions, reflecting the precision of the model to predict wolf locations. Key for this study's geographic component, the Haversine Average Distance calculates the average distance in kilometers of all the observations between predicted and actual locations, factoring in the Earth's curvature to assess the spatial accuracy of the models.

To evaluate which models yielded the most promising results, an average of these metrics was taken across the 16 time series. This approach allowed for a direct comparison of model performance, highlighting strengths and weaknesses in both numeric precision and spatial accuracy.

In addition to numerical metrics, visual comparisons were also employed to assess how well each model captured the actual movement patterns of the wolves. By mapping both the predicted and actual trajectories of the wolves on the same visual plane, we could qualitatively assess whether the models not only achieved favorable scores on the metrics but also convincingly replicated the real-world movements of the wolves.

IV. Dataset Description

The two datasets used in this study—one for wolves' locations [6] and one for elks' [7]—were sourced from Movebank, a database that catalogs animal movement and geolocation data from various research studies. This analysis specifically utilizes data collected from 2001 to 2011, during which both wolf and elk populations in the same geographic region were monitored in Alberta, Canada. The two datasets were aligned to ensure that the timeframe was consistent across both species, allowing for accurate comparative analysis of their movements and interactions. The wolf dataset comprises observations of 16 individual wolves, totaling 54,467 data points, while the elk dataset includes data on 42 elks, amounting to 146,802 observations. The timeframe of observations for the wolves varied, ranging from 2 months to 1.5 years. The locations of these animals were monitored using GPS trackers, which provided location data every 1 to 4 hours and identified each animal by individual ID, ensuring a comprehensive dataset for analysis. We will further explore the detailed movements of both wolves and elks to gain a deeper understanding of their behaviors and interactions within their habitat.

A. Wolves Movements Description

The analysis of the movements of 16 wolves across a vast area of approximately 18,000 square kilometers reveals intriguing insights into their spatial behaviors and interactions. The wolves are organized into five distinct packs, each consisting of 3 to 4 individuals. Notably, wolves within the same pack display highly coordinated movements, typically staying within clearly defined territory boundaries, with minimal overlap with neighboring packs. (see Figure 3). This pack behavior is supported by the observation of a strong positive correlation in movement among members of the same pack. On the provided figure, the green and red crosses on the map symbolize the initial and final observations of the wolves, respectively.

The analysis of travel distances shows that wolves travel extensively, covering anywhere from 200 to 600 kilometers per month. These movements are not random; rather, they are strategically aligned with the nocturnal and crepuscular periods, as wolves are significantly more active between midnight and noon (00:00 to 12:00). This temporal pattern highlights their adaptation to cooler temperatures and possibly to the behaviors of their prey, which are also more active during these times.

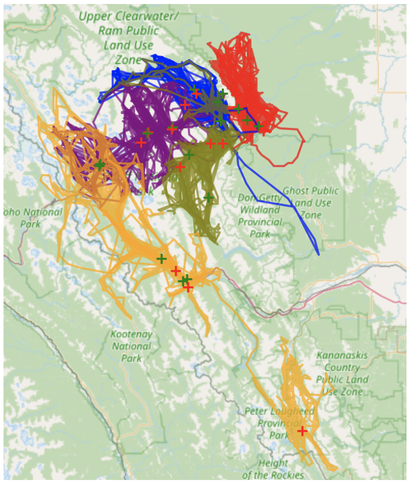

Fig. 3. Wolves Movements Map from 2001 to 2011

B. Wolf and Elk Interaction Description

The interactions between wolves and elks within the studied area demonstrate significant territorial overlap and proximity, indicating a complex ecological relationship. From the proximity analysis detailed in section 3.b, we calculated the proximity between each wolf and the nearest elk clusters at every recorded observation. This metric revealed that wolves and elks from the datasets share closely aligned territories.

Spatial overlap between these species is visually confirmed through a superimposed map (see Figure 4), which combines movement paths of both wolves (red) and elks (green). This visual representation distinctly shows both species navigating the same territories, reinforcing the findings from the proximity metrics. The map serves as a compelling illustration of how intertwined their habitats are.

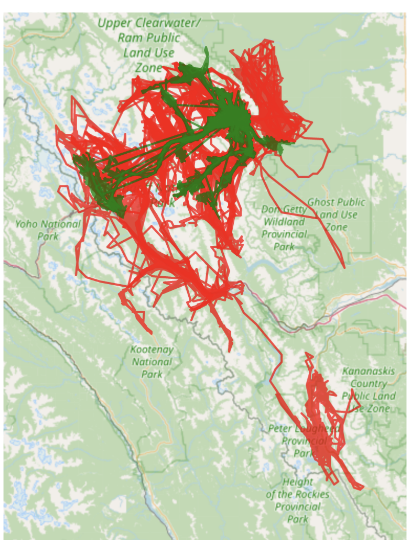

Fig. 4. Wolves and Elks Movements Map from 2001 to 2011

Further analyses using the Spearman correlation technique shed light on the relationship between the movements of wolves and the high-density spots of elks. Although the correlation varied among individual wolves—with some showing significant correlations and others less so—it overall suggests a meaningful ecological interaction.

V. Implementation of the Code

A. Downloading the Data

The datasets for both wolves and elks were sourced from Movebank and initially loaded into VSCode. The first step involved checking for missing values and outliers to ensure data quality before further analysis.

B. Pre-Processing

The pre-processing involved merging the wolf and elk datasets based on matching timestamps, ensuring that each data point from the wolves could be directly compared with contemporaneous data from the elks. Data handling was facilitated by the pandas package, while numpy was used for essential mathematical operations. Additionally, the sklearn library was utilized for implementing K-means clustering, aiding in the categorization and analysis of complex data structures.

C. Data Analysis

The pandas library facilitated an examination of the datasets, helping to categorize aspects such as the number of observations, the number of individual animals, and the specific periods of data collection. Calculations for geographical distances, which take into account the curvature of the Earth, were performed using numpy, leveraging the Haversine formula. For visualization, the Folium library played a key role by generating interactive maps that illustrated the movements and territories of the wolves and elks. These maps effectively highlighted areas of frequent activity and overlap, enriching our understanding of how these species interact within their environments.

D. Model Implementation

The modeling phase involved using a grid search to tune hyper-parameters such as the number of layers, nodes, and learning rates, aiming to optimize each model for its respective time series. Keras, a high-level neural network API, was the primary tool for implementing and training deep learning models. The grid search allowed for systematic testing across various configurations, balancing model complexity and performance to prevent overfitting while maximizing predictive accuracy.

VI. Results

To facilitate a systematic comparison, the average performance of each model was calculated across all 16 wolf time series, focusing on the Root Mean Square Error (RMSE) and the Haversine Distance. By examining the models' performances through these metrics, we will identify which models generally outperform others in terms of prediction accuracy (see Table 2). Following this comparative analysis, we will explore the practical utility of the top-performing models by visually analyzing how accurately their predictions match the actual movements of the wolves.

ARIMA, serving as the baseline model, recorded the highest RMSE and Haversine Distance values, with RMSEs of 0.1794 for latitude and 0.1831 for longitude, and Haversine Distances of 40.61 km for latitude and 40.27 km for longitude. These figures highlight its relative inaccuracy compared to the more sophisticated neural network models.

Despite its suitability for sequential data, the RNN surprisingly under-performed in comparison to other neural networks, posting RMSEs of 0.1271 for latitude and 0.2270 for longitude, with Haversine Distances of 20.04 km and 19.36 km respectively. In contrast, the FNN, with its simpler architecture, achieved unexpectedly strong performance, recording RMSEs of 0.1137 for latitude and 0.2291 for longitude, and Haversine Distances of 18.83 km and 16.98 km respectively, thus outperforming the RNN in this analysis.

GRU and LSTM exhibited similar levels of performance, both significantly outperforming the ARIMA baseline and RNN, similar to FNN. GRU recorded an RMSE of 0.1288 for longitude and Haversine distances of 19.58 km for latitude and 17.78 km for longitude. Meanwhile, LSTM posted RMSEs of 0.1187 for latitude and 0.1985 for longitude, with Haversine Distances of 18.00 km for latitude and 17.50 km for longitude. These results underscore the efficacy of both models in handling complex time series data. The particularly strong performance of FNN may be attributed to its training on more complex architectures compared to LSTM and GRU, due to its lower computational demands. This advantage allowed FNN to explore a wider range of model configurations, potentially enhancing its ability to effectively capture spatial movements.

The enhanced LSTM model, which incorporates proximity metrics between wolves and elks, displayed the best performance overall. With RMSEs of 0.1126 for latitude and 0.1889 for longitude, and Haversine Distances of 16.99 km for latitude and 16.46 km for longitude, LSTM+ not only showed slight improvements over standard LSTM but also stood out as the most accurate model in the comparative analysis. This suggests that integrating external ecological factors might enhance the model's ability to predict more accurately.

TABLE II
Average Model Performance

Model	RMSE Latitude	RMSE Longitude	Haversine Latitude	Haversine Longitude
ARIMA	0.1794	0.1831	40.61	40.27
FNN	0.1137	0.2291	18.83	16.98
RNN	0.1271	0.2270	20.04	19.36
LSTM	0.1187	0.1985	18.00	17.50
GRU	0.1288	0.2159	19.58	17.78
LSTM+	0.1126	0.1889	16.99	16.46

The ranking based on these metrics places LSTM+ at the top for its enhanced accuracy, followed by LSTM and GRU, which both performed similarly well. FNN also demonstrated strong results, potentially benefiting from the flexibility in its training configurations. RNN trailed behind other neural networks, showing some limitations in this analysis, yet it still performed significantly better than ARIMA. This hierarchy highlights the varying effectiveness of each model and underscores not only that neural networks outperform simpler linear models like ARIMA, but also that the inclusion of external variables in complex models like LSTM refines its predictive capabilities. Notably, all neural network models performed better with 20 steps in the overlapping window, except for the RNN, which showed optimal results with 30 steps. Additionally, models with a higher number of nodes (ranging from 64 to 256) and fewer layers (typically three) demonstrated improved performance, therefore suggesting that wider neural network structures are more effective than deeper ones for this particular application.

The predictions made by the models were visually represented on maps to assess their accuracy and plausibility. On these maps, the training set movements were depicted in blue, the actual movements from the test set in green, the predicted movements in red and the last observation of the training set is represented by a purple cross. Each point on the map represents a wolf's location at a given time, with lines connecting sequential observations to illustrate the movement paths.

Despite some models showing better performance than others in terms of metrics, the overall visual assessment indicates that the results across models are mixed. All models generally struggled to accurately recreate the actual movements of wolves across time series. While some models produced reasonable predictions for certain time series, they failed to do so for others, indicating a significant variability in performance.

Notably, some models did manage to produce movement paths that were plausible, aligning with how real wolves might move across the landscape. For instance, among other examples, the LSTM model produced believable movements for wolf B087, and the RNN model captured realistic movements for wolves B065 and JW02 (see Figures 5, 6 and 7). These figures show that despite the predictions not aligning precisely with the test set data, the modeled paths still resembled potential real-life wolf trajectories.

However, many predictions did not convincingly replicate plausible wolf movements. For example, predictions by the LSTM model for wolves B045 and B077 were notably inaccurate and did not reflect typical wolf movement patterns (see Figures 8 and 9). This discrepancy highlights that even the more successful models, such as LSTM and the enhanced LSTM, often fell short of predicting realistic, accurate movements.

This visual analysis underscores that while some models can occasionally predict movements that are superficially plausible, consistently producing accurate and believable predictions across different scenarios remains a significant challenge.

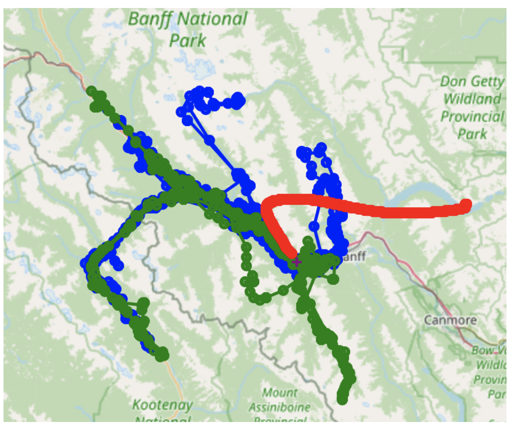

Fig. 5. LSTM Predictions for Wolf B087

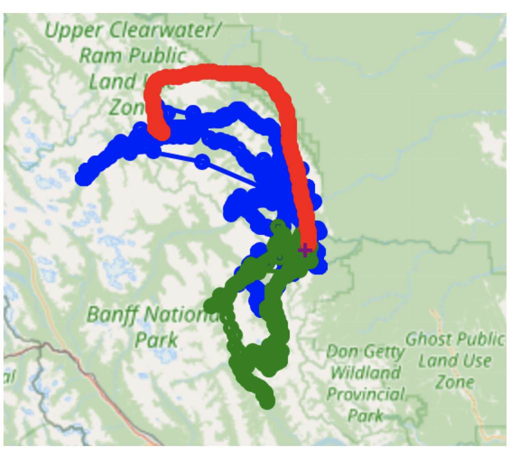

Fig. 6. RNN Predictions for Wolf B065

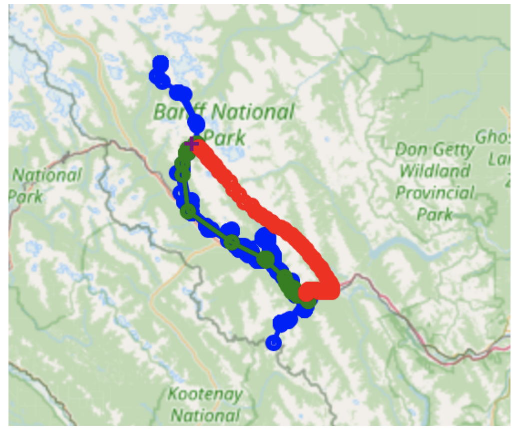

Fig. 7. RNN Predictions for Wolf JW02

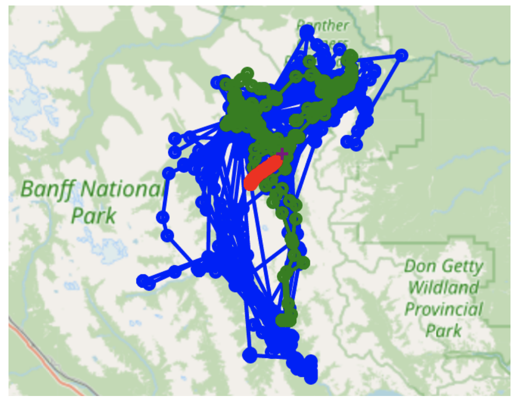

Fig. 8. LSTM Predictions for Wolf B045

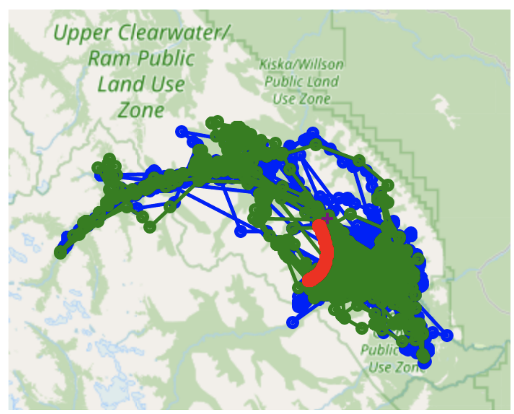

Fig. 9. LSTM Predictions for Wolf B077

VII. Conclusion

Wolf movements, influenced by a variety of environmental factors such as temperature, snow depth, prey movements, and others, exhibit a complexity that poses significant challenges to predictive modeling. This study's deep learning comparative analysis revealed that advanced neural network models like LSTM outperform simpler models such as ARIMA. However, even sophisticated models struggle to deliver consistently plausible and accurate predictions. The inconsistencies and inaccuracies in predictions may stem from an insufficient inclusion of external factors that influence wolf movements. Additionally, a limited timeframe fails to capture the seasonality and long-term trends of wolf behavior, as highlighted for example by the time series of wolf B083, which shows no discernible trend or seasonality (see Figure 10).

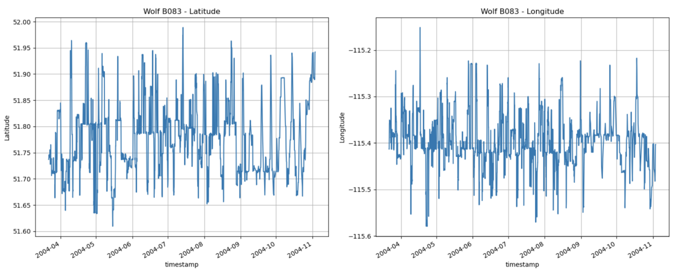

Fig. 10. Time series of wolf B083

With the improvements seen in the enhanced LSTM, this analysis also suggests that incorporating external factors, such as proximity to wolves' prey, can enhance the accuracy of movement predictions. This highlights the inadequacy of relying solely on historical location data to predict wolf movements. The complex nature of these movements, often influenced by factors beyond mere geographical data, suggests that incorporating additional environmental variables such as topography, and weather conditions might enhance predictive accuracy. Moreover, the presence of other prey species like deer and moose could also provide a significant improvement in predicting wolf behavior.

To address these complexities, enhanced computational resources are necessary. With more robust computational support, it would be possible to develop more complex models that can be trained with larger datasets, more epochs, and a broader range of hyperparameters. Such resources would also allow for the implementation of refined methodologies that could accommodate the inclusion of multiple external factors, potentially leading to more accurate and reliable predictions.

Looking forward, future research should focus on integrating these environmental factors into predictive models. Additionally, incorporating observations over a larger timeframe could help better understand trends and seasonality in wolf movements, providing deeper insights into their behavioral patterns. This progression would not only theoretically improve the models but could also provide more practical and actionable insights for wildlife management and conservation efforts.

References

[1] Northwest, C. (2018, November 14). Understanding wolf behavior—for your safety and theirs. Conservation Northwest - Protecting, Connecting and Restoring Wildlands and Wildlife. https://conservationnw.org/understanding-wolf-behavior/

[2] The role of wolves in ecosystems. Washington Department of Fish Wildlife. https://wdfw.wa.gov/species-habitats/at-risk/species-recovery/gray-wolf/influence

[3] Ecology, J. O. A. (2023, September 28). Predicting dispersal and conflict risk for wolf recolonisation in Colorado. The Applied Ecologist. https://appliedecologistsblog.com/2023/09/28/predicting-dispersal-and-conflict-risk-for-wolf-recolonisation-in-colorado/

[4] How Wolves Hunt - Living with Wolves. (2023, July 19). Living With Wolves. https://www.livingwithwolves.org/how-wolves-hunt/

[5] International Wolf Center. (2024, April 18). Hunting Feeding Behavior — International Wolf Center. International Wolf Center — Teaching the World About Wolves. https://wolf.org/wolf-info/basic-wolf-info/biology-and-behavior/hunting-feeding-behavior/

[6] Wolf Data Source : Movebank. https://www.movebank.org/wolf data

[7] Elk Data Source : Movebank. https://www.movebank.org/elk data

[8] Lewis, M. A., White, K. a. J., Murray, J. D., Department of Mathematics, University of Utah, JWB 233, Salt Lake City, UT 84112 USA, Department of Zoology, Downing Street, University of Cambridge, Cambridge CB2 3EJ, UK, Department of Applied Mathematics, University of Washington, Seattle, WA 98195, USA. (1997). Analysis of a model for wolf territories. In J. Math. Biol.

[9] Hebblewhite, M., Merrill, E. H., McDonald, T. L. (2005). Spatial decomposition of predation risk using resource selection functions: an example in a wolf–elk predator–prey system. Oikos, 111(1), 101–111. https://doi.org/10.1111/j.0030-1299.2005.13858.x

[10] Franke, A., Caelli, T., Kuzyk, G. W., Hudson, R. J. (2006). Prediction of wolf (Canis lupus) kill-sites using hidden Markov models. Ecological Modelling, 197(1–2), 237–246. https://doi.org/10.1016/j.ecolmodel.2006.02.043

[11] Ditmer, M. A., Wittemyer, G., Zeller, K. A., Breck, S. W., Fletcher, R. J. J., Crooks, K. R., U.S. Government, USFS Rocky Mountain Research Station, Colorado State University, Aldo Leopold Wilderness Research Institute, USDA National Wildlife Research Center, University of Florida. (2023). Predicting dispersal and conflict risk for wolf recolonization in Colorado. Journal of Applied Ecology, 2327–2339. https://doi.org/10.1111/1365-2664.14504

[12] GeeksforGeeks. (2023a, March 2). Gated recurrent unit networks. GeeksforGeeks. https://www.geeksforgeeks.org/gated-recurrent-unit-networks/

[13] Helping Tool for Coding and Redacting : ChatGPT https://chatgpt.com/?oai-dm=1