Preprints.org

Article

Multi-Sensor Integration and Machine Learning for High-Resolution Classification of Herbivore Foraging Behavior

Bashiri Iddy Muzzo [*] , Kelvyn Bladen , Andres Perea , Shelemia Nyamuryekung'e , Juan J. Villalba

Article

Multi-Sensor Integration and Machine Learning for High-Resolution Classification of Herbivore Foraging Behavior

Bashiri Iddy Muzzo [1,*], Kelvyn Bladen [2], Andres Perea [3], Shelemia Nyamuryekung'e [4] and Juan Villalba [1]

[1] Department of Wildland Resources, Quinney College of Natural Resources, Utah State University, Logan, UT 4322-5230, USA

[2] Department of Mathematics and Statistics, Utah State University, 3900 Old Main Hill, Logan, UT 84322, USA

[3] Department of Animal and Range Sciences, New Mexico State University, Las Cruces, NM 88003, USA

[4] Division of Food Production and Society, Norwegian Institute of Bioeconomy Research (NIBIO), PB 115, N-1431 Ås, Norway

[*] Correspondence: iddy.muzzo@usu.edu; Tel.: (+1) 435 938 1269

Simple Summary: The study used Random Test-Split (RTS) and Cross-Validation (CV) machine learning data partition methods to test different models to classify cattle behavior foraging behaviors states, foraging activities, posture, and activity by posture, using GPS coupled accelerometer data with 12-hour / days continuous recording observation as supporting ground truth. RTS in XGBoost performing best for general activity state classification, while CV in Random Forest excelled in more detailed foraging activities and activity-posture classifications. Key movement indicators like speed, Actindex, and sensor values (x, y, and z) were vital in predicting behaviors, suggesting specific sensors for tracking behaviors of interest to ranchers. The results highlight the benefits of continuous monitoring and advanced data analysis for real-time livestock tracking, leading to better grazing management, improved animal welfare, and more sustainable land use.

Abstract: The study classified cows' foraging behaviors using machine learning (ML) models evaluated through Random Test-Split (RTS) and Cross-Validation (CV) data partition methods. Models included Perceptron, Logistic Regression, Support Vector Machine, K-Nearest Neighbors, Random Forest (RF), and XGBoost (XGB). These models classified activity states (Active vs. Static), foraging behaviors (Grazing (GR), Resting (RE), Walking (W), Ruminating (RU)), posture states (Standing up (SU) vs. Lying down (LD)), and activity-by-posture combinations (RU_SU, RU_LD, RE_SU, RE_LD). XGB achieved the highest accuracy for state classification (74.5% RTS, 74.2% CV) and foraging behavior (69.4% CV). RF outperformed XGB in other classifications, including GR, RE, and RU (62.9% CV vs. 56.4% RTS), posture (83.9% CV vs. 79.4% RTS), and activity-by-posture (58.8% CV vs. 56.4% RTS). Key predictors varied: Speed and Actindex were crucial for GR and W when increasing and for RE and RU when decreasing. X low values were linked to RE_SU and RU_SU, while X and Z influenced RE_LD more. RTS showed higher accuracy in behavioral state and general foraging classification. These results emphasize CV in RF's reliability in managing complex behavioral patterns and the importance of continuous recording devices and movement metrics to monitor cattle behavior accurately.

Keywords: random test-split; cross-validation; random forest; XGBoost; foraging behaviors

1. Introduction

Herbivores, particularly livestock, have an important role in the structure and functioning of arid and semiarid grassland ecosystems characterized by low and unpredictable rainfall. These ecosystems, often dominated by grass monocultures, are ecologically vital and yet vulnerable [1]often facing challenges such as limited forage availability, nutritional quality and seasonal variations in plant productivity [2]. As primary consumers, herbivores are central to nutrient cycling, plant community dynamics, and overall ecosystem health. Their foraging behavior, including activities such as grazing, rumination, walking, resting, and drinking water, directly or indirectly influences vegetation structure, biodiversity, and soil health. The foraging activities of herbivores affect nutrient cycling and soil conditions, which are crucial for maintaining healthy plant communities [3]. For instance, optimal grazing can enhance nutrient availability by depositing dung and urine, which enriches the soil and supports plant growth [4]. However, overgrazing can lead to irreversible damage, such as soil erosion and loss of plant species diversity, ultimately decreasing land productivity [5]. Additional activities, such as grooming and nursing, are also important as they contribute to livestock populations' social structure and overall well-being [6], indirectly affecting their productivity and ecological systems. Understanding these behaviors is vital for improving land management strategies, ensuring sustainability, while mitigating the negative impacts of overgrazing and climate change [7]. Traditional methods of monitoring herbivore grazing behavior, such as fixed-point sampling and direct visual observation, are often impractical and labor-intensive, particularly in the vast and remote locations throughout arid and semiarid grasslands.

Bailey et al. [8] recommended the use of precision livestock management (PLM) tools by real-time monitoring and management approaches that use global position system (GPS) tracking, accelerometers, and other sensor technologies in maintaining rangeland health. Nyamuryekung'e [9]) emphasized the use of these animal movement tracking devices to assess grazing patterns and classify foraging behaviors in these challenging environments. While GPS collars provide precise location data, enabling researchers to track animal movements across large areas, additional data such as identifying grazing hotspots across the landscape is significant. For instance, Nyamuryekung'e et al. [10] used collars to provide insights into grazing, walking, and resting times, and other research groups (e.g., Brennan et al. [11]) have even assessed season-long season livestock grazing behavior using low-cost collars. Accelerometers detect changes in movement patterns with specific activities, such as grazing and standing [12,13]. By combining GPS data with accelerometer information, researchers can more accurately infer foraging behaviors and better understand how herbivores utilize the landscape [14]. Bailey et al. [15] suggested that combinations of GPS tracking coupled with accelerometer monitoring may be more accurate than either device used by itself. However, Ganskopp and Johnson [16] pointed out that these devices have accuracy-related challenges like satellite-related errors and loss of satellite reception owing to atmospheric conditions, topography, canopy, and near infrastructure. Bonneau et al. [17] recommended combining telemetry devices with timestamp cameras, while Aquilani et al. [18] and Bailey et al. [15] further emphasized the use of both cameras with machine learning (ML) to enhance the assessment and understanding of livestock grazing behavior. Cameras can effectively capture foraging and mothering behaviors such as grooming and nursing, and social interactions, such as mating or fighting, as well as responses to environmental factors like vegetation changes, water availability, or the presence of predators. However, camera systems are subject to limitations, especially battery life, during continuous recording in free-ranging grazing systems with no power source [19].

Several studies revealed that ML, a collection of powerful data processing and analysis techniques, can be applied to animal behavioral classification based on data collected using wearable sensors [12,20,21]). Integrating GPS, Accelerometer, and camera data requires advanced data analysis techniques, such as ML algorithms, to process large, multimodal datasets and automatically identify complex behavioral patterns. Supervised machine-learning models have been used successfully to improve classification accuracy and effectively manage non-linear relationships in datasets, making them well-suited for distinguishing behaviors [22] like grazing versus non-grazing. For example,

Augustine and Derner [23] studied grazing behaviors in semiarid rangelands using a two-axis accelerometer and GPS data, identifying distance traveled and head-down posture (assessed by the y-axis sensor) as primary indicators. They applied classification and regression tree (CART) analyses, achieving 87.8% accuracy for grazing and 86.5% for non-grazing activities, with an overall misclassification rate of 12.9%. Cabezas et al. (2022) extended these findings by combining GPS and accelerometer data to analyze herd movement patterns and classify behaviors, thus demonstrating the added value of spatial data for sustainable rangeland use. Despite employing coupled devices, their study used different supervised ML methods for accelerometer-based behavior classification and unsupervised k-medoids ML algorithm for GPS data analysis.

In conclusion, a gap remains in integrating coupled GPS-accelerometer data with continuous behavioral observation systems, such as external field cameras, to test various supervised ML models for livestock behavior classification. While most studies have focused on confined systems, this research targets free-grazing environments, where landscape features and forage availability significantly influence cattle behavior and movement patterns. By continuously observing behaviors like grazing, rumination, and resting through external cameras, this study aims to improve the accuracy of activity classification, which is crucial for optimizing grazing management and enhancing animal welfare. Accurate detection of these behaviors will also support better health monitoring by identifying signs of stress or discomfort. Furthermore, emphasizes the application ofadvanced ML data partition techniques, such as data splitting, cross-validation, and different supervised models, is expected to improve classification accuracy. This study hypothesizes that refining these models with more appropriate data partition technique will lead to more reliable predictions, advancing precision livestock monitoring and promoting sustainable grazing management in diverse rangeland environments.

2. Materials and Methods

2.1. Study Areas

This study was conducted at the USU Richmond Research farm (41.9227° N, 111.8136° W); elevation: 1511 m, annual temperature ranges -8.9°C to 15.6°C, annual precipitation of 525.8 mm and snowfall of 177.8 cm [24]. The experimental pasture, delineated by a five-strand barbed wire perimeter fence, entails 222,578.3 square meters of meadow brome grass (*Bromus inermis*) pasture monoculture (~ 2500 Kg/ha; [25]) with an uneven distribution of intermediate wheatgrass (*Thinopyrum intermedium*). A semi-permanent electric fence was also built at the center of the pasture to divide it into 2 equal blocks. Temporary electric fence, perpendicular to this fence divided the pasture into 6 paddocks of 36,421.74 square meters (Figure 1). The study procedures described herein was approved by the Utah State University Institutional Animal Care and Use Committee (approval number 2566).

2.2. Animals, Sensors, and Camera Deployment

The experiment was conducted from July to September 2024. Twenty-four Angus matured mother cows (body weight [BW] 614 ± 20 kg) and their nursing calves (6 to 8 months of age) (BW = 244 ± 4 kg) were randomly assigned to six paddocks of brome grass monoculture (4 pairs/paddock). Of these, 22 cows were randomly fitted with LiteTrack Iridium 750+ GPS collars coupled with triaxial accelerometers (Lotek Engineering, Newmarket, Ontario, Canada). Each collar weighed approximately 900 grams and was designed for a neck size of 50 cm. The collars featured a buckle for easy placement, and the fit was adjusted to allow a person's fingers to fit comfortably between the cow's neck and the collar. The collars were set to collect one GPS position every 5 minutes. The Standard fix was collar configuration, recording 18 positions, which offers greater detail in tracking animal movements. In contrast, the SWIFT configuration, with only 9 positions, still provides adequate data to monitor foraging behavior at the same interval accurately. The reduced number of GPS positions in SWIFT fixes was offset by the method's increased efficiency and lower power

consumption. Data was transmitted via the Iridium satellite network. The collars operated within a frequency range of 148 MHz to 174 MHz for data transmission, ensuring reliable real-time communication with the satellite even in remote locations. They were designed to withstand extreme environmental conditions, with an operating temperature between -40°C to 70°C. The collars captured data such as GMT Time (time of recording), Latitude and Longitude (geographic location), Altitude (elevation above sea level), Temperature (ambient temperature), Voltage (collar battery level), DOP (Dilution of Precision, reflecting GPS data quality), and Satellites (number of satellites used in the GPS fix).

Additionally, these GPS collars have incorporated a tri-axis accelerometer to track movement on three axes: X (left-right), Y (forward-backward), and Z (up-down), recording pendular beats across all axes every 5 minutes, coinciding with the GPS position fixes. The orientation of the collars was controlled during attachment to the animals to ensure consistent and reliable accelerometer readings across all axes, maintaining data consistency. Ground truth observations were conducted through video recordings using three GoPro Hero 12 cameras from August 27 to 29, 2024, following an initial 12-hour training session on August 26. During this session, the cows were closely followed with cameras to familiarize them with the operator and equipment, ensuring their behavior was not affected during the recording period. Three cows were randomly selected as focal animals from the 22 collared cows, representing three of the six paddocks. To facilitate monitoring, these cows were marked on their backs with white paint (Rust-Oleum) for easy visibility from any angle. Three observers, each equipped with one of the GoPro cameras, were assigned to follow one focal cow per paddock. They maintained a 5 to 10-meter distance to the animal to ensure comprehensive video recording during designated periods. Recording sessions were conducted over three consecutive days (August 27–29), with video footage captured using the cameras mounted with GoPro Volta for extended battery life. Each day, the cows were observed during three intervals: morning (7:00 am–11:00 am), afternoon (12:00 pm–4:00 pm), and evening (5:00 pm–8:00 pm). These intervals were chosen to capture the majority of the activities performed by the cows across daylight: grazing, resting, walking, and ruminating behaviors. These are essential components of the cows' activity patterns and ensure that all behaviors are adequately covered.

2.3. Pre-Processing of GPS, Accelerometer, and Camera Data

The experiment collected GPS and accelerometer data for all 22 cow collars (22GAD) (approximately 4 collars/paddock), which were pre-processed to ensure accuracy and consistency before analysis. Data pre-processing, cleaning, and visualization were conducted using ArcGIS Pro version 3.2.2 and R version 4.4.2. Using ArcGIS Pro, 22 GADS (latitude/longitude in degrees) were converted into a point feature class in WGS_1984 using the "XY Table to Point" tool, and points outside the study area were removed by clipping with the "Richmond study shapefile" for all paddocks. The GPS collars had a positional accuracy of approximately 10 meters, which resulted in some data points crossing into adjacent paddocks, especially since the paddocks were close together and separated by an electric fence. This error caused cows' data points to be detected in the wrong paddocks. To mitigate this, the "Buffer" tool in ArcGIS Pro was used to omit any points affected by this error by buffering a 10-meter distance around each corner of the paddocks. After that, the exported cleaned 22 GAD with GPS geographic coordinates (degrees) column was converted to a projected coordinate system of eastern and northern points (meters) in the Utah State geographical UTM 12 northern zone (WGS84) using R software. The timestamps, initially reported in GMT, were also converted to Utah Mountain Time (Zone-12, UTC-6), the local time zone for each GPS-accelerometer data point. We further adopted Liu et al. [26] data cleaning procedure in our 22 GAD by omitting points with a duration of less than 5 minutes, at least 4 satellites used for data recording to estimate the 3 position, and altitude ranges between 1500 and 1570 m based on the description of the study location. The dataset comprising merged GPS and accelerometer data for the observed three animals (3RC) was extracted from the overall 22 GAD dataset. During the cleaning process, only 1% of data points were omitted from the 3RC dataset. The distances traveled, speed, and tri-axis

accelerometer (X: forward-backward, Y: side-side, and Z: up-down) showing movement patterns by a single cow was visualized in R. The data accurately reflected the cows' movements, with high sensor spikes corresponding to active behaviors such as grazing and walking, while low spikes were associated with static behaviors.

Three trained technicians were assigned to watch each cow's videos and label behavior activities in the prepared datasheet. Five exclusive cow behaviors were distinguished: grazing, walking, resting, rumination, and drinking water. During the transcription process, a behavior change was considered when a cow shifted from one activity to another, with the transition occurring at more than a 3-minute interval. Grazing was defined as the cow actively foraging with the head down to the grass with or without leg step movement; walking referred to cow displacement in space without foraging or ruminating; rumination was stationary behavior identified when the cow was observed chewing a regurgitated bolus of feed and continuing to chew until it was swallowed again; and Resting was defined as any period of immobility, either standing or lying down, with minimal movement, apart from head movements or chewing the cud, and when cows were not grazing, walking, or ruminating; Drinking water referred when cow's mouth was over or in the water trough. The cows' resting or ruminating postures were also noted accordingly, as the animal was standing up or lying down. Nursing was defined as the behavior when a calf suckles from its mother, while grooming referred to actions where cows lick, scratch, or use their heads or tongues to interact with themselves or other cows. An animal was labeled as performing a specific activity if it engaged in the behavior for over 3 minutes. During transitions between activities, intervals lasting less than 3 minutes were considered part of the previous activity, while intervals exceeding 3 minutes were classified as a change into a new activity. These behaviors were further categorized into two primary states: active (AC) and static (ST). Active behaviors included walking, grazing, and drinking water, while static behaviors encompassed ruminating and resting. The cow's standing or lying down posture was also recorded for each static behavior. Grooming and nursing were categorized as static behaviors when performed for prolonged periods but classified as "other" when occurring briefly. This structured approach to behavior categorization enabled detailed tracking of each cow's activity patterns and time allocation.

2.4. Feature Calculations

The accelerometer and GPS data were used to derive different features for input in several Machine learning models to classify previously defined behaviors. The features calculated from accelerometer data were Sum_XYZ, calculated by aggregating the values of the X, Y, and Z axes, reflecting the overall intensity of movement, and Avg_XYZ was determined as the average movement values across the three axes. In addition, absolute values of each axis were computed using the absolute value notation ($|X|$, $|Y|$, $|Z|$), allowing for a more precise representation of movement intensity; Sum_XYZ_absolute involved summing the absolute values ($|X| + |Y| + |Z|$) of the X, Y, and Z axes; Avg_Sum_XYZ_absolute was the average of these summed absolute values. Activity Index (Actindex), conceptually resembling the Magnitude of Acceleration, was calculated by taking the square root of the sum of squared values of accelerations across the X, Y, and Z axes. This index measures the cows' activity levels, where higher values indicate more active behaviors and lower values suggest rest periods. Energy expenditure was computed as the cumulative sum of the Actindex, representing the total energy expended per duration of time particular behavior observed. We also adopted additional features from Versluijs et al. [27] , such as overall dynamic body acceleration (ODBA), Vector of Dynamic Body Acceleration (VEDBA), Magnitude of Acceleration (AMAG), Pitch, and Roll.

$$Pitch = artan(\frac{x}{\sqrt{y^2+z^2}}) \quad (1)$$

$$Roll = artan(\frac{y}{\sqrt{x^2+z^2}}) \quad (2)$$

$$ODBA = |dx|+|dy|+|dz| \quad (3)$$

$$AMAG = \sqrt{x^2 + y^2 + z^2} \quad (4)$$

$$VEDBA = \sqrt{dx^2 + dy^2 + dz^2} \quad (5)$$

From the GPS data, we calculated several metrics to analyze animal movement. Distance (meters) represents the distance traveled between two consecutive GPS points. Speed is derived from the distance traveled over time during the observation period, indicating the movement speed of the cow. The Straightness Index was calculated as the ratio of the straight-line distance between the initial recording and subsequent coordinates of a particular movement for each cow and day. Additionally, the Movement Angle (bearing) was determined using an arc-tangent function with respect to a due east trajectory. This calculates the angle (in radians) between the first and last coordinates, and this angle was subsequently converted to degrees. The derived GPS and accelerometer metrics were multiplied to calculate combined sensor features. This approach enabled a more comprehensive analysis of the cows' movement and behavior patterns, providing deeper insights into their foraging behaviors using combined GPS-accelerometer devices. ML models were tested on different performance parameters using all calculated features to classify cattle foraging behaviors.

A random forest tree-based method was used to select features by aggregating trees and providing average feature importance. Following this, a filter method using the correlation coefficient was applied to measure the relationship between selected features and the target variable, helping to avoid feature redundancy. Thereafter, machine learning models were tested on different performance parameters using all calculated features to classify cattle foraging behaviors.

2.5. Machine Learning Data Partition Strategy

The study applied the Random Train-Test Split (RTS) and 5-fold cross-validation (CV) data partition methods to evaluate various ML models for classifying animal behavior and activities. RTS divides the dataset into training and testing subsets, typically using a 70:30 ratio, and is particularly effective in resource-constrained scenarios [28]. However, the split ratio can influence model performance, and multiple splits have been shown to enhance reliability [29]. In contrast, 5-fold CV partitions the data into five equal subsets, rotating the training and testing roles across all folds [30]. This method minimizes overfitting and provides a more robust performance estimate [31]. Despite the strengths of these approaches, challenges in model selection remain. As Cawley and Talbot [32] emphasized, biased performance evaluation can occur when selection criteria are improperly optimized, underscoring the need for careful model selection practices. The study evaluated six ML models for behavior classification, each employing distinct learning approaches: Perceptron adjusts weights based on errors, Logistic Regression predicts probabilities using a sigmoid function, Support Vector Machine (SVM) separates classes with optimal margins [29], K-Nearest Neighbors (KNN) classifies based on the nearest data points, Random Forest (RF) aggregates multiple decision trees for robust predictions, and XGBoost (XGB) enhances accuracy through gradient-boosted trees [30]. To ensure a fair and reliable assessment of both methods and their impact on classification accuracy, the study prioritized using the models with highest accuracy in RTS as the benchmark for evaluating the CV method.

2.6. Models Performance Assessment

The performance of the models was evaluated based on several metrics. Accuracy was measured as the proportion of correctly classified instances (both positive and negative) out of the total instances: (TP+TN)/(TP+TN+FP +FN). Precision measures the proportion of correctly predicted positive instances out of all predicted positives: TP/(TP+FP). Recall estimates the proportion of correctly predicted positive instances out of the total positive instances: TP/(TP+FN). The F1 score combines the precision and recall scores into a single measure, providing a balance between the two: 2 * (Precision * Recall) / (Precision + Recall).

3. Results

3.1. Data Summary and Challenges

The dataset consists of 1,280 recorded instances, with 632 classified as active and 648 as static. The behavioral states were distributed as follows: 535 cases of grazing (GR), 258 of resting (RE), 390 of ruminating (RU), 83 of walking (W), and 14 of drinking water (DW). Resting behavior was further broken down into 99 instances of resting in a lying down (RE_LD) posture and 159 instances of resting in a standing upright (RE_SU) posture. Ruminating behavior was similarly divided into 209 ruminating in a lying down (RU_LD) posture and 148 ruminating in a standing upright (RU_SU) posture. Posture data showed 925 instances of standing upright (SU) and 308 instances of lying down (LD). Behavioral classification was based on GPS movement patterns, accelerometer thresholds, and direct observations. Some challenges occurred during data collection. Three GPS collars apart from the observed cows failed to collect data throughout the experimental period, which reduced the sample size. Camera overheating after 3 to 4 hours due to high summer temperatures disrupted continuous video recording intended for validation.

3.2. Behavior Classification Using Random Train-Test Split Method

3.2.1. States Classification

The performance of different models is shown in Table 1. Various features were included as input to optimize classification accuracy across machine learning models. The features selected X, Y, Z, Actindex, Distance and Speed achieved the highest classification accuracy. Ensemble learning models achieved higher classification rates and fewer misclassifications than simple classifiers. The Perceptron model exhibited the lowest performance among the models tested, with an accuracy of 63.8%, lower than Logistic Regression (72.4%). Similarly, SVM achieved an accuracy of 71.1%, which is 2.9% lower than KNN (74.0%). However, ensemble models outperformed other approaches, with Random Forest achieving 73.2% accuracy, 9.4% higher than Perceptron, along with consistent Static metrics of 74.0%. XGBoost demonstrated the highest performance, achieving 74.2% accuracy, 10.4% higher than Perceptron, and a Static Precision of 72.0%, Static Recall of 79.0%, and a Static F1 Score of 75.0%. These results highlight the effectiveness of ensemble learning models for this classification, particularly XGBoost, in accurately classifying cows' behavioral states.

Table 1. Table presenting performance metrics for each model on test data, including accuracy, precision, recall, F1 score for active and static states using Random test split method (RTS). The best results for a given metric are bolded to highlight which model was optimal for that analysis or task.

Model	Model Accuracy (%)	STATES Active Precision (%)	Recall (%)	F1 Score (%)	STATES Static Precision (%)	Recall (%)	F1 Score (%)
Perceptron	63.8	68	52	59	62	76	68
Logistic Regression	72.4	76	65	70	70	**79**	74
Support Vector	71.1	73	67	70	70	75	72
K-Nearest Neighbor	74	73	**74**	**74**	**74**	74	74
Random Forest	73.2	73	73	73	73	74	74
XGBoost	**74.2**	**77**	69	73	72	**79**	**75**

3.2.2. Activity Classification

The behavior classification results are presented in Table 2. The same features used for state classification were applied, excluding Sum_XYZ after feature selection. Random Forest achieved the highest overall accuracy in general activities, with accuracies of 65.9% for GR (Grazing), 68% for RE (Resting), and 50% for W (Walking). It excelled in GR and RE, but struggled with W. XGBoost followed closely, with 63.3% accuracy for GR, 67% for RE, and 13% for W, facing similar challenges with Walking. SVM showed the highest Precision for W (100%) but had low Recall (4%), indicating it could identify W when present but missed many instances. Logistic Regression performed poorly with W, with an F1 Score of 0%, making it ineffective for classifying W. For fine activities, Random Forest remained the most reliable model, with an overall accuracy of 59.7%, excelling in GR (67%) and RU (Ruminating) (60%) (Table 2). XGBoost had an accuracy of 61.7% for GR and RU, but struggled with RE, similar to its performance with general activities.

In summary, Random Forest was the most consistent and accurate model for both general and fine activities classification. XGBoost showed solid performance but had limitations with W and RE. SVM had high precision for W but poor recall, while Logistic Regression underperformed overall, especially with W.

Table 2. The table presents the overall performance of different ML models: Grazing (GR), Resting (RE), Walking (W) and Ruminating (RU) using Random test split method (RTS). The best results for a given metric are bolded to highlight which model was optimal for that analysis or task.

Classification	Model	Activity	Precision (%)	Recall (%)	F1 Score (%)	Model Accuracy (%)
General activities	Perceptron	GR	50	64	57	
		RE	54	35	42	45.8
		W	4	8	5	
	Logistic Regression	GR	60	53	56	
		RE	62	76	68	61.2
		W	0	0	0	
	SVM	GR	62	55	58	
		RE	63	77	69	
		W	100	4	8	62.5
	K-Nearest Neighbor	GR	55	64	59	
		RE	65	65	65	60.4
		W	100	4	8	
	Random Forest	GR	**63**	64	64	
		RE	**68**	75	71	**65.9**
		W	**50**	4	7	
	XGBoost	GR	63	62	62	63.3
		RE	67	72	69	
		W	13	8	10	
Fine activities	Perceptron	GR	**70**	63	66	
		RE	25	12	16	53.5
		RU	46	68	55	
	Logistic Regression	GR	62	76	68	56.1
		RE	0	0	0	

	RU	49	66	56	
	GR	66	70	68	
SVM	RE	**53**	12	19	58
	RU	50	**72**	59	
K-Nearest Neighbor	GR	64	71	67	
	RE	37	30	33	54.9
	RU	50	50	50	
Random Forest	GR	67	73	70	
	RE	38	30	34	59.7
	RU	**60**	62	**61**	
XGBoost	GR	67	**78**	72	
	RE	46	**31**	**37**	**61.7**
	RU	59	59	59	

3.3. Behavior Classification Using Cross-Validation Method

3.3.1. States Classification

We adopted a random 5-fold cross-validation technique to assess performance in activity classification. We used features selected from above to identify consistent predictors across methods for classifying states as Static (ST) or Active (AC). Performance metrics showed that XGBoost outperformed other models with 74.2% accuracy, followed by Random Forest at 73.2%, while Logistic Regression had the lowest accuracy at 69% and a 31% misclassification rate. XGBoost hyperparameters included 200 rounds, eta of 0.06, maximum depth of 4, colsample_bytree and min_child_weight of 0.5, and subsample of 0.8. Partial Dependence Plots (PDPs) revealed that Speed and Actindex were the strongest predictors of AC, with Actindex increasing sharply up to 30 and stabilizing, and Speed rising until 50 metres/min before plateauing (Figure 3). X showed a slight positive effect, Y displayed a negative sigmoidal trend, and Z had a positive sigmoidal pattern. These results confirmed that Speed, Actindex, and X were strongly associated with AC, with XGBoost delivering the most precise and confident predictions compared to the broader distributions in Logistic Regression and Random Forest.
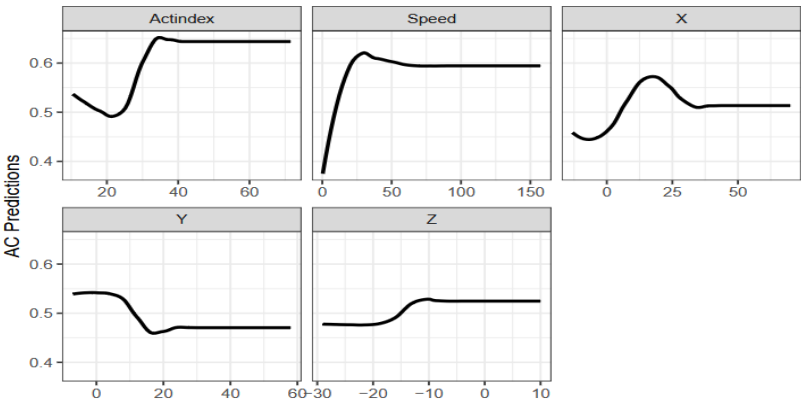

Figure 3. Partial Dependence Plots (PDPs) for predicting active states in cows using different features: X, Y, Z, Speed (metre/minutes), and Actindex.

3.3.2. Activity Classification and by Posture

Similar to the Random Split (RS) method, we initially tested the classification of grazing (GR), resting (RE), and walking (W) using a 5-fold cross-validation approach. The features selected were consistent with those identified in the RS method, suggesting a stable pattern across both techniques

of which the same set of features were used in activity classification and by posture. Since XGBoost and Random Forest were found to perform better, we focused on these models for fine behaviors classification, including posture-based classification in the cross-validation method. After fine-tuning, XGBoost slightly outperformed Random Forest, achieving 69.38% overall accuracy compared to 68.51%. Both models performed well for grazing (GR, ~67%) and resting (RE, ~77%), but struggled with walking (W), where XGBoost (48.3%) outperformed Random Forest (22.2%) (Table 3). After reclassifying behaviors into GR, rumination (RU), and RE, accuracy dropped to 62.38% (RF) and 60.35% (XGB). For precision, Random Forest performed better for RE (47.1% vs. 36.5%), while XGBoost had a slight edge for GR (67.2% vs. 66.6%) and RU (58.9% vs. 59.6%). Both models performed well for posture classification (standing vs. lying down), with Random Forest at 83.94% and XGBoost at 83.7%. RF excelled in LD (79.9%), while XGB performed slightly better in SU (85.1%). For activity-by-posture classification, both models had ~58.8% accuracy. XGBoost was better at detecting RE_LD (47.6%), while Random Forest performed better for RE_SU (46.2%).

Table 3. The table presents the overall performance of different machine learning models: Grazing (GR), Resting (RE), Walking (W) and Ruminating (RU) and Resting in Lying (RE_LD), Resting in Standing (RE_SU), Ruminating in Lying (RU_LD), and Ruminating in Standing (RU_SU) using Cross-validation method (CV). The best results for a given metric are bolded to highlight which model was optimal for that analysis or task.

Classification	Method	Model accuracy (%)	Behaviors	Precision (%)	Recall (%)	F1 (%)
General activities	Random Forest	68.51	GR	65.2	67.9	66.5
			RE	71.8	77.5	74.5
			W	22.2	2.4	4.3
	XGBoost	**69.38**	GR	67.1	67.2	67.2
			RE	72	77.9	74.9
			W	48.3	16.9	25
Fine activities	Random Forest	**62.38**	GR	66.6	80.9	73.1
			RE	47.1	18.6	26.7
			RU	59.6	65.9	62.6
	XGBoost	60.35	GR	67.2	76.4	71.5
			RE	36.5	20.9	26.6
			RU	58.9	64.4	61.5
Posture	Random forest	**83.94**	LD	79.9	47.7	59.8
			SU	84.7	96	90
	XGBoost	83.7	LD	76.4	50.3	60.7
			SU	85.1	94.8	89.7
Activities by posture	Random Forest	**58.87**	RE_LD	30.6	15.2	20.3
			RE_SU	46.2	34	39.1
			RU_LD	50.9	52.6	51.8
			RU_SU	52.2	39.9	45.2
	XGBoost	58.78	RE_LD	47.6	10.1	16.7
			RE_SU	43.1	13.8	21
			RU_LD	52.3	55.5	53.8
			RU_SU	64.4	25.7	36.7

The Partial Dependence Plots (PDPs) show that Actindex and Speed are strong predictors for GR and W, with a sharp increase before stabilizing (Figure 4a). They decrease consistently for RE and RU, indicating a negative effect. X increases sharply for GR and plays a key role for RE and RU when it decreases (Figure 4b). Y shows more variability but is slightly linked to RU. Z drops sharply for RE and increases for GR. For posture classification (Figure 4c), Speed and X are the strongest predictors of standing up (SU) over lying down (LD), with Speed increasing sharply up to 50 before leveling off and X maintaining a positive trend. In posture-by-activity classification, Speed is a strong predictor of resting (RE_LD, RE_SU) and rumination behaviors (RU_LD, RU_SU), with lower values indicating these states. X and Z are secondary predictors, with higher X values linked to RE_LD (Figure 4d). In summary, Speed is a key predictor for GR and W when they increase, and for RE and RU when they decrease. X is strong for GR when it increases, while both X and Y predict RE and RU when they decrease. Z predicts RE with a sharp drop and positively influences GR when it increases. Speed and X are the best predictors for SU over LD, emphasizing their role in classifying foraging behaviors.
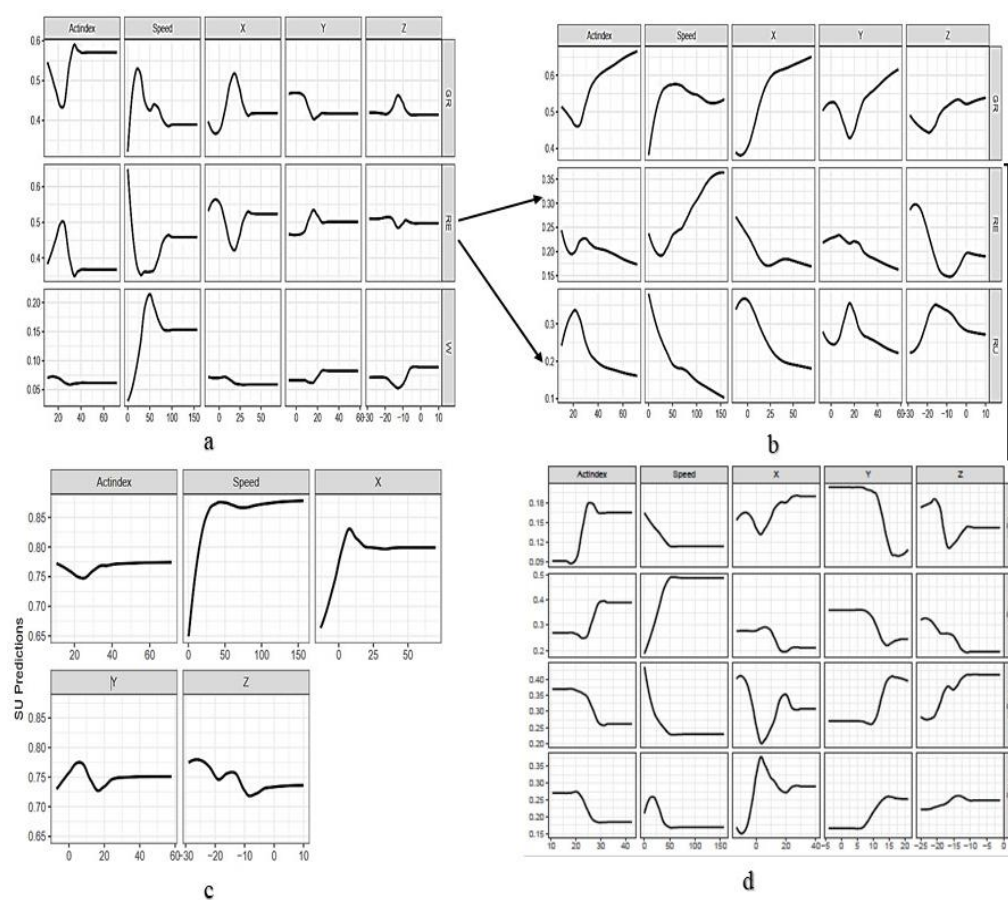

Figure 4. Partial Dependence Plots (PDPs) showing predicted features for (a) general activities Grazing (GR), Resting (RE), and Walking (W); (b) fine activities GR, RE, and Rumination (RU); (c) posture Standing up (SU) versus Lying down (LD); and (d) activity by posture RE_LD (resting lying down), RE_SU (resting standing up), RU_LD (rumination lying down), and RU_SU (rumination standing). Features include speed (meters per minute), Actindex, and X, Y, and Z axes, analyzed using the CV method.

4. Discussion

4.1. State Classification

The classification of cows' behavioral states provided key insights into the performance of the methods tested across various machine learning models. For classifying cows' behavioral states (Active vs. Static), XGBoost achieved the highest accuracy, with 74.5% for RTS and 74.2% for CV. This consistency highlights the model's robustness and aligns with prior findings by Putun and Yilmaz

[33] and Ibrahim et al. [34], who noted XGBoost's exceptional ability to handle complex datasets and accurately classify animal behaviors. Similarly, Random Forest also performed well, improving accuracy from 73.2% in RTS to 74.1% in CV. These results are consistent with the observations of Chakraborty et al. [35] and Wyner et al. [36], who highlighted the model's adaptability and generalization capabilities across diverse datasets. However, despite the strong performance of ensemble models like XGBoost and Random Forest, challenges remain regarding scalability and robustness. Expanding trials to include more observations (cows) and integrating diverse data sources could further enhance these models' performance [37] [38]Simpler models, such as Logistic Regression, were more prone to underfitting. This finding is evident in many PDPs in our present study, which reveal complex relationships between predictors and the response that a sigmoidal function cannot fully capture. Overall, ensemble models such as XGBoost and Random Forest consistently outperformed the simpler classifiers, showcasing their effectiveness in handling complex datasets with higher accuracy and fewer misclassifications.

Partial Dependence Plots (PDPs) revealed the relationships between key features and cows' activity states, with Speed and Actindex emerging as the most significant predictors of active behavior. These findings align with Pütün and Yılmaz [33] and Mladenova et al. [39], who similarly identified these metrics as reliable indicators of cow activity. Additional X, Y, and Z features showed varying effects. Feature X exhibited a slight positive influence, Y displayed a weak negative trend, and Z followed a positive sigmoidal trajectory. Although not as strong as Speed and Actindex, these features collectively contribute to the overall prediction accuracy [33]. Tracking Speed and Actindex allow ranchers to identify active cows effectively while incorporating additional features like X, Y, and Z, which can further refine prediction accuracy. These insights offer actionable data for improving herd management by identifying key behavioral patterns associated with specific features [40]. Furthermore, the results underscore the value of non-parametric ensemble learning models like XGBoost, which demonstrated superior accuracy and robustness. By integrating ML models with meaningful feature selection, ranchers can develop more accurate and efficient livestock monitoring systems, ultimately enhancing animal behavior tracking and herd management practices [41]

4.2. Activity Classification

When comparing the performance of Random Forest (RF) and XGBoost (XGB) across different classification tasks, a clear distinction emerges between the cross-validation (CV) and random test split (RTS) methods. In the Grazing (GR), Resting (RE), and Walking (W) classification task, the performance of XGBoost and Random Forest varied depending on the evaluation method. XGBoost slightly outperformed Random Forest using the cross-validation method, achieving an overall accuracy of 69.38%. However, in the random split method, Random Forest achieved a higher overall accuracy of 65.9%, compared to XGBoost's 63.3%. Both models faced challenges with the Walking (W) classification, but XGBoost achieved a higher accuracy of 48.3%, compared to RF's 22.2%. This variability in performance aligns with findings from various studies that highlight the context-dependent nature of these algorithms. For instance, animal foraging activities can vary due to factors such as forage availability and climate. Walking (W) may be less prevalent in monoculture grassland systems than other behaviors, potentially leading to highly imbalanced data. The finding is supported by Wang [42] who found that XGBoost outperformed Random Forest by better handling class imbalances and complex data structures, achieving higher accuracy. Additionally, XGBoost is particularly effective on structured, medium-sized datasets, as evidenced by its performance in telecommunications customer data analysis [43]. The results align with the Grazing, Resting, and Walking tasks, where the dataset's structure and size may have favored XGBoost in the cross-validation method.

Random Forest (RF) is known for its robustness and ability to handle medium datasets, which is particularly beneficial for classifying complex behaviors, such as those of cattle. When the behavioral activities were reclassified into Grazing (GR), Ruminating (RU), and Resting (RE), cross-validation with Random Forest achieved an overall higher accuracy compared to XGBoost (Table 3).

Regarding precision, Random Forest showed higher precision for GR and RU than RE, indicating its effectiveness in distinguishing between active behaviors like grazing and ruminating but less so for resting, consistent with the findings of Pütün and Yılmaz [44]. The algorithm's overall accuracy was based on 5-minute fixed-sensor data from three cows, continuously recorded over three days (48 hours). Several studies on foraging behavior observation in long-range scenarios have reported higher accuracy, suggesting that integrating multi-sensor systems and ML for high-resolution data, as seen in continuous observations like the present study, could lead to even higher accuracy. Accuracy would likely increase further with extended observation periods over a week or two. Even in a confined dairy experiment conducted by Chen et al.[45] , continuous observation provided a more precise representation of cattle behavior compared to fixed-interval sampling. Their study also highlighted that longer observation periods enhance accuracy, as behaviors fluctuate over time due to environmental factors, resource availability, and individual variability. After removing Walking (W) and Drinking Water (DW), the focus shifted to the more distinguishable behaviors: Grazing (GR), Resting (RE), and Ruminating (RU). RF continued to show superior performance in this updated classification, achieving higher accuracy compared to XGBoost's (Tables 2 and 3). The exclusion of Walking and Drinking Water allowed for better differentiation between the remaining activities, reducing misclassifications. This improved performance highlights RF's ability to better handle the task of distinguishing between complex behaviors like Resting (RE), Grazing (GR), and Ruminating (RU), as seen in studies such as Wang et al. [23]. Several studies have achieved notable accuracy in classifying grazing and non-grazing behaviors. Augustine and Derner [23] reported 87.8% accuracy by observing 5 to 9 cows during the summers of 2008 and 2011, with GPS position recorded at 5-minute intervals and tri-axis accelerometers detecting 255 movements per interval. Similarly, Brennan et al. [11] achieved an 11.2% misclassification rate by observing 2-3 cows from 2016 to 2018 in South Dakota by using low-cost homemade GPS collars equipped with high-frequency 3-axis accelerometers, recording GPS fixes at 1-minute intervals and accelerometer data at 12 Hz. While these studies relied on intermittent observations over several years with extensive ground truth and high-frequency data, they are resource-intensive and less feasible for routine ranching. In contrast, the present study used continuous observations of three cows for 12 hours per day over three days, with GPS and accelerometer data recorded at 5-minute intervals. This approach achieved a classification accuracy of 74.5%, offering a more cost-effective and practical alternative with longer battery life and affordable equipment for ranching operations. However, slightly less accurate, continuous observation at fixed intervals provides a solid foundation for real-time behavioral monitoring. Extending observation periods or increasing the sample size, combined with multi-sensor integration and ML, could further enhance the accuracy of foraging behavior classification, bridging the gap between resource-intensive research and practical applications for ranchers.

In the posture Classification task, where the model classified whether the animal was Standing or Lying Down, RF again outperformed XGBoost, achieving high accuracy compared to XGBoost's slightly lower accuracy. This result emphasizes RF's strength in handling complex feature interactions, particularly in classification tasks that involve posture data. These findings align with Biau [46] who showed that RF excels in non-linear classification tasks, such as those involving subtle distinctions between postures and activities. Kleanthous al. [47] similarly found that RF performed well in tasks involving continuous sensor data with subtle distinctions, such as determining an animal's posture. RF demonstrated superior performance across all sub-categories in the activity by posture classification task. RF achieved high accuracy for grazing (GR), while XGBoost achieved slightly lower accuracy. However, both models struggled to classify Resting (RE) when lying down (RE_LD) and standing (RE_SU), with RF achieving only moderate accuracy in these sub-categories. XGBoost performed slightly better in these sub-categories but still lagged behind RF in overall accuracy for this task. These findings are consistent with Chen and Guestrin [48] who noted that XGBoost performs well with balanced datasets but struggles with imbalanced class distributions, particularly in tasks involving fine-grained categories, such as posture-based activities. In contrast, RF's ability to handle multiple imbalanced categories makes it more robust. In conclusion, while both

RF and XGBoost faced challenges, particularly with the classification of Walking (W) and Drinking Water (DW), Random Forest (RF) consistently outperformed XGBoost across most tasks, especially in more complex classifications like posture Classification and Activity by Posture. RF's ability to manage intricate feature interactions, particularly in tasks involving multiple categories or posture data, gives it a distinct advantage. XGBoost, although slightly more accurate in the Random Split Method after excluding problematic activities, still lagged behind RF in most tasks, particularly in posture and activity-by-posture classification. These results align with previous studies, which suggest that RF is better suited for non-linear feature relationships, while XGBoost excels in linear tasks but struggles with high-dimensional and overlapping datasets. Additionally, the continuous nature of the observation data, combined with GPS and accelerometer sensors that record at 5-minute intervals, could provide similar accuracy while extending battery life and reducing device costs. This makes it more feasible for long-term use in foraging behavior classification and addresses the challenges posed by highly variable and overlapping feature patterns. Understanding animal foraging behaviors becomes crucial for sustaining rangeland ecosystem health in arid and semi-arid regions, which are sensitive to disturbances such as overgrazing and desertification.

The Partial Dependence Plots (PDPs) revealed key features influencing animal behaviors. For Grazing (GR), speed was the most significant predictor, with grazing probability decreasing as speed increased. This finding supports previous research by Jia et al. (2018), who noted that grazing cattle typically exhibit confined movement patterns. The Y-axis showed strong positive correlations, indicating lateral movements characteristic of grazing, while the Z-axis displayed a complex trend, reflecting the vertical head and neck motions during grazing. The findings are consistent with those of Sivakumar et al. [28], who observed similar head and neck movements during grazing. For Resting (RE), speed and Actindex were the main predictors. Low speed and a stable Actindex indicated minimal movement, which aligns with findings from Liu et al. [22], who reported that resting periods in cattle are marked by low variability in movement data, showing minimal activity. In Ruminating (RU), Actindex was the most influential predictor, capturing repetitive activity patterns. The Y-axis showed moderate importance, with lateral movements correlating to rumination, and the Z-axis showed a pattern of vertical head motions, consistent with findings by Tamura et al. [49]. Additionally, speed played a moderate role, with limited but structured movement, while the X-axis contributed spatial movement data, supporting Zhang et al. [50], who found that the X-axis is likely to relevant for sheep ruminating classification. For activity by posture Classification, the X-axis, and mean speed were key features in distinguishing between lying and standing postures. Both the X and Z axes were critical for predicting posture, confirming the findings of Liu et al. [26], who highlighted the importance of accelerometer data for posture classification. Rasmussen et al. (2018) also emphasized the effectiveness of using multiple axes for posture detection.

5. Conclusions

This study examined the use of Random Test-Split (RTS) and Cross-Validation (CV) data partition methods to test different models to classify cattle behavior foraging behaviors states, foraging activities, posture, and activity by posture based based on GPS and accelerometer data. XGBoost outperformed Random Forest (RF) overall state classification either active or static. RF under 5-fold CV excelled in tasks like Grazing, Resting, and Walking (GR, RE, W) as well in more complex classifications, such as posture and activity-by-posture classification. RF's ability to manage intricate feature interactions and imbalanced class distributions, especially in Walking (W) and Resting (RE), made it more effective for these tasks. Partial Dependence Plots (PDPs) highlighted key features like Speed and Actindex which were crucial for predicting active behaviors and grazing patterns, supporting findings from previous research. The continuous sensor data collected at 5-minute intervals over multiple days provided higher temporal resolution, capturing short-term behavior changes often missed in intermittent data collection methods. This approach improved classification accuracy and offered a cost-effective, practical solution for real-time livestock monitoring. In contrast to multi-year, intermittent observation studies, continuous data collection

allows for more detailed insights into behavioral shifts throughout the day. The ability to track behaviors like grazing, resting, and walking in real time makes this approach ideal for long-term monitoring and practical applications in livestock management. The integration of machine learning models with continuous behavioral data has the potential to revolutionize livestock management systems. Offering real-time devices and accurate insights enables better decision-making for ranchers, improving herd management and farm productivity. Future research should focus on scaling these methods, incorporating diverse data sources, and evaluating the long-term effectiveness of continuous monitoring for comprehensive livestock management.

Author Contributions: B.I.M.: Conceptualization, investigation, data curation, writing—original draft preparation, review and editing, K.B.: data curation, Analysis, Review and editing original draft. A.P.: data curation, Analysis, Review and editing original draft, S.N.: data curation, Analysis, Review and editing original draft, J.J.V.: Supervisor, Review and editing original draft. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Strömberg, C.A.E.; Staver, A.C. The History and Challenge of Grassy Biomes. *Science (1979)* 2022, *377*, 592–593, doi:10.1126/SCIENCE.ADD1347.

2. Giridhar, K.; Samireddypalle, A.; Giridhar, K.; Samireddypalle, A. Impact of Climate Change on Forage Availability for Livestock. *Climate Change Impact on Livestock: Adaptation and Mitigation* 2015, 97–112, doi:10.1007/978-81-322-2265-1_7.

3. Teague, R.; Kreuter, U. Managing Grazing to Restore Soil Health, Ecosystem Function, and Ecosystem Services. *Front Sustain Food Syst* 2020, *4*, 534187, doi:10.3389/FSUFS.2020.534187/BIBTEX.

4. Bilotta, G.S.; Brazier, R.E.; Haygarth, P.M. The Impacts of Grazing Animals on the Quality of Soils, Vegetation, and Surface Waters in Intensively Managed Grasslands. *Advances in Agronomy* 2007, *94*, 237–280, doi:10.1016/S0065-2113(06)94006-1.

5. Papanastasis, V.P. Restoration of Degraded Grazing Lands through Grazing Management: Can It Work? *Restor Ecol* 2009, *17*, 441–445, doi:10.1111/J.1526-100X.2009.00567.X.

6. Sanz, A.; Blanco-Penedo, I.; Quintans, G.; Álvarez-Rodríguez, J. Mother-Offspring Bonding Revisited: A Blueprint for the Future of Beef Cattle Farming. *Appl Anim Behav Sci* 2024, *277*, 106346, doi:10.1016/J.APPLANIM.2024.106346.

7. Derner, J.D.; Hunt, L.; Filho, K.E.; Ritten, J.; Capper, J.; Han, G. Livestock Production Systems. 2017, 347–372, doi:10.1007/978-3-319-46709-2_10.

8. Bailey, D.W.; Trotter, M.G.; Knight, C.W.; Thomas, M.G. Use of GPS Tracking Collars and Accelerometers for Rangeland Livestock Production Research 1., doi:10.1093/tas/txx006.

9. Nyamuryekung'e, S. Transforming Ranching: Precision Livestock Management in the Internet of Things Era. *Rangelands* 2024, *46*, 13–22, doi:10.1016/J.RALA.2023.10.002.

10. Nyamuryekung'e, S.; Cibils, A.F.; Estell, R.E.; VanLeeuwen, D.; Spiegal, S.; Steele, C.; González, A.L.; McIntosh, M.M.; Gong, Q.; Cao, H. Movement, Activity, and Landscape Use Patterns of Heritage and Commercial Beef Cows Grazing Chihuahuan Desert Rangeland. *J Arid Environ* 2022, *199*, 104704, doi:10.1016/J.JARIDENV.2021.104704.

11. Brennan, J.; Johnson, P.; Olson, K. Classifying Season Long Livestock Grazing Behavior with the Use of a Low-Cost GPS and Accelerometer. *Comput Electron Agric* 2021, *181*, 105957, doi:10.1016/J.COMPAG.2020.105957.

12. Cabezas, J.; Yubero, R.; Visitación, B.; Navarro-García, J.; Algar, M.J.; Cano, E.L.; Ortega, F. Analysis of Accelerometer and GPS Data for Cattle Behaviour Identification and Anomalous Events Detection. *Entropy 2022, Vol. 24, Page 336* 2022, *24*, 336, doi:10.3390/E24030336.

13. Arcidiacono, C.; Porto, S.M.C.; Mancino, M.; Cascone, G. Development of a Threshold-Based Classifier for Real-Time Recognition of Cow Feeding and Standing Behavioural Activities from Accelerometer Data. *Comput Electron Agric* 2017, *134*, 124–134, doi:10.1016/J.COMPAG.2017.01.021.

14. Hounslow, J.L.; Fossette, S.; Chong, W.; Bali, R.; Tucker, A.D.; Whiting, S.D.; Gleiss, A.C. Behaviour-Specific Spatiotemporal Patterns of Habitat Use by Sea Turtles Revealed Using Biologging and Supervised Machine Learning. *Journal of Applied Ecology* 2023, *60*, 1828–1840, doi:10.1111/1365-2664.14438.

15. Bailey, D.W.; Trotter, M.G.; Tobin, C.; Thomas, M.G. Opportunities to Apply Precision Livestock Management on Rangelands. *Front Sustain Food Syst* 2021, *5*, 611915, doi:10.3389/FSUFS.2021.611915/BIBTEX.

16. Ganskopp, D.C.; Johnson, D.D. GPS Error in Studies Addressing Animal Movements and Activities. *Rangel Ecol Manag* 2007, *60*, 350–358, doi:10.2111/1551-5028(2007)60[350:geisaa]2.0.co;2.

17. Bonneau, M.; Vayssade, J.A.; Troupe, W.; Arquet, R. Outdoor Animal Tracking Combining Neural Network and Time-Lapse Cameras. *Comput Electron Agric* 2020, *168*, doi:10.1016/j.compag.2019.105150.

18. Aquilani, C.; Confessore, A.; Bozzi, R.; Sirtori, F.; Pugliese, C. Review: Precision Livestock Farming Technologies in Pasture-Based Livestock Systems. *Animal* 2022, *16*, 100429, doi:10.1016/J.ANIMAL.2021.100429.

19. Herlin, A.; Brunberg, E.; Hultgren, J.; Högberg, N.; Rydberg, A.; Skarin, A.; Siniscalchi, M. Animal Welfare Implications of Digital Tools for Monitoring and Management of Cattle and Sheep on Pasture. 2021, *11*, 829, doi:10.3390/ani11030829.

20. Russel, N.S.; Selvaraj, A. Decoding Cow Behavior Patterns from Accelerometer Data Using Deep Learning. *Journal of Veterinary Behavior* 2024, *74*, 68–78, doi:10.1016/J.JVEB.2024.06.005.

21. Mao, A.; Huang, E.; Wang, X.; Liu, K. Deep Learning-Based Animal Activity Recognition with Wearable Sensors: Overview, Challenges, and Future Directions. *Comput Electron Agric* 2023, *211*, 108043, doi:10.1016/J.COMPAG.2023.108043.

22. Liu, J.; Bailey, D.W.; Cao, H.; Son, T.C.; Tobin, C.T. Development of a Novel Classification Approach for Cow Behavior Analysis Using Tracking Data and Unsupervised Machine Learning Techniques. *Sensors* 2024, *24*, doi:10.3390/s24134067.

23. Augustine, D.J.; Derner, J.D. Assessing Herbivore Foraging Behavior with GPS Collars in a Semiarid Grassland. *Sensors (Switzerland)* 2013, *13*, 3711–3723, doi:10.3390/S130303711.

24. Climate United States - Monthly Averages Available online: https://www.usclimatedata.com/ (accessed on 11 February 2025).

25. DigitalCommons, U.; All Graduate Theses, U.; Decker, T. Targeted-Grazing as a Fuels Reduction Treatment: Evaluation of Targeted-Grazing as a Fuels Reduction Treatment: Evaluation of Vegetation Dynamics and Utilization Levels Vegetation Dynamics and Utilization Levels. 2018.

26. Liu, Z.; Lin, Y.; Hoover, J.; Beene, D.; Charley, P.H.; Singer, N. Individual Level Spatial-Temporal Modelling of Exposure Potential of Livestock in the Cove Wash Watershed, Arizona. *Ann GIS* 2023, *29*, 87–107, doi:10.1080/19475683.2022.2075935/ASSET/6DE89D3D-E04E-42E8-B217-06996D17C3BF/ASSETS/IMAGES/TAGI_A_2075935_F0011_OC.JPG.

27. Versluijs, E.; Niccolai, L.J.; Spedener, M.; Zimmermann, B.; Hessle, A.; Tofastrud, M.; Devineau, O.; Evans, A.L. Classification of Behaviors of Free-Ranging Cattle Using Accelerometry Signatures Collected by Virtual Fence Collars. *Frontiers in Animal Science* 2023, *4*, 1083272, doi:10.3389/FANIM.2023.1083272/BIBTEX.

28. Sivakumar, M.; Parthasarathy, S.; Padmapriya, T. Trade-off between Training and Testing Ratio in Machine Learning for Medical Image Processing. *PeerJ Comput Sci* 2024, *10*, e2245, doi:10.7717/PEERJ-CS.2245/SUPP-6.

29. Jain, E.; Neeraja, J.; Banerjee, B.; Ghosh, P. A Diagnostic Approach to Assess the Quality of Data Splitting in Machine Learning. 2022.

30. Li, Y.; Shu, H.; Bindelle, J.; Xu, B.; Zhang, W.; Jin, Z.; Guo, L.; Wang, W. Classification and Analysis of Multiple Cattle Unitary Behaviors and Movements Based on Machine Learning Methods. *Animals* 2022, *12*, doi:10.3390/ani12091060.

31. Ghojogh, B.; Ca, B.; Crowley, M.; Ca, M. The Theory Behind Overfitting, Cross Validation, Regularization, Bagging, and Boosting: Tutorial. 2019.

32. Cawley, G.C.; Talbot, N.L.C. On Over-Fitting in Model Selection and Subsequent Selection Bias in Performance Evaluation. *Journal of Machine Learning Research* 2010, *11*, 2079–2107.

33. Putun, A.; Yilmaz, D. Classification of Cattle Behavior Leveraging Accelerometer Data and Machine Learning. *8th International Artificial Intelligence and Data Processing Symposium, IDAP 2024* 2024, doi:10.1109/IDAP64064.2024.10710843.

34. Ibrahim, T.; Isaac, K.B.; Francis, B.; Lule, E.; Hellen, N.; Chongomweru, H.; Marvin, G. Interpretable Machine Learning Techniques for Predictive Cattle Behavior Monitoring. *2nd International Conference on Sustainable Computing and Smart Systems, ICSCSS 2024 - Proceedings* 2024, 1219–1224, doi:10.1109/ICSCSS60660.2024.10625182.

35. Chakraborty, R.; Pramanik, A.; Majumdar, S.; Ghosh, D.; Layek, S.; Ray, H.; Ghosh, A.; Singh, O.K.; Ghosh, T.K. Performance Analysis of Machine Learning Classifier on Different Cow Health Datasets. *Proceedings of the International Conference on Sensing Technology, ICST* 2023, doi:10.1109/ICST59744.2023.10460776.

36. Wyner, A.J.; Olson, M.; Bleich, J.; Mease, D. Explaining the Success of AdaBoost and Random Forests as Interpolating Classifiers. *Journal of Machine Learning Research* 2017, *18*, 1–33.

37. Najafabadi, M.M.; Villanustre, F.; Khoshgoftaar, T.M.; Seliya, N.; Wald, R.; Muharemagic, E. Deep Learning Applications and Challenges in Big Data Analytics. *J Big Data* 2015, *2*, 1–21, doi:10.1186/S40537-014-0007-7/METRICS.

38. Hey, T.; Butler, K.; Jackson, S.; Thiyagalingam, J. Machine Learning and Big Scientific Data. *Philosophical Transactions of the Royal Society A* 2020, *378*, doi:10.1098/RSTA.2019.0054.

39. Mladenova, T.; Valova, I.; Evstatiev, B.; Valov, N.; Varlyakov, I.; Markov, T.; Stoycheva, S.; Mondeshka, L.; Markov, N. Evaluation of the Efficiency of Machine Learning Algorithms for Identification of Cattle Behavior Using Accelerometer and Gyroscope Data. *AgriEngineering 2024, Vol. 6, Pages 2179-2197* 2024, *6*, 2179–2197, doi:10.3390/AGRIENGINEERING6030128.

40. Oshana Iddi Dissanayake, A.; McPherson, S.E.; Allyndrée, J.; Kennedy, E.; Cunningham, P.; Riaboff, L. Accelerometer-Based Multivariate Time-Series Dataset for Calf Behavior Classification. 2024, doi:10.5281/zenodo.13259482.

41. Deepak, D.; D'Mello, D.A.; Divakarla, U. Advancements in Automated Livestock Monitoring: A Concise Review of Deep Learning-Based Cattle Activity Recognition. *10th International Conference on Advanced Computing and Communication Systems, ICACCS 2024* 2024, 321–327, doi:10.1109/ICACCS60874.2024.10717327.

42. Wang, Y.; Yu, R.; Li, X.; Chen, R.; Liu, J. Grazing Regime Rather than Grazing Intensity Affect the Foraging Behavior of Cattle. *Ecol Inform* 2025, *85*, 102996, doi:10.1016/J.ECOINF.2025.102996.

43. Grinsztajn, L.; Oyallon, E.; Varoquaux, G. Why Do Tree-Based Models Still Outperform Deep Learning on Tabular Data? 2022.

44. Putun, A.; Yilmaz, D. Classification of Cattle Behavior Leveraging Accelerometer Data and Machine Learning. *8th International Artificial Intelligence and Data Processing Symposium, IDAP 2024* 2024, doi:10.1109/IDAP64064.2024.10710843.

45. Chen, J.M.; Schütz, K.E.; Tucker, C.B. Technical Note: Comparison of Instantaneous Sampling and Continuous Observation of Dairy Cattle Behavior in Freestall Housing. *J Dairy Sci* 2016, *99*, 8341–8346, doi:10.3168/JDS.2016-11351.

46. Biau, G.; Fr, G.B. Analysis of a Random Forests Model. *Journal of Machine Learning Research* 2012, *13*, 1063–1095.

47. Kleanthous, N.; Hussain, A.; Khan, W.; Sneddon, J.; Mason, A. Feature Extraction and Random Forest to Identify Sheep Behavior from Accelerometer Data. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 2020, *12465 LNAI*, 408–419, doi:10.1007/978-3-030-60796-8_35/TABLES/1

48.   Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System., doi:10.1145/2939672.2939785.

49.   Tamura, T.; Okubo, Y.; Deguchi, Y.; Koshikawa, S.; Takahashi, M.; Chida, Y.; Okada, K. Dairy Cattle Behavior Classifications Based on Decision Tree Learning Using 3-Axis Neck-Mounted Accelerometers. *Animal Science Journal* 2019, *90*, 589–596, doi:10.1111/ASJ.13184.

50.   Zhang, H.; Ma, Y.; Shi, Y.; Qiao, Y.; Wang, M. A Framework for Automatic Analysis of Sheep Ruminating Behavior Based on Keypoint Detection. *M2VIP - Proceedings of the International Conference on Mechatronics and Machine Vision in Practice* 2024, doi:10.1109/M2VIP62491.2024.10746109.