Predicting moose behaviors from tri-axial accelerometer data using a supervised classification algorithm

Theresa M. Kirchner[1*] , Olivier Devineau[1], Marianna Chimienti[2], Daniel P. Thompson[3], John Crouse[3], Alina L. Evans[1], Barbara Zimmermann[1] and Ane Eriksen[1]

Abstract

Background Monitoring the behavior of wild animals in situ can improve our understanding of how their behavior is related to their habitat and affected by disturbances and changes in their environment. Moose (*Alces alces*) are keystone species in their boreal habitats, where they are facing environmental changes and disturbances from human activities. How these potential stressors can impact individuals and populations is unclear, in part due to our limited knowledge of the physiology and behavior of moose and how individuals can compensate for stress and disturbances they experience. We collected data from collar-mounted fine-scale tri-axial accelerometers deployed on captive moose in combination with detailed behavioral observations to train a random forest supervised classification algorithm to classify moose accelerometer data into discrete behaviors. To investigate the generalizability of our model to collared new individuals, we quantified the variation in classification performance among individuals.

Results Our machine learning model successfully classified 3-s accelerometer data intervals from 12 Alaskan moose (*A. a. gigas*) and two European moose (*A. a. alces*) into seven behaviors comprising 97.6% of the 395 h of behavioral observations conducted in summer, fall and spring. Classification performance varied among behaviors and individuals and was generally dependent on sample size. Classification performance was highest for the most common behaviors lying with the head elevated, ruminating and foraging (precision and recall across all individuals between 0.74 and 0.90) comprising 79% of our data, and lower and more variable among individuals for the four less common behaviors lying with head down or tucked, standing, walking and running (precision and recall across all individuals between 0.28 and 0.79) comprising 21% of our data.

Conclusions We demonstrate the use of animal-borne accelerometer data to distinguish among seven main behaviors of captive moose and discuss generalizability of the results to individuals in the wild. Our results can support future efforts to investigate the detailed behavior of collared wild moose, for example in the context of disturbance responses, time budgets and behavior-specific habitat selection.

Keywords Accelerometer, Biologging, Behavior, Cervid, Moose, *Alces alces*, Machine learning, Behavioral analysis, Alaska, Norway

---

*Correspondence:
Theresa M. Kirchner
theresa.kirchner@inn.no
Full list of author information is available at the end of the article

Kirchner *et al. Animal Biotelemetry*   (2023) 11:32

Page 2 of 13

Background

Understanding the behavior of wild animals can facilitate effective conservation and management [1–3]. Such knowledge can be acquired through direct observations of wild animals, which is time-consuming, challenging and expensive [4]. One alternative is to use location data of wild animals to infer their behavior from characteristics of their movement trajectories [5–7]. However, behavioral inference is limited by the spatial and temporal resolution of the location data, which in turn can be influenced by the behavior itself (e.g. by collar position and habitat choice impacting GPS fix rate) [8–11]. Advances in biologging technology alleviate this limitation by enabling the recording of near-continuous data [12, 13]. In particular, animal-attached accelerometers enable a fine-scale biomechanical approach to the study of behavior [13–15].

Tri-axial accelerometers quantify inertial forces along three orthogonal axes [16, 17]. Attached to an animal, they record acceleration that is the result of both static or gravitational acceleration reflecting the posture of the animal relative to the earth's gravitational field, and dynamic or specific acceleration resulting from changes in speed due to movement of the animal [18–20] and vibrations due to effects of tag attachment [21, 22]. The resulting datasets are large (especially at high sampling frequencies) and complex and commonly, machine learning tools are used to classify the accelerometer data into discrete behaviors, using predictor variables that quantify characteristics of the accelerometer traces [14, 23, 24]. Supervised machine learning algorithms are trained by linking behavioral observations to simultaneously recorded accelerometer data, thereby creating a labeled data set, in order to distinguish the observed behaviors based on characteristic differences in the accelerometer traces, allowing for the quantification of model performance [11, 14, 25]. Such behavioral observations are commonly collected on accelerometer-bearing animals in captivity to facilitate the interpretation of accelerometer data collected on wild, unobserved animals [11, 13, 25].

Moose (*Alces alces*) are a keystone species of the boreal forests and tundra in the northern hemisphere [26–28]. Humans highly regard moose for their high cultural significance, for trophy and recreational hunting, and as a food source [29–31]. However, in some areas, browsing damage to commercial forestry plantations and frequent moose–vehicle collisions result in management decisions aimed at limiting population sizes [29, 32]. Throughout much of their range, moose face changes in environmental conditions and disturbances due to human activities [33–35]. The effects of these potential stressors are not yet well-understood due to our limited knowledge of moose physiology and behavior, and of how much behavioral plasticity can compensate for stress and disturbances experienced by individuals [36, 37].

Monitoring the behavior of moose in situ can improve our understanding of how their behavior is affected by disturbances and changes in their environment [38]. Most previous studies aimed at remotely monitoring moose behavior used radio-telemetry or activity counts from dual-axis motion sensors and distinguished only between active and inactive periods lasting several minutes [35, 39, 40]. Ditmer et al. validated activity counts averaged over 1 min with behavioral observations of a single collared captive moose during one season [41]. Resulting behavior-specific activity counts were then used to improve a model predicting the behavior of collared wild moose from year-round GPS data, assigning one of three potential behaviors (resting, foraging, traveling) per 15- or 20-min movement interval [41]. To predict moose behavior in greater detail (i.e. to predict a higher number of behaviors over multiple seasons), it is important to consider the effect of time of year on the motion signatures of behaviors [38, 40]. For example, collar fit can vary over the course of the year [21, 38], the same locomotor behavior can be associated with varying activity counts depending on ground cover including snow [40, 42], the activity count can vary with seasonal changes in insect harassment and resulting movement [40, 43], and different types of food consumed over the course of the year can be associated with different head movements and consequently, activity counts [40, 44, 45]. Furthermore, it is important to account for inter-individual variation in the motion signatures of behaviors [46–48]. Notably, Herberg used behavioral observations conducted on eight collared captive moose during four seasons in combination with dual-axis accelerometer measurements of moose averaged over 5-min intervals as well as GPS-based location data to predict the proportion of time spent resting, foraging or moving within each 5-min interval [38]. Activity within most of their 5-min intervals comprised multiple behaviors associated with behavior-specific variations in energy expenditures [38, 49], and they proposed the use of continuous accelerometer recordings to improve distinction among behaviors and refine the temporal resolution of the behavioral predictions [38]. Increasing the temporal resolution is important because biologically relevant and energetically costly behaviors such as bouts of locomotion or alertness, can occur on time scales that are shorter than the recording intervals of the technology previously used for detecting behaviors [35, 39]. Accelerometer sampling frequency should be at least twice the frequency of the fastest body movement of interest [51–53]. Investigating moose behavior on a finer temporal scale and distinguishing among a higher number of behaviors can facilitate the

early detection of individual responses to changes in the environment resulting from anthropogenic activities, which can serve as foundation for the assessment of population-level responses [54–56].

Our goal was to train a random forest algorithm to classify continuous high-frequency accelerometer data collected from captive moose over several seasons into discrete behaviors, to detect changes in behavior on the temporal scales on which the behaviors can occur. The aim was to enable future studies to quantify fine-scale disturbance responses, behavior-specific habitat selection and detailed time budgets in wild moose.

Methods
Data collection
To study moose behavior, we fitted 12 individuals of subspecies *A. a. gigas* in Alaska (all female) and two individuals of subspecies *A. a. alces* in Norway (one female, one male) with Vectronic Vertex Plus accelerometer-GPS neck collars (Vectronic Aerospace GmBH; Berlin, Germany), which recorded tri-axial accelerometer data at 32 Hz with a sensor range of ± 4 g and a resolution of 8 bit (Additional file 1: Table S1). Accelerometer data were recorded continuously, and accelerometer time stamps were synchronized with GPS time during GPS fixes (every 15 min in collars in Alaska, every 60 min in collars in Norway). We conducted behavioral observations on individual collared moose and distinguished 21 mutually exclusive behaviors, including multiple foraging, locomotor, grooming and inactive behaviors, expanding on Herberg [38] (Additional file 1: Table S2). The protocol for data capture varied between the two locations as described below.

Alaska
Twelve captive female moose at the Kenai Moose Research Center (Alaska Department of Fish and Game, Alaska) were collared without anesthesia three times for data collection over the course of 3 years. Deployment periods were October 1–11 2020, May 7–November 23 2021, and March 24–July 14 2022. Collars were fitted with a 6-cm gap between the collar and the neck to allow for seasonal changes in neck diameter. The moose were kept in two large (2.6 km^2) enclosures with varying terrain and vegetation consisting of boreal and black spruce forest, meadows, bogs and lakes [57]. Supplemental feed was provided from January through April. Supplemental water was provided in one enclosure during June and early July when warm, dry conditions depleted the natural water supplies from wetlands, and in October and November when natural water sources were frozen prior to adequate snow fall. Each animal was observed for at least six hours per observation day during daylight hours.

During the observations, moose were followed on-foot by one of five observers, who logged time-stamped behaviors to the nearest second using GPS time on a tablet running ArcGIS QuickCapture software (Esri, Redlands, CA, USA) and connected to a handheld GPS unit (Bad Elf GPS Pro, Bad Elf, West Hartford, CT, USA).

Norway
One female and one male moose at the Norwegian Moose Center (Inland Norway University of Applied Sciences, Norway) were collared on November 23 2020, following anesthesia with etorphine and xylazine [58]. The moose were kept in a 0.02 km^2 enclosure with vegetation and undulating terrain, a stream and an artificial water station. A salt lick and daily rations of feed pellets were provided, as well as supplemental browse every second day. The moose were filmed from the outside perimeter of the enclosure between November 23 and December 5 2020, using a Canon XA40 (Canon Europe Ltd, Middlesex, U.K.) handheld video camera mounted onto a tripod. The camera was infrared-enabled to film during low-light conditions. On a few occasions, filming without tripod was conducted in order to maintain visibility of active moose during a filming interval. Filming each day was opportunistic and depended on the activity level of the moose, visibility of the moose from the perimeter of the enclosure, and available daylight. Filming took place in approximately 1-h intervals, and the camera was briefly switched off between intervals. At the start and end of each filming interval, the video was synchronized with GPS time by filming the screen of a handheld GPS unit (GPSMAP 64s, Garmin, Southampton, U.K.). Each filming interval focused on one moose, unless both moose were in close proximity to each other. Collars were removed on December 4, 2020 (Mattis, male) and December 9, 2020 (Idun, female) following anesthesia with etorphine and xylazine [58]. Using the software BORIS v.7.9.22 [59], the videos were then transcribed by a single observer with experience in the data collection on Alaskan moose to ensure comparability between the data sets from the two locations. To avoid errors during the transcription process, exclusion criteria for mutually exclusive behaviors were set to ensure the logical sequence of transcribed behaviors (e.g., standing excluded lying).

### Data preparation
Behavioral data
Observation data from Alaska were downloaded from ArcGIS QuickCapture and checked manually. Duplicated entries were removed (e.g. the same button was pressed repeatedly by accident). Within observations, time periods with nonsensical behavioral sequences were excluded

Kirchner *et al. Animal Biotelemetry* (2023) 11:32

Page 4 of 13

from the analysis (e.g. lying followed by running, without any recording of the moose standing up in between). Observations with many errors were entirely excluded from the analysis. Transcribed observation data from Norway were exported from BORIS for further analysis. All behavioral data were imported into R Studio [60] v. 2022.7.2.576 running R [61] v. 4.2.2 for subsequent analysis.

Accelerometer data

The accelerometer data were downloaded from the collars using Vectronic GPS Plus X software v.10.7.2 (Alaska) or v.10.7.1. (Norway), extracted using Vectronic MotionData Monitor software v.1.2.0 and imported into R Studio [60]. Inspection of the data revealed a delay in the date switching of the timestamps after midnight each day. We therefore excluded the first 20 s after midnight for all observations. Inspection of the data also revealed gaps in the accelerometer data of each collar (< 1 min) that occurred at least once per 24-h period, due to rebooting of the unit, as well as inconsistencies in the values of consecutive seconds of time stamps assigned during GPS time synchronization during GPS fixes. Because of these data gaps and the inconsistencies in time stamps assigned during GPS fixes, we summarized the 32-Hz raw accelerometer data in intervals, rather than correcting each individual time stamp, which also facilitated the temporal matching of the behavioral data with the accelerometer data intervals. Based on a preliminary analysis of the data with interval lengths varying from 1 to 10 s, we summarized the accelerometer data in 3-s intervals to maintain a high temporal resolution of individual behaviors (the shortest mean duration of a behavior in our ethogram was two seconds, Additional file 1: Table S2) while maximizing classification performance (i.e. maximizing recall and precision for the largest number of behaviors). Inspection of the data revealed that one collar (Individual: Minnie) recorded at 8 Hz, while the remaining accelerometers recorded at 32 Hz. However, because we summarized our data into intervals, this data was included in the analysis. Opportunistic video recordings revealed that two accelerometer axes were reversed in the collars from Norway compared to Alaska. The data from Norway were adjusted to standardize axis orientation across all collars (Fig. 1).

From the raw accelerometer data, we calculated variables that were frequently used in other studies [14, 23, 25] and did not require continuous time series, to accommodate the aforementioned gaps and inconsistencies in the data. We then summarized the variables in each 3-s interval (Table 1). Most variables described the distribution of raw accelerometer values within each 3-s interval on each axis (X–Z). In addition, pitch (corresponding to
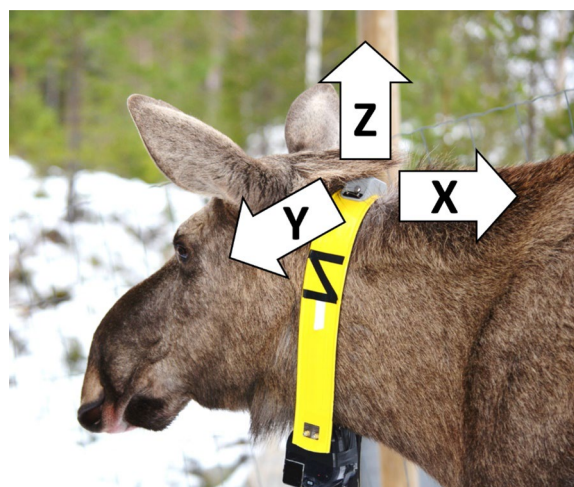
Fig. 1 Accelerometer collar on Idun while standing. Arrows represent axis orientation of the accelerometers mounted in the housing on top of the neck and point towards positive values. X: surge (cranio-caudal axis), Y: sway (medio-lateral axis), Z: heave (ventro-dorsal axis)

Table 1 Predictor variables in the random forest model

Predictor variable	Number of variables per interval
Mean (X, Y, Z, pitch, MSA)	5
Median (X, Y, Z, pitch, MSA)	5
SD (X, Y, Z, pitch, MSA)	5
Min (X, Y, Z, pitch, MSA)	5
Max (X, Y, Z, pitch, MSA)	5
Range (X, Y, Z, pitch, MSA)	5
Interquartile range (X, Y, Z, pitch, MSA)	5
Absolute value of skew (X, Y, Z, pitch, MSA)	5
Kurtosis (X, Y, Z, pitch, MSA)	5
Girth	1
Length	1
Season	1
Sex	1
Subspecies	1

Predictor variables described either the 3-s interval accelerometer data or the time and location of data collection and morphometrics of the collared moose and were used in the random forest model to predict behaviors from the accelerometer data

vertical neck orientation) (Eq. 1) and Minimum Specific Acceleration (MSA) (Eq. 2) were calculated from the raw accelerometer data in each interval. We also included individual metrics which are easy to record in the field: Subspecies, sex, body length, girth and season. Such metrics could improve the generalizability of our model to individuals not seen during model training [25].

Kirchner *et al. Animal Biotelemetry* (2023) 11:32

Page 5 of 13

Calculation of pitch [62]*:*

$$\text{Pitch}(\text{radians}) = a\tan\left(\frac{\text{raw}_x}{\sqrt{\text{raw}_y{}^2 + \text{raw}_z{}^2}}\right) \times \frac{180}{\pi}$$

(1)

Calculation of Minimum Specific Acceleration (MSA) [19]*:*

$$\text{MSA}\left(\frac{m}{s^2}\right) = \left|\sqrt{\text{raw}_x{}^2 + \text{raw}_y{}^2 + \text{raw}_z{}^2} - 1\right|$$

(2)

Labeling of accelerometer data

Visual comparison of the start times of recorded behaviors with the raw accelerometer data for a subset of the data revealed that the recorded start time lagged behind the accelerometer signatures. Therefore, we applied an offset to all behaviors (1 s for data from Alaska, 2 s for data from Norway). The non-overlapping 3-s accelerometer data intervals were labeled with the respective behavior recorded during the observations. Intervals during which more than one behavior was recorded were excluded from analysis. The frequency with which different behaviors were observed varied greatly. Because our goal was to obtain a model that could reliably distinguish the main behaviors of moose, we excluded rare behaviors such as head shaking, scratching and urinating, which represented 2.4% of observations. We summarized all foraging behaviors into a coarser foraging category. To identify when the moose were lying with their head tucked, which has been reported as their energetically

least costly behavior [49], we distinguished between two separate lying behaviors based on their head position: lying with the head down or tucked ("lying_o") and lying with the head up ("lying_u") (Additional file 1: Table S2). Head position of lying moose was assumed to be up unless otherwise noted during the observations (the head position was not recorded for moose in Norway, and therefore whenever these moose were lying, we considered the behavior to be "lying_u"). Our final analysis included 394.7 h of labeled data (380.4 h of observations collected on-foot in Alaska and 14.3 h of annotated video footage from Norway) of the following seven behavioral categories: Foraging, lying_o, lying_u, ruminating, running, standing, walking (Table 2).

Predicting behaviors from accelerometer data

To classify the accelerometer data into behavioral categories, we used a random forest algorithm, which is frequently used for the classification of accelerometer data [23, 47, 63]. A random forest grows many decision trees on bootstrapped subsamples of the data and combines the predictions of all trees to predict the out-of-bag data that were not used to grow the trees, in order to quantify prediction error [64, 65]. Random forest is a comparatively fast supervised classification algorithm that, through the combination of many decision trees and introduced stochasticity in the modeling process, increases classification performance and can process correlated and interacting predictor variables as well as missing values [64–67]. To accommodate the unbalanced nature of our dataset, we assigned weights to the

Table 2 Samples sizes for each individual and behavior

Animal ID	Foraging	Lying_o	Lying_u	Ruminating	Running	Standing	Walking	Total
Stella	12,308	3763	19,247	13,242	18	5408	1661	55,647
Babe	15,732	2791	15,832	9787	89	8462	2590	55,283
Wilma	6970	1383	20,242	15,274	16	9465	1474	54,824
Sky	13,126	1354	15,733	15,999	49	5226	1715	53,202
Shiner	7397	1390	23,916	11,784	9	6359	2176	53,031
Cayenne	13,149	732	12,568	12,927	7	8536	2197	50,116
Roxanne	9392	1386	16,271	10,006	41	7243	3716	48,055
Minnie	6349	1678	22,927	8434	28	5253	1988	46,657
Winnie	4326	99	3426	4078	0	1128	266	13,323
Vicky	3205	49	4034	4451	37	709	236	12,721
Mattis	753	0	2729	2036	50	3897	727	10,192
Idun	865	0	2614	929	13	1894	637	6952
Lily	2832	0	1277	2203	7	362	203	6884
Olivia	2303	10	1392	1930	0	1035	121	6791
Total	98,707	14,635	162,208	113,080	364	64,977	19,707	473,678

Number of labeled 3-s accelerometer data intervals for each behavior and individual moose used to train the random forest model classifying animal-borne accelerometer data into seven discrete behaviors

Kirchner *et al. Animal Biotelemetry*   (2023) 11:32

Page 6 of 13

observations of each behavior that were inversely proportional to the class size of the respective behavior (i.e., we weighed observations so that the weight of observations of behavior X was equal to the number of observations of the rarest behavior divided by the number of observations of behavior X). Assigning greater weight to observations of rare behaviors reduces the error rate of classifications of the rare classes [67]. We used the random forest implementation from H2O through the h2o R package [68] v. 3.38.0.1 with 200 trees. To test the generalizability of our model to new individuals not included during model training, we performed leave-one-individual-out cross-validation, where the model was repetitively trained on all but one of the individuals and evaluated with the labeled data of the remaining, held-out individual [69–71]. We first ran a random forest with the full set of predictor variables. Random forests are capable of handling both correlated and non-informative predictor variables [65–67] and, while a higher number of predictor variables might increase computation time, our priority was to maximize behavioral classification performance. To assess the effect of variable selection on model performance, we then re-ran the model with only those predictor variables that had scored the highest variable importance ($\geq 3\%$) in the full model [67]. To evaluate classification performance, accuracy is a commonly used metric [72]. However, it is a suboptimal metric for evaluating classification performance in imbalanced datasets (such as ours) [72–74]. Thus, modeling with the goal of maximizing accuracy may not be the best procedure for our dataset. Therefore, we focus the discussion of the performance of our model on the metrics recall and precision (but also give accuracy values since this is a common metric used in other studies) [73].

Results
Model performance
Out of 50 predictor variables in the full model, 16 scored a variable importance of at least 3% and were included in the reduced model. Recall and precision of most behaviors in the full model were slightly higher than or equal to recall and precision of the reduced model, except for lying with the head down/tucked and ruminating (Table 3). Therefore, we focus the description of our results and the discussion on the full model.

Across all individuals and behaviors, our model classified 473678 3-s accelerometer data intervals from 14 moose into seven behaviors (Fig. 2) with mean recall of 0.75 (± 0.10) and mean precision of 0.62 (± 0.24) (Table 3).

Across all individuals, classification performance varied by behavior and was generally best for the three most common behaviors (lying with the head up, ruminating, foraging) constituting 79% of our data, with recall and precision

Table 3 Effect of variable selection on model performance

	Full RF		Reduced RF	
	Recall	**Precision**	**Recall**	**Precision**
Foraging	0.86	0.90	0.84	0.89
Lying_o	0.78	0.34	0.75	0.41
Lying_u	0.74	0.84	0.74	0.83
Ruminate	0.79	0.80	0.82	0.78
Run	0.74	0.28	0.73	0.24
Stand	0.55	0.56	0.52	0.53
Walk	0.79	0.62	0.79	0.60

Comparison of model performance between the full random forest model run with all 50 predictor variables and subsequent reduced random forest model run with only the 16 most important variables

ranging from 0.74 to 0.90. Model performance was more variable among the four rarer behaviors constituting the remaining 21% of our data, with recall and precision ranging from 0.28 to 0.79. Among these behaviors, performance was best for walking and lying with the head down/tucked, while standing had the most misclassifications and was most frequently confused with lying behaviors and foraging (Table 4).

Among individuals, classification performance was variable with overall accuracy ranging from 0.38 (Mattis, the only male in our study) to 0.82 (Sky) (Additional file 1: Table S4). Sample sizes among individuals were highly variable, with six moose each contributing less than 3% to the total data in this study, and eight moose each contributing at least 10%. The six individuals with smaller sample sizes scored lower mean recall (mean ± SD: 0.67 ± 0.05) and mean precision (mean ± SD: 0.55 ± 0.06) values than the eight moose with larger sample sizes (mean recall ± SD: 0.75 ± 0.04, mean precision ± SD: 0.64 ± 0.06).

Among individuals, the rarest behaviors (lying_o, running) showed the highest variation in classification performance, particularly precision (Table 5, Additional file 1: Tables S5–S18).

Variable importance
The most important variable in our model was the standard deviation of acceleration along the heave axis with an overall contribution of 5% to the classification performance of the model (Additional file 1: Table S3). Sixteen variables contributed at least 3%, of which five were metrics of pitch, four metrics of surge and three metrics of heave.

Discussion
Animal-borne accelerometers have wide-ranging applications, from investigating the energy budget [75–77] and health status [78, 79] of individuals to identifying behavior-specific habitat use [80, 81]. By facilitating the

Kirchner *et al. Animal Biotelemetry*   (2023) 11:32

Page 7 of 13
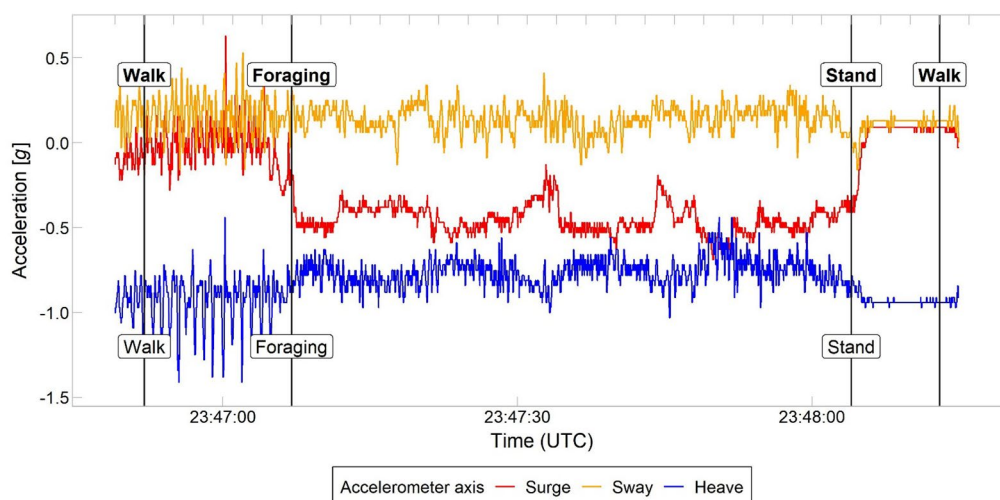
Fig. 2 Example raw accelerometer traces (sampling frequency of 32 Hz) of one captive moose (Stella). Vertical lines indicate the start of a new behavior predicted from the 3-s accelerometer data intervals (bold top labels) and observed during the behavioral data collection (bottom labels). Tick marks on the top axis indicate the start of a new accelerometer data interval

identification of areas important for species conservation [3, 80] and the assessment of effects of disturbances and environmental changes on individual behavior and energy balance [56, 75], this technology can improve wildlife conservation and management efforts. Here, we show that data from animal-borne accelerometers can be used to distinguish among the most common behaviors in moose.

Classification performance
With the three most prevalent behaviors (lying with the head up, ruminating, foraging) scoring the highest recall and precision values between 0.74 and 0.90, classification performance was generally related to class prevalence, which might suggest that the model performed better when the training data contained greater variation in the ways a certain behavior was expressed. While the most prevalent behaviors scored comparable values for both recall and precision, the rarest behaviors (running, lying with the head tucked/down, walking) scored higher recall than precision values. This indicates that our model had fewer false negative predictions of these behaviors, which means that it was able to identify these rare behaviors when the moose were engaging in them, and had a higher number of false positive predictions, which means that it incorrectly predicted these behaviors when other behaviors were occurring. While we assigned greater weights to rare behaviors in order to reduce their classification error [67], it is possible that the weighting was more effective at reducing the number of false negative predictions (and thus increasing recall) than at limiting the number of false positive predictions (and thus increasing precision).

Failing to reduce false positive predictions would lead to a reduction in precision, particularly for behaviors with small numbers of true positive predictions, i.e., behaviors with small sample sizes. Increasing the sample sizes of rare behaviors might improve classification performance for these behaviors but was not feasible in the current study.

Behaviors characterized by little body movement can be difficult to distinguish based on accelerometer data (while predictor variables based on static acceleration might facilitate this distinction, we could not calculate these in the current study), and attempting to distinguish among several inactive behaviors with our model (lying with the head down/tucked, lying with the head up, standing) comes at the risk of reducing the overall classification performance [63, 82]. Nonetheless, we did not group these behaviors together because we wanted to evaluate the performance of our model in distinguishing among these important behaviors. Renecker and Hudson recorded the lowest heart rates in moose lying with the head folded against the abdomen, and an increase in energy expenditure of up to 79% during standing compared to lying with the head tucked [49]. Therefore it was important that our model could distinguish periods of minimal energy expenditure during lying with the head down/tucked from times when moose engage in behaviors associated with increased metabolic rates that serve other functions such as energy gain (ruminating), and increased awareness of and interaction with the surroundings (e.g. during lying with the head up or standing, compared to lying with the head down/tucked). Despite being one of the rarest behaviors in our study, the recall

Table 4 Cross-validation confusion matrix for all individuals

Behavior		Predictions								Recall	Precision	Prevalence (%)
		Lying_u	Ruminating	Foraging	Standing	Lying_o	Walking	Running	Total			
Observations	Lying_u	120,623	17,762	713	11,612	11,003	378	117	162,208	0.74	0.84	34
	Ruminating	10,499	89,245	603	8649	3948	104	32	113,080	0.79	0.80	24
	Foraging	275	553	84,485	5709	544	7063	78	98,707	0.86	0.90	21
	Standing	11,232	4430	5305	35,571	6304	1973	162	64,977	0.55	0.56	14
	Walking	54	50	2860	624	167	15,647	305	19,707	0.79	0.62	4
	Lying_o	1576	131	56	1416	11,423	30	3	14,635	0.78	0.34	3
	Running	2	18	3	2	1	68	270	364	0.74	0.28	0
	Total	144,261	112,189	94,025	63,583	33,390	25,263	967	473,678			

The confusion matrix combines the cross-validation confusion matrixes of the random forest model classifying accelerometer data across all 14 moose observed in captivity in Alaska and Norway. Values in columns represent the number of 3-s accelerometer data intervals predicted for each of the seven behaviors, split into rows based on the behavioral labels of the intervals recorded during the observations. Recall and precision quantify the model classification performance for the respective behavior across all animals in the study. Prevalence indicates the contribution of each behavior to the total sample size of accelerometer data intervals

Kirchner *et al. Animal Biotelemetry*   (2023) 11:32

Page 9 of 13

Table 5 Behavior-specific individual variation in model performance

Performance metric	Behavior						
	Lying_u	Ruminating	Foraging	Standing	Walking	Lying_o	Running
Recall	0.69±0.17	0.74±0.18	0.86±0.04	0.53±0.18	0.73±0.14	0.78±0.15	0.74±0.21
Precision	0.82±0.09	0.78±0.12	0.89±0.07	0.56±0.18	0.57±0.19	0.31±0.31	0.28±0.26
Prevalence (%)	34	24	21	14	4	3	0

Behavior-specific variation in classification performance among 14 individuals of the random forest model classifying seven different behaviors from accelerometer data. Mean and standard deviation of precision and recall are given together with the prevalence of the behaviors in the observational data

of lying with the head down/tucked ranged among the highest values of all behaviors, with 78% of all events that were labeled as lying with the head down/tucked being correctly identified by our model. While the unique neck postures during this inactive behavior might facilitate its distinction, false predictions of this behavior did occur (34% in total) and involved mostly other, more common inactive behaviors (lying with the head up, standing, ruminating), illustrating the challenges of distinguishing inactive behaviors from accelerometer data. We did not distinguish lying with the head down/tucked from the generally much more common behavior lying with the head up during the transcription of videos from Norway and therefore labeled all lying behaviors of these moose as lying with the head up. As a consequence, some data had incorrect labels (the small proportion of data that were labeled as lying with the head up when it should have been labeled as lying with the head down/tucked) that trained the model to incorrectly predict the behavior in these instances as lying with the head up. Similarly, some data with incorrect labels (i.e. lying with the head up) were used to falsify predictions that were actually correct (i.e. lying with the head down/tucked). It is likely that this contributed to the comparatively low precision of our models' predictions of lying with the head down/tucked.

In an accelerometer study on reindeer (*Rangifer tarandus*) that grouped all inactive behaviors (including standing, sleeping and ruminating) into one behavior class, this class had the best classification performance among all behaviors [71], which was better than the classification performance for any of the inactive behaviors in our study. However, the focus of the study on reindeer was the distinction among three foraging behaviors (browsing low, browsing high and grazing) [71]. In contrast, we grouped three foraging behaviors into one overall foraging class, which in turn had a better classification performance than the three foraging behaviors investigated in the study on reindeer (precision of foraging in our study scored higher than precision of all three behaviors in the study on reindeer, and recall of foraging in our study scored higher than recall of two out of the three behaviors in the study on reindeer) [71]. This comparison illustrates

the potential effect of grouping of behaviors on model classification performance and the behavioral inferences that can be drawn from the predictions, emphasizing that behavioral grouping needs careful consideration in studies using supervised classification algorithms to analyze accelerometer data.

Classification performance in our model was comparable to that in Martiskainen et al. classifying accelerometer data from dairy cows [83]. While their model performed better at classifying standing, our model performed better at classifying foraging behavior. Similar to our study, Martiskainen et al. reported misclassifications among less active behaviors (lying, ruminating and standing), which they also suspected was due to the similarities in neck posture of the cows during these behaviors [83]. Their model confused among the behaviors foraging, standing and (lame) walking [83] which is also evident in our predictions. During our observations, we considered a moose to be foraging until it took more than two consecutive steps without bites of food; which prompted a switch to walking. Consequently, some instances where the moose was walking were still recorded as foraging, likely contributing to the misclassifications of these two behaviors. Furthermore, foraging and walking can occur simultaneously in browsing animals, complicating their distinction using accelerometer data.

Model generalizability

Given the goal of classifying unlabeled data in wild animals, cross-validating the model on labeled data from unseen individuals, can provide insights into the generalizability of the model [25, 70, 71]. Therefore, variation in classification performances among individuals is a useful indicator of the generalizability of our model [69–71].

In an effort to maximize model generalizability, we aimed to maximize the amount of variation in our training data by pooling data from as many individuals as possible and including individuals from both sexes and two subspecies [83]. The lowest overall prediction performance (accuracy and mean recall) was observed when our model classified data from the only male moose in our study (Mattis). A possible interpretation is that our model might have limited applicability to male moose.

Morphological differences such as the large weight of the head due to the presence of antlers and resulting increased neck circumference [84] could result in different neck posture and movement of male moose compared to female moose during the same behaviors, precluding the generalizability to male moose of a model that was trained on data from female moose to classify data from neck-mounted accelerometers. This notion is supported by the high total number of false predictions of lying with the head down/tucked for Mattis; a behavior characterized by unique neck postures that is confused mainly with behaviors characterized by limited body movement where neck posture might be an important predictor (standing, lying with the head up and ruminating). However, we did find that these misclassifications also occurred particularly often in Shiner, the female moose with the largest measured chest girth and weight in our study, where a large and heavy head and large neck circumference might have resulted in similar misclassifications to those observed for a (younger and) smaller male with small antlers. This might suggest that the reduced performance of our model in classifying Mattis' data did not stem from a lack of generalizability of our model to (young) male moose with small antlers. Instead, the low sample sizes for several of Mattis' behaviors as well as overall individual variability in model performance, which we discuss below, might have resulted in the comparatively low performance of our model when classifying his data. However; ultimately, due to our small sample size of male moose, we cannot evaluate the generalizability of our model to male moose. European moose constituted only 3.6% of the data, hence their predictions were largely based on data from Alaskan moose. Yet, mean recall and precision of the behavioral classification of the one female European moose in our study, Idun, were higher than the mean values of Alaskan moose with similar sample sizes. While the successful application of our model to Idun's accelerometer data might have been facilitated by the similarities in size between Idun and the yearling Alaskan moose in our study (Babe, Vicky and Winnie), ultimately our sample size is too small to evaluate the generalizability of our model to European moose.

Variation in overall accuracy and behavior-specific recall and precision among individuals with comparable sample sizes (e.g. Shiner and Sky) suggests the influence of factors other than sample size, sex and subspecies on model performance. Such individual differences in classification performance have been observed on a wide range of species from penguins [48] to pinnipeds [25, 47] and caprids [46]. Including individual characteristics as predictor variables might account for some of this individual variation and has been shown to increase the generalizability of classification models [25]. However,

individual length and girth had comparatively low variable importance in our model. Other variables such as age or weight might have been more important [25] but were not included in our model because these metrics are difficult to determine in the field when collaring wild moose. Furthermore, length and girth were not measured on all animals in our study and were inferred from other data for several individuals, potentially confounding the importance of these metrics on the behavioral classification of moose accelerometer data.

Fine-scale differences in placement of the accelerometers among individuals might have contributed to the individual variation in the classification performance of our model [48, 69]. Because most collars were deployed for several months at a time, they were fitted to account for seasonal changes in neck diameter, potentially resulting in changes in how the collars responded to body movement over the course of the deployments, thereby increasing within- and among-individual variation of the data [22, 38, 85]. Because collar fitting in our study was similar to collar fitting on wild moose in the field, our training data included such variation. While this might have reduced the classification performance of our model, it increases generalizability of our model to data from wild animals, where some fine-scale differences in accelerometer placement among individuals can be expected. In our model, season had a variable importance of 2%, suggesting that variation in collar fit over the course of the deployments, or other seasonal variation such as the effect of snow on locomotor behaviors, exerted some influence on the classification in our model. In addition to within- and among-animal variation among collar placement, variation may exist among the accelerometer units themselves [22]. Addressing such variation requires calibration of the units prior to deployment [22, 86, 87]; but calibration data are often not available for existing accelerometer data where collars were deployed in the field without prior calibration [22].

Limitations of our study and recommendations for futures studies

The quality of the time stamps of our accelerometer data prevented a time series analysis of the data at a sub-second level. It was therefore not possible to distinguish between static and dynamic acceleration [17, 18] and analyze the frequency composition of the accelerometer signals [14, 18], to calculate predictor variables that were among the most important for the classification of accelerometer data in other studies [46, 63, 71]. For example, frequency analysis of accelerometer data using fast Fourier transform can facilitate the distinction among simultaneous, rhythmic behaviors such as foraging and walking [63]. In moose, such a frequency

analysis might help distinguish among lying and ruminating, standing, foraging and walking behaviors from accelerometer data. Improving the quality of the time stamps recorded by the accelerometers built into the collars would enable the calculation of these important predictor variables, thus offering a promising way to further improve the performance of behavioral classification models on fine-scale tri-axial accelerometer data of moose.

For the sake of this study, we considered postures (e.g. lying, standing) as separate categories from behaviors (e.g. foraging, ruminating, walking). Postures and behaviors are not mutually exclusive as, for example, a foraging moose is usually standing. Consequently, there was overlap in the accelerometer signatures of the behavioral classes, which we considered exclusive. This could explain some of the misclassifications among these behaviors like for example, foraging and standing, and lying with the head up and ruminating. In future studies, recording posture and behavior separately might facilitate the distinction among these behaviors [63]. However, such a distinction is logistically challenging when logging behaviors in real time in the field.

When applying our model to accelerometer data from wild moose, our model will not be able to classify behaviors that were not included in model training, for example swimming which can occur when moose are foraging on aquatic vegetation [88]. Instead, such behaviors unknown to the model will be misclassified as one behavior (or multiple behaviors) known to the model based on similarity in the accelerometer variables [25]. Increasing the sample size of observations of male and European moose and of rare behaviors would improve the generalizability of our model to new individuals.

Conclusions

We demonstrate the use of accelerometer data to distinguish among seven important behaviors of moose. Potential applications include the quantification of the time budget of wild moose and, by relating behavioral predictions to environmental variables, the investigation of behavior-specific habitat selection as done for other species [80, 81, 89]. Quantifying behavioral responses of moose to changes in their environment can elucidate the effect of disturbances on their time budget. Relating accelerometer data to metabolic rate could elucidate the energetic consequences of behavioral responses of moose to disturbances [15, 56].

Abbreviation
MSA   Minimum specific acceleration

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40317-023-00343-0.

Additional file 1: Table S1. Details on individual moose included in the study. **Table S2**. Ethogram used to record observations of captive moose including sample sizes and durations of all behaviors and grouping of behaviors to label accelerometer data intervals. **Table S3**. Variable importance of the random forest model. Importance of all predictor variables of the random forest model used to predict moose behavior from 3-s intervals of animal-borne accelerometer data. **Table S4**. Summary of the individual classification performance of the random forest model in predicting behavior from accelerometer data for each of the 14 moose in the study. **Tables S5–S18**. Confusion matrixes for the 14 moose. Values in columns represent the number of 3-s accelerometer data intervals predicted for each of the seven behaviors, split into rows based on the behavioral labels of the intervals recorded during the observations. Recall and precision quantify the classification performance of the model for each behavior of each individual. Prevalence indicates the contribution of each behavior to the total sample size of accelerometer data intervals for the respective individual.

Availability of data and materials
Data for this project are available upon reasonable request from the Alaska Department of Fish and Game and subject to a data sharing agreement.

Declarations

Ethics approval and consent to participate
Animal care and handling procedures in Alaska were approved by the Animal Care and Use Committee, Alaska Department of Fish and Game, Division of Wildlife Conservation (protocol number 0086-2020-40). Animal handling in Norway did not require a permit under the Norwegian Food Safety Authority as the collars were deployed during routine hoof trimming and annual veterinary exams.

Consent for publication
All authors agree consent to publish this work to *Animal Biotelemetry*.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Department of Forestry and Wildlife Management, Faculty of Applied Ecology, Agricultural Sciences and Biotechnology, Inland Norway University of Applied Sciences, Campus Evenstad, Anne Evenstads vei 80, 2480 Koppang, Norway. [2]Centre d'Etudes Biologiques de Chizé, UMR 7372, CNRS-La Rochelle

Kirchner *et al. Animal Biotelemetry* (2023) 11:32

Page 12 of 13

Université, La Rochelle, France. [3]Alaska Department of Fish and Game, Kenai Moose Research Center, 43961 Kalifornsky Beach Road, Suite B, Soldotna, AK 99669, USA.

References

1. Roever CL, Beyer HL, Chase MJ, van Aarde RJ. The pitfalls of ignoring behaviour when quantifying habitat selection. Divers Distrib. 2014;20(3):322–33.
2. de Gabriel HM, Karamanlidis A, Grivas K, Krambokoukis L, Papakostas G, Beecham J. Habitat use and selection patterns inform habitat conservation priorities of an endangered large carnivore in southern Europe. Endang Species Res. 2021;44:203–15.
3. Dombroski J, Parks S, Nowacek D. Dive behavior of North Atlantic right whales on the calving ground in the Southeast USA: implications for conservation. Endanger Species Res. 2021;46:35–48.
4. Altmann J. Observational study of behavior: sampling methods. Behaviour. 1974;49(3–4):227–66.
5. Bailey H, Mate B, Palacios D, Irvine L, Bograd S, Costa D. Behavioural estimation of blue whale movements in the Northeast Pacific from state-space model analysis of satellite tracks. Endang Species Res. 2009;10:93–106.
6. Patterson TA, Parton A, Langrock R, Blackwell PG, Thomas L, King R. Statistical modelling of individual animal movement: an overview of key methods and a discussion of practical challenges. Adv Stat Anal. 2017;101(4):399–438.
7. Hooven ND, Williams KE, Hast JT, McDermott JR, Crank RD, Jenkins G, et al. Using low-fix rate GPS telemetry to expand estimates of ungulate reproductive success. Anim Biotelemetry. 2022;10(1):5.
8. Heard DC, Ciarniello LM, Seip DR. Grizzly bear behavior and global positioning system collar fix rates. J Wildl Manag. 2008;72(3):596–602.
9. Mattisson J, Andrén H, Persson J, Segerström P. Effects of species behavior on global positioning system collar fix rates. J Wildl Manag. 2010;74(3):557–63.
10. Shamoun-Baranes J, Bom R, van Loon EE, Ens BJ, Oosterbeek K, Bouten W. From sensor data to animal behaviour: an oystercatcher example. PLoS ONE. 2012;7(5):e37997.
11. Wang Y, Nickel B, Rutishauser M, Bryce CM, Williams TM, Elkaim G, et al. Movement, resting, and attack behaviors of wild pumas are revealed by tri-axial accelerometer measurements. Mov Ecol. 2015;3(1):2.
12. Kooyman GL. Maximum diving capacities of the weddell seal, *Leptonychotes weddelli*. Science. 1966;151(3717):1553–4.
13. Yoda K, Sato K, Niizuma Y, Kurita M, Bost C, Le Maho Y, et al. Precise monitoring of porpoising behaviour of Adelie penguins determined using acceleration data loggers. J Exp Biol. 1999;202(22):3121–6.
14. Watanabe S, Izawa M, Kato A, Ropert-Coudert Y, Naito Y. A new technique for monitoring the detailed behaviour of terrestrial animals: a case study with the domestic cat. Appl Anim Behav Sci. 2005;94(1):117–31.
15. Wilson RP, Börger L, Holton MD, Scantlebury DM, Gómez-Laich A, Quintana F, et al. Estimates for energy expenditure in free-living animals using acceleration proxies: a reappraisal. J Anim Ecol. 2020;89(1):161–72.
16. Whitford M, Klimley AP. An overview of behavioral, physiological, and environmental sensors used in animal biotelemetry and biologging studies. Anim Biotelemetry. 2019;7(1):26.
17. Wilson RP, Shepard ELC, Liebsch N. Prying into the intimate details of animal lives: use of a daily diary on animals. Endang Species Res. 2008;4:123–37.
18. Shepard ELC, Wilson RP, Quintana F, Laich AG, Liebsch N, Albareda DA, et al. Identification of animal movement patterns using tri-axial accelerometry. Endang Species Res. 2008;10:47–60.
19. Simon M, Johnson M, Madsen PT. Keeping momentum with a mouthful of water: behavior and kinematics of humpback whale lunge feeding. J Exp Biol. 2012;215(21):3786–98.
20. Yoda K, Naito Y, Sato K, Takahashi A, Nishikawa J, Ropert-Coudert Y, et al. A new technique for monitoring the behaviour of free-ranging Adelie penguins. J Exp Biol. 2001;204(4):685–90.
21. Dickinson ER, Stephens PA, Marks NJ, Wilson RP, Scantlebury DM. Best practice for collar deployment of tri-axial accelerometers on a terrestrial quadruped to provide accurate measurement of body acceleration. Anim Biotelemetry. 2020;8(1):9.
22. Garde B, Wilson RP, Fell A, Cole N, Tatayah V, Holton MD, et al. Ecological inference using data from accelerometers needs careful protocols. Methods Ecol Evol. 2022;13(4):813–25.
23. Nathan R, Spiegel O, Fortmann-Roe S, Harel R, Wikelski M, Getz WM. Using tri-axial acceleration data to identify behavioral modes of free-ranging animals: general concepts and tools illustrated for griffon vultures. J Exp Biol. 2012;215(6):986–96.
24. Wilson R, Vandenabeele S. Technological innovation in archival tags used in seabird research. Mar Ecol Prog Ser. 2012;11(451):245–62.
25. Ladds MA, Thompson AP, Slip DJ, Hocking DP, Harcourt RG. Seeing it all: evaluating supervised machine learning methods for the classification of diverse otariid behaviours. PLoS ONE. 2016;11(12):e0166898.
26. Kielland K, Bryant JP, Ruess RW. Moose herbivory and carbon turnover of early successional stands in interior Alaska. Oikos. 1997;80(1):25–30.
27. Persson IL, Danell K, Bergström R. Disturbance by large herbivores in boreal forests with special reference to moose. Ann Zool Fenn. 2000;37(4):251–63.
28. Snaith TV, Beazley KF. Moose (*Alces alces americana* [Gray Linnaeus Clinton] Peterson) as a focal species for reserve design in Nova Scotia, Canada. Nat Areas J. 2002;22(3):235–40.
29. Storaas T, Gundersen H, Henriksen H, Andreassen HP. The economic value of moose in Norway—a review. Alces J Devoted Biol Manag Moose. 2001;37(1):97–107.
30. Nelson JL, Zavaleta ES, Chapin FS. Boreal fire effects on subsistence resources in Alaska and adjacent Canada. Ecosystems. 2008;11(1):156–71.
31. Priadka P, Moses B, Kozmik C, Kell S, Popp J. Impacts of harvested species declines on Indigenous Peoples' food sovereignty, well-being and ways of life: a case study of Anishinaabe perspectives and moose. Ecol Soc. 2022. https://doi.org/10.5751/ES-12995-270130.
32. Olaussen JO, Skonhoft A. A cost-benefit analysis of moose harvesting in Scandinavia: a stage structured modelling approach. Resour Energy Econ. 2011;33(3):589–611.
33. Neumann W, Ericsson G, Dettki H. The impact of human recreational activities: Moose as a case study. Alces J Devoted Biol Manag Moose. 2010;47:17–25.
34. Jones H, Pekins P, Kantar L, Sidor I, Ellingwood D, Lichtenwalner A, et al. Mortality assessment of moose (*Alces alces*) calves during successive years of winter tick (*Dermacentor albipictus*) epizootics in New Hampshire and Maine (USA). Can J Zool. 2019;97(1):22–30.
35. Græsli AR, Le Grand L, Thiel A, Fuchs B, Devineau O, Stenbacka F, et al. Physiological and behavioural responses of moose to hunting with dogs. Conserv Physiol. 2020;8:coaa122.
36. Lowe SJ, Patterson BR, Schaefer JA. Lack of behavioral responses of moose (*Alces alces*) to high ambient temperatures near the southern periphery of their range. Can J Zool. 2010;88(10):1032–41.
37. Monteith KL, Klaver RW, Hersey KR, Holland AA, Thomas TP, Kauffman MJ. Effects of climate and plant phenology on recruitment of moose at the southern extent of their range. Oecologia. 2015;178(4):1137–48.
38. Herberg A. Are Minnesota moose warming up to climate change? A validation of techniques for remotely monitoring moose behavior and body temperature. [MSc]. University of Minnesota; 2017. http://conservancy.umn.edu/handle/11299/188796. Accessed 24 Aug 2020.
39. Cederlund G. Activity patterns in moose and roe deer in a north boreal forest. Ecography. 1989;12(1):39–45.
40. Moen R, Pastor J, Cohen Y. Interpreting behavior from activity counters in GPS collars of moose. Alces J Devoted Biol Manag Moose. 1996;32:101–8.
41. Ditmer MA, Moen RA, Windels SK, Forester JD, Ness TE, Harris TR. Moose at their bioclimatic edge alter their behavior based on weather, landscape, and predators. Curr Zool. 2018;64(4):419–32.
42. Fancy SG, White RG. Incremental cost of activity. In: Hudson RJ, White RG, editors. Bioenergetics of wild herbivores. 1st ed. Boca Raton: CRC Press; 2018. p. 143–60.
43. Bunnell FL, Gillingham MP. Foraging behavior: dynamics of dining out. In: Hudson RJ, White RG, editors. Bioenergetics of wild herbivores. 1st ed. Boca Raton: CRC Press; 2018. p. 53–80.

Kirchner *et al. Animal Biotelemetry*   (2023) 11:32

Page 13 of 13

44. Geist V. On the behaviour of the North American Moose (*Alces Alces Andersoni* Peterson 1950) in British Columbia. Behaviour. 1963;20(3–4):377–415.
45. Dorn RD. Moose and cattle food habits in southwest Montana. J Wildl Manag. 1970;34(3):559–64.
46. Dickinson ER, Twining JP, Wilson R, Stephens PA, Westander J, Marks N, et al. Limitations of using surrogates for behaviour classification of accelerometer data: refining methods using random forest models in Caprids. Mov Ecol. 2021;9(1):28.
47. Shuert CR, Pomeroy PP, Twiss SD. Assessing the utility and limitations of accelerometers and machine learning approaches in classifying behaviour during lactation in a phocid seal. Anim Biotelemetry. 2018;6(1):14.
48. Chimienti M, Kato A, Hicks O, Angelier F, Beaulieu M, Ouled-Cheikh J, et al. The role of individual variability on the predictive performance of machine learning applied to large bio-logging datasets. Sci Rep. 2022;12(1):19737.
49. Renecker LA, Hudson RJ. Ecological metabolism of moose in aspen-dominated boreal forests, central Alberta. Can J Zool. 1989;67:1923–8.
50. Gottardi E, Tua F, Cargnelutti B, Maublanc ML, Angibault JM, Said S, et al. Use of GPS activity sensors to measure active and inactive behaviours of European roe deer (*Capreolus capreolus*). Mammalia. 2010;74(4):355–62.
51. Halsey LG, Shepard ELC, Wilson RP. Assessing the development and application of the accelerometry technique for estimating energy expenditure. Comp Biochem Physiol A Mol Integr Physiol. 2011;158(3):305–14.
52. Hounslow JL, Brewster LR, Lear KO, Guttridge TL, Daly R, Whitney NM, et al. Assessing the effects of sampling frequency on behavioural classification of accelerometer data. J Exp Mar Biol Ecol. 2019;512:22–30.
53. Chakravarty P, Cozzi G, Ozgul A, Aminian K. A novel biomechanical approach for animal behaviour recognition using accelerometers. Methods Ecol Evol. 2019;10(6):802–14.
54. Tuomainen U, Candolin U. Behavioural responses to human-induced environmental change. Biol Rev. 2011;86(3):640–57.
55. Pirotta E, Booth CG, Costa DP, Fleishman E, Kraus SD, Lusseau D, et al. Understanding the population consequences of disturbance. Ecol Evol. 2018;8(19):9934–46.
56. Pirotta E, Booth CG, Cade DE, Calambokidis J, Costa DP, Fahlbusch JA, et al. Context-dependent variability in the predicted daily energetic costs of disturbance for blue whales. Conserv Physiol. 2021;9(1):coaa137.
57. Thompson DP, Crouse JA, Barboza PS, Spathelf MO, Herberg AM, Parker SD, Morris MA. Behavior influences thermoregulation of boreal moose during the warm season. Conserv Physiol. 2021;9:coaa130.
58. Evans AL, Fahlman Å, Ericsson G, Haga HA, Arnemo JM. Physiological evaluation of free-ranging moose (*Alces alces*) immobilized with etorphine-xylazine-acepromazine in Northern Sweden. Acta Vet Scand. 2012;54(1):77.
59. Friard O, Gamba M. BORIS: a free, versatile open-source event-logging software for video/audio coding and live observations. Methods Ecol Evol. 2016;7(11):1325–30.
60. RStudio Team. RStudio: integrated development environment for R. Boston: R Studio PBC; 2022.
61. R Core Team. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2022.
62. Chimienti M, Cornulier T, Owen E, Bolton M, Davies IM, Travis JMJ, et al. The use of an unsupervised learning approach for characterizing latent behaviors in accelerometer data. Ecol Evol. 2016;6(3):727–41.
63. Fehlmann G, O'Riain MJ, Hopkins PW, O'Sullivan J, Holton MD, Shepard ELC, et al. Identification of behaviours from accelerometer data in a wild social primate. Anim Biotelemetry. 2017;5(1):6.
64. Breiman L. Random forests. Mach Learn. 2001;45(1):5–32.
65. Ziegler A, König IR. Mining data with random forests: current options for real-world applications. WIREs Data Min Knowl Discov. 2014;4(1):55–63.
66. Cutler DR, Edwards TC, Beard KH, Cutler A, Hess KT, Gibson J, et al. Random forests for classification in ecology. Ecology. 2007;88(11):2783–92.
67. Cutler A, Cutler DR, Stevens JR. Random forests. In: Zhang C, Ma Y, editors. Ensemble machine learning: methods and applications. Boston: Springer; 2012. p. 157–75.
68. LeDell E, Gill N, Aiello S, Fu A, Candel A, Click C, et al. h2o: R interface for the "H2O" scalable machine learning platform. 2022.
69. Rahman A, Smith DV, Little B, Ingham AB, Greenwood PL, Bishop-Hurley GJ. Cattle behaviour classification from collar, halter, and ear tag sensors. Inf Process Agric. 2018;5(1):124–33.
70. Ferdinandy B, Gerencsér L, Corrieri L, Perez P, Újváry D, Csizmadia G, et al. Challenges of machine learning model validation using correlated behaviour data: evaluation of cross-validation strategies and accuracy measures. PLoS ONE. 2020;15(7):e0236092.
71. Rautiainen H, Alam M, Blackwell PG, Skarin A. Identification of reindeer fine-scale foraging behaviour using tri-axial accelerometer data. Mov Ecol. 2022;10(1):40.
72. Tharwat A. Classification assessment methods. Appl Comput Inform. 2020;17(1):168–92.
73. Leoni J, Tanelli M, Strada SC, Berger-Wolf T. Ethogram-based automatic wild animal monitoring through inertial sensors and GPS data. Ecol Inform. 2020;59:101112.
74. Chen C, Liaw A, Breiman L. Using random forest to learn imbalanced data. Berkeley: University of California; 2004 p. 12. Report No.: 666. https://statistics.berkeley.edu/sites/default/files/tech-reports/666.pdf. Accessed 18 Oct 2020.
75. Pagano AM, Atwood TC, Durner GM, Williams TM. The seasonal energetic landscape of an apex marine carnivore, the polar bear. Ecology. 2020;101(3):e02959.
76. Bryce CM, Dunford CE, Pagano AM, Wang Y, Borg BL, Arthur SM, et al. Environmental correlates of activity and energetics in a wide-ranging social carnivore. Anim Biotelemetry. 2022;10(1):1–16.
77. Ste-Marie E, Watanabe YY, Semmens JM, Marcoux M, Hussey NE. Life in the slow lane: field metabolic rate and prey consumption rate of the Greenland shark (*Somniosus microcephalus*) modelled using archival biologgers. J Exp Biol. 2022;225(7):jeb242994.
78. Wilson RP, Grundy E, Massy R, Soltis J, Tysse B, Holton M, et al. Wild state secrets: ultra-sensitive measurement of micro-movement can reveal internal processes in animals. Front Ecol Environ. 2014;12(10):582–7.
79. Tobin C, Bailey DW, Trotter MG, O'Connor L. Sensor based disease detection: a case study using accelerometers to recognize symptoms of Bovine Ephemeral Fever. Comput Electron Agric. 2020;175:105605.
80. Wilson AM, Lowe JC, Roskilly K, Hudson PE, Golabek KA, McNutt JW. Locomotion dynamics of hunting in wild cheetahs. Nature. 2013;498(7453):185–9.
81. Suraci JP, Frank LG, Oriol-Cotterill A, Ekwanga S, Williams TM, Wilmers CC. Behavior-specific habitat selection by African lions may promote their persistence in a human-dominated landscape. Ecology. 2019;100(4):e02644.
82. Graf PM, Wilson RP, Qasem L, Hackländer K, Rosell F. The use of acceleration to code for animal behaviours; a case study in free-ranging Eurasian beavers castor fiber. PLoS ONE. 2015;10(8):e0136751.
83. Martiskainen P, Järvinen M, Skön JP, Tiirikainen J, Kolehmainen M, Mononen J. Cow behaviour pattern recognition using a three-dimensional accelerometer and support vector machines. Appl Anim Behav Sci. 2009;119(1):32–8.
84. Lynch GM, Lajeunesse B, Willman J, Telfer ES. Moose weights and measurements from Elk Island National Park, Canada. Alces J Devoted Biol Manag Moose. 1995;31:199–207.
85. Wilson RP, Rose KA, Gunner R, Holton MD, Marks NJ, Bennett NC, et al. Animal lifestyle affects acceptable mass limits for attached tags. Proc R Soc B Biol Sci. 1961;2021(288):20212005.
86. Aggarwal P. MEMS inertial sensor errors. In: MEMS-based integrated navigation. GNSS Technology and Applications Series. Boston: Artech House; 2010.
87. Won SP, Golnaraghi F. A triaxial accelerometer calibration method using a mathematical model. IEEE Trans Instrum Meas. 2010;59(8):2144–53.
88. Morris KI. Impact of moose on aquatic vegetation in northern Maine. Alces J Devoted Biol Manag Moose. 2002;38:213–8.
89. Spedener M, Tofastrud M, Devineau O, Zimmermann B. Microhabitat selection of free-ranging beef cattle in south-boreal forest. Appl Anim Behav Sci. 2019;213:33–9.

Publisher's Note