

METHODOLOGY

Open Access



A supervised model to identify wolf behavior from tri-axial acceleration

Charlotte Lorand^{1*}, Léa Bouet^{1,2}, Olivier Devineau¹, Marianna Chimienti³, Alina L. Evans¹, Peggy Callahan⁶, Mark Beckel⁶, Timothy G. Laske^{4,5} and Ane Eriksen¹

Abstract

Background In wildlife studies, animal behavior serves as a key indicator of the impact of environmental changes and anthropogenic disturbances. However, wild animals are elusive and traditional GPS studies only provide limited insight into their daily activities. To address this issue, behavior classification models have increasingly been used to detect specific behaviors in wildlife equipped with tri-axial accelerometers. Such models typically need to be trained on data from the target species. The present study focuses on developing a behavioral classification model tailored to the grey wolf (*Canis lupus*) and encompassing a variety of ecologically relevant behaviors.

Methods We collected data from nine captive wolves equipped with collar-mounted tri-axial accelerometers recording continuous acceleration at 32 Hz (“fine-scale”) and averaged acceleration over 5-min intervals (“activity”). Using simultaneous video observations, we trained Random Forest models to classify wolf acceleration data into specific behaviors. We investigated the potential limits to the generalizability of these models to unlabeled data through individual-based cross-validation.

Results We present: (1) a model classifying fine-scale acceleration data (32 Hz) into 12 distinct behaviors (lying, trotting, stationary, galloping, walking, chewing, sniffing, climbing, howling, shaking, digging and jumping) with a class recall of 0.77–0.99 (0.01–0.91 in cross-validation), (2) a model classifying activity data (5-min averages) into 3 behavior categories (static, locomotion and miscellaneous) with a class recall of 0.43–0.91 (0.39–0.92 in cross-validation). Although classification performance decreased following cross-validation, recall scores for lying, trotting, stationary, galloping, walking and chewing individual behaviors (as well as static and locomotion categories) remained above 0.6. Classification performance was consistently poorer for rare behaviors, which constituted less than 1.1% of the training dataset.

Conclusions We demonstrate the use of collar-mounted accelerometer to distinguish between 12 behaviors and 3 behavior categories in captive wolves, at fine-scale (32 Hz) and averaged 5-min resolutions, respectively. We also discuss the generalizability of the two models to free-ranging settings. These models can be employed to support future behavioral studies examining questions such as conflict mitigation, wolf responses to human disturbances, or specific activity budgets.

Keywords *Canis lupus*, Wolf, Random Forest, Classification, Multi-class, Acceleration, Activity, Behavior, Machine learning, Behavior analysis

*Correspondence:

Charlotte Lorand
charlotte.lorand@inn.no

Full list of author information is available at the end of the article



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Background

In wildlife studies across the world, some recurring constraints consistently limit data collection: wild animals are elusive, accessibility and detectability are poor, and field conditions unpredictable [5, 97]. Data on physiology and behavior are especially difficult to obtain at fine temporal and spatial scales, yet they are key indicators of human impact on wildlife [8, 13, 89, 91], both in fundamental research and for implementation of conservation strategies [55]. Over the past decades, bio-loggers have been increasingly used to fill in this gap, with ever-growing precision and versatility, now allowing for the collection of high-resolution spatiotemporal information [2, 8, 15, 16, 65, 66].

In this context, the use of high-resolution data collected from animal-borne accelerometers to detect animal behavior has emerged as a promising solution to several problematics. Location data from GPS trackers have historically been used for behavioral studies, but limitations in device autonomy and available resolutions typically constricted data collection and resulting studies to large-scale habitat use and individual-scale speed [97]. Fine-scale accelerometers are also battery and storage-demanding, which can limit their use to one year at a time under the best scenarios: however, at a comparable scale (e.g., yearly, with resolution as high as 1 min or more), turning to raw acceleration allows for collection of more precise information than with GPS data, provided analysis methods are available. Once calibrated by associating acceleration signal patterns with observed body movements [47, 65], supervised machine learning algorithms such as Random Forests can be used to predict behaviors on unlabeled datasets from the target species [11, 20, 49, 73]. The availability of high-resolution acceleration data thus opens new possibilities for behavioral studies at finer scales [24, 60].

However, even though machine-learning potentially offers a high level of accuracy in such behavioral detection, and has been recently used in various species and fashions [6, 14, 29, 47, 49, 53, 56, 62, 64, 65, 67, 71, 73, 83, 92, 94], they most often remain constricted to few behaviors with a limited level of detail. The present study focuses on the grey wolf (*Canis lupus*), a species that has been extensively studied in various aspects of spatial ecology, from broad landscape use [61] to specific hunting strategies [78], often relying on extensive observations [58] or decades of GPS tracking [13, 79], and aims at identifying 12 ecologically significant behaviors.

In a study from 2022, the Overall Dynamic Body Acceleration (ODBA) from nine captive wolves fitted with 32-Hz accelerometers was used to detect 5 behavior categories (resting, stationary, walking, highly active, and running [12], allowing to quantify day and

night behavior patterns during breeding, pup rearing and recruitment seasons, in Denali National Park. Other studies targeted specific questions such as detection of kill sites [68], or use calibration based on domestic dogs [20]. Wolves are involved in human–wildlife conflict across their whole range [22, 25, 27, 31, 52, 80], which are often fueled by perceived lack of appropriate management response [45, 46, 57, 59, 76, 86, 93]. The ability to detect specific behaviors with ever-growing level of detail holds promises in improving tailored management responses: for instance, preventive prediction of conflict hotspots (which is often used in conservation management practice, see, e.g., [36, 48], has been refined over the last decades by the implementation of connected systems allowing, for instance, real-time alerts of raiding of agricultural areas [75]. Some of these systems now provide early detection through behavioral cues [40, 87] or even use detection-triggered repellents [1, 28, 38, 39, 90], and real-time treatment of acceleration data from an algorithm embedded in a collar to send out alerts based on behavior detection is achievable today [44]. The use of behavioral cues in early conflict detection has proven effective in preventing animal casualties [7], and unequivocal fine-scale behavior detection from acceleration data could further enhance such use, moving toward systems tailored to specific situations.

Provided an accurate and fine-scale behavioral detection algorithm is available, such systems could be used to bring up new insight into wolf ecology, but also adapted to alleviate conflicts related to wolf depredation on domestic animals, since automatic repellents and guarding systems have been demonstrated to be effective on short and long term, respectively [10, 25, 37]. Therefore, in this study we develop a behavioral classification model for detecting a maximum number of ecologically relevant behaviors to accommodate both behavioral research and conservation needs. We first present a model predicting 12 distinct behaviors, calibrated with supervised machine learning using video footage and tri-axial acceleration data sampled at 32 Hz from captive wolves.

However, this fine-scale setup is not the norm in most monitoring frameworks, where devices recording summarized bi- or tri-axial accelerometry (“activity”) are typically used. These devices have been available for longer, are less costly, less power- and storage-hungry, and align with most monitoring or spatial studies only requiring a few positions every day. Hence, we complementarily present a model predicting 3 behavior classes from calibration with such activity data, to be used on datasets commonly collected during routine monitoring, and discuss the need and hindrances to developing replicability between study sites, programs, species, and frameworks.

Methods

Data collection

Data collection took place at the Wildlife Science Center in Minnesota, from October to December 2018. We studied three family groups in separate enclosures, each consisting of a pair of adults and three to six younger wolves. We collected data from three 1.5-year-old wolves in each enclosure, one enclosure at a time (i.e., a total of 9 individuals, see Additional file A.I). The enclosures were flat with no vegetation. All subject animals were immobilized with ketamine and xylazine, weighed, their fur dyed on specific body parts for individual identification, and fitted with GPS collars featuring store-on-board tri-axial accelerometers (Vectronic VERTEX Plus V 2.1, Vectronic Aerospace Inc, here 1.74%–2.61% of body mass). The acceleration modules were set to record (1) tri-axial acceleration (*surge*, horizontal, X; *sway*, lateral, Y; and *heave*, vertical, Z, see Fig. 1) at 32 Hz and 4 g, and (2) summarized tri-axial acceleration over 300 s (Vectronic Aerospace Inc, 2024). Each enclosure was equipped with four IR cameras (Cromorc Wireless Security 1.3 MP, Cromorc Inc.) at the corners.

Data processing

Processing of raw data

For our analysis we used both raw tri-axial acceleration continuously at 32 Hz (hereafter “fine-scale” datasets and corresponding models), as well as tri-axial acceleration averaged across 300 s (hereafter “activity” datasets and corresponding models, Vectronic Aerospace Inc., 2024), and continuous video footage for each enclosure for two to three days [see Additional file A.I]. Since most

previously published ethograms are only relevant for wolves in free-ranging conditions [54], and to account for biologically relevant behaviors displayed by the subject wolves, we designed a specific ethogram based on observations from the video footage, which we used as a reference for labeling [see Additional file A.II]. One observer used software BORIS (v.7.4.1; [33]) to label behaviors with start and stop times, from video segments of 30 min every 3 h throughout the filming periods (total: 8 h per individual), with additional observations of rare behaviors at feeding time. To improve prediction accuracy during model training, we only classified clear, distinct, non-overlapping behaviors with unambiguous identification of the individual. We did not include state-transition behaviors.

Behaviors selection and formatting

The ethogram was composed of 12 behaviors (Fig. 2) based on ecological significance and potential for practical use in wildlife research and management: locomotion (*walking*, *trotting*, *galloping*), interaction with the environment (*jumping*, *climbing*, *shaking*, *digging*), and with food and conspecifics (*chewing*, *sniffing*, *howling*). We partitioned static states into *lying* (lying down immobile), and *stationary* (sitting, lying down while moving the head, or standing), since all of the latter presented high similarity in acceleration metrics. Hereafter, we call “rare” behaviors with a prevalence inferior to 1%: *galloping*, *sniffing*, *climbing*, *howling*, *shaking*, *digging* and *jumping*. Acceleration data and behaviors were matched using their time stamps.

In “fine-scale” models, we compared a model featuring all 12 behaviors (model A, Table 1), to a model in which rare behaviors were grouped into a “miscellaneous” category (model A.1, Table 1). When analyzing the 300-s averaged activity dataset, we tested different ways of grouping the 12 behaviors into fewer ecologically and practically meaningful categories to account for the comparably short duration of individual behaviors (Table 1). For instance, we included a 5-category model (Model C), intending to preserve explicit detection of certain behaviors of ecological interest despite low classification success (namely, separating *running* and *walking*, and including *chewing*). Each 300-s interval was labeled according to the comparative duration of all fine-scale behaviors or categories observed within it, with the label corresponding to the behavior or category that occupied the longest duration.

To correct for sampling imbalance caused by the inclusion of rare behaviors, we incorporated a relative weight calculated as the proportion of each behavior relative to the overall dominant behavior (fine-scale: $P_{\text{lying}} = 0.637$,

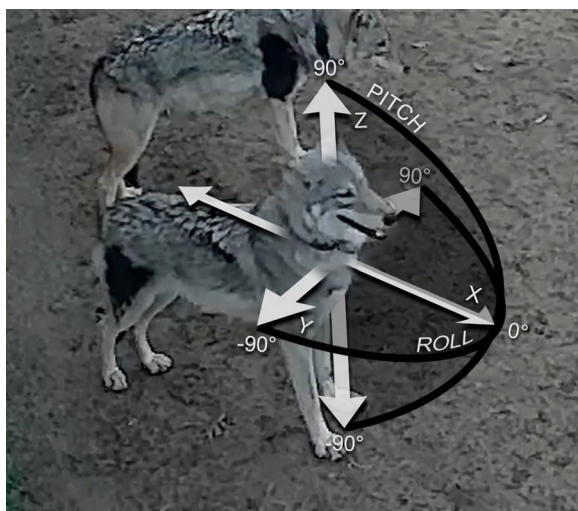


Fig. 1 Orientation of the accelerometer contained in the collar fitted on a subject wolf. X: surge; Y: sway; Z: heave; pitch and roll: vertical and lateral orientation across a 180° angle, respectively

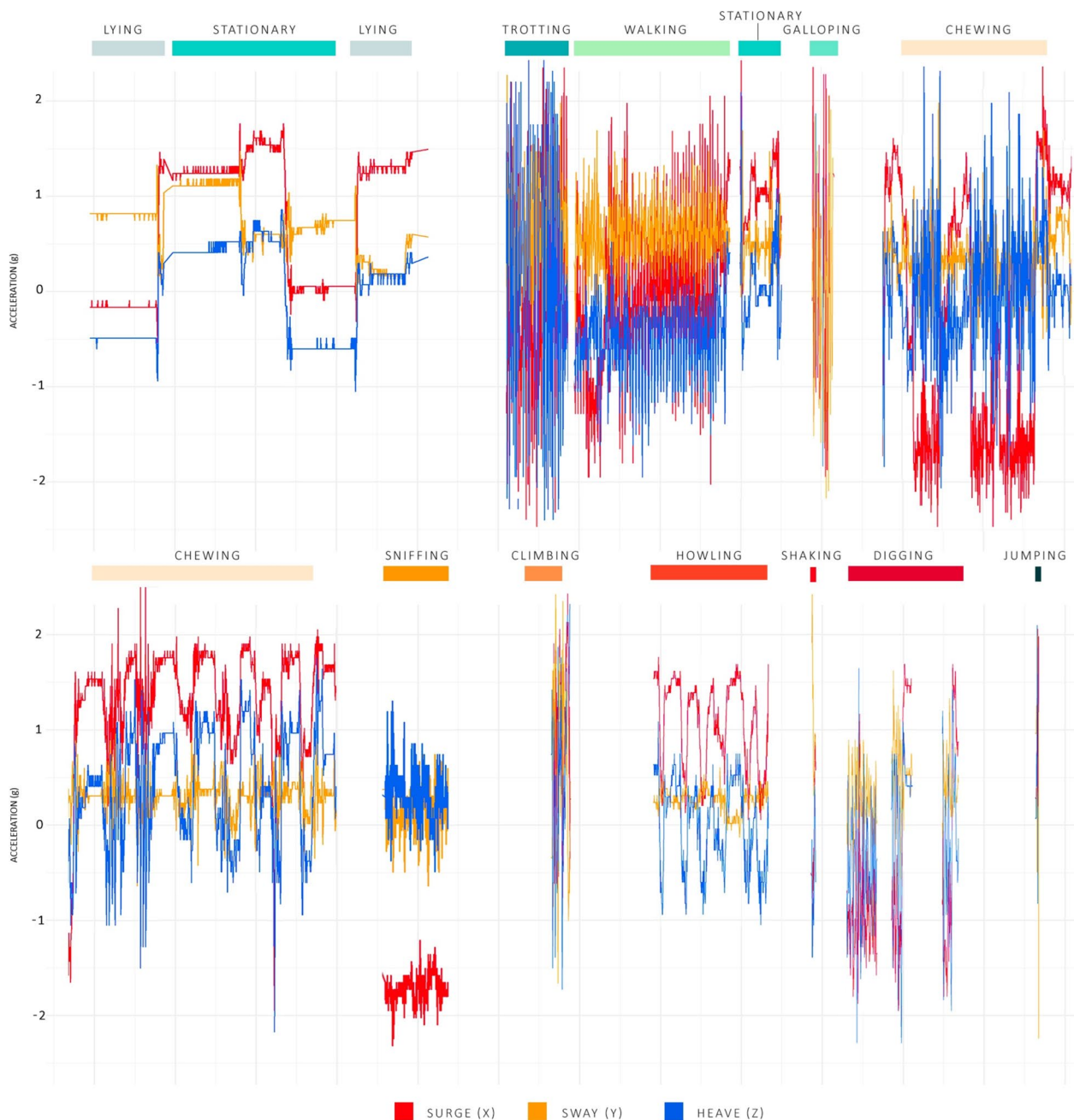


Fig. 2 Acceleration signatures of the 12 fine-scale wolf behaviors targeted for classification using machine-learning algorithms. Raw surge (x, red), sway (y, yellow) and heave (z, blue) are displayed on a main 15-s intervals grid with colored bars indicating the corresponding observed behavior

and activity: $P_{\text{stationary}} = 0.237$, detail of sampling proportions is provided in Additional file A, Table III).

Decision variables

In line with Shepard et al. [83], we used biologically meaningful variables derived from the raw acceleration data, characterizing the orientation and movement of the neck [99]. Aside from standardized mean tri-axial

acceleration values (x , y , z), we included mean *pitch* and *roll* representing vertical and lateral orientation across a 180° angle, respectively (Fig. 1), dynamic tri-axial accelerations (D_x , D_y and D_z), mean overall dynamic body acceleration (ODBA), mean vector of dynamic body acceleration (VEDBA), mean magnitude of acceleration (AMAG) and variance of all previous variables (V) [detail of equations in Additional file B.I]. To calculate these

Table 1 Results of Random Forest model building for classifying wolf behaviors from acceleration data

Model	Behaviors	No. of internal trees	Mean depth	Mean leaves	Recall	Recall ^{CV}	Log-loss	Log-loss ^{CV}
<i>Fine-scale models</i>								
A.1	<i>Lying–trotting–stationary–galloping–walking–chewing–miscellaneous</i>	1400	20	26,997	0.960	0.869	0.338	1.05
A	<i>Lying–trotting–stationary–galloping–walking–chewing–sniffing–climbing–howling–shaking–digging–jumping</i>	2400	20	16,075	0.954	0.862	0.459	4.912
<i>Activity models</i>								
B	<i>Stationary (lying, stationary, howling)–locomotion (walking, trotting, galloping)–miscellaneous</i>	600	17.5	147	0.847	0.841	1.66	1.268
C	<i>Stationary (lying, stationary, howling)–running (trotting, galloping)–chewing–walking–miscellaneous</i>	800	17	130	0.822	0.809	2.02	1.765
D	<i>Stationary (lying, stationary, howling)–running (trotting, galloping)–walking–miscellaneous</i>	800	16.3	113.6	0.821	0.814	4.295	3.578
E	<i>Stationary (lying, stationary, howling)–locomotion (walking, trotting, galloping)–chewing–miscellaneous</i>	1000	16.2	111.2	0.854	0.841	5.86	3.546
F	<i>Lying–trotting–stationary–galloping–walking–chewing–sniffing–climbing–howling–shaking–digging–jumping</i>	2400	11	70.9	0.704	0.674	5.781	12.151

Models are presented in descending order of overall recall scoring. Models retained for potential use in free-ranging settings are in bold. However, only top-models A and B are discussed further in this study. All behaviors not explicitly named are grouped under “miscellaneous”. Only models A and F feature all 12 behaviors separately. Fine-scale: 32-Hz acceleration data; Activity: 300 s summarized acceleration; ^{CV}: cross-validation metrics

acceleration-derived variables, we tested several running means (*S*) ranging from 1 to 30 s. Based on performance during model fitting, the comparative duration of behaviors (e.g., Fig. 2), and the high mobility of the wolf as a species, we used a 1-s running mean for all analyses.

Model building

We built a Random Forest (RF) multi-class classification algorithm [11, 23, 51, 63], using R package *h2o* [34]. We trained a behavioral classifier with 200 trees to balance explanatory power and computational cost [69], using an 80–20 randomized split keeping behavior class proportions for the training and testing sets, respectively [26, 30] and a one-versus-all approach [74]. To alleviate potential overfitting and gain a better understanding of the model's potential in classifying unlabeled data, we performed a leave-one-individual-out cross-validation, with a final performance score averaged across the 9 wolves [4, 100]. All analyses were performed on R version 4.3.3 [70].

Model performance

As we aimed at achieving the best identification of behaviors displayed by unseen free-ranging individuals, we mainly based our performance assessment on individual class recall (i.e., proportion of true positives predicted out of each class) and precision (i.e., proportion of true positives out of all positive predictions) [5]. We also investigated log-loss (quantifying the uncertainty of a classifier

by penalizing false classification) and relative variable importance (based on how much the overall squared error decreased when a variable was selected to split on during the tree building process). We visualized decision variables using packages *plotly* version 4.10.4 [85] and *ggplot2* version 3.5.0 [95], and model results with package *networkD3* version 0.4 [35]. The final results incorporate general training and cross-validation into what we will call the “main models” (“MM”). However, we will focus our discussion on extracted cross-validation results (in the following, “cross-validation” or “CV”) as a reflection of situations with high individual variation, which can limit the potential of the main model.

Results

Behavioral characteristics

Of the 12 behaviors selected for classification, two (*stationary* and *lying*) contributed 87.7% of all 32-Hz acceleration datapoints. The subsequently most represented behaviors were *trotting* (5.7%), *walking* (2.9%) and *chewing* (2.6%). The remaining behaviors were rare and made up 1.1%. Prevalence in activity categories followed the same pattern. Behavior prevalence varied between individuals, and *jumping*, *climbing*, and *digging*, were not displayed at all by 1, 2, and 3 individuals, respectively. Due to technical limitations during footage analysis in one enclosure, 3 individuals contributed less than the others to the overall datasets [see Additional file A.III].

Fine-scale models (32 Hz)

While results of the main 12-behavior model returned a mean per-class recall of 0.88, with scores ranging from 0.77 (*shaking*) to 0.99 (*lying*), the cross-validation results taken separately showed a mean per-class recall of 0.42, ranging from 0.01 (*digging*) to 0.91 (*lying*) (Model A, Fig. 3, details in Table 1). The behaviors classified with the best recall score were also the most represented in the dataset, with *lying*, *stationary*, and *trotting* scoring above 0.75, and *walking*, *galloping* and *chewing* scoring between 0.60 and 0.70. The remaining rare behaviors scored a recall lower than 0.1, except for *sniffing* (0.25). Between individuals, mean recall varied from 0.71 to 0.93. Precision scores were always lower than or comparable to recall, varying from 0.006 (*digging*) to 0.98 (*lying*). Finally, the most important variables to predict behaviors were mean *pitch*, *x*, *ODBA*, *VEDBA*, and *z* (Fig. 3 and Additional file C.VIII). Grouping rare behaviors together in a reduced model (Model A.1, details in Table 1) did not significantly improve recall or precision for the remaining behaviors.

Plotting head *pitch* against overall dynamic body acceleration (*ODBA*), which are among the most important variables, and useful for representing behaviors in a 2D space, shows variable levels of overlap between acceleration datapoints for each behavior (Fig. 4).

Activity model (300 s)

During classification training for activity data, we compared models with various behavior groupings. The simplest 3-category model (*static*, *locomotion* and *miscellaneous*, Model B, Table 1) was the most performant, with per-class recall scores of 0.92, 0.77 and 0.44, respectively (0.92, 0.73 and 0.39 in cross-validation) (Fig. 3). Precision scores were lower than recall, the most prevalent behaviors in the dataset were classified best, and the most important variables for activity models were mean *activity*, *AMAG*, and mean sway (*y*). Model B also strongly relied on mean surge (*x*) [Additional file C.VIII]. Amongst individuals, we obtained mean CV recalls ranging from 0.48 to 0.88. We tentatively performed a leave-one-out approach to enrich this simple 3-category model B with more behaviors, in a 5-category mixed model (Model C, *stationary*, *running*, *walking*, *chewing*, *miscellaneous*, Table 1). Albeit not improving *walking* scores

compared to the full 12-behaviors model F, model C slightly improved *chewing* scores (per-class recall scores of 0.92, 0.81, 0.26, 0.51 and 0, respectively) (0.93, 0.77, 0.16, 0.5 and 0 in cross-validation) [Additional file C.IV], and remained similar to model B in regard to performance metrics.

Discussion

Model performance

From the Random Forest trainings, we obtained models able to discriminate between several combinations of ecologically and practically significant behaviors. We selected two of them for discussion and general use: one model (A) using 32-Hz data to identify 12 behaviors with recalls above 0.77 (of which 6 behaviors with a recall score above 0.6 in cross-validation), and one model (B) using activity data to identify 3 behavior categories, with static and locomotion behaviors scoring recalls above 0.73 in both general model performance and cross-validation. As we want to explore more specifically the potential caveats related to individual variation, in the following, we will only use cross-validation recalls unless otherwise specified.

Fine-scale models

In the fine-scale model, recall scores per behavior were overall positively correlated with prevalence in the training set (Fig. 3), except for the running behaviors (*trotting*, 0.90 and *galloping*, 0.71) which, albeit constituting 6% and 0.4% of the dataset, respectively, scored similarly to static behaviors (*lying*, 0.91 and *stationary*, 0.79) constituting 64% and 24% of the dataset. Other behaviors making for less than 3% of the dataset returned mid-range recall scores (between 0.6 and 0.3). Notably, *howling* (prevalence 0.4%) scored below 0.1, and was frequently confused with being merely stationary, implying that the distinctive back and forth of the head on the *z*-axis when a wolf howls (Fig. 2) was overridden in the model by the comparably upright position of the head. This is consistent with mean *pitch* and *ODBA* being the most important decision variables in the model. Behaviors involving little body movement (e.g., lying or standing in different positions) have been reliably classified in previous studies and various species [29, 41, 47, 65, 67, 92, 94], but can be difficult to differentiate without reducing overall

(See figure on next page.)

Fig. 3 Sankey diagram representations of cross-validation confusion matrices for models A and B, and variable importance. Observed behaviors are featured on the left with relative prevalence indicated in associated circles. Classification results are featured on the right, alongside performance metrics for each classified behavior. Internal links represent classification confusions. The length of external bars on each side and the thickness of internal links represent relative prevalence in sampling, confusions and classification. MM: main model; CV: cross-validation. (all confusion matrices are available in Additional files C.I–VII)

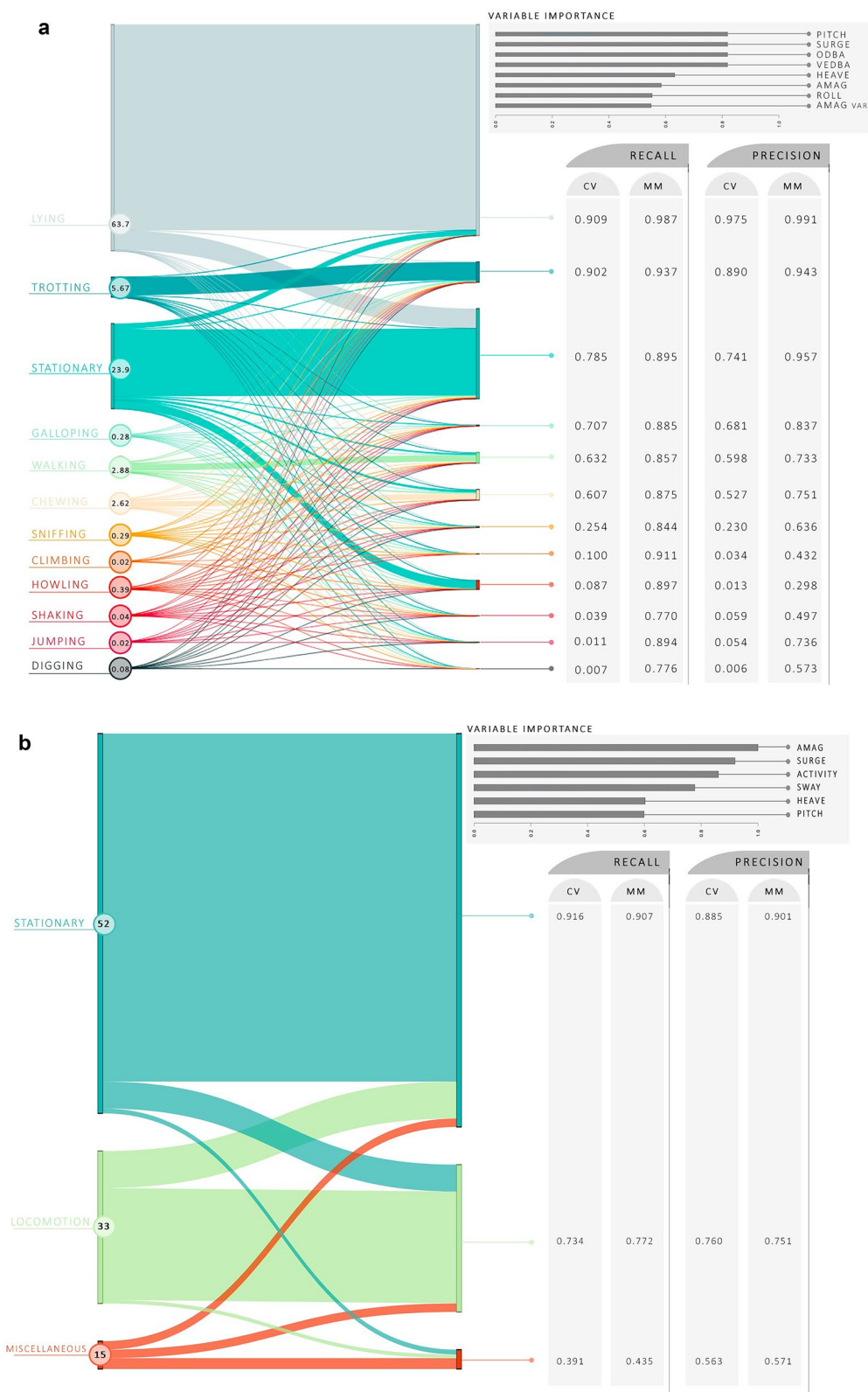


Fig. 3 (See legend on previous page.)

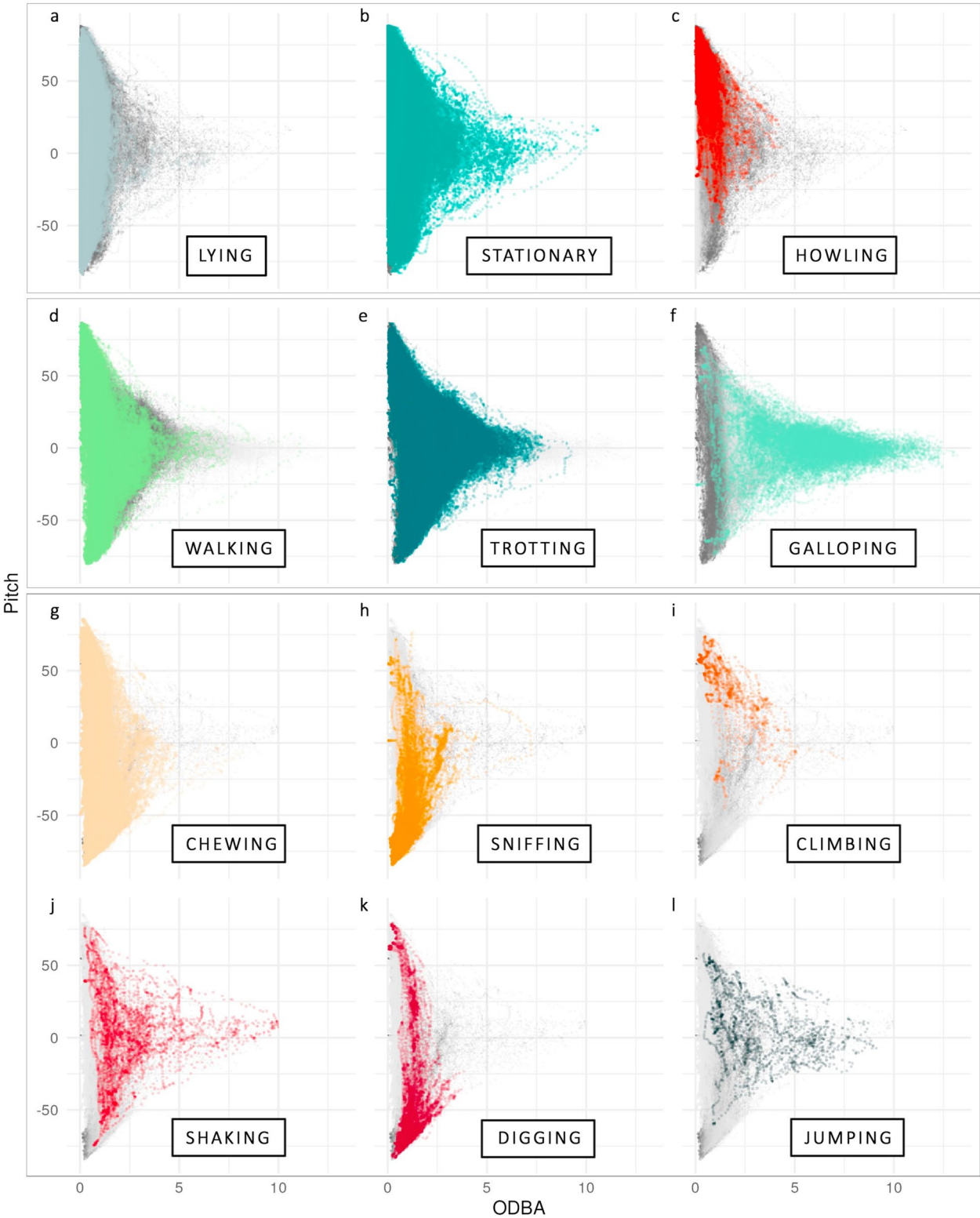


Fig. 4 Distribution of the main decision variables (*pitch*, *ODBA*) for each behavior in the fine-scale model. Distribution of calculated mean pitch (vertical angle of the head of a wolf) and mean ODBA (overall dynamic body acceleration) for static (a–c), locomotion (d–f) and miscellaneous (g–i) behavior categories, for each 32-Hz datapoint. Each behavior within a category is represented in a consistent shade of grey, with target behavior in its associated color, to better highlight specificities and overlaps

classification performance (e.g., by increasing the time window for smoothing running means, [47, 82]). This phenomenon seemed to occur between the three static behaviors, *lying*, *stationary* and *howling*.

Despite the use of a weighting factor, all rare behaviors (prevalence under 0.1%, i.e., *climbing*, *shaking*, *digging* and *jumping*) were poorly classified and frequently confused with other behaviors, which is a common issue in imbalanced datasets. These behaviors also displayed very short duration without distinctive pattern (e.g., *climbing* and *jumping*, Fig. 2). Although it is recommended for the sampling frequency to be at least twice the frequency of the fastest behavior classified [14, 42], which is the case here (0.217 s for *jumping*), these behaviors might be too short compared to the 1-s time window chosen for the running mean and fail in providing the model with enough variation in the decision variables.

On the contrary, running behaviors, albeit displayed sometimes just as rarely (0.28% for *galloping*), showed high recall and precision in classification. These behaviors have longer duration and present more distinctive features (e.g., acceleration patterns or ranges of *ODBA*, see Figs. 2, 4) than comparatively more poorly classified behaviors. The overlaps in pitch/*ODBA* ranges between *static* behaviors or rare behaviors such as *chewing*, *sniffing* and *digging* (Fig. 4g, h, k) seem correlated to the confusions observed in classification (Fig. 3A). The low prevalence and lack of clear characteristics in behaviors such as *shaking* or *jumping* (Fig. 4i, l) is correlated to poor recall. On the other hand, a clear distinction between *ODBA* ranges is visible in the better-performing *locomotion* behaviors. We theorize that, while higher prevalence improves classification power by increasing the available information on behavioral variation during training [47], a longer duration of the target behavior and more distinctive or stereotyped features might have a similar effect to that of a larger sample size. While classification of highly variable behaviors will benefit from larger sample sizes, classification of short behaviors could be improved by combining a classifier with automated detection relying on techniques such as Fast Fourier transform to filter out small body movements [29], KNN clustering identification [5], or techniques such as changepoint analysis [9] or custom Boolean indicators [98] to better detect significant shifts in parameters pointing towards a change in behavior.

Activity models

As the 300-s duration of activity intervals largely exceeds our recommended duration threshold, we did not expect to be able to predict short behaviors with high accuracy from this dataset. Activity models, however, remain important in that they allow us to make the most of

long-term data series, even when data collection started with lower-resolution equipment or when resources such as money, time, battery or storage are limited. This makes interpretation and comparison possible between different types of data, equipment and protocols. As grouping behaviors will affect classification performance and behavior inference [47], we chose to retain the simplest but most performant activity model (Model B), able to adequately classify locomotion states versus static states, as well as another model with clear differentiation between behaviors of ecological interest with the highest potential for reliable classification (Model C, Table 1). Provided that predictions from this last model are backed-up with sensible, contextual interpretation, they can allow for the detection of feeding events from *chewing* behavior, and for the detection of potentially stressful or physically demanding events from *running* behaviors. Similarly, we chose to keep the full fine-scale model (Model A) for applications in free-ranging animals and future improvements rather than performing extensive grouping to improve recall scores (Model A.1, Table 1).

Generalizability

Behavior classification from biologging in free-ranging animals is inevitably confronted with several caveats. First, the definition of a behavior is rather fluid: prior to model training, we defined and classified only non-overlapping behaviors to maximize classification power. However, many behaviors are highly variable in their amplitude, duration, or pattern [18] and are often displayed with additional organic and individual variability, such as partial displays or mixing (e.g., walking and sniffing simultaneously). Realistically, acceleration signatures are also affected by the placement and type of accelerometer device used (e.g., a collar will be biased towards head movements and might be more mobile compared to devices attached to the animal's core such as backpacks or body-mounted tags), the transition between different behaviors, as well as by variations in environmental features. We stress that the data used in this study were recorded in a captive setting, from wolves provided with deer carcasses and kept in flat enclosures with no vegetation and little enrichment (mostly concrete pipes). Notably, behavior prevalence reflects overall low activity levels in captive animals, feeding events, and patrolling of the enclosure, the most common behaviors displayed by social canids in captivity [19, 32, 84]. Consequently, our models lack information on movements influenced by features such as vegetation, topography, environmental obstacles (e.g., rocks or logs), movements through different media (e.g., water or snow), and entire behavior sequences (e.g., hunting and killing of prey). This can lead to poorer performance when using the model on unseen

data. Knowledge about wolf biology, environment, as well as the likely behavior confusions as quantified in this study, are key to ensure proper interpretation of behavior classifications *in situ*.

From the classifier itself, it would be useful to, for instance, extract true class probabilities in terminal leaves and base decisions on the target behavior distribution compared to confused behaviors through a more quantitative approach [21, 72, 101]. Another solution to bridge the gap between model training on captive data, and predictions on unlabeled free-ranging data, is to use the information obtained from models built in captive studies to guide parameter selection, or interpret and generally validate behavior classification obtained from clustering with unsupervised machine learning approaches or Hidden Markov Models [3, 16, 17, 44, 50, 77].

Improvements and recommendations for future work

Continuous development and optimization of existing models in ecology, conservation, and monitoring are an integral part of model building frameworks. In the case of the present study and in behavior classification models in general, increasing the training sample with new comparable data from new individuals and rare behaviors would provide useful information on behavioral variation to improve model recall. Given the vast and variable datasets collected with accelerometer tags available today, it is even more important to carefully plan, select and tailor protocols to the research question, the study species, environment, device location on the body, type of data available, and potential for future studies [81, 82, 92]. Protocols might also be improved, and model validated by using variables derived from other sensors [96], such as conductivity [65], body temperature [43], depth [16], or even footage from collar-mounted cameras or acoustic recordings from embedded microphones. As previously stated, with typically low sample sizes, a supervised model will most likely be undertrained on rare behaviors and on environmental and individual variation, resulting in poorer prediction power in free-ranging settings. Such caveats can only be reliably estimated by multiple trainings on unlabeled data, however, it is noticeable that only training and validation results are consistently reported in the literature. We therefore argue that such models should be systematically cross-validated to avoid overestimation of performance in free-ranging settings, and that individual cross-validation results should be discussed when evaluating performance [4, 44, 100], as they allow an estimation of the potential bias caused by individual variation and can give perspective on the real predictive ability of a model.

Conclusions

Automated behavior identification in free-ranging animals, both in real-time and in delayed studies, has demonstrated usefulness and potential throughout the past decades [7, 39, 40, 48, 87, 88]. In this study, we were able to develop two behavior identification models for wolves, returning adequate classification for common behaviors when tested on captive data. These models can be used to predict behaviors in free-ranging wolves, provided attention is given to the likeliest behavior confusions, especially in the behaviors for which these models show poor predictive ability. It is also necessary to remain aware of the variation in prediction reliability introduced in free-ranging data. The impact of such limitations can be mitigated using contextual clues and knowledge on wolf biology, or the use of complementary models tailored specifically to target more efficiently specific rare behaviors of interest. In further refinement studies, complementary algorithms such as unsupervised clustering models could be used to more adequately identify similar acceleration signatures in highly variable free-ranging situations.

Abbreviations

AMAG	Magnitude of acceleration
DRF	Distributed Random Forest
ODBA	Overall dynamic body acceleration
VEDBA	Vector of dynamic body acceleration

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40317-025-00400-w>.

Supplementary Material 1. Additional files A – Data sampling: Table I. Details of study animals and data collection. Table II. Wolf ethogram for the 12 discrete behaviors, arranged per activity category grouping in the following order: locomotion; miscellaneous; static. Table III. Data overview of acceleration datapoints sampled at 32 Hz and 300-s averaged activity intervals with dominant behavior. Additional files B – Modelling: Table I. Equations and descriptions of decision variables used during model training and of performance metrics used in model assessment. Additional files C – Results: Table I. Confusion matrices from Model A. Table II. Confusion matrices from Model A.1. Table III. Confusion matrix from Model F. Table IV. Confusion matrix from Model C. Table V. Confusion matrix from Model D. Table VI. Confusion matrix from Model E. Table VII. Confusion matrix from Model B. Table VIII. Variable importance

Acknowledgements

We want to extend special thanks to Kristoffer Nordli for collar deployment and help with logistics during data collection. We also thank Elżbieta Król for the latter, as well as Megan Callahan, Noelle Laske, and the WSC staff for assistance with data collection. We finally thank three anonymous reviewers and the editorial team for their time and efforts in improving the manuscript.

Author contributions

AE and ALE conceived first ideas and designed methodology. AE and ALE secured the funding. PC and MB performed captures and immobilization. AE, ALE, PC, MB and TL collected the data. OD and MC collaborated to methodology and analysis at all steps. LB pre-processed the data, performed video labeling, early analysis, reporting and interpretation. CL performed further data treatment, formal data analyses, visualizations, and wrote the drafts. AE,

MC, OD, ALE and TL reviewed the drafts. All authors contributed critically to the drafts and gave final approval for publication.

Funding

Open access funding provided by Inland Norway University Of Applied Sciences. The study was funded by the Hedmark University College, R&D Grant and Regional Research Fund Region Innlandet (RFFINNL), project number 289688. The funders were not involved in the design of the study and collection, analysis, and interpretation of data or in writing the manuscript.

Data availability

The datasets used and in the present study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

The animal use in this study was reviewed and approved by the Institutional Animal Care and Use committee at the Wildlife Science Center, Minnesota USA, protocol number #UMN-005.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Department of Forestry and Wildlife Management, University of Inland Norway, 2480 Evenstad, Norway. ²Université Jean Monnet, 10 Rue Tréfilerie, 42100 Saint-Étienne, France. ³School of Ocean Sciences, Bangor University, Menai Bridge LL59 5AB, UK. ⁴Atrial Fibrillation Solutions, Medtronic plc, Mounds View, MN 55112, USA. ⁵Department of Surgery, University of Minnesota, Minneapolis, MN 55455, USA. ⁶Wildlife Science Center, 22830 Sunrise Road NE, Stacy, MN 55079, USA.

Received: 7 October 2024 Accepted: 9 January 2025

Published online: 29 January 2025

References

- Aloysius N, Geetha M. A review on deep convolutional neural networks. In: Proceedings of the 2017 IEEE International Conference on Communication and Signal Processing, ICCSP 2017, vol. 2018- January, Institute of Electrical and Electronics Engineers Inc.; 2017, p. 588–92. <https://doi.org/10.1109/ICCSP.2017.8286426>.
- Andrzejczak S, Gleiss AC, Lear KO, Pattiaratchi CB, Chapple T, Meekan M. Biologging tags reveal links between fine-scale horizontal and vertical movement behaviours in tiger sharks (*Galeocerdo cuvier*). *Front Mar Sci*. 2019;6:229. <https://doi.org/10.3389/fmars.2019.00229>.
- Battaille BC, Sakamoto KQ, Nordstrom CA, Rosen DAS, Trites AW. Accelerometers identify new behaviors and show little difference in the activity budgets of lactating northern fur seals (*Callorhinus ursinus*) between breeding islands and foraging habitats in the eastern Bering Sea. *PLoS ONE*. 2015;10:e0118761. <https://doi.org/10.1371/journal.pone.0118761>.
- Berrar D. Cross-validation. In: Encyclopedia of Bioinformatics and Computational Biology: ABC of Bioinformatics, vol. 1–3, Elsevier; 2018, p. 542–5. <https://doi.org/10.1016/B978-0-12-809633-8.20349-X>.
- Bidder OR, Campbell HA, Gómez-Laich A, Urgé P, Walker J, Cai Y, et al. Love thy neighbour: Automatic animal behavioural classification of acceleration data using the k-nearest neighbour algorithm. *PLoS ONE*. 2014;9:e0088609. <https://doi.org/10.1371/journal.pone.0088609>.
- Bidder OR, di Virgilio A, Hunter JS, McInturff A, Gaynor KM, Smith AM, et al. Monitoring canid scent marking in space and time using a biologging and machine learning approach. *Sci Rep*. 2020. <https://doi.org/10.1038/s41598-019-57198-w>.
- Blackwell BF, DeVault TL, Fernández-Juricic E, Gese EM, Gilbert-Norton L, Breck SW. No single solution: application of behavioural principles in mitigating human–wildlife conflict. *Anim Behav*. 2016;120:245–54. <https://doi.org/10.1016/j.anbehav.2016.07.013>.
- Bograd SJ, Block BA, Costa DP, Godley BJ. Biologging technologies: new tools for conservation. Introduction. *Endang Species Res*. 2010;10:1–7. <https://doi.org/10.3354/esr00269>.
- Bom RA, Bouten W, Piersma T, Oosterbeek K, van Gils JA. Optimizing acceleration-based ethograms: The use of variable-time versus fixed-time segmentation. *Mov Ecol*. 2014;2:1–8. <https://doi.org/10.1186/2051-3933-2-6>.
- Breck S, Williamson R, Niemeyer C, Shivik JA. Non-lethal radio activated guard for deterring wolf depredation in Idaho: summary and call for research. In: Proceedings of the Vertebrate Pest Conference 2002;20. <https://doi.org/10.5070/v420110182>.
- Breiman L. Random Forests. *Mach Learn*. 2001;45:1–5.
- 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. <https://doi.org/10.1186/s40317-021-00272-w>.
- Carricondo-Sanchez D, Zimmermann B, Wabakken P, Eriksen A, Milleret C, Ordiz A, et al. Wolves at the door? Factors influencing the individual behavior of wolves in relation to anthropogenic features. *Biol Conserv*. 2020;244:108514. <https://doi.org/10.1016/j.biocon.2020.108514>.
- Chakravarty P, Cozzi G, Ozgul A, Aminian K. A novel biomechanical approach for animal behaviour recognition using accelerometers. *Methods Ecol Evol*. 2019;10:802–14. <https://doi.org/10.1111/2041-210X.13172>.
- Chimienti M, Cornulier T, Owen E, Bolton M, Davies IM, Travis MJJ, et al. Taking movement data to new depths: Inferring prey availability and patch profitability from seabird foraging behavior. *Ecol Evol*. 2017;7:10252–65. <https://doi.org/10.1002/ece3.3551>.
- Chimienti M, Cornulier T, Owen E, Bolton M, Davies IM, Travis MJJ, et al. The use of an unsupervised learning approach for characterizing latent behaviors in accelerometer data. *Ecol Evol*. 2016;6:727–41. <https://doi.org/10.1002/ece3.1914>.
- Chimienti M, Desforges JP, Beumer LT, Nabe-Nielsen J, van Beest FM, Schmidt NM. Energetics as common currency for integrating high resolution activity patterns into dynamic energy budget-individual based models. *Ecol Modell*. 2020;434:109250. <https://doi.org/10.1016/j.ecolmodel.2020.109250>.
- 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;12:19737. <https://doi.org/10.1038/s41598-022-22258-1>.
- Clubb R, Mason GJ. Natural behavioural biology as a risk factor in carnivore welfare: How analysing species differences could help zoos improve enclosures. *Appl Anim Behav Sci*. 2007;102:303–28. <https://doi.org/10.1016/j.applanim.2006.05.033>.
- Cutler DR, Edwards TC, Beard KH, Cutler A, Hess KT, Gibson J, et al. Random forests for classification in ecology. *Ecology*. 2007;88:2783–92. <https://doi.org/10.1890/07-0539.1>.
- Dankowski T, Ziegler A. Calibrating random forests for probability estimation. *Stat Med*. 2016;35:3949–60. <https://doi.org/10.1002/sim.6959>.
- Dayer AA, Williams A, Cosbar E, Racey M. Blaming threatened species: media portrayal of human–wildlife conflict. *Oryx*. 2019;53:265–72. <https://doi.org/10.1017/S0030605317000783>.
- Denisko D, Hoffman MM. Classification and interaction in random forests. *Proc Natl Acad Sci U S A*. 2018;115:1690–2. <https://doi.org/10.1073/pnas.1800256115>.
- Dunford CE, Marks NJ, Wilson RP, Scantlebury DM. Identifying animal behaviours from accelerometers: improving predictive accuracy of machine learning by refining the variables selected, data frequency, and sample duration. *Ecol Evol*. 2024;14:e11380. <https://doi.org/10.1002/ece3.11380>.
- van Eeden LM, Eklund A, Miller JRB, López-Bao JV, Chapron G, Cejtin MR, et al. Carnivore conservation needs evidence-based livestock protection. *PLoS Biol*. 2018;16:e2005577. <https://doi.org/10.1371/journal.pbio.2005577>.
- Elsayad AM. Classification of ECG arrhythmia using learning vector quantization neural networks. In: Proceedings - The 2009 International

- Conference on Computer Engineering and Systems, ICCES'09, 2009; p. 139–44. <https://doi.org/10.1109/ICCES.2009.5383295>.
27. Ericsson G, Bostedt G, Kindberg J. Wolves as a symbol of people's willingness to pay for large carnivore conservation. *Soc Nat Resour*. 2008;21:294–309. <https://doi.org/10.1080/08941920701861266>.
 28. Fazil M, Firdhous M. IoT-enabled smart elephant detection system for combating human elephant conflict. In: 2018 3rd International Conference on Information Technology Research, ICITR 2018, Institute of Electrical and Electronics Engineers Inc.; 2018. <https://doi.org/10.1109/ICITR.2018.8736149>.
 29. 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. <https://doi.org/10.1186/s40317-017-0121-3>.
 30. Ferentinos KP. Deep learning models for plant disease detection and diagnosis. *Comput Electron Agric*. 2018;145:311–8. <https://doi.org/10.1016/j.compag.2018.01.009>.
 31. Fernández-Gil A, Naves J, Ordiz A, Quevedo M, Revilla E, Delibes M. Conflict misleads large carnivore management and conservation: Brown bears and wolves in Spain. *PLoS ONE*. 2016;11:e0151541. <https://doi.org/10.1371/JOURNAL.PONE.0151541>.
 32. Frézard A, Le Pape G. Contribution to the welfare of captive wolves (*Canis lupus lupus*): a behavioral comparison of six wolf packs. *Zoo Biol*. 2003;22:33–44. <https://doi.org/10.1002/zoo.10070>.
 33. Friard O, Gamba M. BORIS: a free, versatile open-source event-logging software for video/audio coding and live observations. *Methods Ecol Evol*. 2016;7:1325–30. <https://doi.org/10.1111/2041-210X.12584>.
 34. Fryda T, LeDell E, Gill N, Aiello S, Fu A, Candel A, Click C, Kraljevic T, Nykodym T, Aboyoun P, Kurka M, Malohlava M, Poirier S, Wong W. _h2o: R Interface for the 'H2O' Scalable Machine Learning Platform_. R package version 3.44.0.3; 2023. <https://github.com/h2oai/h2o-3>.
 35. Gandrud C, Allaire JJ, Russell K, Yetman CJ. NetworkD3: D3 JavaScript Network Graphs from R. 2017.
 36. Gastineau A, Robert A, Sarrazin F, Mihoub JB, Quenette PY. Spatiotemporal depredation hotspots of brown bears, *Ursus arctos*, on livestock in the Pyrenees, France. *Biol Conserv*. 2019;238:108210. <https://doi.org/10.1016/j.biocon.2019.108210>.
 37. Gehring T, M, Hawley J, E, Davidson S, J, Rossler S, T, Cellar A, C, Schultz R, N, et al. Are viable non-lethal management tools available for reducing wolf–human conflict? Preliminary results from field experiments. In: Proceedings of the Vertebrate Pest Conference; 2006. p. 22. <https://doi.org/10.5070/v422110098>.
 38. Giefer M, An L. Synthesizing remote sensing and biophysical measures to evaluate human–wildlife conflicts: the case of wild boar crop raiding in rural China. *Remote Sens (Basel)*. 2020;12:618. <https://doi.org/10.3390/rs12040618>.
 39. Giordano S, Seitanidis I, Ojo M, Adami D, Vignoli F. IoT solutions for crop protection against wild animal attacks. In: 2018 IEEE International Conference on Environmental Engineering, EE 2018—Proceedings, Institute of Electrical and Electronics Engineers Inc.; 2018, p. 1–5. <https://doi.org/10.1109/EE1.2018.8385275>.
 40. Hahn NR, Wall J, Denninger-Snyder K, Sairowua W, Goss M, Ndambuki S, et al. Staging behaviours identify spatial and temporal risk of human–wildlife conflict. *J Appl Ecol*. 2023;60:1495–506. <https://doi.org/10.1111/1365-2664.14428>.
 41. Halsey LG, Portugal SJ, Smith JA, Murn CP, Wilson RP. Recording raptor behavior on the wing via accelerometry. *J Field Ornithol*. 2009;80:171–7. <https://doi.org/10.1111/j.1557-9263.2009.00219.x>.
 42. Halsey LG, Shepard ELC, Wilson RP. Assessing the development and application of the accelerometer technique for estimating energy expenditure. *Compar Biochem Physiol A Mol Integr Physiol*. 2011;158:305–14. <https://doi.org/10.1016/j.cbpa.2010.09.002>.
 43. Hetem RS, Mitchell D, De Witt BA, Fick LG, Maloney SK, Meyer LCR, et al. Body temperature, activity patterns and hunting in free-living cheetah: biologging reveals new insights. *Integr Zool*. 2019;14:30–47. <https://doi.org/10.1111/1749-4877.12341>.
 44. Hoffman B, Cusimano M, Baglione V, Canestrari D, Chevallier D, DeSantis DL, Jeantet L, Ladds MA, Maekawa T, Mata-Silva V, Moreno-González V, Trapote E, Vainio O, Vehkaoja A, Yoda K, Zacarian K, Friedlaender A. A benchmark for computational analysis of animal behavior, using animal-borne tags. *Mov Ecol*. 2023. <https://doi.org/10.1186/s40462-024-00511-8>.
 45. Johansson M, Karlsson J, Pedersen E, Flykt A. Factors governing human fear of brown bear and wolf. *Hum Dimens Wildl*. 2012;17:58–74. <https://doi.org/10.1080/10871209.2012.619001>.
 46. Karlsson J, Eriksson M, Liberg O. At what distance do wolves move away from an approaching human? *Can J Zool*. 2007;85:1193–7. <https://doi.org/10.1139/Z07-099>.
 47. Kirchner TM, Devineau O, Chimienti M, Thompson DP, Crouse J, Evans AL, et al. Predicting moose behaviors from tri-axial accelerometer data using a supervised classification algorithm. *Anim Biotelemetry*. 2023;11:32. <https://doi.org/10.1186/s40317-023-00343-0>.
 48. Klees van Bommel J, Badry M, Ford AT, Golumbia T, Burton AC. Predicting human–carnivore conflict at the urban–wildland interface. *Glob Ecol Conserv*. 2020;24:e01322. <https://doi.org/10.1016/j.gecco.2020.e01322>.
 49. 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:e0166898. <https://doi.org/10.1371/journal.pone.0166898>.
 50. Leos-Barajas V, Photopoulou T, Langrock R, Patterson TA, Watanabe YY, Murgatroyd M, et al. Analysis of animal accelerometer data using hidden Markov models. *Methods Ecol Evol*. 2017;8:161–73. <https://doi.org/10.1111/2041-210X.12657>.
 51. Loh WY. Classification and regression trees. *Wiley Interdiscip Rev Data Min Knowl Discov*. 2011;1:14–23. <https://doi.org/10.1002/widm.8>.
 52. Lorand C, Robert A, Gastineau A, Mihoub JB, Bessa-Gomes C. Effectiveness of interventions for managing human-large carnivore conflicts worldwide: Scare them off, don't remove them. *Sci Total Environ*. 2022;838:156195. <https://doi.org/10.1016/j.scitotenv.2022.156195>.
 53. Lyons GN, Pope EC, Kostka B, Wilson RP, Dobrajc Z, Houghton JDR. Tri-axial accelerometers tease apart discrete behaviours in the common cuttlefish *Sepia officinalis*. *J Mar Biol Assoc UK*. 2013;93:1389–92. <https://doi.org/10.1017/S0025315412001245>.
 54. MacNulty D, Mech LD, Smith D. A proposed ethogram of large-carnivore predatory behaviour, exemplified by the Wolf. *J Mammal*. 2007;88:595–605.
 55. Mawdsley JR, O'Malley R, Ojima DS. A review of climate-change adaptation strategies for wildlife management and biodiversity conservation. *Conserv Biol*. 2009;23:1080–9. <https://doi.org/10.1111/j.1523-1739.2009.01264.x>.
 56. McClune DW, Marks NJ, Wilson RP, Houghton JD, Montgomery IW, McGowan NE, et al. Open access tri-axial accelerometers quantify behaviour in the Eurasian badger (*Meles meles*): towards an automated interpretation of field data. *Anim Biotelemetry*. 2014;2:1–6.
 57. McNay ME. Wolf-human interactions in Alaska and Canada: a review of the case history. *Wildl Soc Bull*. 2002;30:831–43.
 58. Mech LD. Newly documented behavior of free-ranging Arctic wolf pups. *Arctic*. 2022;75(2):272–6.
 59. Moena GK, Ordiz A, Kindberg J, Swenson JE, Sundell J, Støen OG. Behavioral reactions of brown bears to approaching humans in fennoscandia. *Ecoscience*. 2019;26:23–33. <https://doi.org/10.1080/11956860.2018.1513387>.
 60. Munden R, Börger L, Wilson RP, Redcliffe J, Brown R, Garel M, et al. Why did the animal turn? Time-varying step selection analysis for inference between observed turning-points in high frequency data. *Methods Ecol Evol*. 2021;12:921–32. <https://doi.org/10.1111/2041-210X.13574>.
 61. Nakamura M, López-Bao JV, Rio-Maior H, Roque S, Gil P, Serronha A, et al. Insights into the dynamics of wolf occupancy in human-dominated landscapes. *Biol Conserv*. 2023;286:110316. <https://doi.org/10.1016/j.biocon.2023.110316>.
 62. 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;215:986–96. <https://doi.org/10.1242/jeb.058602>.
 63. Niculescu-Mizil A, Caruana R. Predicting good probabilities with supervised learning. In: Proceedings of the 22nd international conference on Machine learning (ICML '05), Association for Computing Machinery, New York, NY, USA; 2005, p. 625. <https://doi.org/10.1145/1102351.1102430>.

64. Otsuka R, Yoshimura N, Tanigaki K, Koyama S, Mizutani Y, Yoda K, et al. Exploring deep learning techniques for wild animal behaviour classification using animal-borne accelerometers. *Methods Ecol Evol*. 2024;15:716–31. <https://doi.org/10.1111/2041-210X.14294>.
65. Pagano AM, Rode KD, Cutting A, Owen MA, Jensen S, Ware JV, et al. Using tri-axial accelerometers to identify wild polar bear behaviors. *Endanger Species Res*. 2017;32:19–33. <https://doi.org/10.3354/esr00779>.
66. Pagano AM, Williams TM. Estimating the energy expenditure of free-ranging polar bears using tri-axial accelerometers: a validation with doubly labeled water. *Ecol Evol*. 2019;9:4210–9. <https://doi.org/10.1002/ece3.5053>.
67. Patterson A, Gilchrist HG, Chivers L, Hatch S, Elliott K. A comparison of techniques for classifying behavior from accelerometers for two species of seabird. *Ecol Evol*. 2019;9:3030–45. <https://doi.org/10.1002/ece3.4740>.
68. Petroelje TR, Belant JL, Beyer DE, Svoboda NJ. Identification of carnivore kill sites is improved by verified accelerometer data. *Anim Biotelemetry*. 2020. <https://doi.org/10.1186/s40317-020-00206-y>.
69. Probst P, Boulesteix A-L. To tune or not to tune the number of trees in random forest. *J Mach Learn Res*. 2018;18:1–18.
70. R Core Team. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria; 2024. <https://www.R-project.org/>
71. Rautiainen H, Alam M, Blackwell PG, Skarin A. Identification of reindeer fine-scale foraging behaviour using tri-axial accelerometer data. *Mov Ecol*. 2022. <https://doi.org/10.1186/s40462-022-00339-0>.
72. Reis I, Baron D, Shahaf S. Probabilistic random forest: a machine learning algorithm for noisy data sets. *Astron J*. 2019;157:16. <https://doi.org/10.3847/1538-3881/aaf101>.
73. Resheff YS, Rotics S, Harel R, Spiegel O, Nathan R. AccelRater: a web application for supervised learning of behavioral modes from acceleration measurements. *Mov Ecol*. 2014. <https://doi.org/10.1186/s40462-014-0027-0>.
74. Rifkin R, Klautau A. In defense of one-vs-all classification. *J Mach Learn Res*. 2004;5:101–41.
75. Ronoh EK, Mirau S, Dida MA. Human-wildlife conflict early warning system using the internet of things and short message service. *Eng Technol Appl Sci Res*. 2022;12:8273–7.
76. Røskaft E, Händel B, Bjerke T, Kaltenborn BP. Human attitudes towards large carnivores in Norway. *Wildlife Biol*. 2007;13:172–85. [https://doi.org/10.2981/0909-6396\(2007\)13\[172:HATLCI\]2.0.CO;2](https://doi.org/10.2981/0909-6396(2007)13[172:HATLCI]2.0.CO;2).
77. Sakamoto KQ, Sato K, Ishizuka M, Watanuki Y, Takahashi A, Daunt F, et al. Can ethograms be automatically generated using body acceleration data from free-ranging birds? *PLoS ONE*. 2009. <https://doi.org/10.1371/journal.pone.0005379>.
78. Sand H, Wikenros C, Wabakken P, Liberg O. Effects of hunting group size, snow depth and age on the success of wolves hunting moose. *Anim Behav*. 2006;72:781–9. <https://doi.org/10.1016/j.anbehav.2005.11.030>.
79. Sand H, Zimmermann B, Wabakken P, Andrén H, Pedersen HC. Using GPS technology and GIS cluster analyses to estimate kill rates in wolf—ungulate ecosystems. *Wildl Soc Bull*. 2005;33:914–25. [https://doi.org/10.2193/0091-7648\(2005\)33\[914:ugtagc\]2.0.co;2](https://doi.org/10.2193/0091-7648(2005)33[914:ugtagc]2.0.co;2).
80. Sevilano-Triguero V, Talayero F, López-Bao JV, Estrella-Aguirre S. The social stereotypes of wolves and brown bears. *Hum Dimens Wildl*. 2022;28:1–16. <https://doi.org/10.1080/10871209.2022.2036392>.
81. Sha JCM, Kurihara Y, Tsuji Y, Take M, He T, Kaneko A, et al. Seasonal variation of energy expenditure in Japanese macaques (*Macaca fuscata*). *J Therm Biol*. 2018;76:139–46. <https://doi.org/10.1016/j.jtherbio.2018.07.009>.
82. Shepard ELC, Wilson RP, Halsey LG, Quintana F, Laich AG, Gleiss AC, et al. Derivation of body motion via appropriate smoothing of acceleration data. *Aquat Biol*. 2008;4:235–41. <https://doi.org/10.3354/ab00104>.
83. Shepard ELC, Wilson RP, Quintana F, Laich AG, Liebsch N, Albareda DA, et al. Identification of animal movement patterns using tri-axial accelerometry. *Endanger Species Res*. 2010;10:47–60. <https://doi.org/10.3354/esr00084>.
84. Shivik JA, Palmer GL, Gese EM, Osthaus B. Captive coyotes compared to their counterparts in the wild: Does environmental enrichment help? *J Appl Anim Welfare Sci*. 2009;12:223–35. <https://doi.org/10.1080/10888700902955989>.
85. Sievert C. Interactive Web-based data visualization with R, plotly, and shiny. Chapman and Hall/CRC. ISBN 9781138331457; 2020. <https://plotly-r.com>
86. Slagle KM, Bruskotter JT, Wilson RS. The role of affect in public support and opposition to wolf management. *Hum Dimens Wildl*. 2012;17:44–57. <https://doi.org/10.1080/10871209.2012.633237>.
87. Steen AK. Pattern recognition methods for reduction of human–wildlife conflicts. PhD Thesis. Aarhus University, Denmark; 2014.
88. Steen KA, Therkildsen OR, Karstoft H, Green O. Audio-based detection and recognition of conflict species in outdoor environments using pattern recognition methods. *Appl Eng Agric*. 2014;30:89–96. <https://doi.org/10.13031/aea.30.9817>.
89. Støen O-G, Ordiz A, Evans AL, Laske TG, Kindberg J, Frøbert O, et al. Physiological evidence for a human-induced landscape of fear in brown bears (*Ursus arctos*). *Physiol Behav*. 2015;152:244–8. <https://doi.org/10.1016/j.physbeh.2015.09.030>.
90. Thangavel S, Shokkalingam CS. The IoT based embedded system for the detection and discrimination of animals to avoid human–wildlife conflict. *J Ambient Intell Humaniz Comput*. 2022;13:3065–81. <https://doi.org/10.1007/s12652-021-03141-9>.
91. Thorsen NH, Hansen JE, Støen OG, Kindberg J, Zedrosser A, Frank SC. Movement and habitat selection of a large carnivore in response to human infrastructure differs by life stage. *Mov Ecol*. 2022. <https://doi.org/10.1186/s40462-022-00349-y>.
92. Versluis E, Niccolai LJ, Spedener M, Zimmermann B, Hesse A, Tofastrud M, et al. Classification of behaviors of free-ranging cattle using accelerometer signatures collected by virtual fence collars. *Front Anim Sci*. 2023;4:1083272. <https://doi.org/10.3389/fanim.2023.1083272>.
93. Wam HK, Eldegard K, Hjeljord O. From overlooking to concealed: Predator avoidance in an apex carnivore. *Eur J Wildl Res*. 2012;58:1001–3. <https://doi.org/10.1007/s10344-012-0670-y>.
94. 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. <https://doi.org/10.1186/s40462-015-0030-0>.
95. Wickham H. ggplot2: elegant graphics for data analysis. New York: Springer-Verlag; 2016.
96. Williams HJ, Shepard ELC, Holton MD, Alarcón PAE, Wilson RP, Lambertucci SA. Physical limits of flight performance in the heaviest soaring bird. *Proc Natl Acad Sci*. 2020. <https://doi.org/10.1073/pnas.1907360117/-DCSupplemental>.
97. Wilmers C, Nickel B, Bryce C, Smith J, Wheat R, Yovovich V. The golden age of bio-logging: how animal-borne sensors are advancing the frontiers of ecology. *Ecology*. 2015;96:1741–53.
98. Wilson RP, Holton MD, di Virgilio A, Williams H, Shepard ELC, Lambertucci S, et al. Give the machine a hand: a Boolean time-based decision-tree template for rapidly finding animal behaviours in multisensor data. *Methods Ecol Evol*. 2018;9:2206–15. <https://doi.org/10.1111/2041-210X.13069>.
99. Wilson RP, Shepard ELC, Liebsch N. Prying into the intimate details of animal lives: use of a daily diary on animals. *Endanger Species Res*. 2008;4:123–37. <https://doi.org/10.3354/esr00064>.
100. Yates LA, Aandahl Z, Richards SA, Brook BW. Cross validation for model selection: a review with examples from ecology. *Ecol Monogr*. 2023;93:e1557. <https://doi.org/10.1002/ecm.1557>.
101. Zadrozny B, Elkan C. Obtaining calibrated probability estimates from decision trees and naive Bayesian classifiers. In: Proceedings of the Eighteenth International Conference on Machine Learning (ICML '01), San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.; 2001, p. 609–16.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Terms and Conditions

Springer Nature journal content, brought to you courtesy of Springer Nature Customer Service Center GmbH (“Springer Nature”).

Springer Nature supports a reasonable amount of sharing of research papers by authors, subscribers and authorised users (“Users”), for small-scale personal, non-commercial use provided that all copyright, trade and service marks and other proprietary notices are maintained. By accessing, sharing, receiving or otherwise using the Springer Nature journal content you agree to these terms of use (“Terms”). For these purposes, Springer Nature considers academic use (by researchers and students) to be non-commercial.

These Terms are supplementary and will apply in addition to any applicable website terms and conditions, a relevant site licence or a personal subscription. These Terms will prevail over any conflict or ambiguity with regards to the relevant terms, a site licence or a personal subscription (to the extent of the conflict or ambiguity only). For Creative Commons-licensed articles, the terms of the Creative Commons license used will apply.

We collect and use personal data to provide access to the Springer Nature journal content. We may also use these personal data internally within ResearchGate and Springer Nature and as agreed share it, in an anonymised way, for purposes of tracking, analysis and reporting. We will not otherwise disclose your personal data outside the ResearchGate or the Springer Nature group of companies unless we have your permission as detailed in the Privacy Policy.

While Users may use the Springer Nature journal content for small scale, personal non-commercial use, it is important to note that Users may not:

1. use such content for the purpose of providing other users with access on a regular or large scale basis or as a means to circumvent access control;
2. use such content where to do so would be considered a criminal or statutory offence in any jurisdiction, or gives rise to civil liability, or is otherwise unlawful;
3. falsely or misleadingly imply or suggest endorsement, approval, sponsorship, or association unless explicitly agreed to by Springer Nature in writing;
4. use bots or other automated methods to access the content or redirect messages
5. override any security feature or exclusionary protocol; or
6. share the content in order to create substitute for Springer Nature products or services or a systematic database of Springer Nature journal content.

In line with the restriction against commercial use, Springer Nature does not permit the creation of a product or service that creates revenue, royalties, rent or income from our content or its inclusion as part of a paid for service or for other commercial gain. Springer Nature journal content cannot be used for inter-library loans and librarians may not upload Springer Nature journal content on a large scale into their, or any other, institutional repository.

These terms of use are reviewed regularly and may be amended at any time. Springer Nature is not obligated to publish any information or content on this website and may remove it or features or functionality at our sole discretion, at any time with or without notice. Springer Nature may revoke this licence to you at any time and remove access to any copies of the Springer Nature journal content which have been saved.

To the fullest extent permitted by law, Springer Nature makes no warranties, representations or guarantees to Users, either express or implied with respect to the Springer nature journal content and all parties disclaim and waive any implied warranties or warranties imposed by law, including merchantability or fitness for any particular purpose.

Please note that these rights do not automatically extend to content, data or other material published by Springer Nature that may be licensed from third parties.

If you would like to use or distribute our Springer Nature journal content to a wider audience or on a regular basis or in any other manner not expressly permitted by these Terms, please contact Springer Nature at

onlineservice@springernature.com