Mammalian Biology
https://doi.org/10.1007/s42991-021-00173-8

METHODOLOGY

Predicting the individual identity of non-invasive faecal and hair samples using biotelemetry clusters

Levi Newediuk[1] · **Eric Vander Wal**[1]

Received: 27 November 2020 / Accepted: 21 August 2021

Abstract

Animal diet and health influence fitness, making individual variation in these markers essential for understanding how individuals and populations respond to their environments. Faecal and hair samples provide a record of this information and can be non-invasively collected from animals in the field. However, physiology, diet, and susceptibility to parasitic infections vary within individuals, requiring repeated samples from individuals. We developed a technique using biotelemetry data for individual identification of non-invasive faecal material and hair sampled from female elk (*Cervus canadensis*). We non-invasively collected individually genotyped faecal and hair samples from resting sites, then compared the accuracy of supervised machine learning models to predict the individual identities of the samples. We found both the tightness of global positioning system point clusters and activity level surrounding the sample allowed us to positively identify samples belonging to specific individuals with 77% accuracy. Our approach can be applied to other populations for which biotelemetry data are available and is potentially adaptable for other species. Furthermore, application of our approach will reduce the need for individual identification of non-invasive samples using genetic analysis, which is costly and prone to low recovery success. Increased access to physiological, dietary, and health information obtainable from individual non-invasive samples will strengthen our understanding of animal responses to their environments.

Keywords Biotelemetry · Cluster analysis · Faecal DNA · Individual differences · Non-invasive sampling · Ungulates

Introduction

Diet and health are essential components of fitness that vary among individual animals. Thus, quantifying individual variation in diet and physiological markers is an important part of understanding how individuals and populations interact with and respond to their environments. Animal hair and faecal material contain a record of such information from a point in space and time, allowing them to be linked to the environment experienced by an individual. For example, microhistological analysis of faecal material (Hoy et al. 2019) and stable isotopes of carbon and nitrogen obtained from individual hair samples (Bryan et al. 2013) reveal how changes in environmental conditions affect differences in diet composition between populations. Using parasites shed in faecal material, the prevalence of infections in social species can also be linked to factors such as density and group size (Snaith et al. 2008). Glucocorticoid hormones help to restore homeostasis following acute exposure to stressors like predator encounters (Romero 2004), and thus faecal material with relatively high glucocorticoid concentrations is indicative of populations facing elevated predation risk (Hammerschlag et al. 2017). Because it is not always possible to examine stomach contents or blood, non-invasive sampling of hair and faecal material has become a common approach for obtaining diet (Leighton et al. 2020), parasite load (Snaith et al. 2008), and stress information from wild mammals (Sheriff et al. 2011). However, physiology, diet, and susceptibility to parasitic infections varies within

Handling editors: Elissa Z. Cameron and Leszek Karczmarski.

This article is a contribution to the special issue on "Individual Identification and Photographic Techniques in Mammalian Ecological and Behavioural Research – Part 1: Methods and Concepts" — Editors: Leszek Karczmarski, Stephen C.Y. Chan, Daniel I. Rubenstein, Scott Y.S. Chui and Elissa Z. Cameron.

✉ Levi Newediuk
ljnewediuk@mun.ca

[1] Department of Biology, Memorial University, St. John's Newfoundland A1B 3X9, Canada

Published online: 27 December 2021

populations, a complication that can be resolved by collecting repeated non-invasive samples from individuals.

Individual variation is prevalent in animals (Dingemanse et al. 2010; Guindre-Parker 2020), and at the population level it can reduce the precision of information obtained from non-invasive samples. For example, the existence of dietary specialists within a population of mostly generalists can decouple measurements of stable isotopes from measurements of local food availability (Ramos et al. 2020). Similarly, if a population exhibits a glucocorticoid response to a stressor but variation in stress responses is high, individual responses might not reflect the level of stressor to which the population is exposed (Guindre-Parker et al. 2019). In some cases, high individual variation in glucocorticoid responses can even mask detection of population level responses to the environment (Coppes et al. 2018). Individuals also differ in their susceptibility to parasitic infections depending on age and life history stage, meaning the demographic structure of populations can skew faecal pellet parasite counts (Seeber et al. 2020). Thus, while untargeted collection of non-invasive samples provides a general overview of responses to the environment, repeated samples from individual animals adds context to help interpret how the environment impacts populations.

In addition to distinguishing individual variation from population-level patterns, sampling individual animals strengthens the inferences that can be drawn from non-invasive sampling. For example, pairing individual diet samples with changes in body mass over time makes it possible to track the interacting effects of the environment and individual foraging behaviour on fitness (Giroux et al. 2016). Repeated glucocorticoid samples from individuals can reveal endocrine plasticity, which may be important for individuals to ensure reproductive success in fluctuating environments (Guindre-Parker et al. 2019). While samples from individuals provide more information than non-invasive samples collected at random from the population, planning their collection, and their subsequent assignment to individual animals, is not straightforward.

The individual identities of non-invasively collected samples can be confirmed by comparing individual genotypes of microsatellite loci recovered from hair and faecal material (Bryan et al. 2013; Jesmer et al. 2020). While some collection methods identify individuals relatively successfully (Bach et al. 2022), genetic information recovered from faecal material can be degraded by high temperatures and precipitation, often resulting in low recovery success or restricting sample collection to winter and temperate climates (Rea et al. 2016). Physiological measurements can also be attributed to known individuals by collecting faecal material after observing the individual defecate (Fattorini et al. 2018; Dulude-de Broin et al. 2019; Hunninck et al. 2020). While this individual observation technique mitigates the issue of sample degradation in genetic analysis because exposure to the elements is reduced, the substantial time investment required to observe defecation limits sample size. In comparison to faecal samples the genetic material contained in hair is stable over longer periods of time, but this stability also makes it difficult to determine when the sample was left by the individual (Lukacs and Burnham 2005). Uncertainty in the age of the sample could prevent the physiological information therein from being linked to short-term and transient environmental factors. Thus, an efficient and reliable technique for locating a large number of individually identifiable samples could disentangle the effects of individual differences from the environment.

A potential solution for linking individuals to non-invasively collected hair and faecal samples capitalizes on remote sensing of animal space use. Inference of location and movement characteristics from remotely sensed Global Positioning System (GPS) data has made it possible to identify areas used by individual mammals for parturition (Bonar et al. 2018) and foraging (McNeill et al. 2020). Tracking individual animals to these areas of high use reduces time spent in the field because areas can be prioritized for collection of samples that are more likely to belong to specific individuals. For example, Giroux et al. (2012) linked faecal material to individual GPS-collared white-tailed deer (*Odocoileus virginianus*) by following their foraging tracks. Though genetic analysis or individual observation are the only methods that can conclusively confirm the identities of individual samples (Coppes et al. 2018), the high resolution at which GPS data are now available facilitates the collection of a large number of samples that do not rely on successful extraction of genetic material. However, the performance of these techniques for targeting known individuals must be evaluated on a species-by-species basis.

Here, we developed and demonstrated a technique for individual identification of non-invasively sampled faecal material and hair from female elk (*Cervus canadensis*) using GPS data, motivated by a need to disentangle individual differences in physiology, diet, and parasite load from environmental effects. Elk are well suited for individual collection of non-invasive samples because they must stop moving while ruminating or bedding (Cook 2002), producing discrete GPS location clusters. Particularly during spring and autumn shedding (O'Gara 2002), it is common to find hair at their bedding locations. Furthermore, elk defecate at an average frequency of once every 2 h (Neff et al. 1965), meaning an individual is also likely to leave a faecal sample at location clusters where it has spent at least 2 h. This increases the probability of collecting a sample from a known individual. In spring and summer, we collected fresh faecal material and hair samples from location clusters indicative of bedding sites suspected to belong to known genetic individuals from the population. After confirming the identity of the

individual at each bedding site by genetic analysis of the samples, we compared the predictive performances of supervised machine learning models to distinguish positively and negatively identified samples based on characteristics of the target individual's GPS location clusters.

Methods

Study system

We conducted all field work in southeast Manitoba, Canada, within the traditional lands of the Anishinaabe people (49.134, – 96.557). The study area is characterized by hot summers, with approximately 300 mm of rain and temperatures regularly in excess of 30 °C from May to August. The approximately 150 adult elk in our study population (Manitoba Agriculture and Resource Development, unpublished data) inhabit an area that comprises privately owned agricultural land and public land dominated by marshes, wet hardwood forests, and shrubland. In February 2019, 18 individual females were captured using a net gun fired from a helicopter and each was fit with an Iridium satellite Global Positioning System collar (Vertex Plus 830 g, VECTRONIC Aerospace GmbH, Berlin, Germany). The collars were programmed to collect locations every 30 min from May to August 2019 and had a mean fix success rate of 89.4% (i.e. an average of 3,948 locations per individual). All capture procedures were in accordance with approved animal care protocols (Memorial University of Newfoundland animal use protocol #19-01-EV).

Sample collection and preprocessing

We collected elk faecal pellet and hair samples from May to August 2019, targeting 11 of the 18 collared individuals (hereafter "target individuals"). Collars recorded locations of the elk every 30 min during the period of collection, subjectively allowing us to identify potential bed-sites by looking for areas with a relatively large number of location points close to one another (hereafter "cluster"). After identifying clusters, we searched for recently shed hair and fresh faecal pellets with the appearance of a fresh mucus layer (Le Saout et al. 2016). We collected faecal samples in a sealable plastic bag and stored them in a – 18 °C freezer as quickly as possible following collection, since exposure to warm temperatures, typical of our study area during the sampling period, degrades the genetic material in faecal samples (Rea et al. 2016). All samples were stored in the freezer between 29 and 410 min following collection.

We attempted to increase the odds of sampling a target individual by preferentially collecting samples on or within 5 m of its suspected bed, identifiable as an area of depressed vegetation within the cluster. In many ungulates, a single bedding bout typically only lasts between one and a few hours (Cederlund 1989; Kuzyk and Hudson 2007). However, inactive bedding periods are interspersed with alternating periods of feeding in the same vicinity (Green and Bear 1990; Naylor et al. 2009), and in our study the long periods of time target individuals spent in the vicinity of their samples introduced some uncertainty about which cluster of points was recorded at the time the sample was deposited. To ensure our models were robust to this uncertainty, we specified three location points belonging to the target individual that were closest to each sample as "cluster centres" for use in our machine learning models. Finally, we counted the number of other bed sites within 20 m of the sample as a measure of the number of other elk in the same area (hereafter "activity level"; Fig. 1). We chose a search radius of 20 m as a compromise between selecting a search area larger than the accuracy of the GPS collars (8–15 m; VECTRONIC Aerospace GmbH, Berlin, Germany) and maintaining a small enough area to visually keep track of which beds we had already counted.

During elk captures, whole blood samples were taken from each collared individual to serve as a genetic benchmark. We determined whether non-invasive samples belonged to target individuals by comparing DNA extracted from them to the genetic benchmarks. We sent all faecal pellets, hair samples, and blood samples collected during capture to the Natural Resources DNA Profiling & Forensics Centre in Peterborough, Ontario for extraction and processing of genetic material. Extractions were performed on 10–15 hairs with roots or 1–2 faecal pellets subsampled from each of the original samples using the DNeasy 96 Blood and extraction protocol (Qiagen, Germany). Individuals were then identified by amplifying and sizing between 8 and 9 microsatellite loci (BM4513, BM1009, IGF, AF102257, BM4107, BM1225, BM848, BM5004, BL42) using GeneMarker software (SoftGenetics, USA; data are provided in Supplementary File 1, DNA_data.xlsx). We assigned each faecal pellet and hair sample a classification of either "positive" or "negative" based on whether its genetic material matched the target individual at the cluster.

Supervised machine learning

We used supervised machine learning to build a model that could distinguish which samples belonged to the target individual, i.e. were positively identified. Machine learning is a suite of algorithm-based techniques aimed at making predictions about "testing" data based on observed patterns in "training" data used to build the model. When supervised, the patterns in both training and testing data are known, and the machine learning algorithm seeks to maximize the predictive performance of the model. In classification

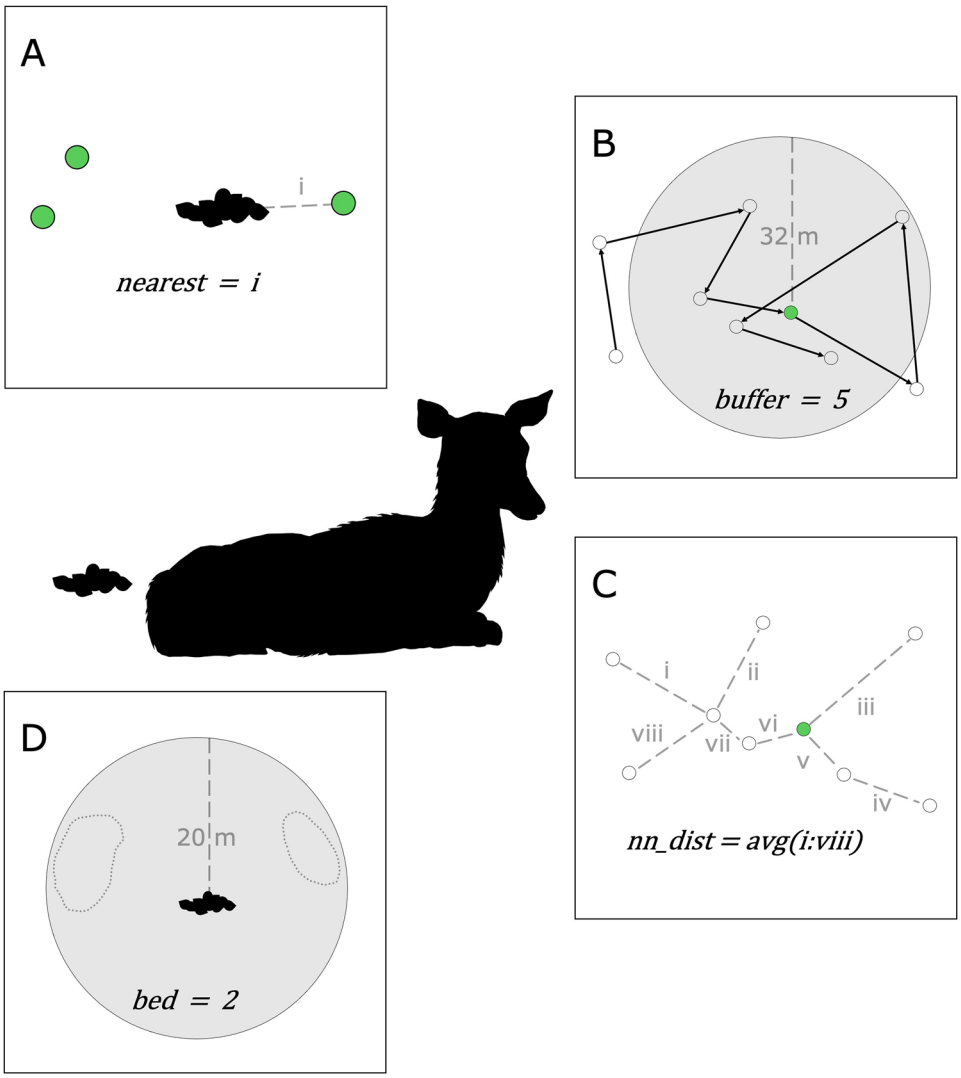

Fig. 1 Predictor variable data collected at sample locations. **A** The three nearest GPS location points to the sample are cluster centres (green circles), and *i* is the distance to the nearest point. **B** Cluster tightness is the number of points in the cluster (white circles) falling within 32 m of each cluster centre. **C** Average nearest-neighbour distance between all points in the cluster. **D** Number of beds within 20 m of the sample

applications, machine learning classifiers predict whether observations fall into one of two or more classes, such as positive and negative identifications.

To prioritize locations for sampling, we wanted to build the best model for classifying positively and negatively identified samples using only predictor variables that could be collected remotely. Based on the duration of bedding bouts in other ungulate species (Cederlund 1989), we anticipated that the use of bedding sites associated with samples would not exceed four consecutive hours. Thus, we used nine 30-min location points centred at each of the potential cluster centres, i.e. four location points before and after each cluster centre, to measure characteristics of the target individual's movement track capable of distinguishing bedding behaviour. First, we measured cluster tightness by determining how many of the eight surrounding location points fell within a 32 m buffer of the cluster centre, including the cluster centre (Fig. 1B; also see Appendix 1 for details about how we determined the buffer radius). We also calculated the average nearest-neighbour distance among each of the nine location points, including the cluster centre (Fig. 1C). In addition to movement characteristics, we used the activity level in the area of the cluster (i.e. the number of bed sites within 20 m of the sample) and proximity of the sample to each cluster centre (i.e. the nearest point to the sample from each cluster) as predictor variables. While both the activity level and proximity to cluster centre predictors must be collected in the field, they could improve the predictive performance of the model. All covariate data are provided in Supplementary File 2 (cluster_data.xlsx).

We applied five different classifiers to the training model to determine which best predicted the two classes of samples in our testing data: linear discriminant analysis (LDA), naïve Bayes (NB), K nearest neighbour (KNN), classification and regression trees (CART), and support vector machines (SVM). We also explored the random forest (RF) classifier

as an alternative regression tree method, owing to its popularity and demonstrated performance for classification (Bahn and McGill 2013). However, performance of the random forest classifier was negligibly different from CART (data not shown), so we considered only CART and not RF in our final analysis. Classifiers can be categorized based on the process by which they assign observations to classes. For example, LDA and NB compute decision boundaries between the classes based on the Bayesian probability that each observation belongs to either class (Casella et al. 2013; Genoud et al. 2020). KNN makes classifications based on the majority class of the K observations closest to a given observation, and CART iteratively builds decision trees based on the attributes of all observations, then assigns observations to classes according to the class to which the majority of observations belong on that branch of the tree (Casella et al. 2013). SVM separates classes by hyperplanes that maximize the distance between classes, with its position dependent on observations termed "support vectors" that occur along the margin of the hyperplane (Casella et al. 2013).

The performance of a machine learning classifier is based on it being able to provide a flexible fit to patterns in the training data without overfitting. Overfitting is a consequence of too much flexibility: the model fits to noise in the training data, ultimately reducing its accuracy when applied to testing data. The machine learning classifiers we used all vary in their levels of flexibility, with nonlinear methods like radial SVM being some of the more flexible, and LDA and linear SVM being some of the least flexible. In addition to selection of the classifier itself, changing the values of constants within the equation used to calculate the probability of an observation belonging to either class provides further control over the degree of flexibility. For each classifier, we tested a range of these constants, termed "tuning parameters", which adjusted the algorithms across a range of flexibility. We selected the optimal tuning parameter values to maximize the classification accuracy of each classifier. We also selected among three types of SVM models at this stage: linear, polynomial, and radial SVM range from least to most flexible based on the shape of the hyperplane that separates classes. Details about the tuning parameters tested in each classifier are provided in the section S2 of the Supplementary Material. After selecting the tuning parameters, we assessed the performance of all combinations of the five classifiers and four predictor variables based on the mean percent accuracy of their classification of the testing data.

We also compared model performance using receiver operating characteristics (ROC). ROC curves convey predictive accuracy by plotting the rate of true positive identifications against false positive identifications predicted by a model. The area under the curve (AUC) quantifies this comparison, with AUC values closer to 1 indicating higher model performance (Fawcett 2006).

Model validation

We used tenfold cross-validation with five repeats to first select the optimal tuning parameters for each algorithm, and then again to select the combination of algorithm and predictor variables that maximized model accuracy. Cross-validation is a resampling method that evaluates the predictive ability of a model by repeatedly testing it on new sets of training and testing data split from the initial data set. In tenfold cross-validation, the data are randomly partitioned into ten "folds", one of which is used for testing and the remaining nine for training. After all ten folds are used successively for both training and testing, ten new folds are partitioned and the process is repeated. Imbalance between observations in the testing and training set—such as in our study, where positively identified samples outnumbered negatively identified samples 3:1—can result in misclassification of the minority class (Liu et al. 2011). Thus, we also undersampled the majority class in each cross-validation fold, matching it to the number of samples in the minority class. We compared performance across the five models according to their mean accuracy from all cross-validation iterations. To ensure neither sample type nor collar relocation frequency influenced model accuracy, we also ran the models again with hair and faecal pellet data separated, and after having rarefied the GPS data to 1-h relocations by removing every second location in the cluster.

Tenfold cross-validation is used for supervised machine learning applications when an independent testing set is not available to test the predictive capabilities of a model (Kindschuh et al. 2016; Sánchez-González et al. 2018). However, when training and testing data are not independent, dependence structures—where nearby observations are more correlated than distant ones—can lead to overly optimistic conclusions about model performance that do not necessarily hold when the model is applied to a novel data set (Roberts et al. 2017; Gregr et al. 2018). As a solution, blocked cross-validation has been shown to produce more realistic assessments of model performance (Roberts et al. 2017). In blocked cross-validation, the data are partitioned by dependence structures rather than randomly into folds. We suspected that the predictor variable data belonging to a single elk may be correlated, and that this dependence structure may bias model performance estimates. Thus, we also performed a separate blocked cross-validation to confirm whether the accuracy of the top model agreed with tenfold cross-validation. We partitioned the 11 elk into folds and used the data from 10 elk as training data to predict the positively identified samples of the remaining elk.

Results

We collected 114 hair and faecal pellet samples between May and August 2019. Of those, 43 yielded recoverable genetic material that we could compare to known individuals in the population. While we recorded three cluster centres for 42 of 43 samples, for one of the samples it was only possible to identify a single cluster, leaving us with a total of 126 clusters in the dataset. After calculating cluster tightness and the nearest neighbour distance between cluster points, we then compared the performance of these two predictors, activity level (i.e. number of bed sites within 20 m of the sample), and distance between the cluster centre and sample for predicting whether samples were positive or negatively identified as belonging to a target individual. In total, we tested 75 different combinations of predictor variables and machine learning classifiers.

Positively and negatively identified samples differed across two of the predictor variables. There was less activity, i.e. there were fewer bed sites within 20 m of the sample, for positively identified versus negatively identified samples ($F_{1,124} = 13.91$, $p = 0.0003$; Fig. 2A). The clusters surrounding positively identified samples were also tighter than those surrounding negatively identified samples, i.e. there were more location points within 32 m of a positively identified cluster centre than those negatively identified ($F_{1,124} = 13.69$, $p = 0.0003$; Fig. 2B). However, neither the average nearest neighbour distance between cluster points ($F_{1,124} = 2.354$, $p = 0.13$; Fig. 2C) nor the proximity of the sample to the cluster centre ($F_{1,124} = 0.389$, $p = 0.53$; Fig. 2D) differed between positively and negatively identified samples.

Overall, cluster tightness and activity level were the best predictor variables for positively identified samples (Table 1). The NB classifier performed best according to tenfold cross-validation, with its predictive accuracy averaging 77% when both cluster tightness and activity level were included as predictor variables in the same model. However, the accuracy of the best model that included only remotely sensed data, i.e. cluster tightness, was less accurate at 71% (Table 1). This was largely due to a decrease in sensitivity of the model, which refers to its ability to correctly classify positive identifications, i.e. its ability to avoid false negatives. When activity level was removed, sensitivity dropped

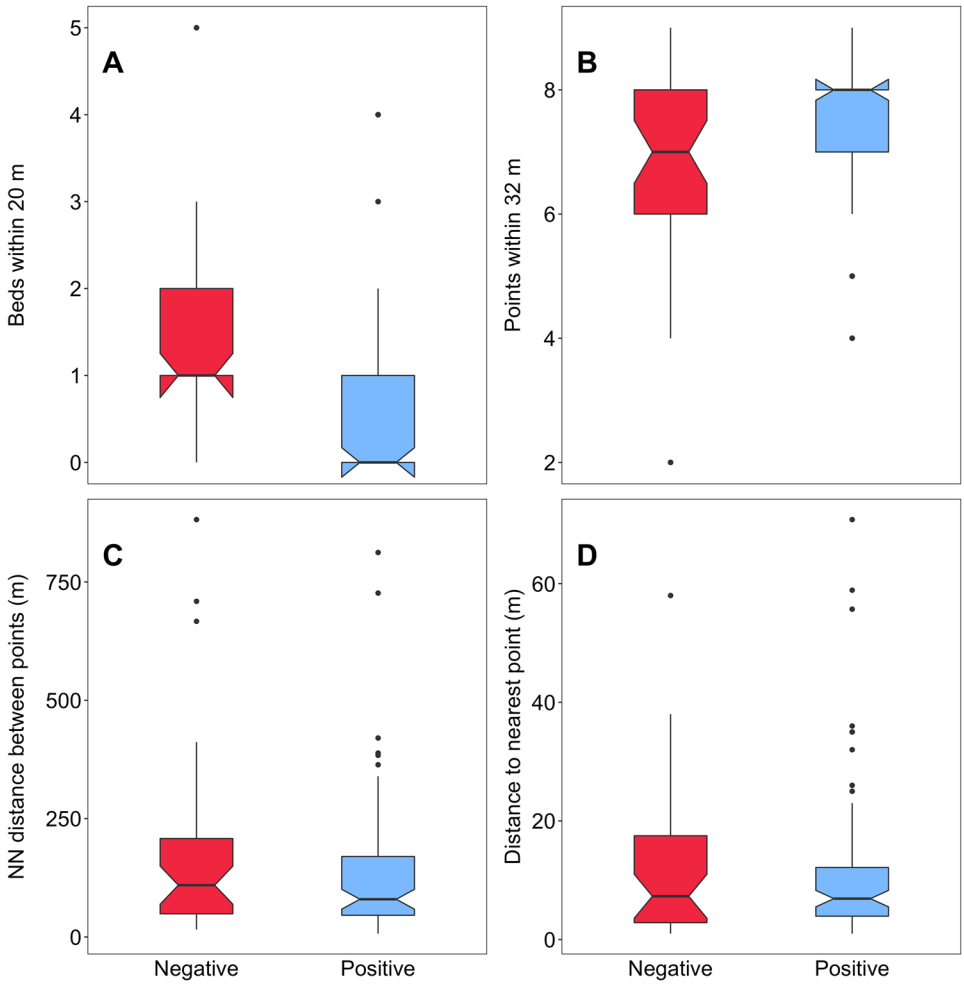

Fig. 2 Box plots showing the median (horizontal black line), quartiles (box ends and vertical lines), and outliers (points) of each predictor variable, separated by positive and negative identification of samples. **A** Compares activity level (the number of other bed sites within 20 m of the sample location), **B** compares cluster tightness (the number of points within a 32 m buffer of the sample location), **C** compares the average nearest neighbour distance among all points in a cluster, and **D** compares the distance between the sample and cluster centre

Table 1 Performance comparison of all combinations of predictor variables for distinguishing positively from negatively identified samples according to tenfold cross-validation, divided into models including both remotely sensed and site-level data versus remotely sensed data only, and ranked in order of mean percent accuracy

Predictor variables[1]	Classifier[2]	Performance metric (95% CI)			
		Percent accuracy	Percent sensitivity	Percent specificity	AUC
Remotely sensed and site level					
buffer + bed	**NB**	**76.6 (72.8, 80.3)**	**64.3 (56.3, 72.2)**	**82.6 (78.3, 86.8)**	**0.73**
Bed	LDA	74.7 (71.5, 77.8)	47.0 (39.4, 54.5)	86.4 (83.2, 89.6)	0.65
buffer + nn_dist + bed	NB	74.6 (71.1, 78.0)	55.4 (47.1, 63.6)	82.0 (78.3, 85.7)	0.69
buffer + bed + nearest	NB	74.3 (71.7, 76.9)	60.9 (54.4, 67.4)	80.5 (77.4, 83.5)	0.69
bed + nearest	LDA	73.8 (70.8, 76.8)	52.3 (45.0, 59.5)	81.5 (77.3, 85.7)	0.65
nn_dist + bed	LDA	73.2 (69.6, 76.8)	59.8 (50.9, 68.7)	79.2 (75.6, 82.9)	0.66
buffer + nn_dist + bed + nearest	LDA	71.5 (68.7, 74.3)	63.1 (54.3, 71.8)	74.2 (70.3, 78.0)	0.71
nn_dist + bed + nearest	SVM	67.7 (64.0, 71.3)	61.1 (52.4, 69.8)	70.7 (65.4, 76.0)	0.69
Remotely sensed only					
Buffer	**NB**	**70.5 (66.7, 74.3)**	**35.7 (28.9, 42.5)**	**84.7 (80.3, 89.1)**	**0.55**
buffer + nearest	NB	68.3 (65.3, 71.3)	39.5 (32.3, 46.6)	81.2 (77.7, 84.8)	0.60
buffer + nn_dist	NB	67.5 (63.7, 71.2)	35.5 (28.2, 42.7)	81.3 (76.7, 85.9)	0.58
buffer + nn_dist + nearest	NB	65.8 (62.0, 69.7)	37.9 (29.2, 46.5)	79.0 (74.3, 83.6)	0.59
Nearest	NB	60.9 (57.0, 64.8)	28.3 (19.7, 36.8)	74.4 (68.1, 80.7)	0.46
nn_dist	NB	60.3 (55.8, 64.8)	27.0 (19.5, 34.4)	76.2 (69.6, 82.8)	0.47
nn_dist + nearest	NB	59.7 (56.7, 62.6)	33.4 (26.5, 40.2)	76.6 (72.8, 80.5)	0.47

The bolded predictor variable combinations correspond to the most accurate model with both remotely sensed and site-level data (*buffer + bed*), and the most accurate model using only remotely sensed data (*buffer*)

[1]Predictor variables include *buffer* = cluster tightness (number of points within 32 m buffer of cluster centre); *nearest* = distance from sample to nearest cluster centre; *nn_dist* = average nearest neighbour distance among points in cluster; *bed* = number of beds within 20 m of sample

[2]Classifiers include naïve Bayes (NB), linear discriminant analysis (LDA), and radial support vector machines (SVM)

from 64 to 37%. However, specificity—which refers to the ability of the model to correctly classify negative identifications, i.e. its ability to avoid false positives, was still high at 85% even without the activity level predictor (Table 1). ROC curves corroborated that the combination of cluster tightness and activity level better balanced specificity and sensitivity (AUC = 0.73 versus AUC = 0.55 for cluster tightness only; Table 1, Fig. 3). For the remaining combinations of predictor variables and classifiers, neither the addition of nearest neighbour distance between cluster points nor the proximity of the sample to the cluster centre improved predictive accuracy (Table 1). Model accuracy declined by only 1% when we rarefied the GPS data to 1-h relocations (Appendix Table A2-1), and models were slightly more accurate (0.7–1%) when we modelled the hair and faecal pellet data separately (Appendix Table A2-2). However, the accuracy of the cluster tightness only model declined by > 10% when we rarefied the data to 1-h relocations.

Results from the blocked cross-validation largely agreed with tenfold cross-validation, with cluster tightness and activity level as the best predictor variables. However, the model lost accuracy and specificity in comparison to tenfold cross-validation, dropping from mean 77% accuracy to 71% accuracy, and mean 83% specificity to 74% specificity when blocked cross-validation was used (Appendix Table A2-3). In contrast, the sensitivity of the blocked cross-validation model increased in comparison to tenfold cross-validation from 64 to 71%. The model with only cluster tightness followed a similar pattern, decreasing from 71% accuracy to 63% accuracy, and 85% specificity to 65% specificity, while sensitivity increased from 36 to 53%. A full list of model predictor variables, classifiers, and their performance metrics are provided in section S2 of the Supplementary Material.

Discussion

We assigned non-invasively collected hair and faecal pellets to individual elk by capitalizing on the characteristics of bed sites identified by their GPS point clusters. This approach can improve how we interpret information from non-invasive samples because stronger inferences can be gained by accounting for variation among individual animals. Indeed, opportunistically collecting multiple samples from the same individuals in the wild is challenging (but see Giroux et al.

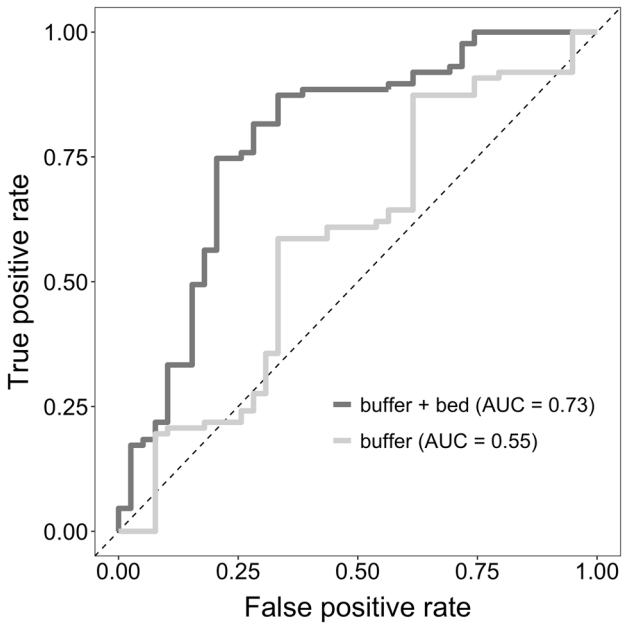

Fig. 3 Receiver operating characteristic (ROC) curves for the most accurate model (buffer + bed), and the most accurate model including only remotely sensed predictor variables (buffer). ROC curves were drawn based on the model's iterative classification of samples as positively and negatively identified. The dashed diagonal line represents a model with a random success rate of classifying positively and negatively identified samples, and curves in the upper left portion of the graph represent models with classification performance better than random (*buffer* = cluster tightness, number of points within 32 m buffer of cluster centre; *bed* = activity level, number of bed sites within 20 m of sample). Area under the curve (AUC) provides a numerical measure of model performance, where AUC = 1 indicates a model with perfect prediction capability

2016; Fattorini et al. 2018; Dulude-de Broin et al. 2019; Hunninck et al. 2020). Furthermore, labour-intensive fieldwork and financial costs associated with non-invasive sample collection from individuals can be prohibitive (Taberlet et al. 1999). In our study, supervised machine learning models including both the tightness of GPS point clusters and activity level, i.e. the number of bed sites within 20 m, at the bed sites allowed us to confirm that samples belonged to a specific individual. We also found that naïve Bayes and linear discriminant analysis outperformed more flexible classifiers like support vector machines and decision trees that are often reported to be highly accurate (Elith et al. 2008; Genoud et al. 2020), which may be the result of the more flexible classifiers overfitting the patterns in our small dataset (Raudys and Jain 1991). While other machine learning approaches have used GPS point clusters to remotely interpret elk behaviour (Van Moorter et al. 2010), ours links individuals to faecal or hair samples from which diet, physiology, and parasite load information might be obtained. We submit our approach is also general enough to adapt for other species with similar GPS point clusters, making it a promising way forward for investigating responses of individuals within animal populations to their environment.

We found that cluster tightness, measured as the number of points occurring within a 32 m buffer of the cluster centre, allowed us to distinguish positively and negatively identified samples with high accuracy. We also expected the nearest-neighbour distance between cluster points to provide another indication of the individual spending time in the vicinity of the sample. However, it did not improve the predictive performance of the model. Others have found distance measures useful for predicting carcass visitation by carnivores that exhibit different movements from bedding ungulates. For example, the maximum distance of nearby non-cluster points to the cluster was predictive of carcass type scavenged by brown bears (*Ursus arctos*; Ebinger et al. 2016). Carcass visitation by carnivores is characterized by multiple and lengthy visits to the same location, interrupted by periods of rest or other unrelated activity (Zimmerman et al. 2007; Ebinger et al. 2016). In contrast, elk and other herbivores forage at multiple locations interspersed with movements between patches and long latency to return to the same patch (Seidel and Boyce 2015). Thus, while nearest neighbour distance within a cluster may indicate a return to the carcass in carnivores, in herbivores like elk it may instead measure directed movements between different foraging patches and resting sites. These differences in their movement behaviour from carnivores would make the number of points within a 32 m buffer of the target point a more consistent predictor of herbivore location clusters than nearest-neighbour distance, and therefore a better predictor of correctly identified samples.

We also expected the distance between the nearest cluster centre and the sample to distinguish positively and negatively identified samples. While this variable did not appear in the most accurate model, we suspect it was excluded because of collar location accuracy rather than elk behaviour. While our mean collar location accuracy according to the manufacturer specifications is 8–15 m (VECTRONIC Aerospace GmbH, Berlin, Germany), the majority of distances between the cluster centre and sample in our study were well under 20 m for both positively and negatively identified samples (Fig. 3D). However, GPS collar locations are only accurate if the distance between subsequent locations is large and exceeds measurement error of the device (Jerde and Visscher 2005). This suggests that any differences in distance between positively and negatively identified samples and cluster centres might have been masked by measurement error. Indeed, Frair et al. (2005) were unable to parse movement behaviour of elk at spatial scales finer than the measurement accuracy of their GPS collars. Interestingly, this measurement error can produce the appearance of spurious 180° angles between subsequent location points even when the collar is stationary (Hurford 2009; Bjørneraas

et al. 2010). Future applications of this approach may be able to use the presence of these 180° turn angles to more precisely pinpoint the location of the target individual relative to the sample.

Our goal was to develop a model that could distinguish correctly from incorrectly identified samples without site-level characteristics to prioritize sampling locations. However, we found that activity level substantially improved model accuracy when paired with cluster tightness. Similarly, cluster models for identification of foraging in GPS-collared animals are also often improved by site-level information such as vegetation productivity (Seidel and Boyce 2015), habitat characteristics (Knopff et al. 2009), and the availability of alternate food sources (Ebinger et al. 2016). However, for those applications models without site-level characteristics are often more important because the objective is to remotely identify foraging behaviour without ground truthing. In contrast, our framework is better suited to the collection of site-level characteristics because each cluster must be visited to obtain a faecal or hair sample. Thus, as a compromise between purely remotely sensed data for prioritization of sampling locations and the need for site-level information to confirm sample identification, we suggest a two-step approach. Areas for sampling can first be targeted using the less accurate model that includes only cluster tightness as a predictor, then the positive identification of those samples was confirmed by the addition of activity level to the model. However, we caution that the accuracy of models using only remotely sensed data should be assessed for relocation frequencies less frequent than 30 min, as our remotely sensed only models were substantially less accurate when we rarefied the data to 1-h relocation frequencies.

While the cluster tightness and activity level model accurately predicted positively identified samples using tenfold cross-validation, accuracy declined when we cross-validated the model blocked by individuals. One explanation for this loss in accuracy is related to data. Unlike tenfold cross-validation where we balanced classes by undersampling, we were unable to do so for individually blocked cross-validation because 6 of the 11 individuals had either all positively or all negatively identified samples. Thus, the class imbalance in blocked cross-validation may have led to a greater number of misclassifications (Liu et al. 2011). Alternatively, the reduction in model accuracy with blocked cross-validation may be explained by the presence of additional individuals at some bed sites that was not captured by our bed site activity measure. In elk, home range size fluctuates with local competition for forage (Barker et al. 2019). However, because of selective encounters among familiar individuals, fine-scale social interactions saturate even as home range overlap continues to increase (Vander Wal et al. 2014). Thus, we may have underestimated the presence of individuals that shared space, but did not bed with our target individuals, particularly if some of the individuals occupied home ranges with higher resource availability and thus a greater density of individuals. Future versions of this analysis could test whether a variable to account for productivity at the sample location, such as habitat type or normalized difference vegetation index, improves classification accuracy.

Though our approach is appropriate for any species with periodic bedding behaviour, we only tested its performance on female elk during the calving season, raising several important considerations for its application to other systems. While female elk only isolate themselves for several days before and after parturition (Altmann 1952), they typically spend weeks following birth of their calves in smaller nursery herds with other female elk (Geist 2002), many of which were also collared in our study. Thus, the probability of our sampling a specific individual was likely different than it would have been during other seasonal periods like winter that are characterized by larger, mixed sex groups. Furthermore, male and female elk differ in both their minimum group size and group dispersion, which depend on increasing population size for female elk (Vander Wal et al. 2013). Thus, the accuracy of the models we tested may differ for male elk. The importance of predictors like activity level may also differ for less gregarious ungulate species like moose, where, for example, adult bed sites are found less frequently in close proximity (McCann et al. 2016). The duration of bedding bouts also varies seasonally and across species (Cederlund 1989; Kuzyk and Hudson 2007), and thus adjusting the temporal period represented by clusters may affect model performance.

We demonstrated that characteristics of GPS point clusters can be used to target individual female elk for collection of hair and faecal samples, providing information about physiology, diet, and parasite load. Accounting for individual differences in physiological markers like hormone levels is essential for correct interpretation of population level responses to stressors (Bonnot et al. 2018). Similarly, tracking individual differences in diet can reveal how populations of herbivores balance competition for food and cope with plant chemical defenses (Jesmer et al. 2020). Linking parasite load to age and sex of individuals can disentangle the influence of life history stage and environmental conditions on their susceptibility (Seeber et al. 2020). Our approach offers an efficient, cost-effective solution for sampling individual elk, and possibly other species fit with biotelemetry collars. For example, our approach is also applicable for other species like moose (McCann et al. 2016) and large carnivores (Knopff et al. 2009; Ebinger et al. 2016) that produce clusters of GPS locations at bedding and feeding sites. This increased access to physiological, dietary, and health information from individuals will strengthen our understanding of animal responses to their environments.

Appendix 1: Method for deciding buffer radius of cluster tightness variable

We wanted to determine which movement track characteristics were most important for distinguishing positively and negatively identified samples. In a similar cluster-based method, Knopff et al. (2009) identified kill sites using the number of cougar (*Puma concolor*) location points within a buffer of the geometric centre of a cluster. Similarly, we used the number of location points within a buffer of some radius from the cluster centre (hereafter "cluster tightness") as a predictor in our models. However, we also wanted to determine the buffer radius that maximized the difference in number of location points between positively and negatively identified samples. Thus, we created a series of buffers with radii ranging from 1 to 100 m by increments of 1 m. We first determined the average number of location points within each buffer that were associated with either positively or negatively identified samples (Fig. A1-1A). We then selected the buffer radius for the cluster tightness predictor that maximized the difference in number of location points between the positively identified samples. We determined this optimum radius to be 32 m (Fig. A1-1B).

See Fig. A1-1.

Appendix 2: Additional tables

See Tables A2-1, A2-2, A2-3.

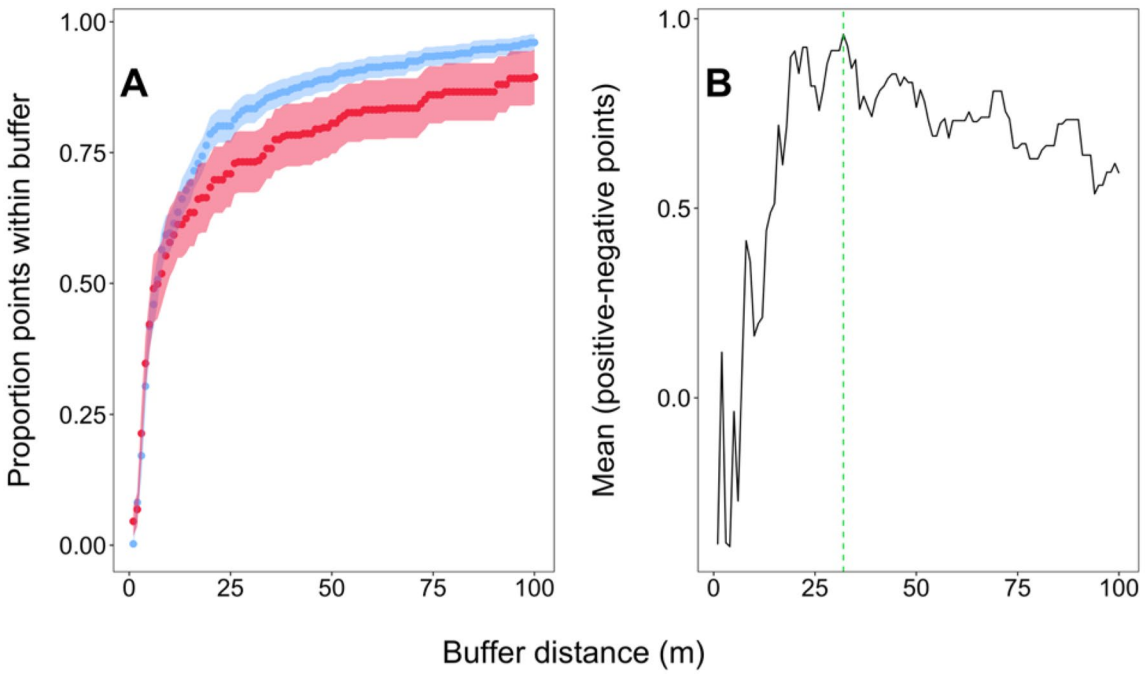

Fig. A1-1 Comparison of cluster tightness between positively and negatively identified samples, where tighter clusters have more points falling within a smaller-radius buffer surrounding the cluster centre. In **A**, the proportion of points falling within a buffer of each radius from 1 to 100 m is shown for positively identified (blue) and negatively identified (red) samples. In **B**, the mean difference in the number of points within the buffer is shown between positively and negatively identified samples. The green dashed line indicates the largest mean difference at buffer radius of 32 m

Table A2-1 Accuracy comparison between 1-h relocation frequency data and 30-min relocation frequency data for distinguishing positively from negatively identified samples

Predictor variables	30-min relocations		1-h relocations	
	Classifier	Percent accuracy (95% CI)	Classifier	Percent accuracy (95% CI)
Remotely sensed and site level				
buffer + *bed*	**NB**	**76.6 (72.8, 80.3)**	NB	75.6 (72.5, 78.8)
Bed	LDA	74.7 (71.5, 77.8)	NB	74.2 (70.9, 77.5)
buffer + *nn_dist* + *bed*	NB	74.6 (71.1, 78.0)	NB	72.3 (68.8, 75.8)
buffer + *bed* + *nearest*	NB	74.3 (71.7, 76.9)	NB	72.1 (68.2, 76.0)
bed + *nearest*	LDA	73.8 (70.8, 76.8)	NB	73.3 (70.2, 76.4)
nn_dist + *bed*	LDA	73.2 (69.6, 76.8)	**SVM**	**(76.9 (73.8, 80.0)**
buffer + *nn_dist* + *bed* + *nearest*	LDA	71.5 (68.7, 74.3)	NB	69.1 (65.4, 72.8)
nn_dist + *bed* + *nearest*	SVM	67.7 (64.0, 71.3)	NB	71.7 (68.1, 75.3)
Remotely sensed only				
Buffer	**NB**	**70.5 (66.7, 74.3)**	NB	60.1 (56.2, 63.8)
buffer + *nearest*	NB	68.3 (65.3, 71.3)	NB	56.1 (52.5, 59.8)
buffer + *nn_dist*	NB	67.5 (63.7, 71.2)	NB	59.9 (56.0, 63.8)
buffer + *nn_dist* + *nearest*	NB	65.8 (62.0, 69.7)	NB	57.2 (53.0, 61.3)
Nearest	NB	60.9 (57.0, 64.8)	NB	59.6 (55.5, 63.7)
nn_dist	NB	60.3 (55.8, 64.8)	**NB**	**61.5 (57.1, 65.8)**
nn_dist + *nearest*	NB	59.7 (56.7, 62.6)	NB	57.3 (52.6, 62.0)

Models include all combinations of predictor variables for distinguishing positively from negatively identified samples according to tenfold cross-validation. The most accurate combination of predictor variables using remotely sensed and site-level predictors, and remotely sensed predictors only, are bolded. Predictor variables include *buffer* = cluster tightness; *nearest* = distance from sample to nearest cluster centre; *nn_dist* = average nearest neighbour distance among points in cluster; *bed* = number of beds within 20 m of sample. Classifiers include naïve Bayes (NB), linear discriminant analysis (LDA), and radial support vector machines (SVM)

Table A2-2 Accuracy comparison of hair and faecal pellet models to full dataset model for distinguishing positively from negatively identified samples

Predictor variables	Full dataset		Faecal pellets		Hair	
	Classifier	Percent accuracy (95% CI)	Classifier	Percent accuracy (95% CI)	Classifier	Percent accuracy (95% CI)
Remotely sensed and site level						
buffer+*bed*	**NB**	**76.6 (72.8, 80.3)**	**NB**	**77.6 (74.2, 81.0)**	**NB**	**77.3 (74.3, 80.4)**
bed	LDA	74.7 (71.5, 77.8)	LDA	74.6 (71.8, 77.3)	NB	75.2 (71.2, 78.4)
buffer+*nn_dist*+*bed*	NB	74.6 (71.1, 78.0)	SVM	74.6 (71.7, 77.5)	NB	73.9 (70.6, 77.3)
buffer+*bed*+*nearest*	NB	74.3 (71.7, 76.9)	NB	73.0 (70.0, 76.2)	NB	74.2 (70.6, 77.7)
bed+*nearest*	LDA	73.8 (70.8, 76.8)	NB	72.4 (69.1, 75.6)	LDA	71.3 (57.4, 75.2)
nn_dist+*bed*	LDA	73.2 (69.6, 76.8)	NB	75.1 (71.8, 78.3)	SVM	76.5 (72.7, 80.3)
buffer+*nn_dist*+*bed*+*nearest*	LDA	71.5 (68.7, 74.3)	NB	73.0 (69.9, 76.2)	NB	71.0 (67.4, 74.6)
nn_dist+*bed*+*nearest*	SVM	67.7 (64.0, 71.3)	LDA	69.9 (66.3, 73.6)	LDA	72.8 (69.2, 76.5)
Remotely sensed only						
buffer	**NB**	**70.5 (66.7, 74.3)**	**NB**	**69.4 (66.1, 72.8)**	**NB**	**71.5 (68.3, 74.7)**
buffer+*nearest*	NB	68.3 (65.3, 71.3)	NB	67.6 (63.9, 71.2)	NB	67.2 (63.8, 70.5)
buffer+*nn_dist*	NB	67.5 (63.7, 71.2)	NB	66.9 (63.1, 70.8)	LDA	66.2 (62.2, 70.2)
buffer+*nn_dist*+*nearest*	NB	65.8 (62.0, 69.7)	NB	66.3 (62.4, 70.2)	NB	64.0 (60.2, 67.9)
nearest	NB	60.9 (57.0, 64.8)	NB	59.7 (55.7, 63.7)	NB	57.2 (52.6, 61.8)
nn_dist	NB	60.3 (55.8, 64.8)	NB	61.7 (57.4, 66.0)	NB	64.1 (60.4, 67.7)
nn_dist+*nearest*	NB	59.7 (56.7, 62.6)	NB	58.1 (53.6, 62.7)	LDA	57.3 (53.5, 61.1)

Models include all combinations of predictor variables for distinguishing positively from negatively identified samples according to tenfold cross-validation. The most accurate combination of predictor variables using remotely sensed and site-level predictors, and remotely sensed predictors only, are bolded. Predictor variables include *buffer*=cluster tightness; *nearest*=distance from sample to nearest cluster centre; *nn_dist*=average nearest neighbour distance among points in cluster; *bed*=number of beds within 20 m of sample. Classifiers include naïve Bayes (NB), linear discriminant analysis (LDA), and radial support vector machines (SVM)

Table A2-3 Performance comparison of all combinations of predictor variables for distinguishing positively from negatively identified samples according to individually blocked cross-validation, divided into models including both remotely sensed and site-level data versus remotely sensed data only, and ranked in order of mean percent accuracy

Predictor variables	Classifier	Performance metric (95% CI)		
		Percent accuracy	Percent sensitivity	Percent specificity
Remotely sensed and site level				
buffer + bed	**KNN**	**72.4 (53.3, 91.5)**	**70.8 (47.4, 94.9)**	**73.9 (53.8, 93.9)**
Bed	LDA	72.2 (51.6, 92.8)	61.2 (33.0, 89.3)	82.2 (60.6, 103.9)
buffer + nn_dist + bed + nearest	LDA	72.0 (56.9, 87.2)	70.5 (47.2, 93.8)	68.3 (49.7, 87.0)
nn_dist + bed	LDA	71.6 (52.3, 90.8)	60.9 (54.4, 67.4)	67.7 (42.5, 92.9)
bed + nearest	NB	70.8 (51.6, 90.0)	72.6 (47.7, 97.6)	79.8 (58.4, 101.1)
buffer + nn_dist + bed	SVM	68.5 (49.9, 87.1)	72.3 (48.4, 96.2)	65.2 (45.5, 85.0)
nn_dist + bed + nearest	LDA	68.2 (48.2, 88.2)	68.8 (41.7, 95.8)	65.2 (40.5, 90.0)
buffer + bed + nearest	LDA	63.9 (50.9, 76.9)	56.4 (33.7, 79.3)	70.2 (51.0, 89.4)
Remotely sensed only				
buffer + nearest	**NB**	**65.7 (52.9, 78.6)**	**55.7 (37.9, 73.4)**	**70.6 (50.9, 90.4)**
nn_dist + nearest	NB	63.1 (50.4, 71.3)	37.9 (26.4, 49.4)	80.4 (67.7, 93.2)
Buffer	NB	62.7 (48.0, 77.4)	41.9 (24.5, 59.3)	71.3 (49.1, 93.5)
buffer + nn_dist + nearest	NB	61.1 (46.4, 75.77)	34.1 (11.3, 56.9)	69.7 (50.0, 89.4)
nn_dist	KNN	52.9 (38.6, 67.1)	42.6 (23.8, 61.5)	59.4 (41.8, 76.9)
Nearest	LDA	52.8 (36.2, 69.3)	29.9 (8.5, 51.4)	63.6 (44.3, 82.9)
buffer + nn_dist	KNN	43.8 (35.1, 52.6)	37.9 (26.4, 49.4)	41.7 (28.6, 54.8)

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s42991-021-00173-8.

Acknowledgements K. Leavesley, K. Rebizant, H. Dettman, and D. Dupont provided logistical support for the project. E. Bishop, R. Huang, K. Kingdon, and J. Turner assisted with sample collection. M. Butler created the silhouette artwork in Fig. 1. The authors would also like to thank B. Seyler and C. Kyle of the Natural Resources DNA Profiling and Forensics Centre in Peterborough, Ontario, Canada and G. Mastromonaco of the Toronto Zoo for their helpful advice on sample collection.

Author contributions EVW and LN conceived of the study. LN led the collection of samples in the field, conducted the analysis, and led the writing of the manuscript. Both authors provided comments on the final version of the manuscript.

Funding This work was supported by a funding agreement from the Manitoba Fish and Wildlife Enhancement Fund (#5951) awarded to E. Vander Wal and L. Newediuk and a Mitacs Accelerate Fellowship (with partner organization Nature Conservancy of Canada) awarded to L. Newediuk. L. Newediuk was supported by a NSERC PGS-D.

Declarations

Conflict of interest None to report.

Ethics approval Memorial University of Newfoundland animal use protocol #19-01-EV.

Consent to participate All authors approve of the contents of the manuscript.

Consent for publication All authors consent to publication of the manuscript.

Data availability Data are available in Supplementary Files 1 & 2. Biotelemetry data are available by request on Movebank (ID 1265606810).

Code availability The code is available on Git Hub (https://github.com/ljnewediuk/elk_bedsite_clusters.git).

References

Altmann M (1952) Social behavior of elk, Cervus canadensis Nelsoni, in the Jackson Hole area of Wyoming. Behaviour 4:116–143

Bach BH, Quigley AB, Gaynor KM, McInturff A, Charles KL, Dorcy J, Brashares JS (2022) Identifying individual ungulates from fecal DNA: a comparison of field collection methods to maximize efficiency, ease, and success. Mamm Biol (Special Issue). https://doi.org/10.1007/s42991-021-00176-5

Bahn V, McGill BJ (2013) Testing the predictive performance of distribution models. Oikos 122:321–331. https://doi.org/10.1111/j.1600-0706.2012.00299.x

Barker KJ, Mitchell MS, Proffitt KM, Devoe JD (2019) Land management alters traditional nutritional benefits of migration for elk. J Wildl Manag 83:167–174. https://doi.org/10.1002/jwmg.21564

Bjørneraas K, Van Moorter B, Rolandsen CM, Herfindal I (2010) Screening global positioning system location data for errors using animal movement characteristics. J Wildl Manag 74:1361–1366. https://doi.org/10.2193/2009-405

Bonar M, Hance Ellington E, Lewis KP, Vander WE (2018) Implementing a novel movement-based approach to inferring parturition and neonate caribou calf survival. PLoS ONE 13:1–16. https://doi.org/10.1371/journal.pone.0192204

Bonnot NC, Bergvall UA, Jarnemo A, Kjellander P (2018) Who's afraid of the big bad wolf? Variation in the stress response among personalities and populations in a large wild herbivore. Oecologia 188:85–95. https://doi.org/10.1007/s00442-018-4174-7

Bryan HM, Darimont CT, Paquet PC, Wynne-Edwards KE, Smits JEG (2013) Stress and reproductive hormones in grizzly bears reflect nutritional benefits and social consequences of a salmon foraging niche. PLoS ONE 8:1–10. https://doi.org/10.1371/journal.pone.0080537

Casella G, Fienberg S, Olkin I (2013) An Introduction to Statistical Learning. Spring Texts Stat. https://doi.org/10.1016/j.peva.2007.06.006

Cederlund G (1989) Activity patterns in moose and roe deer in a North Boreal Forest. Holarct Ecol 12:39–45

Cook JG (2002) North America elk: ecology and management. In: Toweill D, Thomas J (eds) North America Elk: Ecology and Management. Wildlife Management Institute, Washington, DC, pp 259–349

Coppes J, Kämmerle JL, Willert M, Kohnen A, Palme R, Braunisch V (2018) The importance of individual heterogeneity for interpreting faecal glucocorticoid metabolite levels in wildlife studies. J Appl Ecol 55:2043–2054. https://doi.org/10.1111/1365-2664.13140

Dingemanse NJ, Kazem AJN, Réale D, Wright J (2010) Behavioural reaction norms: animal personality meets individual plasticity. Trends Ecol Evol 25:81–89. https://doi.org/10.1016/j.tree.2009.07.013

Dulude-de Broin F, Hamel S, Mastromonaco GF, Côté SD (2019) Predation risk and mountain goat reproduction: evidence for stress-induced breeding suppression in a wild ungulate. Funct Ecol. https://doi.org/10.1111/1365-2435.13514

Ebinger MR, Haroldson MA, van Manen FT, Costello CM, Bjornlie DD, Thompson DJ, Gunther KA, Fortin JK, Teisberg JE, Pils SR, White PJ, Cain SL, Cross PC (2016) Detecting grizzly bear use of ungulate carcasses using global positioning system telemetry and activity data. Oecologia 181:695–708. https://doi.org/10.1007/s00442-016-3594-5

Elith J, Leathwick JR, Hastie T (2008) A working guide to boosted regression trees. J Anim Ecol 77:802–813. https://doi.org/10.1111/j.1365-2656.2008.01390.x

Fattorini N, Brunetti C, Baruzzi C, Macchi E, Pagliarella MC, Pallari N, Lovari S, Ferretti F (2018) Being "hangry": Food depletion and its cascading effects on social behaviour. Biol J Linn Soc 125:640–656. https://doi.org/10.1093/biolinnean/bly119

Fawcett T (2006) An introduction to ROC analysis. Pattern Recognit Lett 27:861–874. https://doi.org/10.1016/j.patrec.2005.10.010

Frair JL, Merrill EH, Visscher DR, Fortin D, Beyer HL, Morales JM (2005) Scales of movement by elk (Cervus elaphus) in response to heterogeneity in forage resources and predation risk. Landsc Ecol 20:273–287. https://doi.org/10.1007/s10980-005-2075-8

Geist V (2002) North America elk: Ecology and management, in: Toweill, Dale E, Thomas, Jack Ward (Eds.), North America Elk: Ecology and Management. Wildlife Management Institute, Washington, DC, pp 389–433

Genoud AP, Gao Y, Williams GM, Thomas BP (2020) A comparison of supervised machine learning algorithms for mosquito identification from backscattered optical signals. Ecol Inform 58:101090. https://doi.org/10.1016/j.ecoinf.2020.101090

Giroux MA, Dussault C, Lecomte N, Tremblay JP, Côté SD (2012) A new way of assessing foraging behaviour at the individual level using faeces marking and satellite telemetry. PLoS ONE. https://doi.org/10.1371/journal.pone.0049719

Giroux MA, Dussault C, Tremblay JP, Côté SD (2016) Winter severity modulates the benefits of using a habitat temporally uncoupled from browsing. Ecosphere. https://doi.org/10.1002/ecs2.1432

Green R, Bear GD (1990) Seasonal cycles and daily activity patterns of Rocky Mountain elk. J Wildl Manag 54:272–279

Gregr EJ, Palacios DM, Thompson A, Chan KMA (2018) Why less complexity produces better forecasts: an independent data evaluation of kelp habitat models. Ecography (cop) 42:428–443. https://doi.org/10.1111/ecog.03470

Guindre-Parker S (2020) Individual variation in glucocorticoid plasticity: considerations and future directions. Integr Comp Biol 60:79–88. https://doi.org/10.1093/icb/icaa003

Guindre-Parker S, McAdam AG, Van Kesteren F, Palme R, Boonstra R, Boutin S, Lane JE, Dantzer B (2019) Individual variation in phenotypic plasticity of the stress axis. Biol Lett 15:1–7. https://doi.org/10.1098/rsbl.2019.0260

Hammerschlag N, Meÿer M, Seakamela SM, Kirkman S, Fallows C, Creel S (2017) Physiological stress responses to natural variation in predation risk: evidence from white sharks and seals. Ecology 98:3199–3210. https://doi.org/10.1002/ecy.2049

Hoy SR, Vucetich JA, Liu R, DeAngelis D, Peterson RO, Vucetich LM, Henderson JJ (2019) Negative frequency-dependent foraging behaviour in a generalist herbivore (Alces alces) and its stabilizing influence on community dynamics. J Anim Ecol. https://doi.org/10.1111/1365-2656.13031

Hunninck L, Palme R, Sheriff MJ (2020) Stress as a facilitator? Territorial male impala have higher glucocorticoid levels than bachelors. Gen Comp Endocrinol 297:113553. https://doi.org/10.1016/j.ygcen.2020.113553

Hurford A (2009) GPS measurement error gives rise to spurious 180° turning angles and strong directional biases in animal movement data. PLoS ONE. https://doi.org/10.1371/journal.pone.0005632

Jerde CL, Visscher DR (2005) GPS measurement error influences on movement model parameterization. Ecol Appl 15:806–810. https://doi.org/10.1890/04-0895

Jesmer BR, Kauffman MJ, Murphy MA, Goheen JR (2020) A test of the Niche Variation Hypothesis in a ruminant herbivore. J Anim Ecol. https://doi.org/10.1111/1365-2656.13351

Kindschuh SR, Cain JW, Daniel D, Peyton MA (2016) Efficacy of GPS cluster analysis for predicting carnivory sites of a wide-ranging omnivore: the American black bear. Ecosphere 7:1–17. https://doi.org/10.1002/ecs2.1513

Knopff KH, Knopff AA, Warren MB, Boyce MS (2009) Evaluating global positioning system telemetry techniques for estimating cougar predation parameters. J Wildl Manag 73:586–597. https://doi.org/10.2193/2008-294

Kuzyk GW, Hudson RJ (2007) Twenty-four-hour activity budgets of mule deer, Odocoileus hemionus, in the Aspen Parkland of East-central Alberta. Can Field Nat 121:299–302. https://doi.org/10.22621/cfn.v121i3.478

Le Saout S, Massouh M, Martin JL, Presseault-Gauvin H, Poilvé E, Côté SD, Picot D, Verheyden H, Chamaillé-Jammes S (2016) Levels of fecal glucocorticoid metabolites do not reflect environmental contrasts across islands in black-tailed deer (Odocoileus hemionus sitkensis) populations. Mammal Res 61:391–398. https://doi.org/10.1007/s13364-016-0294-9

Leighton GRM, Bishop JM, O'Riain MJ, Broadfield J, Meröndun J, Avery G, Avery DM, Serieys LEK (2020) An integrated dietary assessment increases feeding event detection in an urban carnivore. Urban Ecosyst 23:569–583. https://doi.org/10.1007/s11252-020-00946-y

Liu Y, Yu X, Huang JX, An A (2011) Combining integrated sampling with SVM ensembles for learning from imbalanced datasets. Inf Process Manag 47:617–631. https://doi.org/10.1016/j.ipm.2010.11.007

Lukacs PM, Burnham KP (2005) Review of capture-recapture methods applicable to noninvasive genetic sampling. Mol Ecol 14:3909–3919. https://doi.org/10.1111/j.1365-294X.2005.02717.x

McCann NP, Moen RA, Windels SK, Harris TR (2016) Bed sites as thermal refuges for a cold-adapted ungulate in summer. Wildlife Biol 22:228–237. https://doi.org/10.2981/wlb.00216

McNeill EP, Thompson ID, Wiebe PA, Street GM, Shuter J, Rodgers AR, Fryxell JM (2020) Multi-scale foraging decisions made by woodland caribou (Rangifer tarandus caribou) in summer. Can J Zool 98:331–341. https://doi.org/10.1139/cjz-2019-0197

Naylor LM, J. Wisdom M, G. Anthony R, (2009) Behavioral responses of North American Elk to recreational activity. J Wildl Manag 73:328–338. https://doi.org/10.2193/2008-102

Neff DJ, Wallmo OC, Morrison DC (1965) A determination of defecation rate for Elk. J Wildl Manag 29:406–407

O'Gara BW (2002) North America elk: Ecology and management. In: Toweill D, Thomas J (eds), North America Elk: Ecology and Management. Wildlife Management Institute, pp 3–65

Ramos R, Reyes-González JM, Morera-Pujol V, Zajková Z, Militão T, González-Solís J (2020) Disentangling environmental from individual factors in isotopic ecology: A 17-year longitudinal study in a long-lived seabird exploiting the Canary Current. Ecol Indic 111:105963. https://doi.org/10.1016/j.ecolind.2019.105963

Raudys SJ, Jain AK (1991) Small sample size effects in statistical pattern recognition: recommendations for practitioners. IEEE Trans Pattern Anal Mach Intell 13:252–264

Rea RV, Johnson CJ, Murray BW, Hodder DP, Crowley SM (2016) Timing moose pellet collections to increase genotyping success of fecal DNA. J Fish Wildl Manag 7:461–466. https://doi.org/10.3996/112015-jfwm-115

Roberts DR, Bahn V, Ciuti S, Boyce MS, Elith J, Guillera-Arroita G, Hauenstein S, Lahoz-Monfort JJ, Schröder B, Thuiller W, Warton DI, Wintle BA, Hartig F, Dormann CF (2017) Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. Ecography (cop) 40:913–929. https://doi.org/10.1111/ecog.02881

Romero LM (2004) Physiological stress in ecology: lessons from biomedical research. Trends Ecol Evol 19:249–255. https://doi.org/10.1016/j.tree.2004.03.008

Sánchez-González B, Barja I, Piñeiro A, Hernández-González MC, Silván G, Illera JC, Latorre R (2018) Support vector machines for explaining physiological stress response in Wood mice (*Apodemus sylvaticus*). Sci Rep 8:1–14. https://doi.org/10.1038/s41598-018-20646-0

Seeber PA, Kuzmina TA, Greenwood AD, East ML (2020) Effects of life history stage and climatic conditions on fecal egg counts in plains zebras (*Equus quagga*) in the Serengeti National Park. Parasitol Res 119:3401–3413. https://doi.org/10.1007/s00436-020-06836-8

Seidel DP, Boyce MS (2015) Patch-use dynamics by a large herbivore. Mov Ecol 3:1–10. https://doi.org/10.1186/s40462-015-0035-8

Sheriff MJ, Dantzer B, Delehanty B, Palme R, Boonstra R (2011) Measuring stress in wildlife: techniques for quantifying glucocorticoids. Oecologia 166:869–887. https://doi.org/10.1007/s00442-011-1943-y

Snaith TV, Chapman CA, Rothman JM, Wasserman MD (2008) Bigger groups have fewer parasites and similar cortisol levels: a multi-group analysis in red colobus monkeys. Am J Primatol 70:1072–1080. https://doi.org/10.1002/ajp.20601

Taberlet P, Luikart G, Waits LP (1999) Noninvasive genetic sampling: look before you leap. Trends Ecol Evol 14:323–327. https://doi.org/10.1016/S0169-5347(99)01637-7

Van Moorter B, Visscher DR, Jerde CL, Frair JL, Merrill EH (2010) Identifying movement states from location data using cluster analysis. J Wildl Manag 74:588–594. https://doi.org/10.2193/2009-155

Vander Wal E, van Beest FM, Brook RK (2013) Density-dependent effects on group size are sex-specific in a gregarious ungulate. PLoS ONE. https://doi.org/10.1371/journal.pone.0053777

Vander Wal E, Laforge MP, McLoughlin PD (2014) Density dependence in social behaviour: home range overlap and density interacts to affect conspecific encounter rates in a gregarious ungulate. Behav Ecol Sociobiol 68:383–390. https://doi.org/10.1007/s00265-013-1652-0

Zimmerman B, Wabakken P, Sand H, Pedersen HC, Liberg O (2007) Wolf movement patterns: a key to estimation of kill rate? J Wildl Manage 71:1177–1182. https://doi.org/10.2193/2006-306