

RESEARCH

Open Access



Using non-continuous accelerometry to identify cryptic nesting events of Galapagos giant tortoises

Emily Buege Donovan^{1*}, Stephen Blake^{2,3,4,5}, Sharon L. Deem^{4,5}, Patrick D. Moldowan⁵, Ainoa Nieto-Claudín^{4,5}, Freddy Cabrera⁵, Cristian Peñafiel⁵ and Guillaume Bastille-Rousseau¹

Abstract

Background Triaxial accelerometers have revolutionized wildlife research by providing an unprecedented understanding of the behavior of free-living animals. Machine learning is often applied to acceleration data to classify diverse animal behaviors across taxa. However, the high frequency, continuous data collection typically favored for behavioral classification studies often generates very large data sets, which may inhibit remote data acquisition and make data storage challenging. Coarse-frequency sampling or non-continuous bursts of acceleration data reduce these problems. To analyze such data, a suite of variables that summarize key features of the behavior of interest can be generated. These variables can then be used in numerous classification approaches, accommodating variation in data collection methods or sampling regimes. We demonstrate the potential for non-continuous accelerometer data to identify long-duration behavior and employ machine learning to classify the nesting behaviors of the critically endangered eastern Santa Cruz giant tortoise (*Chelonoidis donfaustoi*).

Results We field validated 112 nesting events from 21 giant tortoises. We then derived summary statistics based on accelerometry (e.g., overall dynamic body acceleration, metrics comparing acceleration before and after the probable event) and used them as inputs for Random Forest and Boosted Regression Tree classification algorithms. Our models produced a harmonic mean of precision and sensitivity (F1-score) of 0.91. We tested the generality of our model and found that the model performs well when applied to both novel individuals and years. The most important variable in accurately classifying data sequences was the proportion of acceleration data bursts above an activity threshold followed by the average overall dynamic body acceleration value of the bursts.

Conclusions These results demonstrate the feasibility and efficacy of using non-continuous accelerometer data to identify prolonged, biologically relevant behaviors in free-living wildlife. By using summary variables that do not require continuous sampling, this approach facilitates long-term monitoring of animal behavior. Similar methodology has potential to inform priority questions in ecology and conservation, such as predicting wildlife responses to climate change and identifying critical habitats, with applications across diverse species and behaviors.

Keywords Remote observation, Machine learning, Behavior classification, Chelonian, *Chelonoidis* spp., Reproductive monitoring

*Correspondence:

Emily Buege Donovan
Emily.Donovan@siu.edu

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



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. 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-nc-nd/4.0/>.

Background

The recent integration of GPS tracking devices with complementary sensors has advanced wildlife research, offering unprecedented insights into animal behavior and physiology [1, 2]. Triaxial accelerometers (accelerometers) are particularly useful in addressing questions beyond what can be gleaned from an animal's two-dimensional trajectory [1]. By detecting changes in gravitational or inertial acceleration in three dimensions, accelerometers provide detailed information about the dynamic movement of an animal. These data present opportunities to explore how animals allocate energy resources [3] and how they navigate and respond to their environment in ways previously inaccessible through traditional methods of studying wildlife [4, 5].

Advances in analytical methods such as Machine Learning (ML) have further advanced the utility of accelerometry data by allowing classification of animal behaviors [1, 6]. Effective animal behavior classification models often include supervised learning techniques [7] and have exhibited a wide range of complexities, from simpler algorithms such as K-Nearest Neighbor [8] to ensemble methods (e.g., Random Forest and Gradient Boosted Models [9]) to more complex deep learning methods [4, 10]. The application of ML techniques to accelerometry data interpretation has facilitated the identification of behaviors of free-living animals across a variety of activities and taxa, such as the flight characteristics of birds [11], foraging in mammals [12, 13], and spawning in fish [14].

Most accelerometry-based animal behavioral classification studies have prioritized high-frequency data collection, often obtaining a near-continuous data stream [15–18]. The assumption behind obtaining these fine-scale data to train classification models is that they will increase model accuracy [19], however high sampling frequencies often generate enormous data sets. This can present challenges in data storage on animal-borne devices, particularly when monitoring behavior over prolonged periods. Further, the logistical or budget constraints inherent in wildlife research often preclude such extensive data collection, particularly due to the costs associated with retrieving the data regularly or transmitting large volumes of data through satellite connections.

Alternatively, acceleration data can be collected at lower resolutions to alleviate some of the data storage and transmission issues associated with high sampling frequencies. Low-resolution accelerometry sampling can take two forms: data can be collected continuously, but at a decreased sampling frequency (henceforth “coarse resolution sampling”) or collected in higher resolution bursts with gaps between each burst (henceforth, “burst sampling”). Both sampling regimes have been used to

successfully classify animal behavior. Data are often referenced as coarse resolution when they are collected at a rate of one sample per second (1Hz) or less (e.g., [19]), and such sampling rates have been used to classify behaviors or activity in animals such as snowshoe hares [20], squirrels [21], freshwater turtles [22], and sharks [23]. Burst sampling is generally less common than continuous collection in behavioral identification studies [1]. Nevertheless, non-continuous data have proven useful in applications such as identifying flight modes in a gull species [24], determining activity patterns of bonefish [25], tying changes in hare behavior to landscape use [26], and understanding variation and energetic costs in migrating birds [27, 28]. However, there are inherent challenges with reducing the frequency at which data are collected or introducing gaps into the data stream, and such decisions can potentially affect what behaviors can be identified and the types of questions that the data can be used to answer.

Decisions regarding appropriate acceleration data sampling regimes should consider the duration of the activity of interest [29] in addition to logistical constraints of data collection. To adequately identify a specific motion from acceleration data, it is generally recommended that data are collected at a frequency at least twice that of the duration of the activity itself [30]. While sub-second sampling frequencies (> 1 Hz) may be required to capture behaviors that happen briefly or sporadically, such as changes in body posture (e.g., [31]), not all biologically significant behaviors occur on such short timescales. On a broader scale, accelerometers can be used to examine diel activity patterns, revealing temporal dynamics of energy expenditure across various habitats and contexts [23, 32, 33]. Further, individual movements of an animal often aggregate into larger, more complex events, which may be challenging to discern if only discrete movements are identified. For example, lekking behavior typically involves a series of multiple movements such as motor patterns, vocalizations, and physical interactions, which can span hours [34–36]. This activity may be repeated daily and may persist over multiple months during the mating season [36, 37]. Instead of classifying each movement with a continuous data stream, there is potential to identify the characteristics of acceleration data specific to prolonged behaviors at the event level.

Integrating non-continuous acceleration data into a typical classification approach remains challenging. Summary statistics derived from burst-sampled data offer promise for understanding behaviors of interest. For example, Overall Dynamic Body Acceleration (ODBA; [38]) is a metric commonly used to quantify the activity level of an organism. ODBA is calculated by summing the absolute values of acceleration changes across the three

axes that accelerometers detect over a temporal window. This involves measuring the acceleration in each of the three axes (X, Y, and Z), taking the absolute value of the changes in each axis to consider only the magnitude of the movements, and then summing these values together. By doing this over a specified period, ODBA provides a single value that represents the overall activity level during that time. More refined variables could include the periodicity of a signal or differential effects of a movement across different axes (e.g., [39]). By synthesizing key features from raw data, derived variables may provide an efficient summary of an extended period of interest, which may then be used as inputs for ML algorithms. Furthermore, this approach of summarizing the raw data enables the integration of data from different sampling regimes and may accommodate variation in how accelerometers are mounted or calibrated.

Prolonged activities of high biological importance and conservation value include those related to reproduction, such as mating, nest construction, or parturition. When and where animals choose to undertake these activities can have critical consequences for the reproductive success of individuals and demographics of populations [40]. Nest site selection is particularly critical for oviparous species that do not incubate eggs and for which environmental conditions alone determine egg and hatchling survival [41, 42]. The cryptic nesting habits and often remote habitats of chelonians (turtles) can make gathering reproductive data difficult and costly, especially for long-term or individual-based studies. For instance, a long-term study monitoring population trends of hawksbill sea turtles (*Eretmochelys imbricata*) included over 20,000 h of patrolling a single beach to identify nesting activities and nest locations [43]. A flexible tool for identification of nesting activity and location of nest sites would be invaluable for conservation efforts of populations that require special management (e.g., nest protection or egg collection) and monitoring or those that are challenging to observe for extended periods.

Galapagos giant tortoises (*Chelonoidis* spp.) exemplify these issues. Of 12 extant species, four occur on islands uninhabited by humans [44, 45], and all species nest in remote, rugged, and inaccessible areas [46]. Furthermore, the nesting period is relatively long, generally lasting from June through December, with females depositing up to five clutches in a single season [46]. Galapagos tortoise populations are still recovering from the negative impacts from previous centuries of over-exploitation by humans [47] and are currently threatened by invasive species, land use and climate change, pollution, and disease [48–53]. Recent research has broadly identified the critical role of the timing of nesting and nest site selection in recruitment success [41]. However, the evolutionary

ecology and conservation implications of interactions between nesting and environmental conditions under climate change are poorly known. Using remote monitoring tools to determine the timing and location of nesting of Galapagos tortoises could offer critical insights into their reproductive behaviors and support conservation efforts such as locating nests for physical protection from threats, including feral pigs [48]. Additionally, the nesting activity of Galapagos giant tortoises provides an excellent case study in identifying prolonged behavior from accelerometer data, as successful nesting attempts can last as long as 8–12 h [54, 55].

We obtained accelerometry data collected by data loggers mounted on the carapaces of Galapagos giant tortoises, from which we identified tortoise nesting behavior. From the sequences of tortoise activity, we derived summary variables, which served as inputs for ensemble machine learning methods to accurately predict tortoise nesting activity. This approach can overcome the logistical challenges associated with direct observation, but also lends support to the applicability of similar methodologies to other species and behaviors.

Methods

Study area and tortoise movement tracking

The Galapagos Archipelago is a group of volcanic islands located approximately 1000 km west of continental Ecuador, straddling the equator [56]. Across the archipelago, a hot-wet season is typically observed from January to May, followed by a cool-dry season from June to December [57]. Santa Cruz (986 km²) is a centrally located, human-populated island which rises to an elevation of 860 m [58]. The island's elevational gradient consists of arid lowlands and cooler, wetter upland areas, where much of the land has been converted to agriculture [59]. Santa Cruz hosts two distinct populations of giant tortoise, *Chelonoidis porteri* in the west and *C. donfaustoi* in the east, commonly referred to as the western and eastern Santa Cruz giant tortoise, respectively [60]. While both populations are of management concern, *C. donfaustoi* is listed as Critically Endangered on the IUCN Red List with a population estimated to consist of less than 600 individuals [61, 62].

Between August 2019 and September 2022, as part of the Galapagos Tortoise Movement Ecology Program [63], 21 custom-built GPS/accelerometer trackers (*e-obs* GmbH, Munich, Germany) were deployed on free-living adult female *C. donfaustoi* tortoises in and around the most intensely used known nesting zone of this population [64, 65]. Figure 1b depicts a free-roaming tortoise nesting within this intensely used nesting zone. To increase sample size, from July to September 2022, we also collected data from two captive *C. niger* x *C. becki*

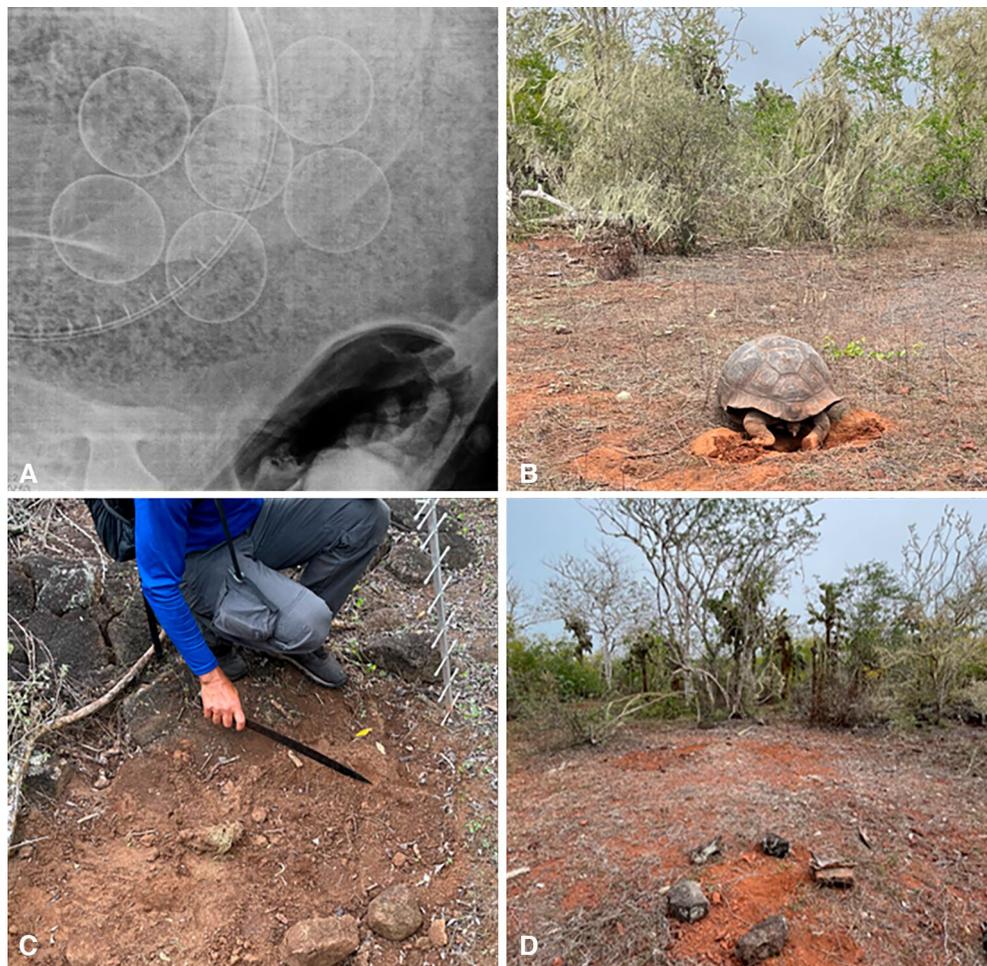


Fig. 1 **a** A typical radiograph of a gravid female *C. donfaustoi* carrying six calcified eggs in the right caudal quadrant of the coelomic cavity. **b** An unmarked female *C. donfaustoi* observed constructing a nest near the study site. **c** Researcher Freddy Cabrera identifying the cap of a potential nest site by examining variation in soil consistency and the presence of integrated fecal material. **d** A validated nest site marked by a loose ring of stones

hybrid individuals maintained at the Fausto Llerena Breeding Center in the town of Puerto Ayora on Santa Cruz Island. Devices were affixed to the front of the carapace using nontoxic plumber's epoxy (Fix-It Stick Epoxy Putty, Oatey, Cleveland, OH, USA). The trackers logged a GPS coordinate once every hour with an accuracy of approximately 10 m. The data loggers also housed a tri-axis accelerometer, which was programmed to record bursts of acceleration data at one of two sampling schedules: a 5.4-s, 10-Hz burst of data every 5 min, or a 2.0-s, 20-Hz burst every 10 min.

Nesting training data collection

During the 2022 and 2023 nesting seasons, the acceleration and movement data of tracked tortoises were remotely downloaded to a hand-held base station approximately once per week from late June through

early December. Data can be downloaded to the base station at distances up to 2 km if unobstructed, however rugged terrain in the field generally restricts downloads to occur within several hundred meters or less.

Gravidity assessments

Radiography has been used successfully to assess fecundity in other free-living chelonians [66–69]. Approximately once every 3 weeks throughout the duration of the nesting season, females were located using the VHF or UHF radio frequency emitted from the data logger. Upon location, each tortoise was radiographed using a portable X-ray generator (MinXRay, Northbrook, IL) to identify the presence of oviductal eggs (Fig. 1a). This schedule minimized radiation exposure to the tortoises while ensuring that a clutch would not be missed between radiography events. For female tortoises laying

multiple clutches, interclutch intervals ranged from 27 to 69 days, with an average of 38.9 days (E.B. Donovan, unpublished data).

Identification of nesting events

Possible windows of nesting activity for each female were identified by comparing sequential radiographs. For example, if the first radiograph of a particular female showed a clutch of fully shelled eggs, and the subsequent radiograph showed no eggs or only follicles, then the female nested between imaging dates. We then visually examined the acceleration data for possible nesting events based on overall activity from the accelerometer and GPS relocations. We identified nights with possible

nesting activity by looking for persistent nocturnal activity (i.e., a large proportion of accelerometer bursts with variable raw values, Fig. 2) and minimal movement in geographical space according to GPS relocations. For possible nest locations, we averaged latitudes and longitudes of the points associated with increased accelerometer activity. To exclude activity associated with movement to and from the nesting site, a nesting event was considered to have begun when sequential points were less than 10 m apart and ended when subsequent points were greater than 10 m apart.

Points identified as potential nesting locations were then validated in the field by navigating to each estimated nest location and searching within a 5-m buffer. Nests

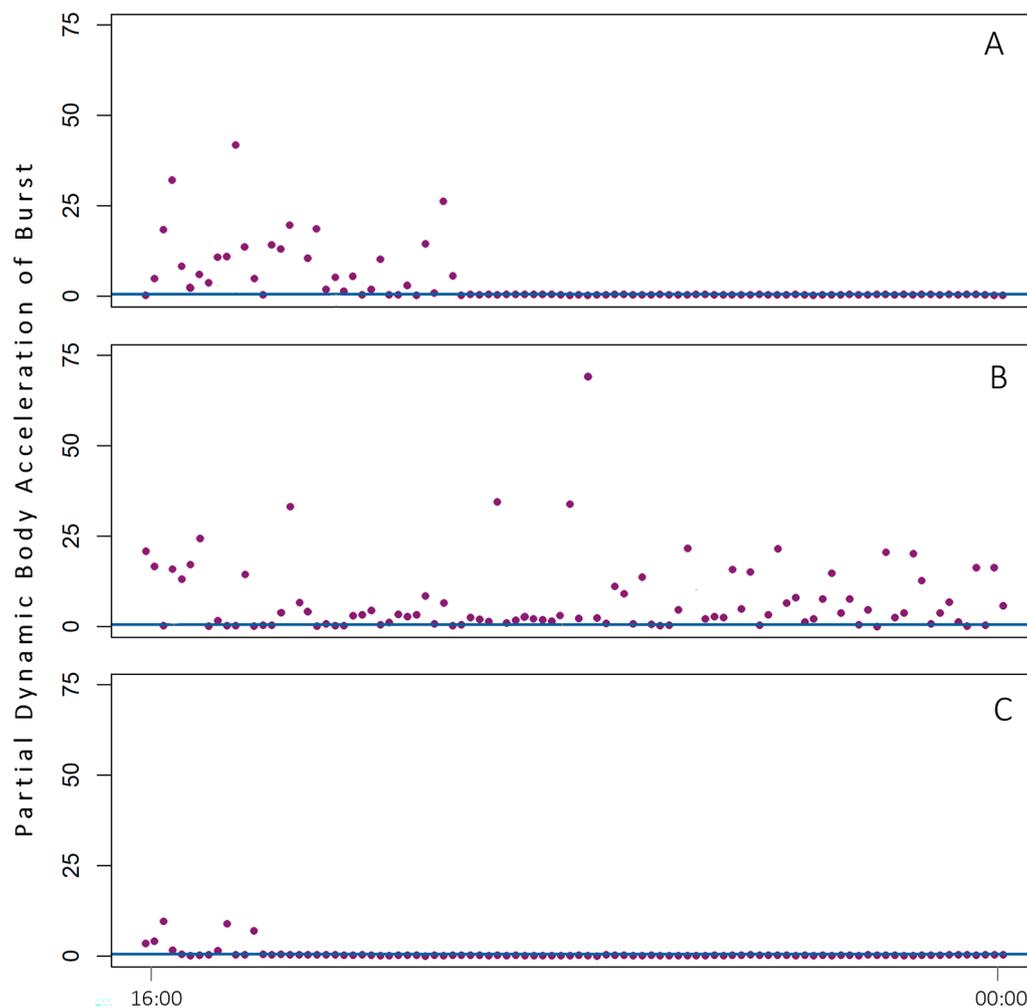


Fig. 2 Graphical representation of one axis of accelerometer data from a data logger attached to a female Galapagos giant tortoise over three consecutive nights. Bursts of acceleration are 5.4 s in length every 5 min from 16:00 to 00:00 and are plotted sequentially across the X-axis. The Partial Dynamic Body Acceleration corrected for the number of recorded values (cPDBA) of each burst is represented by purple points with values referencing the Y-axis. The threshold cPDBA value determined to separate active from inactive bursts is represented by the horizontal blue line at 0.584. This sample of graphical data includes a confirmed nesting event (panel **B**), along with the evenings preceding (panel **A**) and following (panel **C**) the event

were identified by the presence of a nest cap, a compaction of soil above the egg chamber which the female creates by mixing the substrate with urine and feces [46]. At each site, the cap was first identified visually, then confirmed by tapping the ground with a machete handle to listen for differences in soil density (Fig. 1c). Confirmed nest sites were marked with a loose ring of stones and a wooden identification marker (Fig. 1d). All nests were opened upon first detection, around the time when incubation was estimated to be complete, or both to verify clutch size [41]. We also compared the number of eggs present and, if applicable, any abnormalities of eggs (i.e., eggs of atypical shapes or sizes, conjoined eggs) in the nest to the eggs observed in radiograph images to confirm the maternal identity of a clutch.

Algorithm development

We trialed nesting event detection algorithms using Random Forest [70] and Boosted Regression Tree [71] classifiers. Random Forest (RF) is a machine learning technique which utilizes multiple decision trees. RF allows the model to generalize by first bootstrapping the dataset and then building a decision tree for each bootstrapped dataset using a subset of randomly selected variables. Boosted Regression Trees (BRT) is another ensemble learning method that combines multiple weak learners, typically decision trees, to create a strong learner. While RFs build multiple decision trees independently and combine their predictions through averaging or voting, BRTs train trees sequentially, with each new tree focusing on the instances that the previous ones misclassified. This iterative process adjusts the model's emphasis on misclassified data points, gradually improving overall performance. However, boosting tends to be more sensitive to outliers and can potentially overfit the training data. RFs are more commonly applied to classification of animal behavior from accelerometer data ([72–74] but see [9]).

Summary variable derivation

Historical observation in the field has shown that female Galapagos tortoises typically begin nesting approximately 2 h before sunset (sunset is between 17:45 and 18:05 during the nesting season) and persist into nighttime hours, with events lasting as long as 8 to 12 h in total [54, 55]. Based on the estimated start and end times of confirmed nesting events in the field, our empirical data generally support these figures (E. Donovan, unpublished data). We calculated the summary statistics for six time periods each day (15:00–21:00, 16:00–21:00, 16:00–22:00, 16:00–23:00, 16:00–0:00, and 16:00–1:00) and compared RF and BRT outputs to determine which yielded the most

accurate model. Herein, a given night of tortoise activity being analyzed is referred to as a “period of interest”.

We developed a suite of features to be used as predictor variables based on a priori assumptions of their relevance and predictive power for the machine learning model. These metrics were primarily based on the assumption that dynamic acceleration would be increased for more prolonged periods when a tortoise was engaged in nesting activity and that overall tortoise activity prior to nesting would vary from activity after a successful nesting event [75–78]. A complete list of the 11 features used as predictor variables and a description of each is found in Table 1. All of these summary statistics are based on Overall Dynamic Body Acceleration (ODBA, see [38]), which is the sum of the absolute values of the dynamic movements across all three axes (X, Y, and Z). Because of the variation in frequencies at which data were collected (see above; section *Study area and tortoise movement tracking*), we corrected ODBA by dividing this metric by the total number of raw accelerometer values collected over the period of interest (cODBA, Supp Equation SE1). To determine whether axes are differentially affected by digging activity, we included Partial Dynamic Body Acceleration for each axis (cPDBA_x, cPDBA_y, cPDBA_z, Supp Equation SE2), which were also corrected for the number of accelerometer values included in the period of interest.

To provide a measure of consistency of activity throughout the period, we determined the cODBA for each acceleration data burst (Supp Equation SE3.1). We then classified the burst as being associated with activity or inactivity through a gaussian-mixture model using Package *mclust* in R [79]. Mixture-models are a method used to identify natural groupings (clusters) within data. We determined a threshold value over which the burst was considered active by calculating the point where the probability of belonging to one of the two Gaussian-distributed clusters becomes equal to or switches from one to the other, using their means and variances. A histogram showing the distribution of burst-wise cODBA values and the threshold can be found in Supp Fig. 1. We then calculated the proportion of bursts within the period of interest that were active (active_prop, Supp Equation SE5).

Digging activity may not always lead to the successful completion of a nest. A female may abandon attempts to excavate a nest, either because of site unsuitability (e.g., she encounters too many rocks) or because of an external disturbance [54, 80]. We therefore included metrics which compared the cODBA of the period of interest to the cODBA for periods before and after with the intention of reducing the likelihood that the algorithm will erroneously identify an abandoned nesting attempt.

Table 1 Feature variables used to predict tortoise nesting activity

Variable	Description	Equations
cODBA	Overall dynamic body acceleration (ODBA) of a given night of tortoise activity ("period of interest"), corrected for the number of values recorded	SE1
cPDBA_x	Partial dynamic body acceleration (PDBA) of the period interest on the x-axis only, corrected for the number of values recorded	SE2
cPDBA_y	Partial dynamic body acceleration (PDBA) of the period interest on the y-axis only, corrected for the number of values recorded	SE2
cPDBA_z	Partial dynamic body acceleration (PDBA) of the period interest on the z-axis only, corrected for the number of values recorded	SE2
burst_avg	Average cODBA for each burst of accelerometer activity within the period of interest	SE3.1, SE3.2
burst_sd	Standard deviation of ODBA of accelerometer activity bursts	SE3.1, SE4
active_prop	Proportion of accelerometer bursts where cODBA exceeds a threshold value and are classified as active	SE3.1, SE5
greatest_ODBA	Binary variable indicating whether the cODBA of the period of interest is the greatest when compared to 7 nights prior and 7 nights after	
greatest_ODBA_aft	Binary variable indicating whether the cODBA of the period of interest is the greatest when compared to 7 nights after	
avg_ODBA_aft	Average cODBA of the 7 nights following the period of interest	
ODBA_ratio	Ratio of average cODBA 7 nights prior to the evening of interest to the average cODBA 7 nights after	

Feature variables that were extracted from the acceleration data to predict tortoise nesting activity through Random Forest and Boosted Regression Tree machine learning methods. Relevant equations are referenced in the "Equations" column and can be found in the supplementary materials

These metrics included whether the period of interest exhibited the greatest cODBA among the 7 days before and 7 days after (*greatest_ODBA*) and whether it exhibited the greatest cODBA only in the 7 days after (*greatest_ODBA_aft*). We also compared the averages of the cODBA values for these time periods (*ODBA_ratio*). Finally, we included the average cODBA value over the 7 days after the period of interest (*avg_ODBA_aft*). Because the interclutch interval in our study was a minimum of 27 days (average 38.9 days), accelerometry data from periods of 7 days prior to or following periods of interest served as a non-nesting data baseline against which to compare accelerometry data.

Model training and evaluation

We used the program R (version 4.2.3, R Core Team [81]) to extract the predictor variables from the acceleration data and to run the models. RF models were constructed using package 'randomForest' [82] with 500 trees, and BRT models were built in package 'gbm' [83] with 500 trees and a learning rate of 0.1. We extracted all predictor variables for each nesting event validated in the field and gave them a value of 1 for the classification category. To train the model, we extracted the same variables 14 nights before and 14 nights after the successful nesting event and gave those events a value of 0. We chose to draw the training data from the 28-day period surrounding each nesting event because, given the 27-day minimum interclutch interval, the period is unlikely to include a second nesting event for a given female. This

also ensured that tortoise activity both leading up to and following a nesting event were included as training data.

Four confirmed nests (one laid in 2022 and three laid in 2023) were excluded from the training data, as the female nested outside of the period of interest, finishing nest construction prior to the start of all time windows considered for the model. An additional four presumed nesting events (three laid in 2022 and one laid in 2023) were excluded from the training data sets because the presence of a nest was not validated in the field. All eight instances were retained as potential validation data.

We tested three different methods for separating training and validation data sets: temporally [84], by individual [85], and randomly [86]. For temporal validation, we used one year of data to train the model and reserved the other year for validation. Since we had two years of nest data, this method involved two iterations—one for each year. In the training data set for a given year, we retained only the confirmed nesting events and the 14 days before and after each event. For the validation data set, we kept all data from the nesting season (June 1 to December 31) of that year, without the same restriction to nesting events. For individual validation, we randomly selected 70% of individuals to train the model and used the remaining 30% for validation. The training data set for individual validation only included the nesting events and the associated 14 days before and after. The validation data set contained all available data from the nesting season for the remaining individuals. To form the data set for random validation, we randomly selected 70% of all nesting events, along with their corresponding 14 days

before and after. All remaining dates from the nesting season in the overall data set that were not used for training were assigned to the validation set. Since individual tortoises often nested more than once, nesting events from the same tortoise could appear in both the training and validation sets. Individual and random validations were run with five iterations each.

To determine the best model, we calculated three metrics:

1. Precision, or the proportion of true positive predictions among all positive predictions made by the model:

$$\text{Precision} = \frac{\text{True positives}}{\text{True positives} + \text{False positives}}.$$

2. Sensitivity, also known as recall or true positive rate, measures the ability of a model to correctly identify positive instances out of all actual positive instances in the dataset:

$$\text{Sensitivity} = \frac{\text{True positives}}{\text{True positives} + \text{False negatives}}.$$

3. F1-score, which is the harmonic mean of precision and sensitivity, was used as a measure for best overall performance. The F1-score ranges from 0 to 1, with values closer to 1 being more predictive (i.e., better model fit):

$$F1 = \frac{2 \times \text{Precision} \times \text{sensitivity}}{\text{Precision} + \text{sensitivity}}.$$

Using the average F1-score across validation methods, we determined which algorithm (RF or BRF) and time window were overall the most successful in classifying nesting events. For that model, we assessed the importance of predictor variables using the mean decrease in Gini impurity index (“Mean Decrease Gini”) across all decision trees in the ensemble. The Gini impurity index measures how often a randomly chosen element would be incorrectly classified based on the distribution of labels in the data subset used for training [87]. For BRT models, we used a similar metric which is referred to as Variable Importance [71].

For the top-performing model, we examined misclassified events and identified commonalities among them. To assist with assigning potential reasons for misclassification by the algorithm, we examined the estimated start and end times of the event, as well as the variables that were used as predictors in the model. Ultimately, we identified seven categories. For data misclassified as a result of Type I errors (false positives, non-nesting event

identified as nesting by the model), the categories were “abandoned nesting attempt” and “undetected at time of sampling.” For Type II errors, we identified three categories that were related to the timing of nesting: the event began after the start of the period of interest, the event concluded early in the period of interest, or the event occurred entirely outside of the period of interest. The two remaining categories for Type II errors were “unable to locate” (in which the event was identified as nesting based on radiographs and manual review of the acceleration data but was not subsequently validated in the field) and “unknown.” We presented the frequency of Type I and Type II errors in a confusion table.

Results

Using a combination of radiographs and tri-axis accelerometer data, we identified a potential 116 unique nesting events across two field seasons ($n=53$ in 2022, $n=63$ in 2023), where 112 resulted in a validated nest site. Validation of the nest occurred an average of 14 days after it was laid (median=10 days, range=1–63 days). Among the 112 validated nests, 109 were laid by free-living *C. donfaustoi* tortoises and three were laid by tortoises in human care at the Fausto Llerena Breeding Center. Of the 21 free-living female tortoises that were actively tracked, 19 produced one or more nests that were validated in the field during the two-year study period. The remaining two free-living tortoises were either absent from the nesting zone entirely or did not produce a clutch of eggs while present.

The F1-score (harmonic mean of precision and sensitivity) for all classification algorithms was between 0.70 and 0.91 when averaged across iterations (Table 2). The temporal window that performed best varied by validation method, where the best period was 16:00–01:00 for temporal validation, 16:00–23:00 for individual validation, and 16:00–00:00 for random validation (Table 2). Notably, the performance across these three time windows (16:00–11:00, 16:00–00:00, and 16:00–01:00) exhibits considerable overlap when considering all model iterations. This overlap suggests that the observed differences between these time windows may be marginal. As such, the observed differences between these three time windows may arise in part from random variability. When the F1-scores were combined across all validation methods, the RF model from 16:00 to 23:00 was the best model overall. RF models systematically outperformed BRT models when trained and tested on the same data set (Table 2), but the two model types differed little in their sensitivities (Additional file 1: Table S1). Both RF and BRT models generally had a higher precision than sensitivity. On average, 93% of the events classified as nesting events were indeed nesting events (precision)

Table 2 Classifier output comparison for the tortoise nest detection algorithm

Time period	Algorithm	F-1 Score			Overall average
		Temporal	Individual	Random	
15:00–21:00	BRT	0.78 (0.77–0.78)	0.83 (0.78–0.87)	0.70 (0.55–0.78)	0.770
	RF	0.84 (0.79–0.88)	0.85 (0.81–0.89)	0.75 (0.65–0.79)	0.813
16:00–21:00	BRT	0.85 (0.81–0.89)	0.84 (0.81–0.86)	0.73 (0.67–0.82)	0.810
	RF	0.87 (0.80–0.94)	0.90 (0.84–0.94)	0.78 (0.74–0.83)	0.837
16:00–22:00	BRT	0.86 (0.84–0.87)	0.88 (0.84–0.93)	0.78 (0.75–0.82)	0.827
	RF	0.86 (0.81–0.91)	0.90 (0.88–0.93)	0.82 (0.78–0.86)	0.860
16:00–23:00	BRT	0.88 (0.87–0.88)	0.88 (0.81–0.94)	0.77 (0.72–0.83)	0.843
	RF	0.89 (0.86–0.91)	0.91 (0.85–0.95)	0.81 (0.77–0.86)	0.867
16:00–00:00	BRT	0.89 (0.88–0.89)	0.90 (0.85–0.93)	0.81 (0.78–0.86)	0.860
	RF	0.89 (0.88–0.91)	0.90 (0.86–0.93)	0.85 (0.79–0.90)	0.883
16:00–01:00	BRT	0.89 (0.88–0.91)	0.85 (0.75–0.89)	0.82 (0.79–0.86)	0.870
	RF	0.90 (0.88–0.93)	0.86 (0.83–0.90)	0.84 (0.81–0.86)	0.880

Boosted Regression Tree (BRT) and Random Forest (RF) classifier output comparison for the tortoise nest detection algorithm from triaxial accelerometer data. Training data include 112 confirmed nesting events by 21 female Galapagos giant tortoises. We report the F-1 scores (harmonic mean of Precision and Sensitivity) for data summarized over six biologically relevant time periods and three validation methods: withholding a year of data ("Temporal"), withholding data from random individuals ("Individual"), and withholding a random subset of data ("Random"). The value reported is the average across iterations with the range in brackets. The highest performing model for each validation method is shown in bold

and 84% of all nesting events were properly identified by the model (sensitivity) (Additional file 1: Supp Table 1).

Of the summary variables that were extracted from the acceleration data to predict tortoise nesting activity, the proportion of acceleration data bursts above the activity threshold (active_prop) was consistently and substantially the most important (Fig. 3). The average ODBA value of the bursts (burst_avg) was the second most important, followed by the average ODBA for evenings after the event (avg_ODBA_aft). The corrected ODBA value (cODBA) for the night of interest was the fourth most important, and cODBA_y was the most important single-axis partial dynamic body acceleration metric over cODBA_x and cODBA_z. Comparative ODBA metrics were marginally important (ODBA_ratio, greatest_ODBA and greatest_ODBA_aft).

For the best model (16:00–00:00), we identified the events which were misclassified across all three validation methods we tested (both iterations of temporal validation and five iterations each for individual and random validation). For each validation method, models generated more Type II errors (nesting events classified as non-nesting events) than Type I errors (non-nesting events erroneously classified as nesting events) (Table 3). There were 26 events that were misclassified (19 Type II errors, 7 Type I errors, Fig. 3); 58% of the Type I misclassification instances came from two discrete events that were outside of the period in which radiographs were conducted in 2022. The remaining Type I errors were attributed to nesting abandoned attempts by females leading up to a successful event. In two of the five misclassifications, the

time that elapsed between the misclassified attempt and the successful event was 14 days, which is outside of the windows used to identify training data for each nesting event. Among the Type II errors, ten events were likely misclassified due to the event occurring at an atypical time of day. Most frequently, these events ended relatively early in the time window, between 17:00 and 19:00. Other events were misclassified because they occurred entirely outside of the period of interest, generally in the morning hours. A single event began late in the period.

Discussion

We derived a suite of summary variables from non-continuous accelerometer data and used a machine learning framework to reliably identify nesting activity in Galapagos giant tortoises. Despite being derived from accelerometer data collected using a burst sampling regime, the summary statistics we used as inputs for the Random Forest (RF) and Boosted Regression Tree (BRT) models produced high precision and sensitivity in distinguishing nesting from non-nesting behavior. The high F1-scores and adequate temporal and individual cross-validation performance of our models suggest that this could be a valuable method for classifying long-duration behaviors from coarse or non-continuous acceleration data.

We found that RF algorithms generally outperformed BRT in our application (Table 1). One possible explanation for this is that RF models are less prone to overfitting, especially when dealing with noisy data [88]. Given an increase in the size of the training data set, the performance of the two methods we employed here may begin

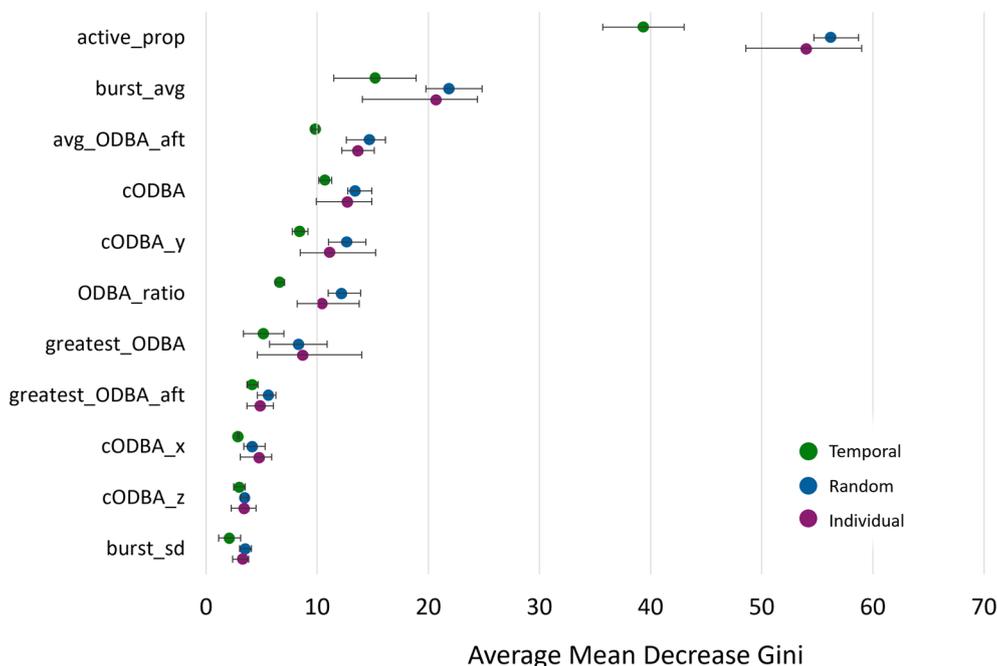


Fig. 3 Variable importance of the top-performing Random Forest model in classifying nesting activity of 21 Galapagos giant tortoises from triaxial acceleration. Values are reported as Mean Decrease Gini, a measure of the importance of each variable in predicting classifications across the ensemble of trees. Variables of a higher decrease Gini value indicate the variable made a greater contribution to model fit. Averages across iterations are shown with horizontal bars indicating the range. Validation methods include withholding a year (“Temporal”), random individuals (“Individual”), and a random subset of data (“Random”). Refer to Table 1 for detailed covariate descriptions

Table 3 Confusion matrix for the top-performing tortoise nest detection algorithm

Predicted		Temporal		Individual		Random	
		0	1	0	1	0	1
Actual	0	3536	3	2495	2	4704	4
	1	9	49	6	36	7	32

Confusion matrix for the top-performing Random Forest model in classifying giant tortoise nesting activity from triaxial acceleration data using three validation strategies. Strategies include withholding a year (“Temporal”), random individuals (“Individual”), and a random subset of data (“Random”). Reported values represent averages across iterations. Values of 0 indicate non-nesting activity and values of 1 indicate nesting events

to converge, and the best technique is likely to vary based on the details of a data set and the type of behavior being classified. While deep learning neural networks can also be powerful and popular classification tools, the relatively small number of nesting events we obtained to train the model influenced the decision not to consider these approaches. Moreover, deep learning generally requires increased computational capacity, yet simpler ensemble ML methods often produce similar or better classification in animal behavior in studies where both techniques are included for comparison [4, 89, 90]. On the opposite end of the model complexity spectrum, simpler approaches such as k-means clustering [8, 91] or using thresholds based on metrics including ODBA [20, 92]

have been successfully applied in other studies to distinguish between behaviors. However, we found that multiple metrics (e.g., active_prop, avg_ODBA_aft; Fig. 3) were important in distinguishing nesting from non-nesting sequences. This was perhaps especially important in distinguishing abandoned nesting attempts from successful events. Thus, relying on a single threshold-based variable would likely be insufficient. Nevertheless, we recommend that researchers assess multiple approaches or algorithms to determine which yields optimal performance for a specific case [4, 9].

The top-performing model identified the proportion of active acceleration bursts (active_prop) during the period of interest as the most important predictor of

tortoise nesting (Fig. 4). Galapagos tortoises are generally assumed to be largely inactive at night (and therefore inactive during typical nesting hours [63]). Although we initially assumed that a metric like cODBA would effectively capture the increased movement rate associated with nesting and readily identify a nesting event, early testing revealed that increased movement alone was insufficient. This may be because a mean-based metric like ODBA can be misleading, as it does not distinguish between persistent, small movements and occasional large movements. The proportion of active bursts was therefore crucial in identifying the type of prolonged, continuous activity associated with nesting and distinguishing it from short-term, high-intensity movement that a tortoise might also be doing during that time. We are uncertain about the nature of these short-term, high-intensity movements, as we did not validate other behaviors in the current study. However, these signatures in the accelerometry data could include movement over rugged terrain in the last hours of the day that overlap with the beginning of the temporal window we used.

Not only are there energetic costs to nesting [75], but there are also physiological constraints regarding the time to subsequently produce a new clutch of eggs. This can lead to marked variations in activity patterns before, during, and after oviposition [76]. Increased nocturnal behavior preceding oviposition, but not after, has been documented in ornate box turtles (*Terrapene ornata*; [77]). Indeed, we observed increased nocturnal activity in the nights preceding nesting in our study based on both GPS and acceleration data, but tortoises resumed

nocturnal inactivity following nesting (see Fig. 2). Thus, the average ODBA for evenings after the period of interest (avg_ODBA_aft) emerged as an important predictor, along with the other comparative ODBA metrics (ODBA_ratio, greatest_ODBA, and greatest_ODBA_aft). These metrics may be especially useful in distinguishing the actual nesting event from abandoned nesting attempts in which a tortoise starts nesting but abandons the nest prior to oviposition.

While our models generally performed with high accuracy, misclassifications among novel data may present significant ramifications. In conservation applications, erroneously identified nesting locations may increase the field efforts required to validate nests, recover eggs, or protect nests. Our models generally had higher precision than sensitivity (Supp Table 1), which suggests that falsely identifying a nest (i.e., a false positive) is less likely than missing one entirely (i.e., a false negative). Given the species' conservation status, failure to protect nests may outweigh the additional fieldwork required to address false positives. As a result, the model may need to be adjusted to prioritize sensitivity. Moreover, errors in identifying true nesting events may affect estimates of fecundity or nesting phenology. Research or conservation goals will ultimately determine whether the model's sensitivity is adequate or if these estimated errors can be mitigated through additional modeling. One potential solution for differentiating true and false positives could be applying a biological filter based on the physiological limitations of the species. For example, we determined from radiographs and field validation of nesting events

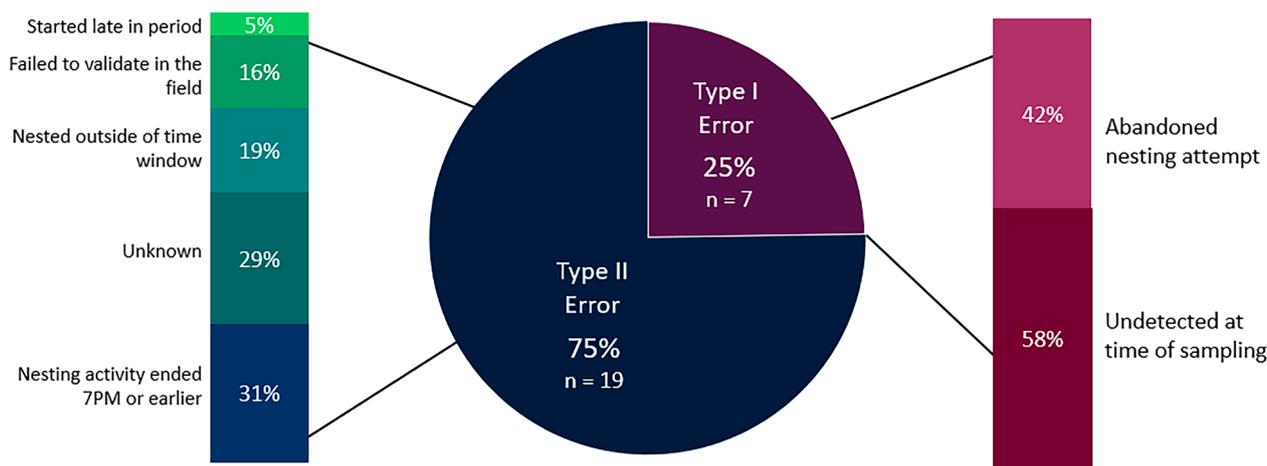


Fig. 4 Presumed reasons for misclassification of giant tortoise nesting events for the top Random Forest models constructed from triaxial accelerometer data. Type I errors were events that were not confirmed as nesting events in the field, but were classified as nesting events by the models. Type II errors consist of events identified as nesting activity through manual review of acceleration data which were classified as non-nesting by the models. There were 26 unique events included in the misclassified data, but events may have been misclassified multiple times across model iterations. Percentages include the total number of instances that a given event was misclassified.

that *C. donfaustoi* requires an average of 36 days to re-clutch. This information allowed for reclassification of several Type I errors as “abandoned nesting attempts” (Fig. 4). Conversely, there were two events that the algorithms consistently classified as nesting, yet they were not confirmed during field validation. Considering factors such as the minimum times required for re-clutching and the geographical range occupied by the females during the events, there is a strong likelihood that these unconfirmed events were indeed nesting activities that were overlooked during the sampling period.

One considerable limitation to the design of this study is that, given the importance of metrics which compare activity before and after nesting events, there is limited opportunity to validate how well the model distinguishes between non-nesting, pre-nesting, and post-nesting behaviors. In assigning the presumed reason for misclassified events, five were identified as probable abandoned nesting attempts. However, we did not validate this behavior and cannot presently quantify how commonly this behavior occurs. While our models successfully differentiate between nesting and non-nesting behaviors, it may conflate these additional, closely related behavioral categories due to the overlap in behavioral signatures. Depending on research objectives, practitioners should consider accounting for these distinctions, perhaps by expanding validation efforts to include these nuanced behaviors. Identification of digging motions within the accelerometer data (e.g., [93]) and further exploration into the differences in activity levels before and after known nesting events could be avenues to shed light on pre- or post-nesting activities and associated energy expenditure [76, 78].

To examine generalizability of the models, we trialed three cross-validation strategies: temporal, individual, and random. Of the three strategies, random validation had the poorest predictive performance. This result was unexpected considering random selection of validation data often improves error estimates, even inflating the estimates among structured ecological data [85, 94]. However, given that a common objective of training behavioral classifier models is to enable observation of animals without labeled data, evaluation of the model’s performance on novel individuals is inherently more meaningful. This is of particular interest for application among species that have remote sub-populations, as training data can be collected from individuals that are more accessible for observation and the model can subsequently be applied to unlabeled data. The models built using temporally delimited data also performed well. This provides promising applications for long-term monitoring of the same individuals. Labeled data can be collected until sufficient to train the models, after which

reproductive monitoring can proceed without the need for on-the-ground observation. This approach offers the opportunity to assess behavior through time and various climatic conditions, and ultimately identify responses to climate change. Both the temporal and individual generalizability of our model present opportunities to explore inter- and intra-individual variation.

While continuous, coarse-frequency sampling and burst sampling both address challenges related to data storage and transmission, the choice of regime can influence the types of behaviors that can be identified. Because burst data typically still collected at high-frequency resolutions, these data can provide more confidence in the discrete behaviors that occur at the time of sampling. Therefore, burst data could be used to assign fine-scale behaviors (e.g., walking, foraging) to specific bursts, potentially leading to more options in future analyses. In our study system, long-term accelerometer data have been used to determine circadian and circannual activity patterns [32], but also have the potential to be used in the identification of shorter-duration behaviors. In the present study, the degree to which metrics not based directly on the ODBA of the event of interest (e.g., `active_prop`, `avg_ODBA_aft`) underscores the potential for burst sampling regimes to distinguish behaviors in other species, even when overall movement rates are low or the animal exhibits minimal change in body posture. However, burst data could also miss critical aspects of animal’s behavior if collection is too sparse. Yu et al. [95] demonstrated the diminishing accuracy of behavioral classifications with increasing intervals between bursts, especially for more infrequent behaviors. Coarse-frequency data collection might help integrate rarer behavior that burst sampling can miss. Overall, the approach based on summary variables are flexible enough to be applied to both types of data.

One possible avenue for bridging the gap between the demand for high-resolution continuous acceleration data and data storage limitations is on-board summarization of accelerometer metrics. On-board processing of acceleration data has previously been used to inform GPS fix rates such that fewer fixes are acquired while the animal is presumed to be resting [5]. This technique has been used to derive summarized metrics of body acceleration for windows of time [39, 95, 96] and also to directly classify behavior [97]. One advantage to this method is that the storage requirements and satellite relay capabilities for summarized values is much smaller than that of raw accelerometry values. However, this approach may be most effective for long-term data storage if the metrics for identifying behaviors of interest are known prior to deployment of the device. If the raw values are erased as they are processed to maximize storage potential, the

flexibility of what can be derived from the data is limited. Yu et al. [97] successfully alleviated storage capacity issues by relying on Bluetooth and cellphone towers, though this infrastructure is more challenging to use in remote locations with limited cellphone reception.

While there are a number of studies which have successfully derived activity budgets or identified short-duration animal behavior from coarse-frequency [1, 9, 98, 99] or burst-sampled [1, 11, 100–102] accelerometer data, relatively few have applied these sampling regimes to longer-duration or aggregated behaviors. Nevertheless, several studies have shown the utility of these methods in detecting prolonged animal behaviors. For instance, Schreven et al. [103] combined GPS tracking and accelerometry to detect nesting attempts in Arctic-breeding pink-footed geese (*Anser brachyrhynchus*), successfully applying non-continuous accelerometer data to the remote identification of nest sites, incubation behavior, and nesting success. In a similar application in Greenland white-fronted geese (*Anser albifrons flavirostris*), Ozsanlav-Harris et al. [104] tested the sensitivity of incubation behavior models to reductions in the frequency of both GPS fixes and accelerometer bursts. Accelerometer-only models constructed with the smallest interval between acceleration bursts (6 min) and the largest interval tested (144 min) both obtained a precision greater than 0.9, although models using shorter intervals were generally more predictive than longer intervals [104]. Non-continuous accelerometer data have also previously been used to successfully identify behaviors associated with lekking in little bustards (*Tetrax tetrax*; [35]). Considering the long reproductive season of many lekking species, adequate monitoring through human observation can be costly and potentially affect animal behavior or lek attendance [105]. Remote identification of incubation or lekking behavior could be useful in identifying the effect of climatic conditions or landscape changes on these important reproductive activities [35] or assist in the location of previously unknown lekking grounds or nesting areas for conservation purposes.

Another potential application for burst or coarse-frequency sampling could be in examining long-duration behaviors of animals that spend extended periods underground or in a dense structure (e.g., a beaver's lodge). Biologging devices deployed on these animals may have dramatically decreased fix success rates [106], and thus data storage and transmission may be of concern. Prolonged, biologically relevant activities, such as burrow construction, could potentially be detected and analyzed using similar techniques as we employed here. Accelerometer data could aid in differentiating energetically demanding activities occurring in these cryptic locations from other vital long-duration behavior

exhibited by many animals such as rest. Using accelerometer data a continuous sampling frequency of 20 Hz, Mortlock et al. [107] were able to detect sleep in wild boar (*Sus scrofa*) with 98% accuracy. However, the ability to decipher body posture during periods of low accelerometer activity was key in differentiating sleep from wakeful rest. Further exploration into the efficacy of the model on down-sampled data could reveal how gaps in the data like those in burst sampling would affect estimates of sleep.

We recommend that practitioners seeking to apply similar methodologies consider the biology and ecology of their species of interest. In the case of identifying chelonian nesting activity, the time window approach we employed here assumes that the species displays some predictability in nesting phenology. This would work well for other chelonians with similar daily activity patterns and nesting ecologies (e.g., sea turtles which nest nocturnally [108]). However, some species typically nest diurnally [109], meaning that other activities, such as foraging, temporally overlap with nest construction. In these instances, it may be necessary to adjust data sampling frequency, the choice in predictor variables, or the inclusion of data from paired sensors, such as GPS, to disentangle these behaviors. Regardless, the effort required to identify the ideal input parameters for a remote nest identification tool for novel chelonian species may be worthwhile, as chelonians are among the most rapidly declining vertebrate groups [110].

Conclusions

Our study presents a successful example of classifying biologically meaningful behavior using summary statistics derived from non-continuous accelerometer data. Through the application of ensemble machine learning models, we reliably identified nesting behavior, a critical determinant of reproductive success in giant tortoises and many other species of birds, reptiles and fish. We recommend further exploration of the generalizability of these models and techniques across different species and populations. Elucidating long-term patterns and variation in reproductive activity can help predict how animals may respond to environmental variation, including anthropogenic disturbance. Reliable identification of important locations for reproductive success in space and time can enable managers to improve the efficacy of conservation actions for giant Galapagos tortoises and potentially numerous other wildlife species.

Abbreviations

BRT	Boosted Regression Trees
ML	Machine Learning

ODBA Overall Dynamic Body Acceleration
RF Random Forest

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40317-024-00387-w>.

Additional file 1.

Acknowledgements

We thank the Galapagos National Park Directorate for permission to conduct this study. This publication is contribution number 2685 of the Charles Darwin Foundation for the Galapagos Islands. We thank Gislayne Mendoza and Anne Guezou for their assistance with fieldwork.

Author contributions

EBD, SB, SLD, and GBR conceived the idea and designed the methodology. FC and CP led data collection efforts in the field and were assisted by all other authors. EBD conducted the analyses and wrote the initial manuscript. All authors contributed to the development of the manuscript and approved of it for publication.

Funding

This research was supported by Southern Illinois University, the Max Planck Institute for Animal Behaviour (Radolfzell, Germany), the National Geographic Society (CRE grant No. WWW-048R-17 awarded to SB), e-obs GMBH, the Saint Louis Zoo Institute for Conservation Medicine, the Houston Zoo, and the Galapagos Conservation Trust.

Data availability

All tortoise GPS and accelerometry data used in this study are publicly available through www.movebank.org within the Galapagos Tortoise Movement Ecology Programme study.

Declarations

Ethics approval and consent to participate

This research was conducted under Southern Illinois University at Carbondale Animal Care and Use Protocol #21-021. This work was also properly permitted through the Galapagos National Park (permit numbers PC-25-22 and PC-37-23).

Consent for publication

Not applicable.

Competing interests

We declare that the authors have no competing interests.

Author details

¹Cooperative Wildlife Research Laboratory, Southern Illinois University, Carbondale, IL, USA. ²Department of Biology, Saint Louis University, Saint Louis, USA. ³Max Planck Institute for Animal Behavior, Radolfzell, Germany. ⁴Saint Louis Zoo, Saint Louis, USA. ⁵Charles Darwin Foundation, Santa Cruz, Galapagos, Ecuador.

Received: 18 July 2024 Accepted: 14 October 2024
Published online: 11 November 2024

References

- Brown DD, Kays R, Wikelski M, Wilson R, Klimley AP. Observing the unwatchable through acceleration logging of animal behavior. *Anim Biotelemetry*. 2013;1:1–16.
- Shepard EL, Wilson RP, Quintana F, Laich AG, Liebsch N, Albareda DA, et al. Identification of animal movement patterns using tri-axial accelerometry. *Endanger Spec Res*. 2008;10:47–60.
- Halsey LG, Shepard EL, Wilson RP. Assessing the development and application of the accelerometry technique for estimating energy expenditure. *Comp Biochem Physiol A Mol Integr Physiol*. 2011;158(3):305–14.
- Nathan R, Spiegel O, Fortmann-Roe S, Harel R, Wikelski M, Getz WM. Using tri-axial acceleration data to identify behavioral modes of free-ranging animals: general concepts and tools illustrated for griffon vultures. *J Exp Biol*. 2012;215(6):986–96.
- Brown DD, LaPoint S, Kays R, Heidrich W, Kummeth F, Wikelski M. Accelerometer-informed GPS telemetry: reducing the trade-off between resolution and longevity. *Wildl Soc Bull*. 2012;36(1):139–46.
- Wang G. Machine learning for inferring animal behavior from location and movement data. *Eco Inform*. 2019;49:69–76.
- Bergen S, Huso MM, Duerr AE, Braham MA, Schmuecker S, Miller TA, et al. A review of supervised learning methods for classifying animal behavioural states from environmental features. *Methods Ecol Evol*. 2023;14(1):189–202.
- 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(2):e88609.
- Ladds MA, Thompson AP, Slip DJ, Hocking DP, Harcourt RG. Seeing it all: evaluating supervised machine learning methods for the classification of diverse otariid behaviours. *PLoS ONE*. 2016;11(12):e0166898.
- Browning E, Bolton M, Owen E, Shoji A, Guilford T, Freeman R. Predicting animal behaviour using deep learning: GPS data alone accurately predict diving in seabirds. *Methods Ecol Evol*. 2018;9(3):681–92.
- Williams HJ, Shepard E, Duriez O, Lambertucci SA. Can accelerometry be used to distinguish between flight types in soaring birds? *Anim Biotelemetry*. 2015;3:1–11.
- Iwata T, Sakamoto KQ, Takahashi A, Edwards EW, Staniland IJ, Trathan PN, et al. Using a mandible accelerometer to study fine-scale foraging behavior of free-ranging Antarctic fur seals. *Mar Mamm Sci*. 2012;28(2):345.
- Harvey-Carroll J, Carroll D, Trivella C-M, Connelly E. Classification of African ground pangolin behaviour based on accelerometer readings: validation of bio-logging methods. *Anim Biotelemetry*. 2024;12(1):22.
- Clarke TM, Whitmarsh SK, Hounslow JL, Gleiss AC, Payne NL, Huveneres C. Using tri-axial accelerometer loggers to identify spawning behaviours of large pelagic fish. *Mov Ecol*. 2021;9(1):26.
- Fehlmann G, O'Riain MJ, Hopkins PW, O'Sullivan J, Holton MD, Shepard EL, et al. Identification of behaviours from accelerometer data in a wild social primate. *Anim Biotelemetry*. 2017;5:1–11.
- 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(6):3030–45.
- Vázquez Diosdado JA, Barker ZE, Hodges HR, Amory JR, Croft DP, Bell NJ, et al. Classification of behaviour in housed dairy cows using an accelerometer-based activity monitoring system. *Animal Biotelemetry*. 2015;3:1–14.
- Wang Y, Nickel B, Rutishauser M, Bryce CM, Williams TM, Elkaim G, et al. Movement, resting, and attack behaviors of wild pumas are revealed by tri-axial accelerometer measurements. *Mov Ecol*. 2015;3:1–12.
- Tatler J, Cassey P, Prowse TA. High accuracy at low frequency: detailed behavioural classification from accelerometer data. *J Exp Biol*. 2018;221(23):jeb184085.
- Studd EK, Boudreau MR, Majchrzak YN, Menzies AK, Peers MJ, Seguin JL, et al. Use of acceleration and acoustics to classify behavior, generate time budgets, and evaluate responses to moonlight in free-ranging snowshoe hares. *Front Ecol Evol*. 2019;7:154.
- Studd EK, Landry-Cuerrier M, Menzies AK, Boutin S, McAdam AG, Lane JE, et al. Behavioral classification of low-frequency acceleration and temperature data from a free-ranging small mammal. *Ecol Evol*. 2019;9(1):619–30.
- Auge A-C, Blouin-Demers G, Murray DL. Developing a classification system to assign activity states to two species of freshwater turtles. *PLoS ONE*. 2022;17(11):e0277491.
- Whitney NM, Papastamatiou YP, Holland KN, Lowe CG. Use of an acceleration data logger to measure diel activity patterns in captive

- whitetip reef sharks, *Triaenodon obesus*. *Aquat Living Resour.* 2007;20(4):299–305.
24. Shamoun-Baranes J, Bouten W, Van Loon EE, Meijer C, Camphuysen C. Flap or soar? How a flight generalist responds to its aerial environment. *Philos Trans R Soc B Biol Sci.* 2016;371(1704):20150395.
 25. Murchie KJ, Cooke SJ, Danylchuk AJ, Suski CD. Estimates of field activity and metabolic rates of bonfish (*Albula vulpes*) in coastal marine habitats using acoustic tri-axial accelerometer transmitters and intermittent-flow respirometry. *J Exp Mar Biol Ecol.* 2011;396(2):147–55.
 26. Ullmann W, Fischer C, Kramer-Schadt S, Pirhofer Walzl K, Eccard JA, Wevers JP, et al. The secret life of wild animals revealed by accelerometer data: how landscape diversity and seasonality influence the behavioural types of European hares. *Landscape Ecol.* 2023;38(12):3081–95.
 27. Weegman MD, Bearhop S, Hilton GM, Walsh AJ, Griffin L, Resheff YS, et al. Using accelerometry to compare costs of extended migration in an arctic herbivore. *Current zoology.* 2017;63(6):667–74.
 28. Flack A, Nagy M, Fiedler W, Couzin ID, Wikelski M. From local collective behavior to global migratory patterns in white storks. *Science.* 2018;360(6391):911–4.
 29. Yu H, Muijres FT, te Lindert JS, Hedenström A, Henningsson P. Accelerometer sampling requirements for animal behaviour classification and estimation of energy expenditure. *Animal Biotelemetry.* 2023;11(1):28.
 30. Chen KY, David R Basset J. The technology of accelerometry-based activity monitors: current and future. *Med Sci Sports Exerc.* 2005;37(11):5490–500.
 31. Aulsebrook AE, Jacques-Hamilton R, Kempnaers B. Quantifying mating behaviour using accelerometry and machine learning: challenges and opportunities. *Anim Behav.* 2024;207:55–76.
 32. Ellis-Soto D. Determining activity patterns of Galápagos tortoises: an intra and inter-island comparison through space and time: University of Konstanz; 2017.
 33. Ryan MA, Whisson DA, Holland GJ, Arnould JP. Activity patterns of free-ranging koalas (*Phascolarctos cinereus*) revealed by accelerometry. *PLoS ONE.* 2013;8(11):e80366.
 34. Thery M. The evolution of leks through female choice: differential clustering and space utilization in six sympatric manakins. *Behav Ecol Sociobiol.* 1992;30:227–37.
 35. Gudka M, Santos CD, Dolman PM, Abad-Gómez JM, Silva JP. Feeling the heat: elevated temperature affects male display activity of a lekking grassland bird. *PLoS ONE.* 2019;14(9):e0221999.
 36. Rintamäki PT, Karvonen E, Alatalo RV, Lundberg A. Why do black grouse males perform on lek sites outside the breeding season? *J Avian Biol.* 1999;199:359–66.
 37. Cestari C, Loïselle BA, Pizo MA. Trade-offs in male display activity with lek size. *PLoS ONE.* 2016;11(9):e0162943.
 38. Wilson RP, White CR, Quintana F, Halsey LG, Liebsch N, Martin GR, et al. Moving towards acceleration for estimates of activity-specific metabolic rate in free-living animals: the case of the cormorant. *J Anim Ecol.* 2006;75(5):1081–90.
 39. Nuijten RJ, Gerrits T, Shamoun-Baranes J, Nolet BA. Less is more: On-board lossy compression of accelerometer data increases biologging capacity. *J Anim Ecol.* 2020;89(1):237–47.
 40. Clutton-Brock TH. Reproductive success: studies of individual variation in contrasting breeding systems: University of Chicago Press; 1988.
 41. Blake S, Cabrera F, Cruz S, Ellis-Soto D, Yackulic CB, Bastille-Rousseau G, et al. Environmental variation structures reproduction and recruitment in long-lived mega-herbivores: Galapagos giant tortoises. *Ecol Monogr.* 2024;94:e1599.
 42. Wilson DS. Nest-site selection: microhabitat variation and its effects on the survival of turtle embryos. *Ecology.* 1998;79(6):1884–92.
 43. McIntosh I, Goodman K, Parrish-Ballentine A. Tagging and nesting research on Hawksbill Turtles (*Eretmochelys imbricata*) at Jumbay Bay, Long Island, Antigua, West Indies. Annual Report Wider Caribbean Sea Turtle Network, University of Georgia, Athens, Georgia, USA. 2003.
 44. Gibbs JP, Goldspiel H. Population biology. Galapagos giant tortoises: Elsevier; 2021. p. 241–60.
 45. Jensen EL, Gaughran SJ, Fusco NA, Poulakakis N, Tapia W, Sevilla C, et al. The Galapagos giant tortoise *Chelonoidis phantasticus* is not extinct. *Commun Biol.* 2022;5(1):546.
 46. Kubisch E, Iburgüengoytia NR. Reproduction. Galapagos Giant Tortoises Elsevier; 2021. p. 157–73.
 47. Cayot LJ. The history of Galapagos tortoise conservation. Galapagos giant tortoises: Elsevier; 2021. p. 333–53.
 48. Cayot LJ, Campbell K, Carrión V. Invasive species: impacts, control, and eradication. Galapagos Giant Tortoises. Elsevier; 2021. p. 381–99.
 49. Charney ND. Galapagos tortoises in a changing climate. Galapagos Giant Tortoises: Elsevier; 2021. p. 317–30.
 50. Flanagan JP. Tortoise health. Galapagos Giant Tortoises: Elsevier; 2021. p. 355–80.
 51. Ramon-Gomez K, Ron SR, Deem SL, Pike KN, Stevens C, Izurieta JC, et al. Plastic ingestion in giant tortoises: an example of a novel anthropogenic impact for Galapagos wildlife. *Environ Pollut.* 2024;340:122780.
 52. Blake S, Cabrera F, Rivas-Torres G, Deem SL, Nieto-Claudin A, Zahawi RA, et al. Invasion by *Cedrela odorata* threatens long distance migration of Galapagos tortoises. *Ecol Evol.* 2024;14(2):e10994.
 53. Nieto-Claudin A, Deem SL, Rodríguez C, Cano S, Moity N, Cabrera F, et al. Antimicrobial resistance in Galapagos tortoises as an indicator of the growing human footprint. *Environ Pollut.* 2021;284:117453.
 54. Bacon JP. Some observations on the captive management of Galapagos tortoises. REPRODUCTIVE BIOLOGY AND DISEASES OF CAPTIVE REPTILES JB Murphy; JT Collins, eds Society for the Study of Amphibians and Reptiles. 1980:97–113.
 55. MacFarland CG, Villa J, Toro B. The Galapagos giant tortoises (*Geochelone elephantopus*) part II: conservation methods. *Biol Cons.* 1974;6(3):198–212.
 56. Jackson MH. Galápagos: a natural history: University of Calgary press; 1993.
 57. Trueman M, d'Ozouville N. Characterizing the Galapagos terrestrial climate in the face of global climate change. 2010.
 58. Snell HM, Stone PA, Snell HL. A summary of geographical characteristics of the Galapagos Islands. *J Biogeogr.* 1996;23(5):619–24.
 59. Laso FJ, Benítez FL, Rivas-Torres G, Sampedro C, Arce-Nazario J. Land cover classification of complex agroecosystems in the non-protected highlands of the Galapagos Islands. *Remote Sens.* 2019;12(1):65.
 60. Poulakakis N, Edwards DL, Chiari Y, Garrick RC, Russello MA, Benavides E, et al. Description of a new Galápagos giant tortoise species (*Chelonoidis*; Testudines: Testudinidae) from Cerro Fatal on Santa Cruz Island. *PLoS ONE.* 2015;10(10):e0138779.
 61. Cayot LJ, Gibbs JP, Tapia W, Caccone A. *Chelonoidis donfaustoi*: The IUCN Red List of Threatened Species 2017 [Available from: <https://doi.org/10.2305/IUCN.UK.2017-3.RLTS.T90377132A90377135.en>].
 62. Sevilla C, Málaga J, Gibbs JP. Tortoise populations after 60 years of conservation. Galapagos giant tortoises: Elsevier; 2021. p. 401–32.
 63. Blake S, Yackulic CB, Cabrera F, Deem SL, Ellis-Soto D, Gibbs JP, et al. Movement ecology. Galapagos Giant Tortoises: Elsevier; 2021. p. 261–79.
 64. Blake S, Yackulic CB, Cabrera F, Tapia W, Gibbs JP, Kummeth F, et al. Vegetation dynamics drive segregation by body size in Galapagos tortoises migrating across altitudinal gradients. *J Anim Ecol.* 2013;82(2):310–21.
 65. Bastille-Rousseau G, Gibbs JP, Yackulic CB, Frair JL, Cabrera F, Rousseau LP, et al. Animal movement in the absence of predation: environmental drivers of movement strategies in a partial migration system. *Oikos.* 2017;126(7):1004–19.
 66. Gibbons JW, Greene JL. X-ray photography: a technique to determine reproductive patterns of freshwater turtles. *Herpetologica.* 1979;1979:86–9.
 67. Mueller JM, Sharp KR, Zander KK, Rakestraw DL, Rautenstrauch KR, Lederle PE. Size-specific fecundity of the desert tortoise (*Gopherus agassizii*). *J Herpetol.* 1998:313–9.
 68. Loehr VJ, Henen BT, Hofmeyr MD. Reproduction of the smallest tortoise, the Namaqualand speckled padloper. *Homopus Signatus Signatus Herpetologica.* 2004;60(4):444–54.
 69. Lovich JE, Puffer SR, Agha M, Ennen JR, Meyer-Wilkins K, Tennant LA, et al. Reproductive output and clutch phenology of female Agassiz's desert tortoises (*Gopherus agassizii*) in the Sonoran Desert region of Joshua Tree National Park. *Curr Herpetol.* 2018;37(1):40–57.
 70. Breiman L. Random forests. *Mach Learn.* 2001;45:5–32.
 71. Elith J, Leathwick JR, Hastie T. A working guide to boosted regression trees. *J Anim Ecol.* 2008;77(4):802–13.
 72. Shuert CR, Pomeroy PP, Twiss SD. Assessing the utility and limitations of accelerometers and machine learning approaches in classifying behaviour during lactation in a phocid seal. *Animal Biotelemetry.* 2018;6(1):1–17.

73. Hanscom RJ, DeSantis DL, Hill JL, Marbach T, Sukumaran J, Tipton AF, et al. How to study a predator that only eats a few meals a year: high-frequency accelerometry to quantify feeding behaviours of rattlesnakes (*Crotalus* spp.). *Animal Biotelemetry*. 2023;11(1):20.
74. 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. *Animal Biotelemetry*. 2023;11(1):32.
75. Congdon JD, Gatten Jr RE. Movements and energetics of nesting *Chrysemys picta*. *Herpetologica*. 1989;94–100.
76. Marchand T, Le Gal A-S, Georges J-Y. Fine scale behaviour and time-budget in the cryptic ectotherm European pond turtle *Emys orbicularis*. *PLoS ONE*. 2021;16(10):e0256549.
77. Tucker CR. Use of automated radio telemetry to detect nesting activity in Ornate Box Turtles. *Terrapene Ornata Am Midl Nat*. 2014;171(1):78–89.
78. Auge A-C, Blouin-Demers G, Murray DL. Differences in activity between reproductive and non-reproductive freshwater turtles during the nesting season. *Herpetol Notes*. 2024;17:153–9.
79. Fraley C, Raftery AE, Scrucca L, Murphy TB, Fop M, Scrucca ML. Package 'mclust'. Gaussian Mixture Modelling for Model Based Clustering, Classification, and Density Estimation. 2012.
80. Márquez C. The natural history of the Galápagos giant tortoise. CreateSpace Independent Publishing Platform 2019.
81. Team RC. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria 2023.
82. Liaw A, Wiener M. Classification and regression by randomForest. *R news*. 2002;2(3):18–22.
83. Greenwell B, Boehmke B, Cunningham J, Developers G, Greenwell MB. Package 'gbm'. R package version. 2019;2(5).
84. Bergmeir C, Benítez JM. On the use of cross-validation for time series predictor evaluation. *Inf Sci*. 2012;191:192–213.
85. Ferdinandy B, Gerencsér L, Corrieri L, Perez P, Újváry D, Cszizmadia G, et al. Challenges of machine learning model validation using correlated behaviour data: evaluation of cross-validation strategies and accuracy measures. *PLoS ONE*. 2020;15(7):e0236092.
86. Vabalas A, Gowen E, Poliakoff E, Casson AJ. Machine learning algorithm validation with a limited sample size. *PLoS ONE*. 2019;14(11):e0224365.
87. Han H, Guo X, Yu H, editors. Variable selection using mean decrease accuracy and mean decrease Gini based on random forest. 2016 7th IEEE international conference on software engineering and service science (icse); 2016: IEEE.
88. Krauss C, Do XA, Huck N. Deep neural networks, gradient-boosted trees, random forests: Statistical arbitrage on the S&P 500. *Eur J Oper Res*. 2017;259(2):689–702.
89. Yu H, Deng J, Nathan R, Kröschel M, Pekarsky S, Li G, et al. An evaluation of machine learning classifiers for next-generation, continuous-ethogram smart trackers. *Mov Ecol*. 2021;9:1–14.
90. Resheff YS, Rotics S, Harel R, Spiegel O, Nathan R. Accelerater: a web application for supervised learning of behavioral modes from acceleration measurements. *Mov Ecol*. 2014;2:1–7.
91. Watanabe S, Sato K, Ponganis PJ. Activity time budget during foraging trips of emperor penguins. *PLoS ONE*. 2012;7(11):e50357.
92. 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. *Animal Biotelemetry*. 2022;10:1–16.
93. Barbuti R, Chessa S, Micheli A, Pucci R. Localizing tortoise nests by neural networks. *PLoS ONE*. 2016;11(3):e0151168.
94. Roberts DR, Bahn V, Ciuti S, Boyce MS, Elith J, Guillera-Aroita G, et al. Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography*. 2017;40(8):913–29.
95. Yu H, Klaassen CA, Deng J, Leen T, Li G, Klaassen M. Increasingly detailed insights in animal behaviours using continuous on-board processing of accelerometer data. *Mov Ecol*. 2022;10(1):42.
96. Cox SL, Orgeret F, Gesta M, Rodde C, Heizer I, Weimerskirch H, et al. Processing of acceleration and dive data on-board satellite relay tags to investigate diving and foraging behaviour in free-ranging marine predators. *Methods Ecol Evol*. 2018;9(1):64–77.
97. Yu H, Deng J, Leen T, Li G, Klaassen M. Continuous on-board behaviour classification using accelerometry: a case study with a new GPS-3G-Bluetooth system in Pacific black ducks. *Methods Ecol Evol*. 2022;13(7):1429–35.
98. Wang J, He Z, Zheng G, Gao S, Zhao K. Development and validation of an ensemble classifier for real-time recognition of cow behavior patterns from accelerometer data and location data. *PLoS ONE*. 2018;13(9):e0203546.
99. Jin Z, Shu H, Hu T, Jiang C, Yan R, Qi J, et al. Behavior classification and spatiotemporal analysis of grazing sheep using deep learning. *Comput Electron Agric*. 2024;220:108894.
100. 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.
101. Fuchs NT, Caudill CC. Classifying and inferring behaviors using real-time acceleration biotelemetry in reproductive steelhead trout (*Oncorhynchus mykiss*). *Ecol Evol*. 2019;9(19):11329–43.
102. Clermont J, Woodward-Gagné S, Berteaux D. Digging into the behaviour of an active hunting predator: arctic fox prey caching events revealed by accelerometry. *Mov Ecol*. 2021;9:1–12.
103. Schreven KH, Stolz C, Madsen J, Nolet BA. Nesting attempts and success of Arctic-breeding geese can be derived with high precision from accelerometry and GPS-tracking. *Animal Biotelemetry*. 2021;9:1–13.
104. Ozsanlav-Harris L, Griffin LR, Weegman MD, Cao L, Hilton GM, Bearhop S. Wearable reproductive trackers: quantifying a key life history event remotely. *Animal Biotelemetry*. 2022;10(1):24.
105. Roy CL, Coy PL. Lek attendance and disturbance at viewing blinds in a small, declining Sharp-tailed Grouse (*Tympanuchus phasianellus*) population. *Avian Conserv Ecol*. 2021;16(2):1.
106. Pitman JB III, Bastille-Rousseau G. Retention time and fix acquisition rate of glued-on GPS transmitters in a semi-aquatic species. *Animal Biotelemetry*. 2023;11(1):24.
107. Mortlock E, Silovský V, Güldenpfennig J, Faltusová M, Olejarz A, Börger L, et al. Sleep in the wild: the importance of individual effects and environmental conditions on sleep behaviour in wild boar. *Proc R Soc B*. 2023;2024(291):20232115.
108. Troëng S, Rankin E. Long-term conservation efforts contribute to positive green turtle *Chelonia mydas* nesting trend at Tortuguero. *Costa Rica Biol Conserv*. 2005;121(1):111–6.
109. Kuchling G. The reproductive biology of the Chelonia: Springer Science and Business Media; 2012.
110. Rhodin AG, Stanford CB, Van Dijk PP, Eiseberg C, Luiselli L, Mittermeier RA, et al. Global conservation status of turtles and tortoises (order Testudines). *Chelonian Conserv Biol*. 2018;17(2):135–61.

Publisher's Note

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