Development of a multi-sensor biologging collar and analytical techniques to describe high-resolution spatial behavior in free ranging terrestrial mammals

Michael S Painter (mpainter@barry.edu)
Václav Silovský
Justin Blanco
Mark Holton
Monika Faltusová
Rory Wilson
Luca Börger
Liza Psotta
Fabian Ramos Almodovar
Luis Estrada
Lukas Landler
E. Pascal Malkemper
Vlastimil Hart
Miloš Ježek

Research Article

Keywords: accelerometer, behavioral classification, biologging, dead-reckoning, GPS, machine learning, magnetometer, magnetic compass heading

Posted Date: January 19th, 2024

DOI: https://doi.org/10.21203/rs.3.rs-3874325/v1

Additional Declarations: The authors declare no competing interests.

Abstract

1. Biologging has proven to be a powerful approach to investigate diverse questions related to movement ecology across a range of spatiotemporal scales and increasingly relies on multidisciplinary expertise. Advancements in sensor design and analytical techniques continue to push the boundaries of this emerging discipline. However, the growing variety of animal-borne equipment, coupled with little consensus regarding analytical approaches to interpret complex datasets presents challenges and makes comparison between studies and study species difficult.

2. Here, we present a combined hardware and analytical approach for standardizing the collection, analysis and interpretation of multi-sensor biologging data. We develop (i) a custom-designed integrated multi-sensor collar (IMSC), which was field tested on 71 free-ranging wild boar (*Sus scrofa*) over 2 years; (ii) a machine learning behavioral classifier capable of identifying six behaviors in free-roaming boar, validated across individuals equipped with differing collar designs; and (iii) laboratory and field-based calibration and accuracy assessments of animal heading measurements derived from raw magnetometer data.

3. The durability and capacity of IMSCs exceeded expectations, with a 94% collar recovery rate and a 75% cumulative data recording success rate across all collars deployed, with a maximum data logging duration of 421 days. The behavioral classifier had an overall accuracy of 85% in identifying the six behavioral classes across all collar designs and improved to 90% when tested on data from the IMSC only. Both laboratory and field tests of magnetic compass headings were in precise agreement with expectations, with overall median magnetic headings deviating from ground truth observations by 1.7° and 0°, respectively.

4. Here we present the development of the IMSC coupled with an analytical framework verified by ground truth data for identifying core behaviors and spatial orientation in free roaming boar. We highlight the potential of additional analyses available using this commercially produced system that can be adapted for use in future studies on terrestrial mammals.

Introduction

In recent decades, animal-borne sensors designed to monitor physiology, behavior, movement, and environmental conditions have revolutionized studies of animal ecology in diverse taxa across a range of spatiotemporal scales (Ropert-Coudert & Wilson, 2005; Rutz & Hays, 2009; Wilmers et al., 2015). This has been made possible due to advances in sensor technology, data management, and analytical techniques, which now underpin both theoretical and applied research on wild animals (Cooke et al., 2012; Rattenborg et al., 2016; Wilmers et al., 2015). However, the emergence of novel biologging techniques require a multidisciplinary approach, often relying on diverse expertise in areas beyond wildlife ecology (Kays et al., 2022; Portugal & White, 2018; Tuia et al., 2022; Wild et al., 2023). Furthermore, animal-borne electronics and datasets are increasingly tailored to a particular study or research group, making access to, and comparison between, biologging studies challenging.

Triaxial accelerometers and magnetometers form the bedrock of biologging studies, providing high-resolution data on animal movement and orientation (Shepard et al., 2008; Williams et al., 2017; Wilson et al., 2008), with recent studies applying various machine learning techniques to identify behaviors from raw accelerometer and/or magnetometer profiles (Wang, 2019), as well as alternative approaches, such as template matching(Walker et al., 2015) and user-defined algorithms for behavior (Wilson et al., 2018). Performance of such models varies due to factors such as the frequency at which data are recorded and the degree of behavioral variation *within* and *between* the behavioral classes attempting to be identified. To date, no consensus has been reached on a single behavioral classification technique across biologging studies, further hindering comparison between studies and species.

Magnetometer data, used in conjunction with accelerometers, can enhance machine learning performance by providing additional orientation data (Williams et al., 2020), and in some contexts, triaxial magnetometer data alone can be used to successfully identify behavior in free-roaming animals (Chakravarty et al., 2019; Williams et al., 2017). In addition, triaxial magnetometers can provide magnetic heading orientation (Matsumura et al., 2011), although extracting compass headings from raw data is not trivial and depends on sensor calibrations and accelerometer-based tilt-compensation corrections (Bidder et al., 2015). Unsurprisingly, calibration techniques are now commonplace in studies that report magnetic heading measurements derived from raw magnetometer data (Gutzler & Watson III, 2022; Logan et al., 2023; Martín López et al., 2016); however, few (Wilson et al., 2007) have provided ground truth validation of magnetic compass accuracy and reliability across ecologically realistic movement dynamics or behaviors.

Integration of GPS technology with accelerometer and magnetometer data has further enhanced the accuracy and depth of spatial information in animal tracking studies and is reflected in the widespread deployment of GPS technology across a range of animal studies over the past three decades (Kays et al., 2015). Beyond its utility in providing reliable positional fixes, GPS is now used to improve the performance (e.g., mitigate drift and heading error) of dead-reckoning path reconstruction that rely on vector integration obtained from synchronized accelerometer and magnetometer data(Gunner et al., 2021) and further underscores the importance of assessing the accuracy of magnetic heading measurements obtained from raw data. Engineering multi-sensor collars (e.g. GPS, accelerometers, magnetometers) capable of recording and storing large volumes of data over months or years that comply with animal welfare standards remains an additional challenge in biologging research (Cook et al., 2017; Holton et al., 2021; Kenward, 2000; Wilson et al., 1986, 2021).

Here we present the development of a multi-sensor biologging collar equipped with GPS and triaxial accelerometer and magnetometer sensors that has been extensively tested in free-ranging wild boar (*Sus scrofa*). In tandem, we have developed a novel method for classifying ecologically relevant behaviors from raw accelerometer data in wild boar using machine learning techniques and provide a detailed assessment of magnetic compass performance based on raw magnetometer data across a range of behavioral contexts. Our findings suggest that both the collars and analytical techniques are robust, adaptable, and suitable for long-term studies with terrestrial mammals, and we discuss the broader applications of this work for wildlife research.

Methods

Study site and Subjects

Field testing the IMSC was carried out in unrestricted, natural habitats throughout the Czech Republic. Boar were captured in corral traps, sedated using methods described below (see also Supplemental Materials) and fitted with the IMSC, then released into the surrounding environment. All data used to develop the behavioral classifier and evaluate magnetic compass performance, were collected at a wildlife reserve (49°57'52.7"N 14°50'14.7"E) owned by Czech University of Life Sciences. Inside the reserve, a semi-natural enclosure (~ 38 m x ~ 46 m), made from non-magnetic wood fencing was used to collect ground truth behavioral data (hereafter 'behavioral enclosure') from six adult wild boar between October 2017 – December 2018 (Fig. S1). Boar were captured opportunistically using dart tranquilizer methods (see Supplemental Materials), then were transported inside the behavioral enclosure and fitted with one of two biologging collar designs (see below). Four infrared game cameras (UOVision UM 565) were installed within the enclosure (Fig. S1) to record ground truth data used for behavioral classifier and magnetic heading analyses (see below).

Trapping, handling, and collaring protocols were performed in accordance with the Ethics Committee of the Ministry of the Environment of the Czech Republic number MZP/2019/630/361 and following ARRIVE guidelines (Percie du Sert et al., 2020). See Supplemental Materials for additional study site information.

Biologging Collar Development

Two collar systems were designed in this study; 'integrated multi-sensor collars' (IMSCs) and 'single-tag collars' (STCs), both fitted with Wildbyte Technologies Daily Diary data loggers (http://www.wildbytetechnologies.com/). Loggers were equipped with tri-axial accelerometers and tri-axial magnetometers (LSM303DLHC, ST Microelectronics) programmed to record continuously at a sample rate of 10 Hz across all six sensors aligned along three orthogonal axes corresponding to the major axes of the boars' bodies (Fig. 1).

Integrated Multi-sensor Collar (IMSC)

The IMSC included a 'Thumb' Daily Diary tag (18 x 14 x 5 mm) with triaxial accelerometer and magnetometer sensors (LSM9DS1, ST Microelectronics), as well as a Vertex Plus GPS collar, scheduled to record GPS fixes at 30 min intervals. Collars were equipped with an integrated 'drop-off mechanism' and VHF beacon to enable collar recovery from the field. All collar electronics were powered from a single battery pack (4-D cell) and total deployment weight was 716 g. The Daily Diary tag was protected by a custom-designed polyurethane housing (40 mm x 25 mm x 12 mm) positioned on the outside of the plastic collar belt. The orientation of the tag relative to the collar, as well as the orientation of collar relative to the animal, remained fixed for all IMSC deployments (Fig. 1, Table 1).

Single-tag Collars (STCs)

All STCs were equipped with the 'Square' Daily Diary tag (27 x 26 x 10 mm) and recorded data to a removable 32 GB MicroSD card. The logger was powered with a single cell 3.6 V lithium battery (SAFT, LS17500CNR) and was oriented and levelled within a 12 cm x 4.8 cm dia PVC-U cylindrical tube housing secured to a plastic collar belt. Total STC weight was 250 g. All STC housings were positioned ventrally at the base of the animal's neck (Fig. 1A, B). However, logger orientation was rotated in one STC deployment (Fig. 1B) to test the positional robustness of the behavioral classifier (see below).

See supplemental material for additional information regarding collar specifications and deployments.

Data collection

Field testing the IMSC involved 71 collar deployments over a two-year period on adult (> 12 months, > 40 kg) free roaming wild boar (52 females, 18 males, 1 unidentified). Collars were evaluated for robustness, capacity, and functionality over 6,001 tracking days, cumulatively across all deployments.

Behavioral classifier and magnetic compass performance data was collected from six free roaming individuals inside the behavioral enclosure. Before collaring, calibration data used for hard- and soft-iron magnetometer corrections (Gunner et al., 2021; Williams et al., 2017) were collected by rotating the collars through three-dimensional space for 5 min within the immediate area of the behavioral enclosure. The resulting accelerometer 'calibration signature' was also used to time-sync biologging data with ground truth recordings from each game camera. Upon data retrieval, raw data files were uploaded to DDMT software (Wildbyte Technologies - Swansea University, Singleton Park, Swansea, UK, SA2 8PP), for further processing, including magnetometer calibrations. A summary of data collection and performance evaluations for each collar design is provided in Table 1.

Behavioral Classifier Development

Training Dataset Construction

Triaxial (x,y,z) accelerometer data from three individuals fitted with STCs were used to develop the behavioral classifier (Table 1). Six broad behavioral classes ('Continuous Walk', 'Foraging', 'Resting', 'Running', 'Standing', and 'Other') were established using the criteria listed in the Supplemental Materials. Behaviors were identified using video records and corresponding accelerometer profiles were located by matching video timestamps with synced timestamps in the DDMT software. Profiles were then extracted to create behavioral ethograms composed solely of triaxial accelerometer data falling into one of the six behavioral classes. To facilitate future refinement of the classifier, 'Foraging', 'Running', and 'Standing' classes were further subdivided to produce three additional, 'higher resolution' behavioral categories:

'Rooting', 'Trotting', and 'Vigilance', respectively, resulting in a total of nine behavioral classes. The higher-resolution behavioral classes were collapsed into their parent classes for initial classifier evaluation.

Each marked behavioral epoch was subdivided into 4-second non-overlapping windows to generate baseline observations for classifier training (i.e., entities to be classified following feature extraction). The 4-second observation window was chosen in consideration of two factors: the shortest-duration behavior desirable to detect, and the minimum acceptable detection latency. In total, there were 13,461 training observations (14.96 hours of marked data), with the following breakdown of observations and training percentage for the six 'core' behavioral classes: 'Continuous Walk' (1445, 11%), 'Foraging' (2601, 19%), 'Resting' (6345, 47%), 'Running' (1042, 8%), 'Standing' (1668, 12%), and 'Other' (360, 3%) (Table 2, Table S1). Training data for the 'higher resolution' behavioral subclasses are provided in Table 2, Table S1. The proportions of observations used to train the behavioral classifier were selected *a priori* to reflect the frequencies of these behavioral classes thought to occur in natural contexts (VS, MJ personal observations). The training dataset was constructed from three individuals (B3, B6 male; B5, female), all fitted with STCs with identical tag orientations (Fig. 1A, Table 1).

Feature Extraction

Eighteen features were extracted from each 4-second raw-data observation. These features were: the estimated 'signal power' in each of four frequency bands (0-2.5 Hz, 2.5-5 Hz, 5-7.5 Hz, and 7.5–10 Hz; four features), the signal median (one feature), and the signal variance (one feature), for each of the three accelerometer axes. The power features were derived from the Welch method of power spectral density estimation (2s windows with 1s overlap, 64-point Discrete Fourier Transforms), by integrating the output in the designated frequency ranges. All features guaranteed to be non-negative (i.e., all except the median features) were log-transformed to a decibel-proportional scale prior to further processing. Finally, features were z-scored and Principal Components Analysis was performed, retaining a number of components required to preserve 95% of the total data variance (eight components). The resultant 13,461 x 8 matrix served as the training data for a 5-nearest neighbor classifier with cityblock distance as the metric.

Behavioral Classifier Evaluation

Performance of the behavioral classifier was evaluated using continuous accelerometer recordings collected from three individuals (B4, B7, B30) not used in classifier training (Table 1). Prior to evaluation, behaviors were verified using ground truth video recordings and corresponding accelerometer profiles were identified as described above. Because the tag orientation was not identical across training and test boar (Fig. 1, Table 1), test data x,y,z-acceleration vectors at every time step were multiplied by the 3D rotation matrix required to map them to the coordinate frame used for training data. After 18-dimensional feature extraction, test data observations were transformed using the training data mean and standard deviation vectors before being projected onto the 8-dimensional principal component space of the training data for classification. Initial test data classifications were made at every possible time step using a 4-second symmetric, non-causal sliding window.

Post-processing

Initial classifications were smoothed with a nonlinear filter; specifically, the class at each time-step was replaced with the modal class of a 1-second forward-looking window. This filtering step resulted in a set of candidate behavioral events, each delimited by a starting and ending time, which were then subject to two pre-determined heuristic criteria to yield the final set of classifications. The first was that each candidate behavioral event was required to be of a minimum duration: 'Foraging' (5s; 'Rooting' 3s), 'Resting' (120s), 'Running' (3s; 'Trotting' 3s), 'Standing' (2s; 'Vigilance' 2s), and 'Other' (1s). Any candidate event not meeting its minimum duration was reassigned to the next most likely class for which the duration criterion could be met. Class likelihoods were determined using the relative class-proportions among the five nearest training set-neighbors corresponding to each time-step in the candidate event. Class-proportions were summed across time-steps and sorted to produce a rank-ordered likelihood for the classes. Candidate events for which this procedure failed to yield a valid alternate class assignment were merged with the subsequent event[1].

The second heuristic was that any candidate 'Standing' event flanked by 'Resting' activity was reassigned to the 'Resting' class. Specifically, this reassignment was made if the majority of a 120s window on either side of the candidate 'Standing' event was classified as 'Resting'.

Magnetometer Data

To assess magnetic compass heading accuracy and reliability magnetometer data were collected from four collars under two conditions: a controlled laboratory environment (hereafter, 'lab evaluation') designed to test the precision of the magnetometer, and from three free-roaming boar inside the behavioral enclosure (hereafter, 'field test') (Table 1).

During the lab evaluation, the tag was levelled and centered inside an electromagnetic enclosure containing four Helmholtz's coils used to manipulate the strength and alignment of an experimentally generated magnetic field. Two orthogonally aligned coils were used to cancel the residual horizontal component of the Earth's magnetic field (+/- 0.1%) and to adjust the vertical component of the magnetic field to match that of an Earth strength vertical field (~ 45,000 nT). Two inner orthogonally aligned coils were used to generate Earth-strength magnetic fields (total strength ~ 50,000 nT) that could be rotated into alignment into one of four cardinal compass alignments corresponding to topographic North, South, East, and West (Kirschvink, 1992). The tag was oriented such that one end of the x-axis was aligned toward topographic North which was then defined as the 'heading direction' in DDMT for analysis. Tag orientation remained static, whereas the horizontal component of the magnetic field was rotated by 90° increments into alignment with each of the four cardinal compass directions for a period of 10 sec in each alignment. Magnetic heading measurements calculated by DDMT were plotted relative to the four expected cardinal compass directions using the gghistogram function in the R package ggpubr (Kassambara, 2020).

Field tests of the magnetometer performance were carried out concurrently with data collected for the behavioral classifier within the behavioral enclosure on three free-roaming individuals (Table 1). Video recordings provided ground truth magnetic headings and a spatial array of 'magnetic landmarks' were installed within the camera's field of view to provide known magnetic references to better estimate magnetic headings of focal subjects. Magnetic landmarks were either non-magnetic cables tethered between trees or the non-magnetic fence-line forming the behavioral enclosure (Fig. S1). A total of 45 independent behavioral epochs from all five core behavioral classes, totalling 5:27 (min:sec), were selected to test the precision of the magnetic heading data. Heading predictions were made by two investigators not involved in data collection and blind to all raw magnetometer data. Using only video records, investigators predicted boar magnetic heading using the available magnetic landmarks described above. For each prediction, the average magnetic heading was estimated over the duration of the segment identified. When investigator predictions differed by less than 20° (n = 40), they were averaged to establish the final magnetic heading, whereas when predictions differed by more than 20° (n = 5), investigators determined a final prediction after reevaluating the recording together. A third investigator blind to the magnetic predictions extracted the magnetic heading data from DDMT for further analysis.

Results

IMSC Field Performance: durability, capacity, lifetime

Between 2019 and 2022, 67 of the 71 total collars (~ 94%) deployed on free-ranging boar were recovered and data recording durations ranged from 9 to 421 days. The remaining four collars (~ 6%) experienced an unknown GPS malfunction and remain unrecovered. Of the 67 collars retrieved, 51 (76%) were fully functional and no appreciable damage was noted, while 11 (16%) exhibited mechanical damage likely due to physical stresses associated with boar behavior, and the remaining 5 collars (7%) failed prematurely due to an unexpected electrical fault. Of the fully functional subset, 35 (69%) collars recorded data until retrieval, whereas 9 collars (18%) recorded data for > 50% of deployment period and the remaining 7 collars (14%) recorded data for < 50% of deployment period. Overall, free-ranging boar equipped with IMSCs were tracked for 6,001 days and a total of 4,547 days of biologging data were recorded, corresponding to 75% of the cumulative deployment duration.

Behavioral Classifier

Classifier performance was evaluated using accelerometer data from 2,100 independent ground truth behavioral epochs (i.e., independent behaviors falling into one of the six behavioral classes) across three individuals, totalling 08:28:15 (HH:mm:ss) of data (Table 3, Table S2). Classifier performance was evaluated on an event-by-event basis, (i.e. per 0.1 sec sample). Overall behavioral classifier performance was 85.1% across all behaviors from all three individuals (Table 4) and includes data from the STC and IMSC collar designs with different tag positions and orientations. Of the five behavioral classes of interest (i.e., excluding 'Other' which was composed of heterogeneous behaviors only identified by the classifier when a behavior did not fall into any of the five core behavioral categories), the likelihood that any given prediction matched the ground truth class label (i.e., precision), ranged from 77.1% ('Walking' and 'Standing') to 96.5% ('Resting') (Table 4). Classifier recall, i.e., the proportion of behavioral epochs correctly identified by the classifier, ranged from 74.7% (‘Running’) to 91.8% (‘Resting’) (Table 4). Classifier performance was consistent between the three deployments, ranging from 83.5% (B4) to 89.9% (B30), and surprisingly, the collar with the highest performance (B30, IMSC) was least similar to those used to train the classifier (Table 5). All possible pairs of the 8 principal components used to identify the six behavioral classes are plotted, along with histograms corresponding to each component in isolation, to illustrate the collective and relative contribution of the principal components towards class-separability (Fig. 2). Precision and recall metrics were substantially lower when tested on the three expanded behavioral classes, reflecting their similar acceleration profiles relative to their respective parent classes. However, overall classifier performance remained robust, with an accuracy of 78.4%, although there was larger variation in performance between collar designs when tested on the expanded classes (Table 6).

Magnetic Heading: Lab Evaluation

Following calibration procedures describe above, median magnetic heading measurements calculated by DDMT were in agreement with each of the experimentally generated magnetic field alignments: N = 2.99°, S = 179.16°, East = 88.21°, W = 268.66° (Fig. 3), with an overall median heading error of 1.7° relative to expected.

Magnetic Heading: Field test

Across all 45 magnetic heading samples, the median discrepancy between DDMT magnetic compass heading measurements and ground truth predictions was 0° (CI: -3.1° and 6.9°) (Fig. 4B). Median bootstrapped 95% confidence intervals relative to predictions were calculated using the function boot from the boot package (Canty & Ripley, 2020). Discrepancy between DDMT heading and corresponding ground truth prediction ranged from − 30° to 21° (Fig. 4B). As shown in Fig. 4A, the distribution of compass headings obtained is evenly distributed across all possible magnetic heading alignments and the error in the DDMT magnetic compass heading measurements compared to predictions was uniform (i.e., error was unbiased across the range of magnetic directions) as indicated by the manova model previously described in Landler et al., 2022 (model results: intercept: approx. $F = 0.65$, $p = 0.53$, error proportion: approx. $F = 0.79$, $p = 0.46$) (Fig. 4C). The ‘error proportion' was calculated as the angular deviation between the DDMT measurement and the ground truth prediction divided by the total angular deviation. The cosine and sine of the magnetic heading in radians were used as the response variables and the error proportion as a linear covariate. The intercept of this model was used to test for a significant departure from uniformity (Landler et al., 2022). Importantly, the accuracy of magnetic compass headings was consistent across all three individuals evaluated (Table 7), fitted with different collar designs, biologger positions and orientations (Fig. 1, Table 1), as well as across all behavioral classes, including behaviors characterized by large acceleration amplitudes and variation (e.g., ‘Foraging’, ‘Walking’, ‘Running’).

Discussion

Animal-borne telemetry systems have emerged as a powerful tool to further characterize animal movement, behavior, and ecology. The availability of reliable collar systems equipped with a range of sensor technologies adaptable across multiple studies and species is valuable for several reasons, including that it eliminates the need to develop and test novel equipment, and that datasets collected from a standardized system may catalyse additional collaboration, data sharing, and advance progress in analytical techniques[2]. The IMSC developed here, equipped with triaxial accelerometer and magnetometer sensors, GPS technology, as well as a variety of additional sensors not used in the current study, has proven to be highly reliable under the harsh demands imposed by wild boar. Across the 71 IMSC deployments, 94% of deployed collars were recovered resulting in recorded biologging data for 75% of the cumulative deployment duration. While the maximum recording duration was an impressive 421 days on one individual, the majority of IMSCs (72%) were terminated prematurely due to hunting or automobile collisions, which does not reflect collar capacity. In a separate study, 36 IMSCs identical to those described above, were deployed on free-ranging red deer (*Cervus elaphus*) and had an average and maximum data recording duration of 203 and 529 days, respectively. Given the standardization, durability, and functionality of the IMSC, these collars are well-suited for long-term studies in terrestrial mammals, and we hope they will be adopted for use in future biologging studies.

Concurrent with the IMSC development, we have built a behavioral classifier capable of identifying ecologically relevant behaviors from six behavioral classes in wild boar. The classifier had an overall performance of 85% and, of the five core classes, identified ‘Resting’ with the highest precision, and ‘Standing’ had the lowest precision and was most often misclassified as ‘Resting’, likely due to the similar acceleration profiles between these behaviors. Classification recall performance was highest in ‘Resting’ and lowest in ‘Running’. The majority of undetected ‘Runs’ were misclassified as ‘Forage’, a class that includes ‘Rooting’ characterized by large and variable x-axis acceleration amplitudes, like those associated with ‘Running’ accelerometer profiles. Importantly, the test dataset for core and expanded behavioral classes reflected the proportions of behaviors used in classifier training, which in turn, approximated the overall behavioral repertoire of wild boar in natural contexts.

The classifier exhibited the best overall performance (89.9% accuracy) when tested with data collected from the IMSC, despite being trained on data exclusively from STCs, suggesting that the classifier has an inherent plasticity and is capable of classifying behaviors from biologging tags attached in various orientations and positions. As expected, classifier performance on the expanded suite of behavioral classes was not as robust, largely due to the similarities between the parent class and higher resolution classes. To explore this further, we build upon the framework detailed inWilson et al., 2018 using DDMT’s *Behavior Builder* and *Time Series* functions in an attempt to distinguish between behavioral classes with similar acceleration profiles, such as ‘Standing’ and ‘Vigilance’ behaviors. Applying these post-classification techniques to a subset of our current dataset drastically improved ‘Vigilance’ resulting in > 50% precision and recall metrics (see Fig. S2). Although encouraging, a more detailed investigation using larger datasets across multiple behaviors is needed. Furthermore, we expect that a similar improvement in classification performance could also be achieved in the pre-processing stages of classifier development by creating new features that capture subtle differences in accelerometer signatures between similar classes, like those identified between 'Standing' and 'Vigilance' behaviors.

The classifier was trained and tested solely from triaxial accelerometer data, an important *a priori* consideration. Because spatial features of the behavioral enclosure remained consistent throughout the study (e.g., location of water source and feeding area, shaded areas used as bedding sites), including locations and viewing angles of the cameras used to collect ground truth videos, it was important to exclude magnetometer data from classifier training and testing, as behaviors under these circumstances cannot be assumed to be randomly oriented. For example, in our study 'Resting' alignment was biased due to limited shaded areas in the enclosure. Had magnetometer data been included in the behavioral analysis, the classifier would likely identify 'Resting' using biased magnetometer data that have no relevance beyond the confines of the behavioral enclosure and would result in artifactually positive classifications that artificially inflate precision and recall metrics. We acknowledge that magnetometer data can be valuable for behavioral identification under certain contexts (Chakravarty et al., 2019; Williams et al., 2017); however, it remains unclear if studies that incorporate magnetometer data into machine learning analyses could be predisposed to such biases, as the relative contribution of magnetometer data used for behavioral identification are rarely provided.

None-the-less, triaxial magnetometer data can provide a wealth of opportunities for exploring movement ecology in greater detail, such as dead-reckoning analyses (Gunner et al., 2020) and studies of magnetic alignment (Begall et al., 2013; Červený et al., 2017). Given the salience of magnetometer data in biologging research, it is surprising that few studies have validated the precision of magnetic compass headings calculated from raw triaxial magnetometer data (Wilson et al., 2007). Therefore, we provide a detailed characterization of magnetic heading measurements under laboratory and natural contexts with magnetometer sensors mounted in different positions and orientations. Magnetic headings calculated by DDMT were consistent with ground truth predictions, with an overall median deviation from expected of 1.7° and 0° in the laboratory and field test, respectively. These data confirm that the magnetometer calibrations (i.e., soft- and hard-iron corrections) and tilt-compensation algorithms applied in DDMT are well-suited for extracting high-frequency magnetic compass bearings from raw magnetometer data.

Importantly, the field test carried out on free-roaming boar included magnetic measurements that were obtained from the five core behavioral classes. 'Running' had the largest average deviation relative to expected (9°) and may be due to the large variation in acceleration amplitudes that introduce 'noise' into accelerometer-dependent tilt-compensation calculations and/or the (in)ability of observers to accurately predict magnetic alignment from more erratic behavioral classes, such as 'Running'. Unintuitively however, magnetic headings obtained from 'Resting' behavior, characterized by little-to-no variation in acceleration profile, also had a relatively high deviation from expected (8°) and was likely due to an obstructed view of the animal's head alignment caused by a dense canopy covering the bedding area where boar would exclusively rest. It is noteworthy that magnetic compass performance remained accurate across all five core behaviors, and compass performance was evaluated across a representative range of all possible magnetic directions (i.e., 0° – 359°). This is the first study to our knowledge that has provided a detailed characterization of magnetic compass performance in free-roaming animals using ground truth data.

Of particular interest is the implementation of dead-reckoning to reconstruct high resolution movement traces in free roaming mammals. As a proof-of-concept, we take advantage of three important elements made possible by the IMSC presented in the current study: (i) the behavioral classifier capable of identifying ecologically relevant behaviors in free-roaming boar, (ii) a reliable stream of magnetic heading data recorded at sub-second intervals, and (iii) GPS fixes recorded at 30 min intervals. Dead-reckoning relies on vector integration, where vectors depend on speed (or distance travelled) and heading estimates derived from raw biologging data (for details see Bidder et al., 2015; Gunner et al., 2021). Deriving speed from biologging data is notoriously difficult (Cade et al., 2018) and previous work has assigned speed coefficients to manually labelled behavioral classes to estimate vector lengths for dead-reckoning path reconstruction (Bidder et al., 2015). We build upon this approach by using machine learning to identify behavioral classes from large volumes of continuous biologging data, which were then assigned speed coefficients based on ground truth observations. Coupling our behavioral classification techniques with the accuracy of our verified magnetic heading data yielded high-resolution track reconstruction that was further refined by 'anchoring' tracks to the landscape using time-synced GPS fixes (Fig. 5). The tortuosity of the reconstructed track in Fig. 5 that explicitly avoids boundaries and obstacles highlights the precision of these methods compared to using GPS data alone and offers a powerful approach to investigate movement ecology over multiple spatiotemporal scales.

Although the emergence of biologging techniques has revolutionized studies of animal ecology, a growing set of challenges accompanies these technologies, requiring multidisciplinary expertise. The IMSC developed here, coupled with a robust behavioral classifier and a detailed verification of magnetic heading performance provides a comprehensive system that can be adopted and adapted for future studies on terrestrial mammals.

References

1. Begall, S., Malkemper, E. P., Červený, J., Němec, P., & Burda, H. (2013). Magnetic alignment in mammals and other animals. *Mammalian Biology*, *78*, 10–20.
2. Bidder, O. R., Walker, J. S., Jones, M. W., Holton, M. D., Urge, P., Scantlebury, D. M., Marks, N. J., Magowan, E. A., Maguire, I. E., & Wilson, R. P. (2015). Step by step: reconstruction of terrestrial animal movement paths by dead-reckoning. *Movement Ecology*, *3*(1), 1–16.
3. Cade, D. E., Barr, K. R., Calambokidis, J., Friedlaender, A. S., & Goldbogen, J. A. (2018). Determining forward speed from accelerometer jiggle in aquatic environments. *Journal of Experimental Biology*, *221*(2), jeb170449. https://doi.org/10.1242/jeb.170449
4. Canty, A., & Ripley, B. (2020). *Boot: Bootstrap R (S-Plus) Functions (1.3-25)*.
5. Červený, J., Burda, H., Ježek, M., Kušta, T., Husinec, V., Nováková, P., Hart, V., Hartová, V., Begall, S., & Malkemper, E. P. (2017). Magnetic alignment in warthogs Phacochoerus africanus and wild boars

Sus scrofa. *Mammal Review*, *47*(1), 1–5.

6. Chakravarty, P., Maalberg, M., Cozzi, G., Ozgul, A., & Aminian, K. (2019). Behavioural compass: Animal behaviour recognition using magnetometers. *Movement Ecology*, *7*(1). https://doi.org/10.1186/s40462-019-0172-6
7. Cook, T. R., Gubiani, R., Ryan, P. G., & Muzaffar, S. B. (2017). Group foraging in Socotra cormorants: A biologging approach to the study of a complex behavior. *Ecology and Evolution*, *7*(7), 2025–2038.
8. Cooke, S. J., Hinch, S. G., LuCaS, M. C., & Lutcavage, M. (2012). Biotelemetry and biologging. *Fisheries Techniques, 3rd Edition. American Fisheries Society, Bethesda, Maryland*, 819–860.
9. Gunner, R. M., Holton, M. D., Scantlebury, M. D., van Schalkwyk, O. L., English, H. M., Williams, H. J., Hopkins, P., Quintana, F., Gómez-Laich, A., Börger, L., Redcliffe, J., Yoda, K., Yamamoto, T., Ferreira, S., Govender, D., Viljoen, P., Bruns, A., Bell, S. H., Marks, N. J., … Wilson, R. P. (2021). Dead-reckoning animal movements in R: a reappraisal using Gundog.Tracks. *Animal Biotelemetry*, *9*(1), 23. https://doi.org/10.1186/s40317-021-00245-z
10. Gunner, R. M., Wilson, R. P., Holton, M. D., Scott, R., Hopkins, P., & Duarte, C. M. (2020). A new direction for differentiating animal activity based on measuring angular velocity about the yaw axis. *Ecology and Evolution*, *10*(14), 7872–7886.
11. Gutzler, B. C., & Watson III, W. H. (2022). Measurements of American lobster heart rate and movements in the wild using a low-cost open source datalogger. *Marine Biology*, *169*(4), 52.
12. Holton, M. D., Wilson, R. P., Teilmann, J., & Siebert, U. (2021). Animal tag technology keeps coming of age: an engineering perspective. *Philosophical Transactions of the Royal Society B*, *376*(1831), 20200229.
13. Kassambara, A. (2020). *ggpubr:'ggplot2'based publication ready plots. R package version 0.4. 0*.
14. Kays, R., Crofoot, M. C., Jetz, W., & Wikelski, M. (2015). Terrestrial animal tracking as an eye on life and planet. *Science*, *348*(6240), aaa2478.
15. Kays, R., Davidson, S. C., Berger, M., Bohrer, G., Fiedler, W., Flack, A., Hirt, J., Hahn, C., Gauggel, D., & Russell, B. (2022). The Movebank system for studying global animal movement and demography. *Methods in Ecology and Evolution*, *13*(2), 419–431.
16. Kenward, R. E. (2000). *A manual for wildlife radio tagging*. Academic press.
17. Kirschvink, J. L. (1992). Uniform magnetic fields and double wrapped coil systems: Improved techniques for the design of bioelectromagnetic experiments. *Bioelectromagnetics*, *13*(5), 401–411. https://doi.org/DOI: 10.1002/bem.2250130507
18. Landler, L., Ruxton, G. D., & Malkemper, E. P. (2022). The multivariate analysis of variance as a powerful approach for circular data. *Movement Ecology*, *10*(1), 21. https://doi.org/10.1186/s40462-022-00323-8
19. Logan, R. K., Luongo, S. M., Vaudo, J. J., Wetherbee, B. M., & Shivji, M. S. (2023). Hunting behavior of a solitary sailfish Istiophorus platypterus and estimated energy gain after prey capture. *Scientific Reports*, *13*(1), 1484. https://doi.org/10.1038/s41598-023-28748-0

20. Martín López, L. M., Aguilar de Soto, N., Miller, P., & Johnson, M. (2016). Tracking the kinematics of caudal-oscillatory swimming: a comparison of two on-animal sensing methods. *Journal of Experimental Biology*, *219*(14), 2103–2109.
21. Matsumura, M., Watanabe, Y. Y., Robinson, P. W., Miller, P. J. O., Costa, D. P., & Miyazaki, N. (2011). Underwater and surface behavior of homing juvenile northern elephant seals. *Journal of Experimental Biology*, *214*(4), 629–636. https://doi.org/10.1242/jeb.048827
22. Percie du Sert, N., Hurst, V., Ahluwalia, A., Alam, S., Avey, M. T., Baker, M., Browne, W. J., Clark, A., Cuthill, I. C., Dirnagl, U., Emerson, M., Garner, P., Holgate, S. T., Howells, D. W., Karp, N. A., Lazic, S. E., Lidster, K., MacCallum, C. J., Macleod, M., … Würbel, H. (2020). The ARRIVE guidelines 2.0: Updated guidelines for reporting animal research. *PLOS Biology*, *18*(7), e3000410. https://doi.org/10.1371/journal.pbio.3000410
23. Portugal, S. J., & White, C. R. (2018). Miniaturization of biologgers is not alleviating the 5% rule. *Methods in Ecology and Evolution*, *9*(7), 1662–1666.
24. Rattenborg, N. C., Voirin, B., Cruz, S. M., Tisdale, R., Dell'Omo, G., Lipp, H.-P., Wikelski, M., & Vyssotski, A. L. (2016). Evidence that birds sleep in mid-flight. *Nature Communications*, *7*(1), 12468.
25. Ropert-Coudert, Y., & Wilson, R. P. (2005). Trends and perspectives in animal-attached remote sensing. *Frontiers in Ecology and the Environment*, *3*(8), 437–444.
26. Rutz, C., & Hays, G. C. (2009). *New frontiers in biologging science*. The Royal Society London.
27. Shepard, E. L. C., Wilson, R. P., Quintana, F., Laich, A. G., Liebsch, N., Albareda, D. A., Halsey, L. G., Gleiss, A., Morgan, D. T., & Myers, A. E. (2008). Identification of animal movement patterns using tri-axial accelerometry. *Endangered Species Research*, *10*, 47–60.
28. Tuia, D., Kellenberger, B., Beery, S., Costelloe, B. R., Zuffi, S., Risse, B., Mathis, A., Mathis, M. W., van Langevelde, F., & Burghardt, T. (2022). Perspectives in machine learning for wildlife conservation. *Nature Communications*, *13*(1), 792.
29. Walker, J. S., Jones, M. W., Laramee, R. S., Holton, M. D., Shepard, E. L. C., Williams, H. J., Scantlebury, D. M., Marks, N. J., Magowan, E. A., & Maguire, I. E. (2015). Prying into the intimate secrets of animal lives; software beyond hardware for comprehensive annotation in 'Daily Diary'tags. *Movement Ecology*, *3*, 1–16.
30. Wang, G. (2019). Machine learning for inferring animal behavior from location and movement data. *Ecological Informatics*, *49*, 69–76. https://doi.org/https://doi.org/10.1016/j.ecoinf.2018.12.002
31. Wild, T. A., Wikelski, M., Tyndel, S., Alarcón-Nieto, G., Klump, B. C., Aplin, L. M., Meboldt, M., & Williams, H. J. (2023). Internet on animals: Wi-Fi-enabled devices provide a solution for big data transmission in biologging. *Methods in Ecology and Evolution*, *14*(1), 87–102.
32. Williams, H. J., Holton, M. D., Shepard, E. L. C., Largey, N., Norman, B., Ryan, P. G., Duriez, O., Scantlebury, M., Quintana, F., Magowan, E. A., Marks, N. J., Alagaili, A. N., Bennett, N. C., & Wilson, R. P. (2017). Identification of animal movement patterns using tri-axial magnetometry. *Movement Ecology*, *5*(1). https://doi.org/10.1186/s40462-017-0097-x

33. Williams, H. J., Taylor, L. A., Benhamou, S., Bijleveld, A. I., Clay, T. A., de Grissac, S., Demšar, U., English, H. M., Franconi, N., & Gómez-Laich, A. (2020). Optimizing the use of biologgers for movement ecology research. *Journal of Animal Ecology*, *89*(1), 186–206.
34. Wilmers, C. C., Nickel, B., Bryce, C. M., Smith, J. A., Wheat, R. E., & Yovovich, V. (2015). The golden age of bio-logging: How animal-borne sensors are advancing the frontiers of ecology. *Ecology*, *96*(7), 1741–1753.
35. Wilson, R., Holton, M., Virgilio, A. di, Williams, H., Shepard, E., Lambertucci, S., Quintana, F., Sala, J. E., Balaji, B., & Lee, E. S. (2018). Give the machine a hand: A Boolean time-based decision-tree template for rapidly finding animal behaviours in multisensor data. *Methods in Ecology and Evolution*, *9*(11), 2206.
36. Wilson, R. P., Grant, W. S., & Duffy, D. C. (1986). Recording devices on free-ranging marine animals: does measurement affect foraging performance? *Ecology*, *67*(4), 1091–1093.
37. Wilson, R. P., Liebsch, N., Davies, I. M., Quintana, F., Weimerskirch, H., Storch, S., Lucke, K., Siebert, U., Zankl, S., & Müller, G. (2007). All at sea with animal tracks; methodological and analytical solutions for the resolution of movement. *Deep Sea Research Part II: Topical Studies in Oceanography*, *54*(3–4), 193–210.
38. Wilson, R. P., Rose, K. A., Gunner, R., Holton, M. D., Marks, N. J., Bennett, N. C., Bell, S. H., Twining, J. P., Hesketh, J., & Duarte, C. M. (2021). Animal lifestyle affects acceptable mass limits for attached tags. *Proceedings of the Royal Society B*, *288*(1961), 20212005.
39. Wilson, R. P., Shepard, E. L. C., & Liebsch, N. (2008). Prying into the intimate details of animal lives: Use of a daily diary on animals. *Endangered Species Research*, *4*(1–2), 123–137. https://doi.org/10.3354/esr00064

Footnotes

1. If no subsequent event was available - only possible at the end of a file - the candidate event was discarded as unclassifiable. A small number of time steps at the beginning of each file were similarly discarded, since a fixed amount of time must accumulate before the classifier can make its first decision. Such edge-effects have negligible impact on classifier evaluation.
2. We do not imply that the field of animal-borne telemetry and biologging is not collaborative, and indeed, would argue the opposite. However, we suggest that increased overlap between methodologies may encourage further collaboration and promote the growth of this emerging discipline.

Tables

Tables 1 to 4 are available in the Supplementary Files section.

Table 5. Behavioral Classifier Performance Summary. Precision and recall percentages are shown for all six behavioral classes, partitioned by individual, as well as overall classifier accuracy (%) per individual.

Behavioral Classifier Performance Summary

Behavioral Class	B4		B7		B30	
	Precision (%)	Recall (%)	Precision (%)	Recall (%)	Precision (%)	Recall (%)
Walk	77.8	86.4	74.2	91.2	87.1	85.6
Other	34.1	58.0	53.7	72.3	4.3	17.4
Rest	93.5	85.6	98.3	95.0	98.6	100.0
Forage	96.1	88.6	62.9	39.1	94.6	96.2
Run	92.3	78.4	94.1	75.2	75.0	60.2
Stand	58.7	72.4	85.9	83.0	95.3	67.3
Overall Accuracy (%)	**83.5**		**84.2**		**89.9**	

Table 6 and 7 are available in the Supplementary Files section.

Figures

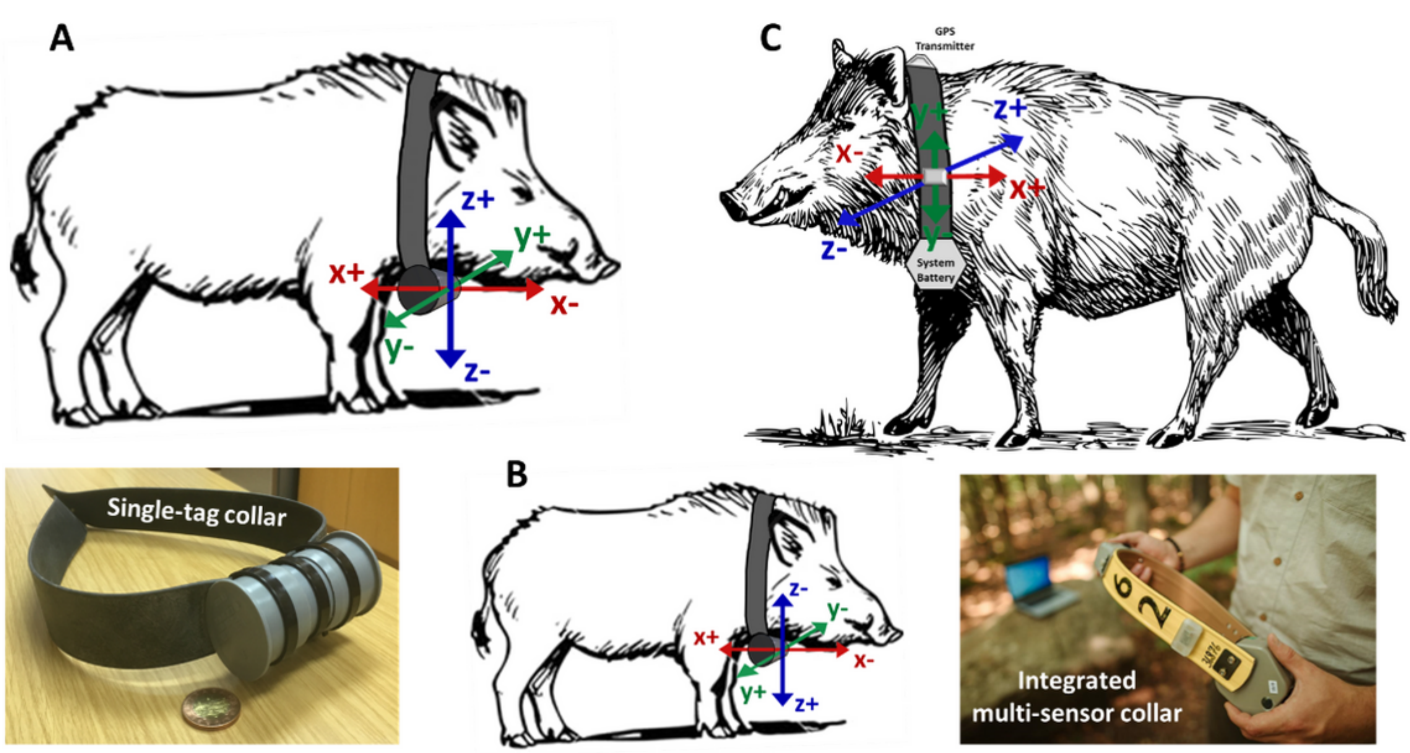

Figure 1

Biologging collars used throughout the study. Accelerometer axes orientation is superimposed on the logger and axis polarity indicates the acceleration value as the axis is pointed toward gravity. Note the different axis alignments between STC designs (A, B). Both the logger position and logger orientation used in all IMSCs (C) differs from STC logger position and orientations. Photographs both collar designs are shown below their respective schematics.

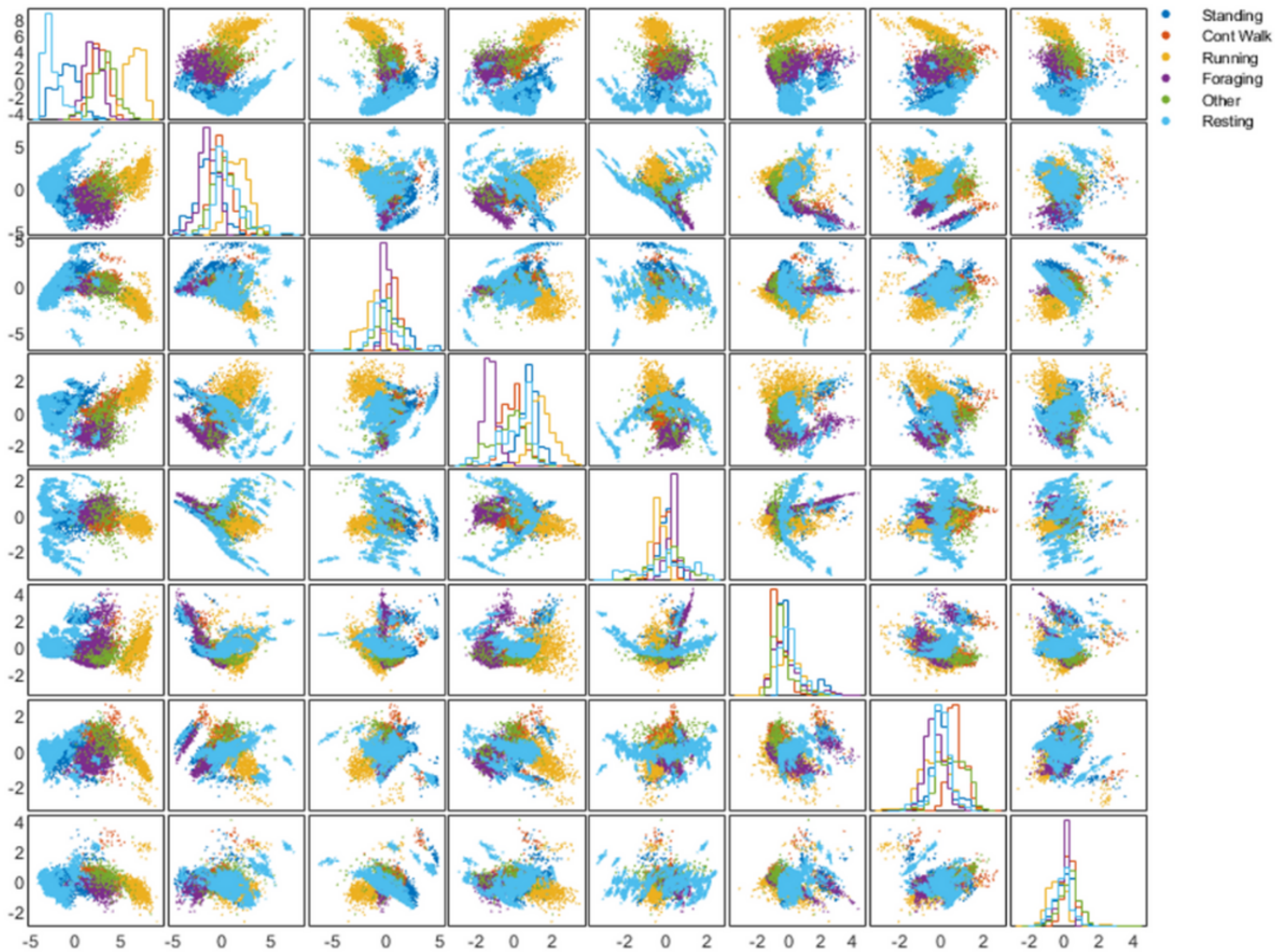

Figure 2

Matrix showing plots of all possible pairs of the 8 principal components (PCs) that were used in classification. Points correspond to training observations (n=13461 in each plot) and are colored according to behavioral class. Numbering columns and rows each from 1-8, respectively, beginning at the top left corner of the matrix, the column number corresponds to the PC plotted on the horizontal axis and the row number to the PC plotted on the vertical axis. For example, the plot in row 3, column 2 has the second PC plotted on the horizontal axis and the third PC plotted on the vertical axis. The plots along the diagonal are histograms, colored by class, for each of the 8 PCs. (Plots mirrored across the diagonal show the same two PC's with the axes swapped).

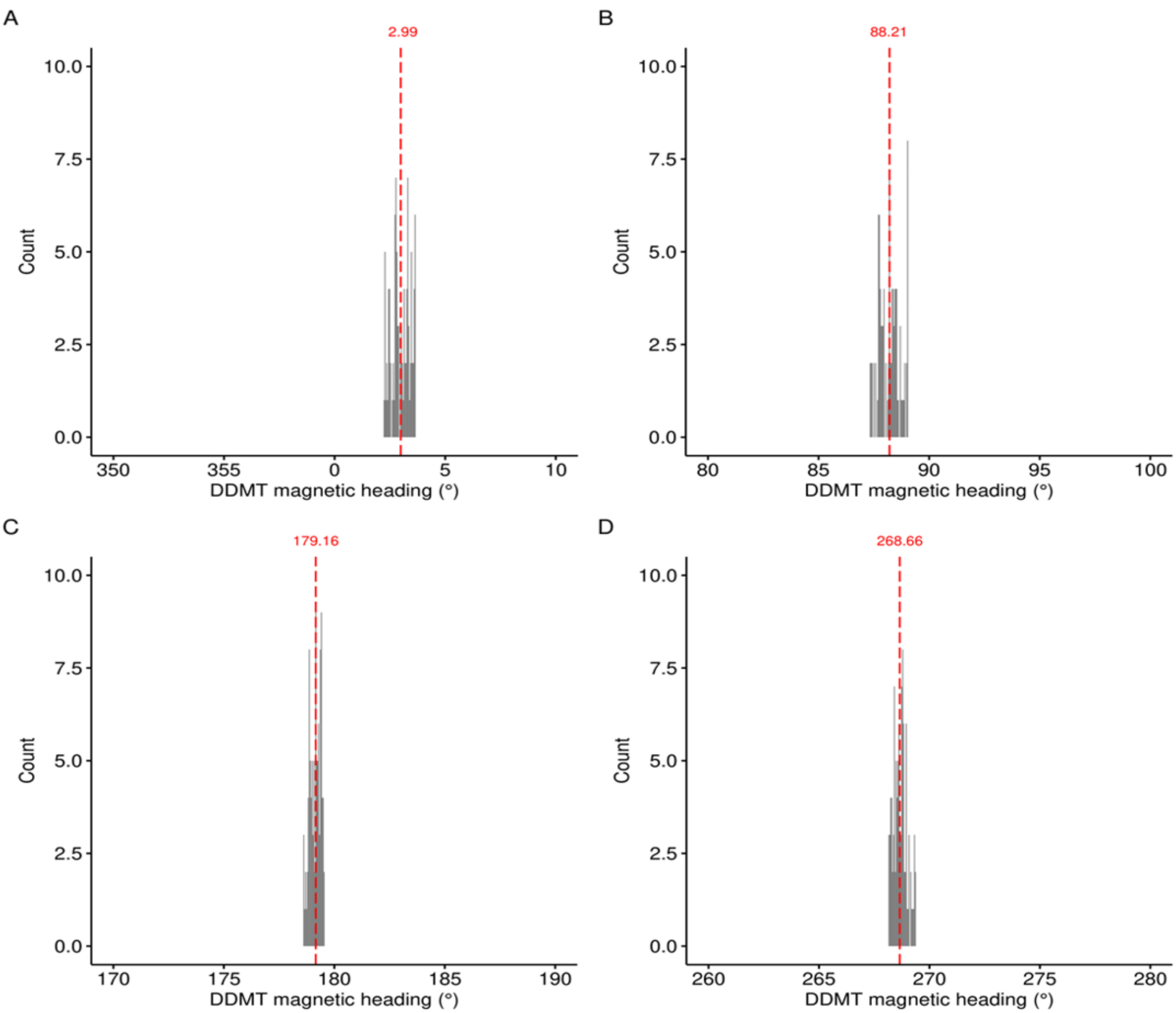

Figure 3

Lab test of triaxial magnetometer data used to calculate magnetic heading measurements in DDMT software. Histograms plot the total count of 100 samples (10 Hz x 10 sec) recorded in each magnetic field alignment (i.e. mN = topoN, E, S, W), relative to magnetic heading bearings calculated in DDMT after magnetometer calibration procedures. Plots A-D correspond to experimentally generated Earth-strength magnetic fields aligned at: North (0°), East (90°), South (180°), West (270°), respectively. Median values for each magnetic field alignment are shown in red.

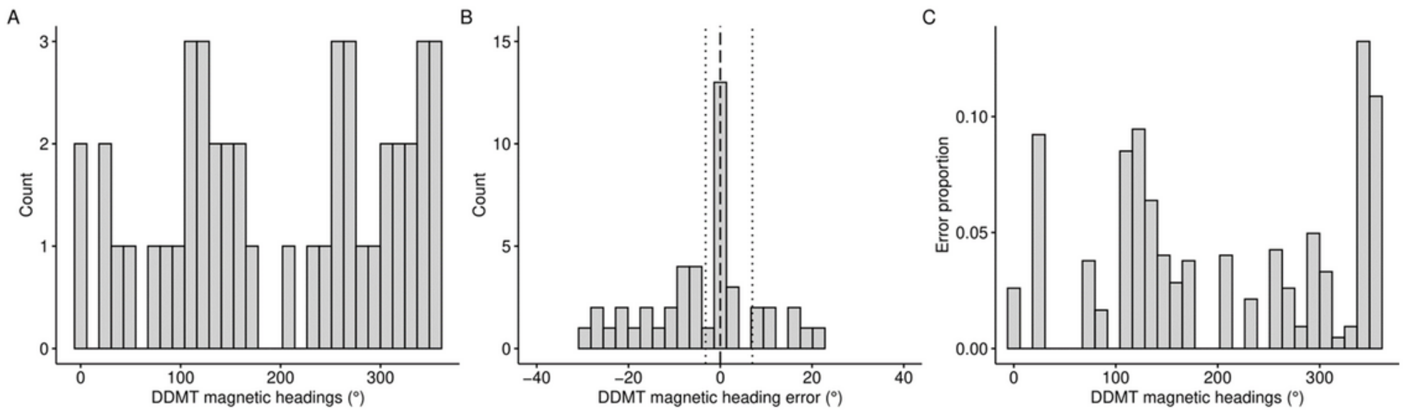

Figure 4

Results from the magnetometer field test collected from free roaming individuals equipped with STC and IMSC designs. A total of 45 samples were evaluated and compared to ground-truth predictions of magnetic heading. A) Histogram of the overall distribution of magnetic compass measurements produced by DDMT shows that samples were obtained from the range of possible compass directions. B) The discrepancy between DDMT magnetic compass measurements and ground-truth recordings, i.e., DDMT magnetic heading output error (median error = 0°, black dashed line; bootstrapped 95% CI: -3.1° and 6.9°). C) The error produced by DDMT was uniform across the range of possible magnetic compass headings.

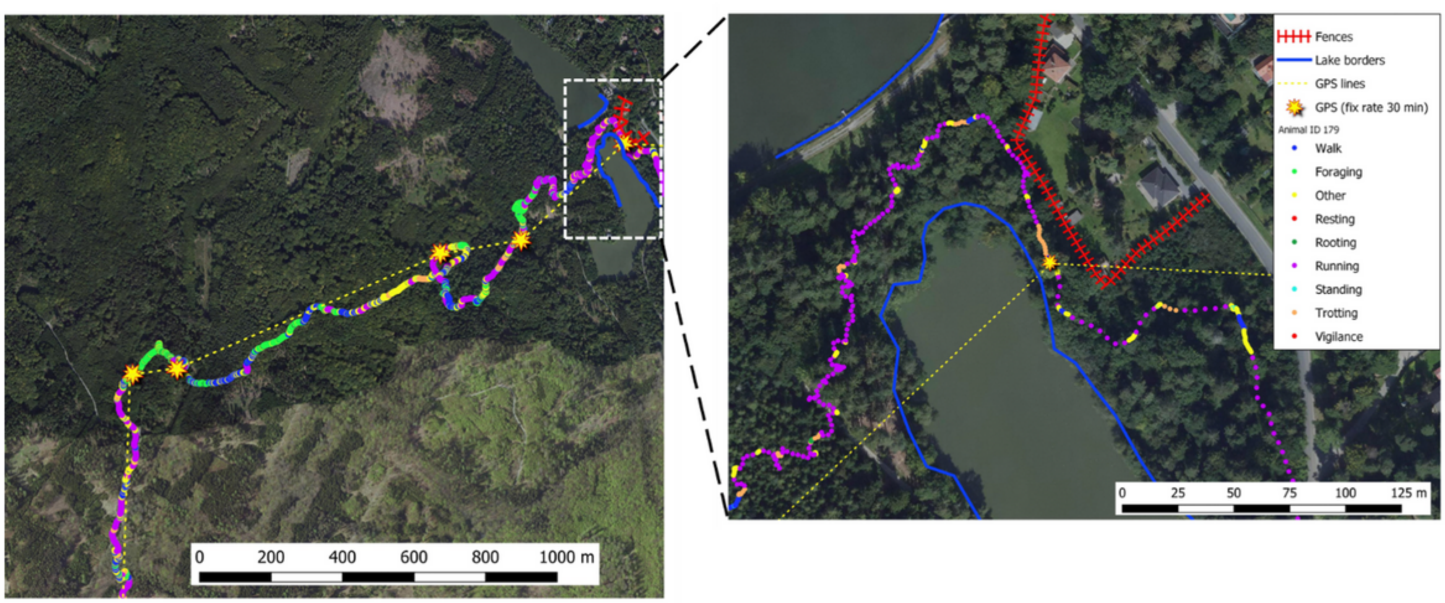

Figure 5

Example of dead-reckoning path reconstruction in a free ranging wild boar equipped with the IMSC. Behaviors were first identified from continuous accelerometer data using the behavioral classifier and were then uploaded to DDMT. User-defined speed coefficients were assigned to each behavioral class which was then then integrated with time-synced magnetic heading data to reconstruct movement traces between GPS fixes, which anchored the tack to the landscape. Right insight shows a detailed track segment highlighting the added value of dead-reckoning analyses compared to using GPS data alone and lends further credibility to the precision of the dead-reckoning methodology used in this study.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- BoarMSSupplementaryMaterialsFinal.docx
- Tables.docx