Using accelerometer-based behavioral classification to enhance scavenger conservation

Gideon Vaadia*[1], Marta Acácio*[1,2,3,4], Tal Agassi[5], Nili Anglister[1], Yigal Miller[6], Ohad Hatzofe[6], Patricia Mateo-Tomás[7], Jorge Rodríguez-Pérez[7], María Fernández-García[7]; Pedro P. Olea[8,9], Ignacio Otero[10], Noa Pinter-Wollman[11], Moni Shahar[5] & Orr Spiegel[1]

* Equal contributions

Corresponding author

Marta Acácio - marta.acacio@gmail.com; martaoc@tauex.tau.ac.il

Published version:

Vaadia, G., Acácio, M., Agassi, T., Anglister, N., Miller, Y., Hatzofe, O., Mateo-Tomás, P., Rodríguez-Pérez, J., Fernández-García, M., Olea, P. P., Otero, I., Pinter-Wollman, N., Shahar, M., & Spiegel, O. (2025). Using accelerometer-based behavioural classification to enhance scavenger conservation. Journal of Applied Ecology, 00, 1–18. https://doi.org/10.1111/1365-2664.70180

[1] School of Zoology, Faculty of Life Sciences, Tel Aviv University, Tel Aviv, Israel

[2] CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, InBIO Laboratório Associado, Campus de Vairão, University of Porto, 4485-661 Vairão, Portugal

[3] CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, InBIO Laboratório Associado, Instituto Superior de Agronomia, University of Lisboa, 1349-017 Lisboa, Portugal

[4] BIOPOLIS Program in Genomics, Biodiversity and Land Planning, CIBIO, Campus de Vairão, 4485-661 Vairão, Portugal

[5] Center of AI and Data Science, Tel Aviv University (TAD), Tel Aviv, Israel

[6] Israeli Nature and Parks Authority, Am Ve’Olamo 3, Jerusalem, Israel

[7] Biodiversity Research Institute (CSIC – University of Oviedo – Principality of Asturias), Mieres, Spain

[8] Terrestrial Ecology Group (TEG_UAM), Departamento de Ecología, Facultad de Ciencias, Universidad Autónoma de Madrid (UAM), Madrid, Spain

[9] Centro de Investigación en Biodiversidad y Cambio Global (CIBC-UAM), Universidad Autónoma de Madrid (UAM), Madrid, Spain

[10] GREFA (Grupo de Rehabilitación de la Fauna Autóctona y su Hábitat), Majadahonda, Spain

[11] Department of Ecology and Evolutionary Biology, University of California Los Angeles, Los Angeles, California, USA

Abstract

1. Human activities are endangering animal species globally and implementing effective conservation strategies requires understanding animal behavior and ecology. Advancements in GPS tracking technology, accelerometry, and machine learning algorithms are allowing the *in-situ* study of animal movement and behavior remotely. However, the challenge of building supervised machine learning algorithms and collecting the large datasets required to train them is hampering the widespread use of these tools. Additionally, the reliability of these models in classifying unobserved behaviors is rarely validated, resulting in possible classification errors.
2. We built a supervised accelerometer-based behavioral classification model for griffon vultures (*Gyps fulvus*). Similarly to most other avian scavenger populations worldwide, griffons are critically endangered in Israel and neighboring countries, mostly due to feeding on poisoned carcasses. Thus, identifying this scavenger's feeding behavior and foraging areas is crucial for their conservation.
3. We trained a Random Forest model on acceleration data of 14 captive and 17 free-roaming griffons. We classified 5783 behavioral observations into 6 classes: feeding, lying, standing, other ground behaviors, flapping and soaring flight. The model performed well (0.96 accuracy, 0.89 precision and 0.82 recall) and, importantly, feeding behaviors were accurately classified (0.87 precision, 0.92 recall). We calculated an observation-specific confidence score and demonstrated its effectiveness in identifying true- and false-positive classifications, in both captive and free-roaming individuals. Finally, we used our model to reliably identify feeding hotspots, where vultures can be at higher risk of poisoning.

4. *Synthesis and applications.* We provide a tool to help identify vulture feeding hotspots, supporting carcass management efforts to prevent poisoning. Integrated with near real-time tracking, our model can support global efforts to combat scavenger poisoning. The training dataset, model and codes are provided in a user-friendly platform, along with a conceptual framework, to encourage use by ecologists and conservation practitioners.

Resumen

1. Las actividades humanas suponen una amenaza para la fauna a escala global e implementar estrategias de conservación efectivas al respecto requiere entender el comportamiento y la ecología animal. Avances en tecnologías como el seguimiento GPS, los acelerómetros y los algoritmos de aprendizaje automático facilitan el estudio remoto in situ de los movimientos y comportamientos animales. Sin embargo, elaborar algoritmos supervisados de aprendizaje automático y recopilar la gran cantidad de datos necesarios para entrenarlos son retos que dificultan el uso generalizado de estas herramientas. Además, la fiabilidad de estos modelos al clasificar comportamientos no observados no suele validarse, lo que pueda dar lugar a errores de clasificación.
2. Elaboramos un modelo supervisado de clasificación de comportamientos de buitre leonado (Gyps fulvus) con datos de acelerómetro. En consonancia con la mayoría de las poblaciones de carroñeros en el mundo, los buitres están críticamente amenazados en Israel y países vecinos, principalmente por consumir carroñas envenenadas. Por tanto, identificar comportamientos y áreas de alimentación de estos carroñeros es clave para su conservación.
3. Entrenamos un modelo “Random Forest” con datos de 14 buitres en cautividad y 17 en libertad, clasificando 5.783 comportamientos observados en 6 clases: alimentación, tumbado, posado, otros comportamientos en suelo, aleteo y planeo. El modelo funcionó bien (exactitud: 0,96, precisión: 0,89 y sensibilidad:

0,82), identificando con exactitud los comportamientos de alimentación (precisión: 0,87 y sensibilidad: 0,92). Calculamos un índice de confianza específico de cada observación, que discriminó con solvencia verdaderos y falsos positivos tanto en buitres en cautividad como en libertad. Finalmente usamos el modelo para identificar zonas de alimentación, dónde el riesgo de envenenamiento de los buitres pudiera incrementarse.

4. Síntesis y aplicaciones: Proporcionamos una herramienta útil para identificar zonas de alimentación de buitres, apoyando así tareas de gestión para prevenir su envenenamiento. Junto con el seguimiento casi en tiempo real, nuestro modelo puede ayudar en la lucha contra el envenenamiento de fauna carroñera. Los datos de entrenamiento, el modelo y los códigos, junto con un marco conceptual, están disponibles en una plataforma de fácil uso para fomentar su utilización por ecólogos y profesionales de la conservación.

Keywords (up to 8):

Accelerometer, Behavior classification, Random Forest, Griffon Vulture, Poisoning, Conservation, Biotelemetry, Supervised machine learning.

Introduction

Anthropogenic activities are endangering animals around the world (Venter et al., 2016). To combat the current wave of species extinction, we need to understand animal behavior and ecology to minimize threats and conflicts, and to implement effective conservation strategies (Fehlmann et al., 2023; van Eeden et al., 2018). Over the last couple of decades, technological advances have provided tremendous insights into animal ecology and behavior (Kays et al., 2015; Nathan et al., 2022), often with direct implications for conservation (Tuia et al., 2022). The use of GPS-tracking technology, for instance, has contributed to a deeper understanding of animal movements and space use, which can inform the design and administration of protected areas (Hays et al., 2019). GPS-tracking has also helped identifying the locations of animal mortality (Sergio et al., 2019) and location-specific causes of mortality (Serratosa et al., 2024). Uncovering the location and spatial extent of animal threats is critical for managing endangered species (Kane et al., 2022; Olea & Mateo-Tomás, 2014) and for mitigating human-induced mortality (Serratosa et al., 2024).

While GPS tracking provides valuable insights into where and when animals use particular habitats, it does not provide direct information on the specific behaviors animals perform within those habitats. For instance, a site where an individual stops could be used for resting, foraging, or other behaviors - which cannot be distinguished using GPS data alone. Complementing GPS-tracking data with additional sensors offers insights into the behavior and energy use of elusive and cryptic animals (Shepard et al., 2008; Smith & Pinter-Wollman, 2021; Spiegel et al., 2015; Tuia et al., 2022; Williams et al., 2020). Tri-axial accelerometers (ACC) are widely used in behavioral research, among other sensors (Brown et al., 2013; Kays et al., 2015; Nathan et al., 2012). These devices measure acceleration in three orthogonal axes (sway – x; surge – y; and heave - z) that change according to the animal's posture and locomotion. These measurements can be recorded either continuously or in "bouts" (i.e., sampling units) of a few seconds at varying resolution (i.e., frequency, in Hz) and intervals (for example, recording for 5 seconds at 20Hz, every 10 minutes). Different acceleration signatures enable the measurement of movement-related energy expenditure (Gleiss et al., 2011; Halsey et al., 2009) and can be used to distinguish among different behaviors (Shepard et al., 2008), for example, for estimating flight duration in small migratory passerines (Bäckman et al., 2017).

Machine learning algorithms are used to classify raw acceleration bouts into different behavioral classes (Nathan et al., 2012; Resheff et al., 2014; Valletta et al., 2017; Wang, 2019; Yu et al., 2021). These algorithms can operate in an unsupervised manner, identifying similarities in acceleration data to produce unlabeled clusters of similar measurements that subsequently need to be manually classified into specific behaviors (Chimienti et al., 2016; Wang, 2019). Alternatively, supervised learning involves training an algorithm with a dataset in which each behavior is labeled, allowing the algorithm to 'learn' the distinctive acceleration patterns of different behaviors (Nathan et al., 2012; Wang, 2019; Yu et al., 2021). However, depending on the level of detail required and on how distinctive the behaviors are, assembling a training dataset can be laborious, as it typically requires direct observations of animals in the wild or in captivity, synchronized with the ACC measurements (Campbell et al., 2013; Dickinson et al., 2021). Despite these difficulties, supervised machine learning algorithms have been successfully used to classify behaviors across diverse animal groups, including baboons (Fehlmann et al., 2023), large pelagic fish (Clarke et al., 2021), sea turtles (Jeantet et al., 2020), condors, and other vultures (Rast et al., 2024; Spiegel et al., 2013; Williams et al., 2015). Commonly used algorithms include artificial neural networks, extreme gradient boosting, and random forests (Resheff et al., 2014; Yu et al., 2021). Random forests have the advantage of being able to model complex interactions between the, often correlated, predictor variables, therefore not requiring the pre-processing and filtering of variables (Cutler et al., 2007), and simplifying behavioral classification.

Once a machine learning model is trained, it can classify new, unseen accelerometer data into the trained behavioral classes, invariably resulting in some classification errors (Glass et al., 2020; Jeantet et al., 2020). Errors emerge from a few, non-mutually exclusive, processes. First, acceleration bouts, particularly long ones, may include transitions among behaviors resulting in a mixture of different acceleration signatures (Resheff et al., 2024). Second, rare behaviors may be underrepresented or missing from the limited training dataset (e.g., seasonal and rare behaviors such as copulation). Third, the behavioral repertoire of some individuals may be broader than what the algorithm is trained for. Because some behaviors might be difficult to observe in captivity (e.g., flight behaviors, Williams et al., 2015), this last error is particularly relevant for algorithms trained on captive individuals that are used to predict the behaviors of wild animals (Dickinson et al., 2021). Still, the models must choose the best fitting behavioral class among the available options, even if none provides a particularly good fit. These errors demand a mechanism to verify the accuracy of each behavioral classification, allowing the model to distinguish between true-positive and false-positive classifications (Bidder et al., 2014; Glass et al., 2020). While some studies offer guidance on how to best use and analyze large acceleration datasets (e.g., Leos-Barajas et al., 2017; Resheff et al., 2014; Williams, Taylor, et al., 2020), the complexity of these tools remains a barrier for non-experts, hindering their use in conservation science and practice.

Here we develop an accelerometer-based behavioral classification tool and validate its real-world application in ecology and conservation, using griffon vultures (*Gyps fulvus*) as a case study. As obligate scavengers, vultures support key ecosystem functions by consuming carcasses and recycling nutrients (Buechley & Şekercioğlu, 2016). Yet, around the world, 70% of vulture species are in danger of extinction, with poisoning driven by consuming carcasses containing toxic substances being one of the leading causes for population declines (Ives et al., 2022; Ogada et al., 2012; Plaza et al., 2019). Poisoning can be either intentional or unintentional. For instance, poachers may lace carcasses with poison to prevent these raptors from alerting environmental authorities of poached wildlife (Mateo-Tomás & López-Bao, 2020; Ogada et al., 2016), and farmers may do so for combating pests and mammalian carnivores. Anti-inflammatory drugs used to treat cattle are also lethal to vultures, leading to poisoning at these carcasses (López-Bao & Mateo-Tomás, 2022; Ogada et al., 2012; Plaza et al., 2019). Several hundred vultures may quickly gather to eat at a single carcass, increasing their vulnerability to mass poisoning events (McNutt & Bradley, 2014). Other scavenger species also feed on carcasses (Olea et al., 2019), exposing them to similar risks of poisoning (Katzner et al., 2024; López-Bao & Mateo-Tomás, 2022). Early detection of carcasses might facilitate their proper management to, for example, prevent vultures and other wild species from feeding on carrion with toxic substances by removing these from the environment. Moreover, when poisoning events are promptly detected, vultures and other animals can undergo medical intervention (Acácio et al., 2023; Anglister et al., 2023). Considering the vulture's large roaming areas (Spiegel et al., 2015), tracking technology and behavioral classification are essential tools to identify vultures' feeding events. Vultures can therefore act as sentinels, facilitating carcass detection, and maximizing prompt intervention efforts to reduce detrimental effects associated with consumption of contaminated carrion. For instance, by reducing the number of fatalities at a poisoning event to avoid long-term effects on species' populations (Acácio et al., 2023; Slabe et al., 2022).

In this study, our goals are to 1) develop an ACC-based behavioral classification algorithm, which, together with the training dataset and a conceptual framework of the methodological workflow, is made freely available to conservationists and ecologists; 2) validate the algorithm's classifications, by comparing the confidence scores of true-positive and false-positive classifications, using both the training dataset and data from free-roaming vultures; and 3) apply our novel algorithm to real-life scenarios with important conservation implications – i.e., rapid carcass detection to prevent vulture poisoning. Ultimately, our goal is to combine technological advancements in GPS and accelerometry to improve wildlife conservation efforts, and to develop a tool that is easily transferable to other systems.

Materials and Methods

Study system

The study took place in Israel, where griffon vultures are critically endangered (Mayrose et al., 2017). A historical population of thousands of griffons is currently declining; three decades ago, there were only 400 griffons in this population, and fewer than 200 individuals remain today (Hatzofe, 2020). Pesticide poisoning from consuming laced carcasses is the leading cause of griffon mortality, accounting for 45% of documented deaths between 2010-2021 in this region (Anglister et al., 2023). Lead poisoning and ingestion of animals treated with anti-inflammatory drugs each contribute to 6% of mortality events, posing additional threats to this population (Anglister et al., 2023). To prevent the local extinction of this species, the Israel Nature and Parks Authority (INPA) runs an intricate management program, including the provisioning of contaminant-free food at supplementary feeding stations (Spiegel et al., 2013, 2015), the release of captive-bred and translocated griffons (Efrat et al., 2020), and individually tracking vultures using GPS-Accelerometer transmitters, to identify poisoning events and other threats. When wild carcasses are detected in a random location within areas of known pastoral activity and poisoning history, or when vultures exhibit minimal movement, suggesting they are unwell, rangers are sent to the field to remove the carcasses and/or transport affected individuals to a wildlife hospital, underscoring the critical role of GPS-tracking data for the conservation of this population (Acácio et al., 2023).

Every year, approximately 100 free-roaming griffons are captured by the INPA using a cage trap, to identify the individuals with metal and color rings and with patagial tags. In these trapping events, a few individuals are fitted with a GPS-ACC transmitter (Ornitela OT-50) using a Teflon harness in a leg-loop configuration. The transmitters are equipped with solar panels that recharge the batteries, and transmit the collected data via the GSM network, eliminating the need for recapturing individuals to retrieve information. The current study did not require additional ethical approval, since it uses tracking data collected from these GPS-ACC tags and no captures of free-roaming vultures were conducted specifically for this study. The capture and tagging procedures were approved by ethics committee of the Israel Nature and Parks Authority (permit number 42166).

Conducting observations to build an ACC training dataset

Between January 2020 and February 2022, we deployed GPS-ACC transmitters on 31 griffon vultures, 14 captive vultures and 17 free-roaming individuals. The captive vultures were housed in in 4 breeding programs, rehabilitation or wildlife facilities, in Israel and in Spain: Ramat Hanadiv (Israel, n = 4), Hai-Bar Carmel (Israel, n = 4), Cabárceno Wildlife Park (Spain, n = 3) and GREFA wildlife hospital (Spain, n = 3). In each cage, there were 6 to 12 vultures. Additional behavioral data was collected in Israel for 17 free-roaming griffon vultures. One individual dropped his transmitter and was deployed with another device (thus there were 31 individuals but 32 transmitters). In Israel, the transmitters were deployed using a leg-loop harness and in Spain the loggers were deployed using a backpack harness.

The transmitters were programmed to collect GPS and ACC at independent schedules, and differently for captive and free-roaming griffons. The transmitters of captive griffons were programmed to collect tri-axial accelerometer data at 20Hz almost continuously (10-min-long periods, with a 1 sec interval in between). These 10min bouts were parsed into 5 sec bouts to match the free-roaming dataset. Bouts of 5 sec at 20Hz were recorded for free-roaming griffons every 10 min, depending on the transmitter's battery charge (see Supplementary Material for details).

To classify each 5 sec ACC bout as a particular behavior, we conducted direct observations and video recordings of the tagged griffons, both in captivity and in the wild. In total, we performed observations for 79 days. Direct observations of captive and wild griffons were conducted with a spotting scope (Swarovski ATX spotting scope 85mm), ensuring a sufficient distance to not disturb the vulture's natural behavior. The video recordings were captured using a camera mounted on a wall support in Spain, and with nest cameras at the captive breeding facilities, or live streaming nest cameras at wild nests in Israel (BirdLife Israel, 2022). The direct observations of wild vultures were performed at roosting sites, at approximately 250 of the individuals.

We recorded six ecologically important behavioral classes: 'Standing' - vulture is resting upright (could be roosting, and may include minor preening and changes in body posture); 'Lying' - vulture is lying parallel to the ground, either resting or incubating; 'Feeding' – vulture is either directly eating from a carcass, or engaged in intense social interactions next to the carcass (e.g., fighting or posturing towards other vultures before eating); 'Ground' - includes all other active ground behaviors which are not directly related to feeding or resting (e.g., walking, running, hopping, etc.); 'Flapping' - active flight with wingbeats; and 'Soaring' - passive flight (e.g., thermal soaring, gliding, etc.). Because long flights do not occur in captivity, we used GPS-ACC data from 17 free-roaming griffons in southern Israel to classify 'Soaring' (passive) and 'Flapping' (active) flight behaviors. We identified segments of continuous flight using the GPS ground speed (ground speed >4m/sec) and plotted the acceleration measurements taken during these flights. The acceleration signatures of soaring and passive flights are so distinctive (Figure 1B,C, Williams et al., 2015) that there was no need to ground-truth these behaviors with visual sightings (which would be challenging, considering their large roaming areas).

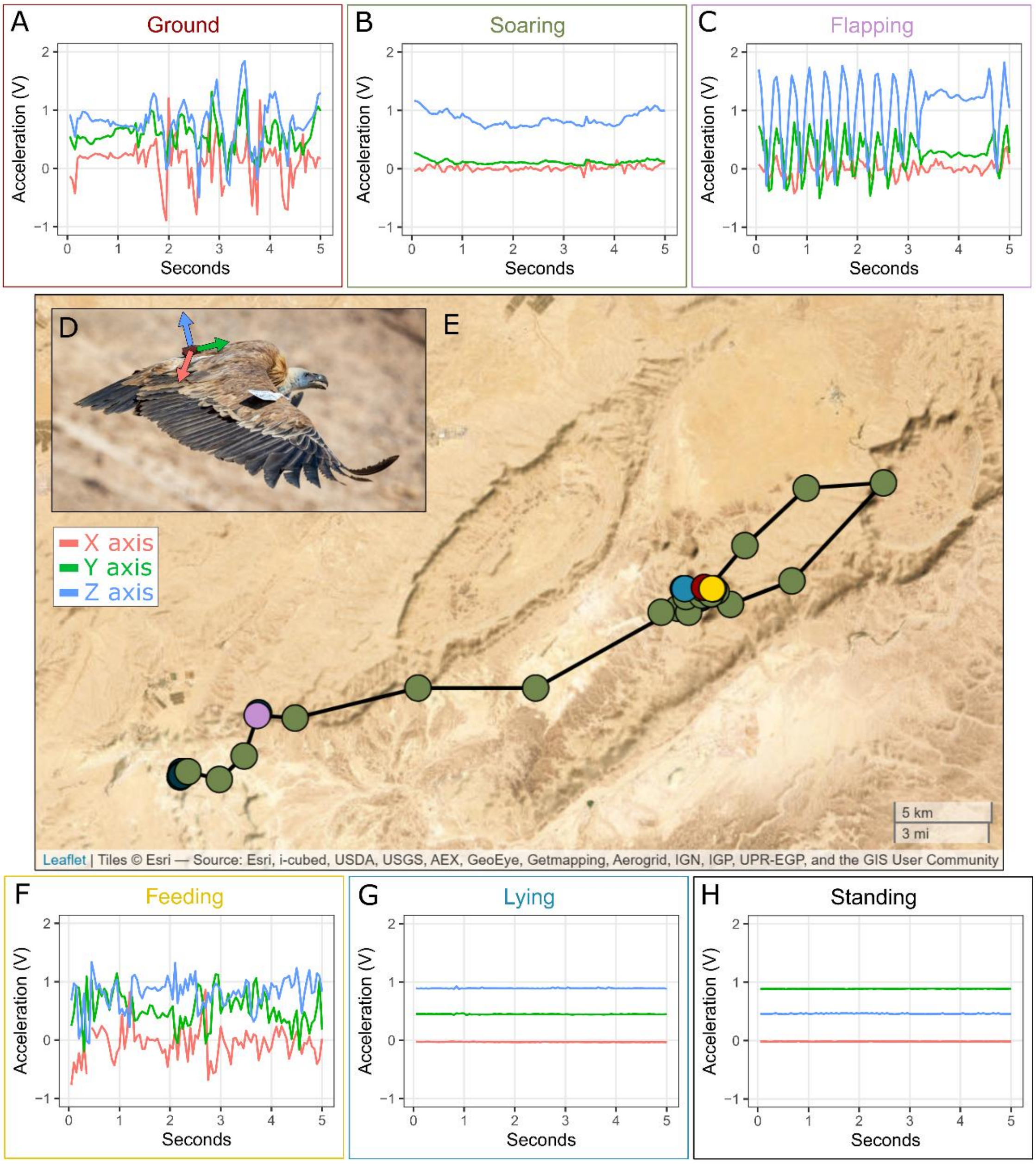


Figure 1 - Examples of accelerometer-based classification of griffon vulture's behaviors. Acceleration measurements of bouts classified as: (A) 'Ground, (B) 'Soaring' flight, (C), 'Flapping' flight, (F) 'Feeding', (G) 'Lying' and (H) 'Standing'. The acceleration data was collected at 20Hz during 5 seconds for three orthogonal axes (D): sway – X

(red), surge – Y (green), and heave - Z (blue). (E) GPS-tracking of a griffon vulture over one day. The colors of the GPS locations match the behaviors recorded on that location: ‘Ground’ – red; ‘Soaring’ – green; ‘Flapping’ – violet; ‘Feeding’ – yellow; ‘Lying’ – light blue; ‘Standing’ – black. This illustrates the large daily movements of griffon vultures, emphasizing the logistical challenges associated with surveillance in the desert study area. Photo credit: Yacov Ben Bunan.

Pre-processing the ACC data and model training

Before deployment on the griffons, 50 transmitters were calibrated on a leveled surface, in all six possible perpendicular orientations. This calibration allowed us to obtain a transmitter-specific instrument error for translating raw acceleration data (in mV) into acceleration units (m/sec^2). For 14 transmitters (out of 32) without specific error values, we used the average error across the measured transmitters (n = 50). The calibration values used are publicly available on GitHub.

We identified the start and end of each accelerometer bout and excluded from the ACC behavioral dataset all bouts shorter than 5 sec, as well as all bouts that matched more than one behavioral class during the 5 sec period. Each acceleration bout was summarized into 47 statistical features commonly used in other studies using machine learning algorithms to perform behavioral classifications of ACC data (e.g., Nathan et al., 2012; Yu et al., 2021). For a full list of features, see Supplementary Table S1. All analyses were performed in R (R Core Team, 2023).

Using the R packages *ranger* (Wright & Ziegler, 2017) and *parsnip* (Kuhn & Vaughan, 2024), we built a random forest model to classify behaviors using the annotated acceleration bouts. We started by splitting this dataset into ‘training’ (67%) and ‘testing’ (33%) subsets, an *ad hoc* measure commonly found in other machine-learning applications (e.g., (Jeantet et al., 2020). Using the ‘training subset’, we built a random forest model and we evaluated the performance of our model using the ‘testing’ subset. We built a confusion matrix and calculated three performance metrics, for the full model and for each behavior: (*i*) accuracy; (*ii*) precision; and (*iii*) recall. The equations and descriptions for each metric can be found in Table 1. For example, a model may have 0.90 accuracy (i.e., 90% of all behaviors were predicted correctly), 0.85 precision for a specific behavior (e.g., 85% of all ‘Feeding’ predictions were indeed ‘Feeding’ observations and 15% were a different behavior and wrongly identified as ‘Feeding’), and 0.80 recall of a specific behavior (e.g., 80% of ‘Feeding’ observations were correctly predicted as ‘Feeding’ and 20% were wrongly classified as another behavior).

After training and evaluating the performance of the algorithm with the split annotated dataset, we built the final random forest model using the full dataset for training, likely improving the performance of the algorithm. This full algorithm was then used to classify unobserved accelerometer bouts to identify feeding in free roaming vultures (see below).

More details of the model building sequence can be found in Supplementary Material 2, and a full description of the model building process can be found in Figure 2. All the training data and the code necessary to train and build the algorithm are publicly available on Zenodo (Acácio et al., 2025) and GitHub. The repository includes a tutorial suitable for two types of users: those who may wish to apply our (already-trained) model to their own data (e.g., researchers and conservationists working on similar vulture/raptor species), and those wishing to use our pipeline for training and building their own model (e.g., researchers and conservationists working on other species, or with different sampling protocols). With these tutorials, our main goal is to bridge the gap between researchers and practitioners.

Table 1 – Performance metrics used to evaluate the random forest model performance, considering the true-positive (TP), true-negative (TN), false-positive (FP) and false-negative (FN) predictions.

Performance metric	**Equation**
Accuracy	$\frac{TP + TN}{TP + TN + FP + FN}$
Precision	$\frac{TP}{TP + FP}$
Recall	$\frac{TP}{TP + FN}$

<u>Calculating confidence scores to validate model predictions</u>

Using the training dataset, we calculated a confidence score for each behavioral classification (i.e., for every bout). This confidence score is the level of consensus amongst the different decision trees within the random forest (i.e., the proportion of trees that agree on the highest scoring prediction). For example, if the model classifies a given bout as 'Feeding' with a confidence score of 0.7, then 70% of the trees agreed on that classification. To determine the validity of this score as an indicator of the behavioral classification's reliability, we compared the scores of correctly identified behaviors (true-positives) and of incorrectly identified behaviors (false-positives) in the testing subset. We then used a generalized linear mixed model (GLMM) with an ordered beta distribution and a logit link to compare scores of the two groups. The confidence score (range 0 to 1) was the response variable, and the explanatory variables were the Boolean correctness of the model prediction (categorical; true-positive or false-positive), the predicted behavior (categorical), and their interaction. The model included device ID as a random intercept. The GLMM was built using *glmmTMB* R package (Brooks et al., 2017), and the fit of the model and residuals were evaluated using *DHARMa* R package (Hartig, 2022).

To understand the effect of the harness configuration (backpack or leg-loop) on the confidence scores, we compared the confidence scores of true-positive and false-positive classifications of behaviors recorded with the two different harnesses. We performed this comparison for the two behaviors with most observations: standing and feeding. We built two separate GLMMs for each behavior. Each GLMM included the confidence score as a response variable, and the Boolean correctness of the model prediction (true-positive or false-positive), the predicted behavior, and their interaction, as explanatory variables. We also included device ID as a random intercept. To further explore the influence of harness type on the behavioral classification, we trained a new random forest model using only the leg-loop data (n = 3428) and tested it on the backpack dataset (n = 714). This model was trained on a subset of bouts that included only the three behavioral categories present in both datasets (‘Standing’, ‘Ground’ and ‘Feeding’).

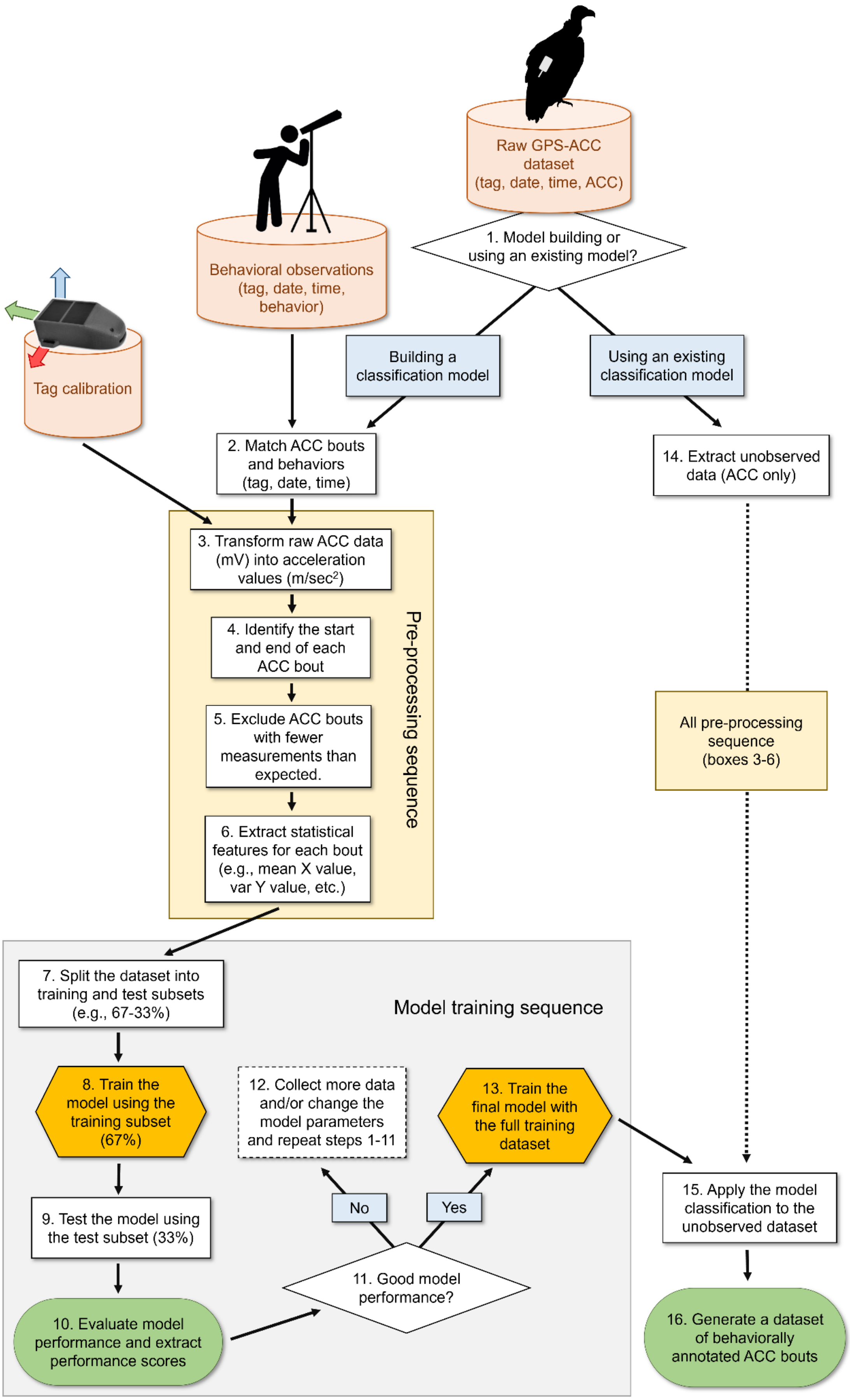

Raw GPS-ACC dataset (tag, date, time, ACC)
Behavioral observations (tag, date, time, behavior)
Tag calibration
1. Model building or using an existing model?
Building a classification model
Using an existing classification model
2. Match ACC bouts and behaviors (tag, date, time)
3. Transform raw ACC data (mV) into acceleration values (m/sec²)
4. Identify the start and end of each ACC bout
5. Exclude ACC bouts with fewer measurements than expected.
6. Extract statistical features for each bout (e.g., mean X value, var Y value, etc.)
Pre-processing sequence
14. Extract unobserved data (ACC only)
All pre-processing sequence (boxes 3-6)
7. Split the dataset into training and test subsets (e.g., 67-33%)
Model training sequence
8. Train the model using the training subset (67%)
12. Collect more data and/or change the model parameters and repeat steps 1-11
13. Train the final model with the full training dataset
9. Test the model using the test subset (33%)
No
Yes
11. Good model performance?
10. Evaluate model performance and extract performance scores
15. Apply the model classification to the unobserved dataset
16. Generate a dataset of behaviorally annotated ACC bouts

Figure 2 - Methodological workflow outlining the process of identifying behaviors of wild animals based on accelerometer data. All the training data, code and tutorials necessary to follow this workflow are available on GitHub (www.github.com/Orrslab/ACC_behavior_classification) and at Acácio et al., 2025.

Using the confidence score to validate 'Feeding' predictions of free-roaming vultures

To assess the reliability of our algorithms at classifying unobserved data, we validated predicted 'Feeding' behaviors of free-roaming vultures. We focused on this behavior due to its importance for identifying poisoning events, the main cause of vulture mortality in our study area (Anglister et al., 2023). This validation is important because even a highly accurate model introduces classification errors. For example, considering a transmitter collecting 72 bouts a day, with 10 of those classified as 'Feeding', and a model with 90% precision for 'Feeding'. Over the course of one week, the device would collect 504 bouts, 70 of which classified as 'Feeding'. Considering the model's precision, 7 of these 'Feeding' classifications would be false-positives which, extrapolating for a population of 50vultures, this would correspond to approximately 350 false-positive feeding predictions per week.

We combined information about the location of supplementary feeding stations, satellite imagery, and GPS positions from griffon-borne transmitters to assess the likelihood that the unobserved vultures' ACC-predicted 'Feeding' behavior represents a true feeding event. Between November-December 2020, we collected GPS and accelerometer data from 7 tagged free-roaming griffons in southern Israel (transmitter schedule described in the Supplementary Material). These individuals were selected because they provided consistent high-resolution data throughout this two-month period, making them suitable for the fine-scale analysis of feeding behavior. We matched a GPS position to an accelerometer bout if they were recorded within 5 min of each other. We designated four situations with decreasing probability of representing real feeding events based on the GPS location and GPS ground speed: ‘Station’ – if the ACC identified a feeding event within 250m from a supplementary feeding station, it likely represents a true feeding event (likely true-positive predictions); ‘Open area’ – if the ACC identified a feeding event at a GPS position that is at an open landscape (but not on a cliff), where naturally-occurring food is sometimes available (wildlife or livestock carcasses), it may represent a true feeding event ; ‘Cliff’ – if the ACC identified a feeding event on cliff faces, where food is largely absent but where vultures spend a large proportion of their time roosting, it is not likely to be a true feeding event; ‘Flight’ – if the GPS ground speed was >4m/sec the vulture was probably flying, and therefore it is likely a false feeding event. . The 250 m radius around the feeding station accounted for the vultures’ behavior of standing nearby and overlooking the station before feeding. Given the potential 5-minute offset between GPS fixes and accelerometer bouts, this buffer allows for the possibility that a griffon could move to the station and begin feeding within that time window. All points were mapped on satellite images of the study area and were visually examined after this classification to confirm the assignment to each situation (for example, to confirm vulture presence on a cliff, in an open area or near a feeding station).

To determine if the confidence score of the classification can be used to identify false-positives in free-roaming griffons, we compared the algorithm’s confidence scores of ‘Feeding’ predictions at ‘Stations’ (i.e., high probability of true-positives) with ‘Feeding’ predictions at ‘Cliffs’ or during ‘Flight’ (i.e., high probability of false-positives). We omitted the ‘Open area’ situation since it could represent a mixture of feeding and non-feeding behaviors and were therefore less conclusive for this comparison. We used a GLMM, with a beta distribution and a logit link, in which the response variable was the algorithm's confidence score, and the explanatory variable was the classification accuracy according to the GPS location (likely true-positive or likely false-positive). We included device ID as a random intercept.

Case study: identification of vulture feeding hotspots to prevent poisoning

To demonstrate the applicability of the ACC algorithm to a real-world conservation problem, we used it to identify griffons' feeding hotspots outside supplementary feeding stations (i.e., places where safe carcasses are provided to vultures). Considering the high risk of carcasses outside feeding stations being contaminated with toxic substances for vultures (e.g., pesticides or NSAIDs, Anglister et al., 2023), their rapid detection and removal from the field is a priority for wildlife authorities in Israel (Acácio et al., 2023). Accordingly, mapping those areas where vultures are feeding on potentially contaminated carcasses may guide management actions.

In November 2022, we collected one month of GPS and accelerometer data for 51 free-roaming griffons in Southern Israel, aiming to identify the locations of feeding events that occurred outside supplementary feeding stations (events that present higher risk of poisoning). After applying the random forest algorithm to this dataset, we matched the accelerometer 'Feeding' bouts with a GPS location using three criteria. First, if they were collected within 5 min of each other, and if the GPS ground speed was below 4m/sec (indicating the bird was not flying). Second, if no GPS position matched these criteria, we matched ACC bouts with GPS locations if they were collected within 11 min of each other (while maintaining the ground speed criteria), to account for a possible delay in the time to acquire a position by the GPS. If no GPS position matched these criteria, the 'Feeding' bout was discarded from further analysis, because we could not infer where the feeding event took place.

Using the results of the previous analyses, where we assessed if the confidence score could be used to minimize the number of false-positives, we excluded bouts with confidence scores below 0.5. This conservative threshold was chosen to avoid eliminating true-positives, as failing to detect feeding areas posed a greater risk for griffon conservation than including false-positive observations. However, this threshold is system- and data-specific and is expected to be different for other species and systems. We also excluded ‘Feeding’ bouts that occurred within supplementary feeding stations, at known roost sites (the latter likely represent false-positives), and outside the study area (southern Israel and Jordan). With the remaining locations (n = 264), we created a 2D kernel (grid size = 1000, bandwidth = bandwidth.nrd function from MASS R package, Venables & Ripley, 2002), portraying the density of locations, using *bkde2D* function of *KernSmooth* R package (Wand, 2024). On this density map, we overlayed the information of known carcasses independently identified in the field, outside feeding stations, during this same time period (n = 5). The carcasses were located by local rangers, either through reports from farmers, chance encounters during field patrols, or via an alert system that flags unusual landings of tracked raptors based on GPS data. This system uses location data from several species to identify potential poisoning events, independent of accelerometer-based behavioral classifications used in this study.

Finally, to assess if the behavioral classification impacted the designation of feeding hotspots compared to a mapping based on GPS metrics alone, we built an additional density map using all GPS locations from the same dataset, without filtering for “feeding-only” locations. We excluded locations that occurred within supplementary feeding stations, at known roost sites, or where the ground speed exceeded >4m/sec.

Results

Behavioral classification

We collected 5783 behavioral observations for 14 captive and 17 free-roaming griffon vultures (a total of 31 individuals), during 57 days (18 days for captive individuals and 39 for free-roaming individuals). The most common behavior was ‘Standing’ (3488 observations), and the least common behavior was ‘Ground’ (67 observations, Supplementary Table S2). When training the random forest model with 67% of the observed (i.e., ground-truthed) dataset, we achieve an overall accuracy of 0.96, precision of 0.89 and recall of 0.82. Specifically, the model predicted ‘Feeding’ behaviors with precision of 0.87 and recall of 0.92 (Figure 3, Supplementary Table S2). ‘Ground’, on the other hand, had the poorest predictions (precision = 0.57, recall = 0.15). The indirectly inferred behaviors ‘Soaring’ and ‘Flapping’ were well predicted by our model (‘Soaring’: precision = 0.99, recall = 0.99, ‘Flapping’: precision = 0.98, recall = 0.95).

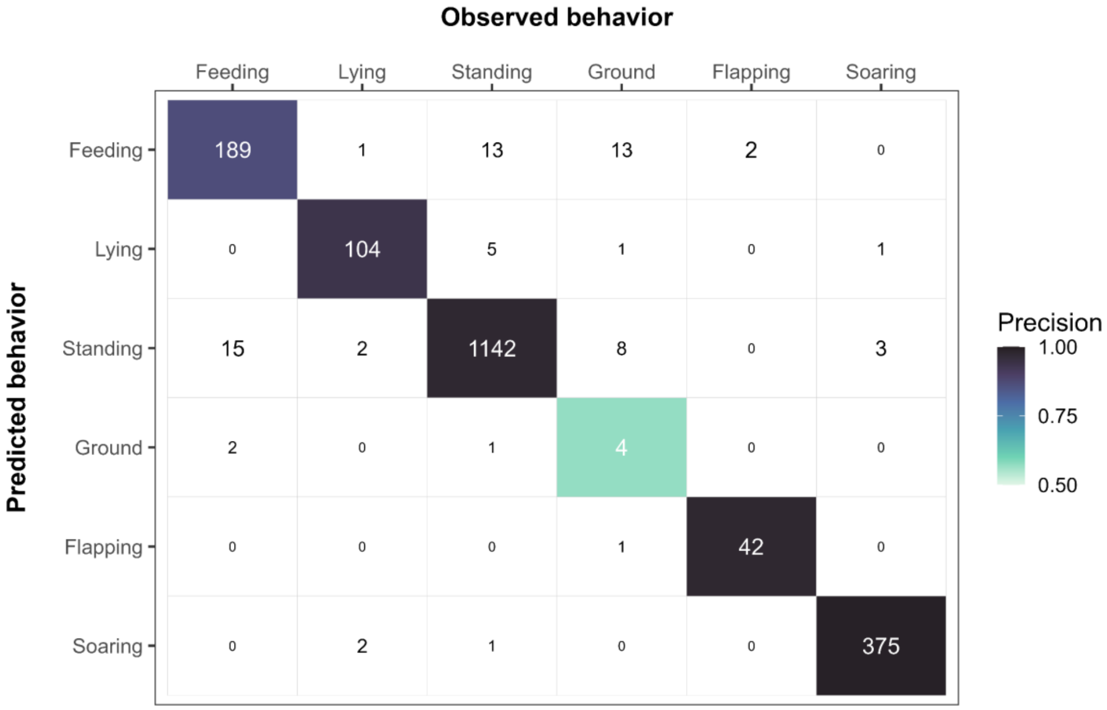


Figure 3 - Confusion matrix of the random forest model to classify vulture behavior based on accelerometer data. Rows represent the behavior predicted by the algorithm we developed, and columns represent the behaviors we observed directly. The colors in the diagonal show the precision for each behavior, with darker colors indicating higher precision. The size of the text outside the diagonal indicates the proportion of false-positives in each behavioral category, with larger numbers indicating a larger proportion

of false-positives. For example, more ‘Feeding’ bouts were wrongly classified as ‘Standing’ than as ‘Ground’, and none of the ‘Feeding’ bouts were wrongly classified as ‘Soaring’, ‘Flapping’, or ‘Lying’.

Performance of the confidence score in validating model predictions

Overall, the confidence scores of correctly-identified behaviors (true-positives) were significantly higher than the scores of inaccurately-identified behaviors (false-positives; GLMM: model estimate ± se = 0.876 ± 0.195, p-value < 0.001, Figure 4, Supplementary Table S3). ‘Ground’ behaviors, which had the lowest number of observations (n=67), were the exception, with significantly higher confidence scores of false-positives compared to true-positives (Figure 4). ‘Ground’ true-positive behaviors also had the lowest confidence scores (mean confidence score ± sd = 0.41 ± 0.03). ‘Soaring’, ‘Lying’, ‘Standing’ and ‘Flapping’ were the behaviors with highest true-positive confidence scores (mean confidence score ± sd; ‘Soaring’ = 0.99 ± 0.08, ‘Lying’ = 0.98 ± 0.06, ‘Standing’ = 0.98 ± 0.08, ‘Flapping’ = 0.93 ± 0.12). ‘Feeding’ had on average a high confidence score but also a large variation (mean confidence score ± sd = 0.82 ± 0.15, Figure 4, Supplementary Table S2, Supplementary Table S3).

There were no significant differences between the confidence scores of behaviors recorded with backpack or leg-loop harnesses for the two tested behaviors: Standing’ (GLMM: estimate ± se = 0.374 ± 0.283, p-value = 0.283) and ‘Feeding’ (GLMM: estimate ± se = -0.725 ± 0.798, p-value = 0.364, Supplementary Figure S1 and Supplementary Table S4). The model trained on leg-loop data and tested on backpack data had high overall accuracy (accuracy = 0.86) and performed well at classifying ‘Standing’ behaviors (‘Standing’ precision = 0.99; recall = 0.91). All true ‘Feeding’ behaviors were correctly identified as such (‘Feeding’ recall = 1). However, most ‘Ground’ behaviors were misclassified as ‘Feeding’, which reduced the precision of the

'Feeding' category ('Feeding' precision = 0.52). Similarly to the full model, 'Ground' behaviors had the poorest performance (Supplementary Table S5).

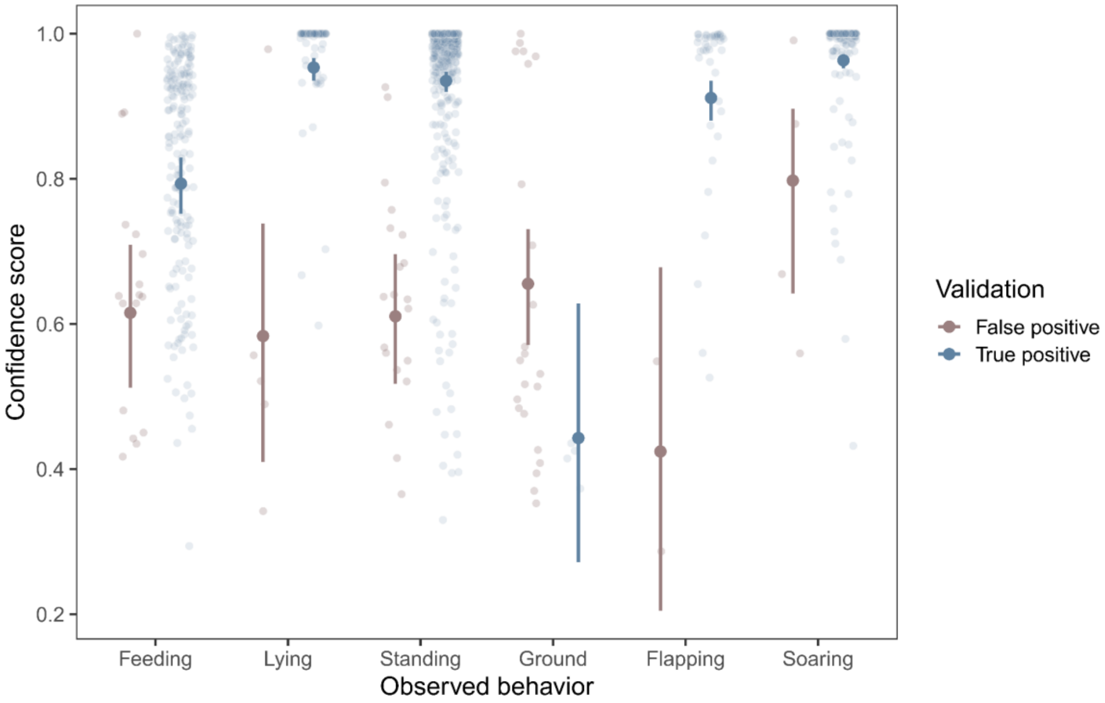


Figure 4 – Model predicted confidence scores of true-positive (blue) and false-positive (dark red) behavioral classifications. The darker points and error bars indicate the model predicted confidence scores and 95% confidence intervals for true-positive and false-positive behavioral classifications. The lighter points show the raw data.

Confidence score to validate 'Feeding' predictions of free-roaming vultures

We used the GPS locations to validate 175 'Feeding' bouts from 7 free-roaming vultures in Southern Israel. Overall, 126 'Feeding' bouts (72%) occurred within a supplementary feeding station ('Station' – likely true-positives), 22 bouts (13%) were located on 'Open areas' (likely a mix of true- and false-positives), 20 bouts (11%) were on 'Cliffs' (likely false-positives), and 7 bouts (4%) were in 'Flight' (likely false-positives, Figure 5). The relatively high proportion of feeding bouts identified on cliffs likely reflects the considerable amount of time griffons spend in these areas. Overall, of all ‘Feeding’ bouts identified by the algorithm, 72-85% (all ‘Station’ bouts + at least part of the ‘Open area’ bouts) were likely real feeding events. Furthermore, after removing ‘Cliff’ and ‘Flight’ bouts (easily identifiable using only the GPS location, satellite imagery, and ground speed), 85-100% of the ‘Feeding’ predictions (all ‘Station’ bouts + at least part of the ‘Open area’ bouts) were indeed likely feeding events.

Importantly, the confidence scores of ‘Feeding’ bouts likely to be true-positives were higher (mean ± sd: ‘Station’ = 0.75 ± 0.16) than the scores of bouts likely to be false-positives (‘Cliff’ and ‘Flight’ = 0.56 ± 0.19). This comparison was statistically significant (GLMM: estimate ± se = -0.805 ± 0.167, p-value <0.001, Supplementary Table S5). When considering solely the ‘Feeding’ bouts with a confidence score over 0.5, 114 bouts (88.4%) occurred within a supplementary feeding station and were likely true-positives. This threshold maximizes the number of true-positive predictions, while minimizing the number of false-positives (Figure 5).

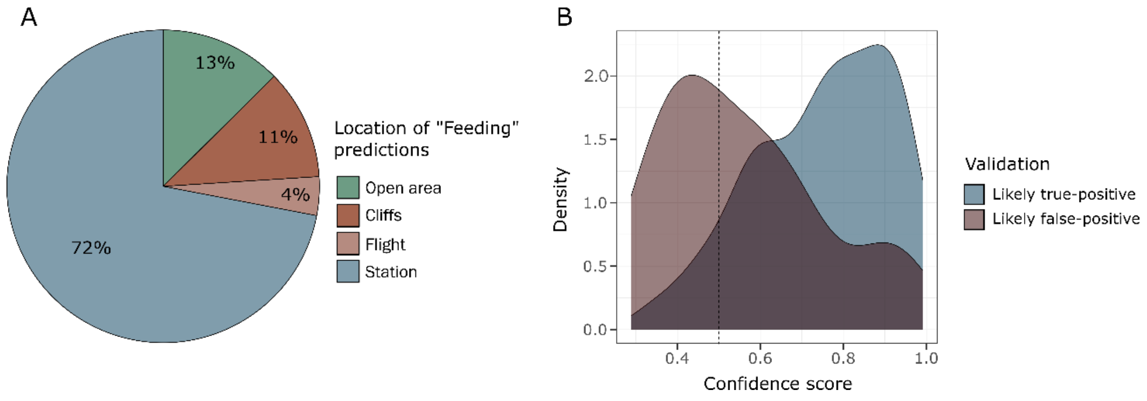


Figure 5 - Validation of ‘Feeding’ behaviors using data from free-roaming griffons. A - Percentage of ‘Feeding’ predictions (n = 175) located within a supplementary feeding station (‘Station’ - blue), on open landscape (‘Open area’ - green), on cliffs (‘Cliffs’ - dark brown) or in flight (‘Flight’ - light brown). B - Distribution of the confidence scores of

'Feeding' bouts likely to be true-positives (located within a feeding station, in blue) and likely to be false-positives (located on cliffs or in flight, in dark red). The dashed line indicates the confidence threshold of 0.5, a conservative threshold that reduces the number of false-positives, while including nearly all true-positive predictions.

Mapping vulture's feeding hotspots to facilitate poisoning identification

In November 2022, we collected 4595 'Feeding' bouts of 51 griffon vultures in our study area. After sequentially removing the bouts without a GPS location (n = 586), bouts inside feeding stations (n = 2534), outside Southern Israel and Jordan (n = 157), bouts located in known roosts (n = 979), and bouts with a confidence score below 0.5 (n = 60, Figure 5), we retained 264 bouts of 31 vultures, that allowed us to map their feeding hotspots.

We built a KDE with the remaining 264 'Feeding' bouts and detected a hotspot of feeding events in the Judean desert. This hotspot is consistent with the INPA ranger's reports for this same period, where 4 out of 5 reported carcasses outside feeding stations were within the KDE (Figure 6). This density map differed substantially from the one based solely on GPS-derived metrics (Supplementary Figure S2), containing 1,938 potential feeding locations. Notably, the GPS-only KDE failed to identify a key feeding hotspot in Jordan. This confirms that the algorithm can be used to identify areas with high probability of vultures' feeding on potentially contaminated carcasses and highlights the added value of the behavioral classification.

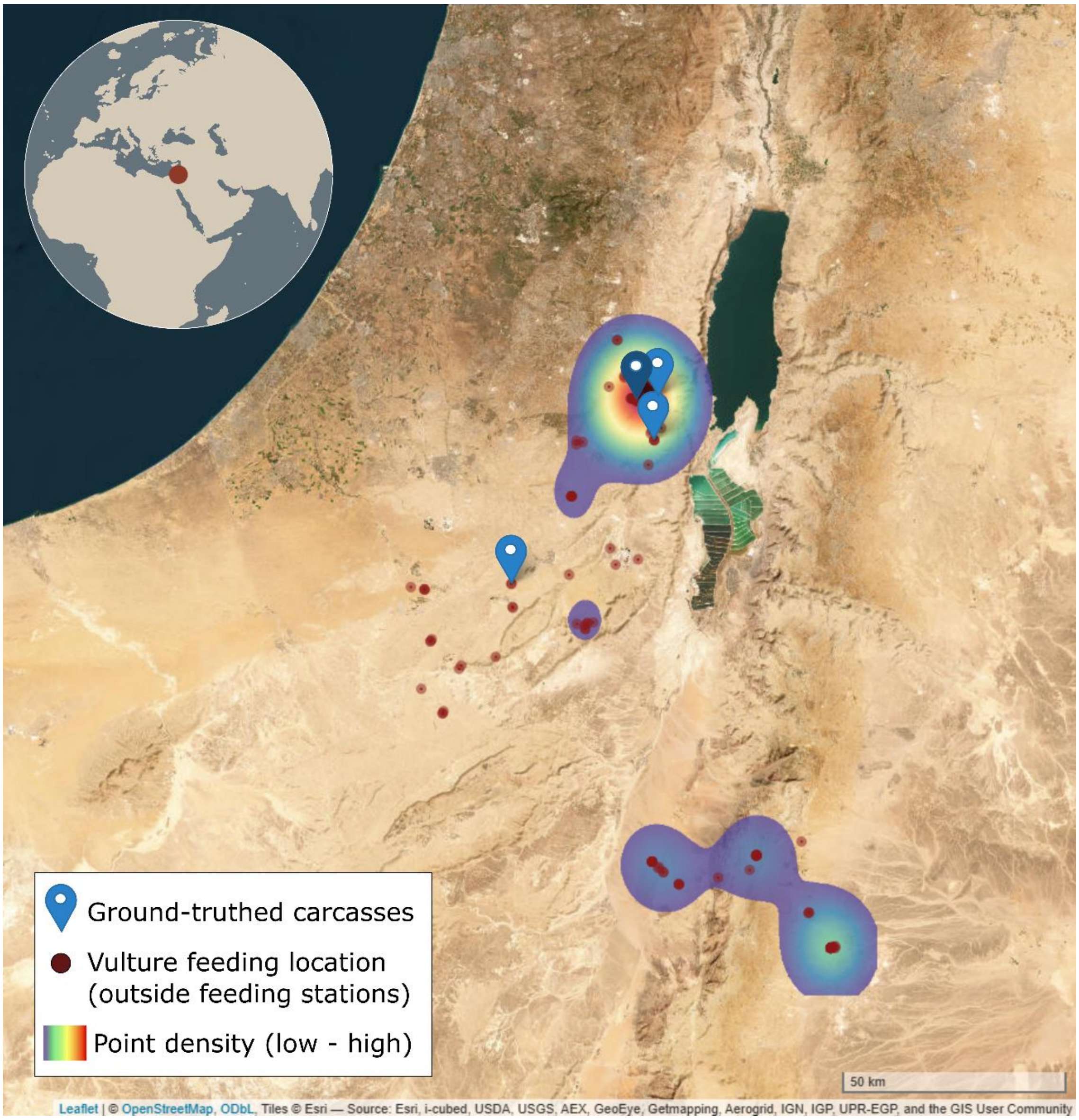


Figure 6 - Acceleration-based behavioral classification as a tool to identify griffon vultures' feeding hotspots outside supplementary feeding stations. The red points show the vulture feeding locations in Southern Israel and Jordan over November 2022, identified using the random forest algorithm. The blue markers indicate the location of ground-truthed carcasses outside feeding stations; the darker blue indicates two carcasses in approximately the same location. The polygons indicate the density of vulture feeding locations, showing the areas where vultures are at greater risk of poisoning. The polygon colors indicate the density of vulture feeding locations, with blue

areas having lower density and red areas higher density. The inset shows the location of the study area in the world.

Discussion

Recent advancements in tracking technology and analytical tools are enhancing our understanding of animal ecology and behavior and improving its applications for biodiversity conservation (Tuia et al., 2022; Williams et al., 2020). In this study, we add to this body of literature by developing a machine learning algorithm to classify griffon vultures' behaviors, thoroughly validating the behavioral classifications, and using them to inform conservation efforts, namely carcass detection to prevent poisoning. Our model accurately predicted griffons' behaviors, allowing for the identification of potential feeding events outside feeding stations and the mapping of feeding hotspots where vultures and other scavengers can engage in risky behaviors such as the consumption of poisoned carcasses (Peters et al., 2023). These maps may become fundamental tools for monitoring-effort prioritization and for optimizing on-the-ground actions for the conservation of vultures and other scavengers (e.g., the detection of poisoning events, Rast et al., 2024). Another major contribution of this study is the use and validation of the algorithm's confidence in each behavioral classification, showing the utility of this approach for other behaviors and contexts. Assessing the degree of confidence in this manner is rarely done in ecological studies (Bidder et al., 2014), but we highlight how this approach may minimize misclassifications (e.g., false-positives) especially when resources for ground-truthing are limited. Our algorithm and training dataset are made accessible to other researchers and conservationists studying vultures and similar species. Moreover, they can be easily adapted to classify the behaviors of other species in diverse study systems. Importantly, to further promote this usability, we provide a methodological workflow to guide potential users in the process of identifying behaviors of wild animals based on accelerometer data.

Accelerometer-based behavioral classification as a tool for vulture conservation

With our behavioral classification model, we were able to identify vulture feeding hotspots in Southern Israel. Indeed, the areas where our tracked vultures displayed 'Feeding' behaviors matched the locations of known 'wild' carcasses (i.e., outside feeding stations) during this same period. This case study used only a single month worth of high-resolution data embedded within a long-term lower resolution tracking effort (Acácio et al., 2024; Spiegel et al., 2013), but it exemplifies how GPS and accelerometer data can be used to direct conservation efforts. The use of GPS tracking has been instrumental for vulture conservation in Israel (Spiegel et al., 2013), particularly for the detection of poisoning events (Acácio et al., 2023; Anglister et al., 2023). The local government environmental agency, INPA, developed a near-real time alert system that warns rangers whenever a vulture lands at a suspicious area and when vultures are either moving very little or are suspected to be dead (Nemtzov et al., 2021). Rangers then actively respond to these alerts by inspecting the area and removing the carcasses; therefore, reducing the number of false alarms is important – both to reduce costs and workload, as well avoiding erosion of rangers' responsiveness.

A similar near real-time alert system, using GPS data, has also been used for the monitoring of African elephants (*Loxodonta Africana,* Wall et al., 2014) and to track California condors (*Gymnogyps californianus*) in the vicinity of wind farms (Sheppard et al., 2015). It has also been suggested as an anti-poaching tool to prevent the extinction of large mammals (O'Donoghue & Rutz, 2016). We propose that all these systems could be improved by using accelerometer data to remotely identify animal behavior and risky events sooner and more reliably. Indeed, our results show that relying solely on GPS-derived filters to identify potential feeding hotspots failed to identify a key feeding hotspot in Jordan and produced over seven times more data points, many of which were likely false positives. Such an overload of low-quality alerts could lead to reduced responsiveness by the rangers, ultimately undermining conservation efforts on the ground.

Combining maps of feeding hotspots (either fixed or season-specific ones) with similar near real-time alert systems may be crucial for vulture management and conservation. For example, the feeding areas that griffons use systematically throughout the year should be prioritized in terms of surveillance and sanitation efforts, to prevent vultures (and other scavengers) from accessing carcasses contaminated with toxic substances. Additional management actions could be implemented, such as establishing new supplementary feeding stations in these areas, or increasing carcass supply at existing stations, either all year-round or during particular seasons, to match potential seasonal changes in vulture's activity areas. Additionally, the hotspots could be used to implement geofences where data collection and transmissions would be at higher frequency. This increased resolution may be critical in poisoning events, where the actual feeding may be quite fast (sometimes consuming a carcass within minutes) and vultures may perish quickly, depending on the type and amount of toxic substance ingested. Then, information regarding the griffon's location and behavior is obtained and communicated faster: when a griffon lands in these areas and only if it feeds there (as indicated by the ACC classification), an alert should be sent to the rangers for immediate carcass inspection. While the system should also trigger alerts for any feeding events detected outside feeding stations (to allow for carcass inspection and potential removal to reduce the risk of poisoning), identifying risky hotspots can help optimize resource allocation and prioritize conservation actions in high-risk areas.

Around the world, an increasing number of individuals of multiple vulture species are being tracked with GPS-Accelerometer devices, showing that they roam exceptionally large areas on their daily movements (Kane et al., 2022; Spiegel et al., 2015). Considering that about 70% of vulture species are endangered (Ives et al., 2022; Ogada et al., 2012; Plaza et al., 2019), surveillance systems that combine GPS tracking with accelerometry may be a useful tool to improve management actions in their large roaming areas to combat major threats such poisoning. For instance, such ACC-based systems will enhance existing applications of GPS-tagged vultures to inform on-ground actions against illegal wildlife persecution (Rast et al., 2024; Rodríguez-Pérez et al., in press) or to improve regulations for carrion disposal to feed vultures and other wild species (Mateo-Tomás et al., 2023). Future studies could apply our methodology to publicly available GPS-ACC datasets to identify high-risk areas for vultures and guide targeted conservation interventions at a larger geographic scale.

Our thoroughly validated training dataset can also be used to classify the behavior of griffons in other populations, as well as other old and new-world vulture species, particularly in Africa and Asia, where vulture populations continue to decrease. For example, our algorithm and training dataset could be used to predict the behaviors of endangered *Gyps* species in Africa and Asia (e.g., *Gyps africanus*, *Gyps coprotheres*, *Gyps bengalensis*, amongst others), or even other vulture species (e.g., *Torgos* and *Trigonoceps* species), considering their morphological and behavioral similarities with the griffon vulture. The use of surrogate species to identify accelerometer-based behaviors has been examined in other systems, with a variety of results. For example, the behaviors of domestic dogs were good predictors of the behaviors of dingoes and cheetahs (Campbell et al., 2013), but the behavior of domestic caprids did not predict well the behavior of their wild counterparts (Dickinson et al., 2021). Therefore, we recommend caution when using our trained model to classify the behavior of other vulture species. In addition, our algorithm and modelling pipeline can be easily adapted for other, not related, animal species, as long as researchers provide their own training dataset for their study species.

Validating the accuracy of predictions of unobserved behaviors

Tri-axial accelerometers and classification algorithms have increasingly been used to obtain fine-scale behavior of wild animals (Nathan et al., 2012; Resheff et al., 2014; Wang, 2019; Yu et al., 2021). However, after training and testing the model on a validated dataset, the model must classify unobserved and, sometimes, unknown behaviors. In this case, the model then matches the unknown behavior with the best fitting known acceleration signature, resulting in misclassifications (Glass et al., 2020). Most ecological studies fail to acknowledge this limitation and do not provide a metric of how likely a particular classification is to be true (Glass et al., 2020). Here we tackle this methodological gap and calculate a confidence score, which allows us to distinguish between true-positive and false-positive classifications. Our approach is computationally simple to implement and does not require running more complex classification models. In addition to the confidence scores, we used biologically relevant information to validate observations classified as 'Feeding'. For this subset of data, 15% of the observations were likely misclassifications because they occurred on cliffs (where there is no food in our case) or in flight. Filtering out observations based on easily accessible metrics (here, the topography, knowledge of the behavior of the species, spatial position, and GPS ground speed) as well as any observations with a confidence score below a relevant threshold, increases the accuracy of behavioral classifications.

Selecting filtering thresholds is always a balance between two types of errors. Here, we considered a conservative threshold of 0.5 to distinguish true-positive and false-positive 'Feeding' predictions of free-roaming griffons, at the risk of including some false-positive predictions in our dataset (Type I error). However, in this case, the risk of not including part of the true-positive predictions (Type II error) is higher than including some false-positives; not including all true-positives could mean that some feeding hotspots would not be identified, potentially compromising sanitary management and overlooking potential feeding and poisoning events. We encourage other researchers to use a similar approach whenever possible, combining confidence scores with ground-truthing information, to improve the accuracy of their conclusions.

In addition, we note that different behaviors, or even the same behavior in different species or with different accelerometer devices, may have different confidence scores distributions. For example, with our dataset a threshold of 0.9 could have been suitable for distinguishing flapping and soaring flight behaviors, to study, for instance, flight biomechanics. Therefore, the threshold of confidence should consider the underlying distribution of confidence scores for the behaviors in mind and should be defined according to this and the study objectives, balancing the risks of data loss with the costs of including false-positives in the dataset.

Finally, quantifying temporal correlations between behaviors could also help improve model performance or assist in post-processing filtering of the classifications (Supplementary Material 3, Supplementary Figure S3). For instance, it is possible to combine a correlation matrix of the behaviors with the confidence scores. In our dataset, ‘Feeding’ is often followed by other ‘Feeding’ behaviors (Supplementary Figure S3C). Thus, if a high-confidence 'Feeding' behavior is followed by a low-confidence 'Feeding' prediction, the strong positive dependency between these two behaviors could support treating the second ‘Feeding’ as likely true-positive. Future studies could also implement more complex models that allow for the incorporation of the correlation matrix within the model.

Challenges and considerations of accelerometer-based behavioral classification

Different tag placement and different attachment methods can greatly influence accelerometer signatures, and consequently the behavioral classifications (Garde et al., 2022). Nevertheless, our results show that our algorithm is reliable for more than one attachment method, further increasing its usefulness. These non-significant differences may result from the similarity in logger placement between the two attachment types (about 3cm difference), as well as from the limited spinal flexibility of griffon vultures. Still, the large confidence intervals in this comparison (due to the small sample size for backpack harnesses), as well as difficulty of our algorithm in classifying ‘Ground’ behaviors recorded with a backpack harness shows that this topic deserves further investigation with larger sample sizes and with other species.

After building the random forest, it is crucial to validate its accuracy, precision, and recall. In our study, the overall model was highly accurate, yet behaviors differed in their precisions. As expected, ‘Ground’ was consistently the behavior with the poorest predictions across all our validations. This behavioral category included several, quite distinct, ground behaviors (like walking, running, hopping, etc.), in an attempt to account for all possible behaviors a griffon may display and to minimize misclassifications at the cost of losing accuracy (Glass et al., 2020). ‘Feeding’, on the other hand, was accurately classified by our model, however despite a relatively large number of bouts in the training dataset (n = 587), the confidence scores of this behavior had large variance. A possible reasoning is that griffon’s feeding behavior is highly complex and may include rapid shifts between fighting, posturing (spreading the wings), as well as eating per se (Bosè & Sarrazin, 2007) – all inseparable within a 5s timeframe. Including so many different postures in a single behavioral category results in high variation of confidence scores.

In addition, the number of conspecifics within a feeding event may further influence the behaviors that individuals display while foraging (Bosè et al., 2012), increasing within-individual variability for both wild and captive vultures. To mitigate the effect of within-individual variability in our training dataset, we ensured that multiple captive individuals were feeding at the same carcass, to replicate the wild feeding conditions. Finally, it is likely that individuals differ in their behavior while foraging (e.g., dominant vs subordinates, Bosè et al., 2012; Bosè & Sarrazin, 2007), emphasizing the need to improve behavioral classification models and account for individual differences in behavior (Kirchner et al., 2023). In general, we suggest that future models can improve accuracy and precision by further splitting our six classes into subclasses that reflect more homogenous elementary behaviors (e.g., pecking, tearing meat apart, fighting). Merging ‘Ground’ and ‘Feeding’ categories could also potentially improve the model’s accuracy; however, since these two behaviors are not necessarily linked, this would come at the cost of decreased resolution in detecting true feeding events, compromising management and conservation applications. For specific applications focused solely on identifying feeding activity, an alternative approach could involve merging all feeding and all non-feeding behaviors in a binary classification. This could simplify the interpretation, and we suggest that this option is worth exploring in a future study

Despite the potential of accelerometer-based behavioral classification, collecting such large volumes of data can be costly, both in terms of data transmission and storage, as well as in terms of device memory and battery (Hounslow et al., 2019). Short sampling intervals (2-3 secs) at high resolution may reduce the probability of having multiple behaviors within a single bout, but may drain batteries faster, which can result in incomplete sampling designs and lower the device’s lifespan. Integrating low-frequency accelerometry with additional sensors (e.g., time-depth recorders for marine species, Jeantet et al., 2020), may still effectively allow the study of animal behavior without significantly increasing costs or reducing device longevity (Hounslow et al., 2019). In addition, analyzing such large volumes of data can also be challenging, so we emphasize the need for collaboration in between fields of knowledge, with ecologists and data scientists working together for the conservation of biodiversity (Tuia et al., 2022). Lastly, as human activities are increasingly impacting the planet and driving species towards extinction, it is critical to harness technological advances for effective conservation and to safeguard the future of our planet's species and ecosystems.

Conclusions

In this study, we showed the potential of accelerometer-based behavioral classification to improve the management and conservation of endangered scavengers. By reliably identifying feeding behaviors and mapping feeding hotspots, our approach can help the detection of poisoning events earlier and optimize management resources to high-risk areas. We further show that combining the algorithm’s confidence score with simple GPS-derived filters, can greatly improve the reliability of the identification of feeding hotspots. Finally, our workflow, training dataset, and model are provided in an open-access platform, to facilitate the adoption of this framework in the global management and conservation of endangered scavengers.

Acknowledgments

We thank the workers and rangers of the INPA for their hard work and dedication, especially Yaniv Levy-Paz, Sappir Simchi, Avishai Bar-On, Elad Zisso, Asher Perez, Elya Maatuf (Elika), Zehava Sigal, Amit Cohen, and Amir Arnon. We also thank all members of the Movement Ecology and Individual Behavior Laboratory at Tel Aviv University, in particular Assaf Uzan for all support, and thank Kaija Gahm for helpful discussions on earlier versions of this work. We thank S. Borragán and A. Ruiz at Cabárceno Wildlife Park of the Cantabria Government for allowing us to tag and record vultures in this facility.

Conflict of interest statement

Patricia Mateo-Tomás is an Associate Editor of Journal of Applied Ecology, but took no part in the peer review and decision-making processes for this paper.

Funding

Funding for this work was provided by the NSF-BSF grant: NSF IOS division 2015662/BSF 2019822 to NPW and OS. GV was supported by the Ecology MSc

Fellowship (School of zoology, Tel Aviv University). MA was supported by the George S. Wise Postdoctoral Fellowship (Tel Aviv University). PMT, JPR and MFG were supported by the Regional Government of Asturias (GRUPIN grant: IDI/2021/000075), and the Spanish Ministry of Science and Innovation (TED2021-132519B-I00 and PID2023-149634OB-I00), the Complementary Plans (PCB1005_AST), and “Fundación Biodiversidad” of the Spanish Ministry of Ecological Transition and Demographic Challenge (WINDWILDNET project) through PRTR and NextGenerationEU Funds. PMT, MFG and PPO were supported by BBVA Foundation Grants to Scientific Research Teams, Call 2018.

Data accessibility

All the training data, the code necessary to train and build the algorithm and a tutorial are publicly on Zenodo Repository https://doi.org/10.5281/zenodo.16152589 (Acácio et al., 2025) and on GitHub (www.github.com/Orrslab/ACC_behavior_classification).

Author contributions

Gideon Vaadia, Orr Spiegel, Moni Shahar and Marta Acácio conceived the ideas and designed the methodology. Gideon Vaadia, Orr Spiegel, Nili Anglister, Yigal Miller, Ohad Hatzofe collected the data. Patricia Mateo-Tomás and María Fernández-García tagged the Spanish vultures, whose behaviors were recorded by Patricia Mateo-Tomás, Jorge Rodríguez-Pérez, María Fernandéz-García, Ignacio Otero and Pedro P. Olea, and processed by Patricia Mateo-Tomás and Jorge Rodríguez-Pérez. Gideon Vaadia, Tal Agassi and Marta Acácio analyzed the data. Marta Acácio and Gideon Vaadia wrote the first version of the manuscript, with critical input from Noa Pinter-Wollman and Orr Spiegel. All authors contributed critically to drafts of the manuscript and gave approval for publication.

References

Acácio, M., Anglister, N., Vaadia, G., Harel, R., Nathan, R., Hatzofe, O., & Spiegel, O. (2023). A lifetime track of a griffon vulture: The moving story of Rehovot (Y64). *Ecology*, *104*(4), e3985. https://doi.org/10.1002/ecy.3985

Acácio, M., Gahm, K., Anglister, N., Vaadia, G., Hatzofe, O., Harel, R., Efrat, R., Nathan, R., Pinter-Wollman, N., & Spiegel, O. (2024). Behavioral plasticity shapes population aging patterns in a long-lived avian scavenger. *Proceedings of the National Academy of Sciences*, *121*(35), e2407298121. https://doi.org/10.1073/pnas.2407298121

Acácio, M., Vaadia, G., Agassi, T., Anglister, N., Miller, Y., Hatzofe, O., Mateo-Tomás, P., Rodríguez Pérez, J., Fernández-García, M., Olea, P. P., Otero, I., Pinter-Wollman, N., Shahar, M., & Spiegel, O. (2025). *Data for: "Using accelerometer-based behavioral classification to enhance scavenger conservation"* [Dataset]. Zenodo. https://doi.org/10.5281/zenodo.16152589

Anglister, N., Gonen-Shalom, S., Shlanger, P., Blotnick-Rubin, E., Rosenzweig, A., Horowitz, I., Hatzofe, O., King, R., Anglister, L., & Spiegel, O. (2023). Plasma cholinesterase activity: A benchmark for rapid detection of pesticide poisoning in an avian scavenger. *Science of The Total Environment*, *877*, 162903. https://doi.org/10.1016/j.scitotenv.2023.162903

Bäckman, J., Andersson, A., Pedersen, L., Sjöberg, S., Tøttrup, A. P., & Alerstam, T. (2017). Actogram analysis of free-flying migratory birds: New perspectives based on acceleration logging. *Journal of Comparative Physiology A*, *203*(6), 543–564. https://doi.org/10.1007/s00359-017-1165-9

Bidder, O. R., Campbell, H. A., Gómez-Laich, A., Urgé, P., Walker, J., Cai, Y., Gao, L., Quintana, F., & Wilson, R. P. (2014). Love Thy Neighbour: Automatic Animal Behavioural Classification of Acceleration Data Using the K-Nearest Neighbour Algorithm. *PLOS ONE*, *9*(2), e88609. https://doi.org/10.1371/journal.pone.0088609

BirdLife Israel. (2022). *BirdLife Israel live nest cameras* [Computer software]. https://www.birds.org.il/en

Bosè, M., Duriez, O., & Sarrazin, F. (2012). Intra-specific competition in foraging Griffon Vultures Gyps fulvus: 1. Dynamics of group feeding. *Bird Study*, *59*(2), 182–192. https://doi.org/10.1080/00063657.2012.658639

Bosè, M., & Sarrazin, F. (2007). Competitive behaviour and feeding rate in a reintroduced population of Griffon Vultures Gyps fulvus. *Ibis*, *149*(3), 490–501. https://doi.org/10.1111/j.1474-919X.2007.00674.x

Brooks, M., Kristensen, K., van Benthem, K., Magnusson, A., Berg, C., Nielsen, A., Skaug, H., Maechler, M., & Bolker, B. (2017). glmmTMB Balances Speed and Flexibility Among Packages for Zero-inflated Generalized Linear Mixed Modeling. *The R Journal*, *9*(2), 378–400. https://doi.org/doi:10.32614/RJ-2017-066

Brown, D. D., Kays, R., Wikelski, M., Wilson, R., & Klimley, A. P. (2013). Observing the unwatchable through acceleration logging of animal behavior. *Animal Biotelemetry*, *1*(1), 20. https://doi.org/10.1186/2050-3385-1-20

Buechley, E. R., & Şekercioğlu, Ç. H. (2016). The avian scavenger crisis: Looming extinctions, trophic cascades, and loss of critical ecosystem functions. *Biological Conservation*, *198*, 220–228. https://doi.org/10.1016/j.biocon.2016.04.001

Campbell, H. A., Gao, L., Bidder, O. R., Hunter, J., & Franklin, C. E. (2013). Creating a behavioural classification module for acceleration data: Using a captive surrogate for difficult to observe species. *Journal of Experimental Biology*, *216*(24), 4501–4506. https://doi.org/10.1242/jeb.089805

Chimienti, M., Cornulier, T., Owen, E., Bolton, M., Davies, I. M., Travis, J. M. J., & Scott, B. E. (2016). The use of an unsupervised learning approach for characterizing latent behaviors in accelerometer data. *Ecology and Evolution*, *6*(3), 727–741. https://doi.org/10.1002/ece3.1914

Clarke, T. M., Whitmarsh, S. K., Hounslow, J. L., Gleiss, A. C., Payne, N. L., & Huveneers, C. (2021). Using tri-axial accelerometer loggers to identify spawning behaviours of large pelagic fish. *Movement Ecology*, *9*(1), 26. https://doi.org/10.1186/s40462-021-00248-8

Cutler, D. R., Edwards Jr., T. C., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J., & Lawler, J. J. (2007). Random Forests for Classification in Ecology. *Ecology*, *88*(11), 2783–2792. https://doi.org/10.1890/07-0539.1

Dickinson, E. R., Twining, J. P., Wilson, R., Stephens, P. A., Westander, J., Marks, N., & Scantlebury, D. M. (2021). Limitations of using surrogates for behaviour classification of accelerometer data: Refining methods using random forest models in Caprids. *Movement Ecology*, *9*(1), 28. https://doi.org/10.1186/s40462-021-00265-7

Efrat, R., Hatzofe, O., Miller, Y., & Berger-Tal, O. (2020). Determinants of survival in captive-bred Griffon Vultures Gyps fulvus after their release to the wild. *Conservation Science and Practice*, *2*(12), e308. https://doi.org/10.1111/csp2.308

Fehlmann, G., O'Riain, M. J., Kerr-Smith, C., Hailes, S., Holton, M., Hopkins, P., & King, A. J. (2023). Using behavioral studies to adapt management decisions and reduce negative interactions between humans and baboons in Cape Town, South Africa. *Conservation Science and Practice*, *5*(7), e12948. https://doi.org/10.1111/csp2.12948

Garde, B., Wilson, R. P., Fell, A., Cole, N., Tatayah, V., Holton, M. D., Rose, K. A. R., Metcalfe, R. S., Robotka, H., Wikelski, M., Tremblay, F., Whelan, S., Elliott, K. H., & Shepard, E. L. C. (2022). Ecological inference using data from accelerometers needs careful protocols. *Methods in Ecology and Evolution*, *13*(4), 813–825. https://doi.org/10.1111/2041-210X.13804

Glass, T. W., Breed, G. A., Robards, M. D., Williams, C. T., & Kielland, K. (2020). Accounting for unknown behaviors of free-living animals in accelerometer-based classification models: Demonstration on a wide-ranging mesopredator. *Ecological Informatics*, *60*, 101152. https://doi.org/10.1016/j.ecoinf.2020.101152

Gleiss, A. C., Wilson, R. P., & Shepard, E. L. C. (2011). Making overall dynamic body acceleration work: On the theory of acceleration as a proxy for energy expenditure. *Methods in Ecology and Evolution*, *2*(1), 23–33. https://doi.org/10.1111/j.2041-210X.2010.00057.x

Halsey, L. G., Green, J. A., Wilson, R. P., & Frappell, P. B. (2009). Accelerometry to Estimate Energy Expenditure during Activity: Best Practice with Data Loggers. *Physiological and Biochemical Zoology*, *82*(4), 396–404. https://doi.org/10.1086/589815

Hartig, F. (2022). *_DHARMa: Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression Models_* (Version 0.4.6) [R Package]. https://CRAN.R-project.org/package=DHARMa

Hatzofe, O. (2020). *Summary of the griffon vulture counts in Israel* (pp. 1–5) [Internal report]. Israel Nature and Parks Authority.

Hays, G. C., Bailey, H., Bograd, S. J., Bowen, W. D., Campagna, C., Carmichael, R. H., Casale, P., Chiaradia, A., Costa, D. P., Cuevas, E., Bruyn, P. J. N. de, Dias, M. P., Duarte, C. M., Dunn, D. C., Dutton, P. H., Esteban, N., Friedlaender, A., Goetz, K. T., Godley, B. J., … Sequeira, A. M. M. (2019). Translating Marine Animal Tracking Data into Conservation Policy and Management. *Trends in Ecology & Evolution*, *34*(5), 459–473. https://doi.org/10.1016/j.tree.2019.01.009

Hounslow, J. L., Brewster, L. R., Lear, K. O., Guttridge, T. L., Daly, R., Whitney, N. M., & Gleiss, A. C. (2019). Assessing the effects of sampling frequency on behavioural classification of accelerometer data. *Journal of Experimental Marine Biology and Ecology*, *512*, 22–30. https://doi.org/10.1016/j.jembe.2018.12.003

Ives, A. M., Brenn-White, M., Buckley, J. Y., Kendall, C. J., Wilton, S., & Deem, S. L. (2022). A Global Review of Causes of Morbidity and Mortality in Free-Living Vultures. *EcoHealth*, *19*(1), 40–54. https://doi.org/10.1007/s10393-021-01573-5

Jeantet, L., Planas-Bielsa, V., Benhamou, S., Geiger, S., Martin, J., Siegwalt, F., Lelong, P., Gresser, J., Etienne, D., Hiélard, G., Arque, A., Regis, S., Lecerf, N., Frouin, C., Benhalilou, A., Murgale, C., Maillet, T., Andreani, L., Campistron, G., … Chevallier, D. (2020). Behavioural inference from signal processing using animal-borne multi-sensor loggers: A novel solution to extend the knowledge of sea turtle ecology. *Royal Society Open Science*, *7*(5), 200139. https://doi.org/10.1098/rsos.200139

Kane, A., Monadjem, A., Aschenborn, H. K. O., Bildstein, K., Botha, A., Bracebridge, C., Buechley, E. R., Buij, R., Davies, J. P., Diekmann, M., Downs, C. T., Farwig, N., Galligan, T., Kaltenecker, G., Kelly, C., Kemp, R., Kolberg, H., MacKenzie, M. L., Mendelsohn, J., … Kendall, C. J. (2022). Understanding continent-wide variation in vulture ranging behavior to assess feasibility of Vulture Safe Zones in Africa: Challenges and possibilities. *Biological Conservation*, *268*, 109516. https://doi.org/10.1016/j.biocon.2022.109516

Katzner, T. E., Pain, D. J., McTee, M., Brown, L., Cuadros, S., Pokras, M., Slabe, V. A., Watson, R. T., Wiemeyer, G., Bedrosian, B., Hampton, J. O., Parish, C. N., Pay, J. M., Saito, K., & Schulz, J. H. (2024). Lead poisoning of raptors: State of the science and cross-discipline mitigation options for a global problem. *Biological Reviews*, *99*(5), 1672–1699. https://doi.org/10.1111/brv.13087

Kays, R., Crofoot, M. C., Jetz, W., & Wikelski, M. (2015). Terrestrial animal tracking as an eye on life and planet. *Science*, *348*(6240), aaa2478. https://doi.org/10.1126/science.aaa2478

Kirchner, T. M., Devineau, O., Chimienti, M., Thompson, D. P., Crouse, J., Evans, A. L., Zimmermann, B., & Eriksen, A. (2023). Predicting moose behaviors from tri-axial accelerometer data using a supervised classification algorithm. *Animal Biotelemetry*, *11*(1), 32. https://doi.org/10.1186/s40317-023-00343-0

Kuhn, M., & Vaughan, D. (2024). *parsnip: A Common API to Modeling and Analysis Functions* (Version 1.2.1) [R]. https://CRAN.R-project.org/package=parsnip

Leos-Barajas, V., Photopoulou, T., Langrock, R., Patterson, T. A., Watanabe, Y. Y., Murgatroyd, M., & Papastamatiou, Y. P. (2017). Analysis of animal accelerometer data using hidden Markov models. *Methods in Ecology and Evolution*, *8*(2), 161–173. https://doi.org/10.1111/2041-210X.12657

López-Bao, J. V., & Mateo-Tomás, P. (2022). Wipe out highly hazardous pesticides to deter wildlife poisoning: The case of carbofuran and aldicarb. *Biological Conservation*, *275*, 109747. https://doi.org/10.1016/j.biocon.2022.109747

Mateo-Tomás, P., & López-Bao, J. V. (2020). Poisoning poached megafauna can boost trade in African vultures. *Biological Conservation*, *241*, 108389. https://doi.org/10.1016/j.biocon.2019.108389

Mateo-Tomás, P., Rodríguez-Pérez, J., Fernández-García, M., García, E. J., Valente e Santos, J. P., Gutiérrez, I., Olea, P. P., Rodríguez-Moreno, B., & López-Bao, J. V. (2023). Wildlife as sentinels of compliance with law: An example with GPS-tagged scavengers and sanitary regulations. *Journal of Applied Ecology*, *60*(10), 2188–2198. https://doi.org/10.1111/1365-2664.14487

Mayrose, A., Labinger, Z., Steinitz, O., & Leader, N. (2017). *The red book of birds in Israel* (2017th ed.). INPA, SPNI.

McNutt, J. W. 'Tico,' & Bradley, J. (2014). Report on Kwando (Botswana) Vulture poisoning investigation 16 November 2013. *Vulture News*, *66*, 35–41.

Nathan, R., Monk, C. T., Arlinghaus, R., Adam, T., Alós, J., Assaf, M., Baktoft, H., Beardsworth, C. E., Bertram, M. G., Bijleveld, A. I., Brodin, T., Brooks, J. L., Campos-Candela, A., Cooke, S. J., Gjelland, K. Ø., Gupte, P. R., Harel, R., Hellström, G., Jeltsch, F., … Jarić, I. (2022). Big-data approaches lead to an increased understanding of the ecology of animal movement. *Science*, *375*(6582), eabg1780. https://doi.org/10.1126/science.abg1780

Nathan, R., Spiegel, O., Fortmann-Roe, S., Harel, R., Wikelski, M., & Getz, W. M. (2012). Using tri-axial acceleration data to identify behavioral modes of free-ranging animals: General concepts and tools illustrated for griffon vultures. *Journal of Experimental Biology*, *215*(6), 986–996. https://doi.org/10.1242/jeb.058602

Nemtzov, S., Hatzofe, O., Steinitz, O., & Vine, G. (2021). *An Innovative Automatic Location-Based Real-Time Alert System to Prevent Wildlife Poisoning Using GPS-Tagged Griffon Vultures has Led to Better Conservation of Endangered Species in Protected Areas* [Computer software]. https://panorama.solutions/en/solution/innovative-automatic-location-based-real-time-alertsystem-prevent-wildlife-poisoning-using

O'Donoghue, P., & Rutz, C. (2016). Real-time anti-poaching tags could help prevent imminent species extinctions. *Journal of Applied Ecology*, *53*(1), 5–10. https://doi.org/10.1111/1365-2664.12452

Ogada, D., Keesing, F., & Virani, M. Z. (2012). Dropping dead: Causes and consequences of vulture population declines worldwide. *Annals of the New York Academy of Sciences*, *1249*(1), 57–71. https://doi.org/10.1111/j.1749-6632.2011.06293.x

Ogada, D., Shaw, P., Beyers, R. L., Buij, R., Murn, C., Thiollay, J. M., Beale, C. M., Holdo, R. M., Pomeroy, D., Baker, N., Krüger, S. C., Botha, A., Virani, M. Z., Monadjem, A., & Sinclair, A. R. E. (2016). Another Continental Vulture Crisis: Africa's Vultures Collapsing toward Extinction. *Conservation Letters*, *9*(2), 89–97. https://doi.org/10.1111/conl.12182

Olea, P. P., & Mateo-Tomás, P. (2014). Living in risky landscapes: Delineating management units in multithreat environments for effective species conservation. *Journal of Applied Ecology*, *51*(1), 42–52. https://doi.org/10.1111/1365-2664.12176

Olea, P. P., Mateo-Tomás, P., & Sánchez-Zapata, J. A. (Eds.). (2019). *Carrion Ecology and Management* (Vol. 2). Springer International Publishing. https://doi.org/10.1007/978-3-030-16501-7

Peters, N., Kendall, C. J., Davies, J. G., Bracebridge, C., Nicholas, A., Mgumba, M. P., & Beale, C. M. (2023). Identifying priority locations to protect a wide-ranging endangered species. *Biological Conservation*, *277*, 109828. https://doi.org/10.1016/j.biocon.2022.109828

Plaza, P. I., Martínez-López, E., & Lambertucci, S. A. (2019). The perfect threat: Pesticides and vultures. *Science of The Total Environment*, *687*, 1207–1218. https://doi.org/10.1016/j.scitotenv.2019.06.160

R Core Team. (2023). *R: A language and environment for statistical computing* [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/

Rast, W., Portas, R., Shatumbu, G. I., Berger, A., Cloete, C., Curk, T., Götz, T., Aschenborn, O., & Melzheimer, J. (2024). Death detector: Using vultures as sentinels to detect carcasses by combining bio-logging and machine learning. *Journal of Applied Ecology*, *61*(12), 2936–2945. https://doi.org/10.1111/1365-2664.14810

Resheff, Y. S., Bensch, H. M., Zöttl, M., Harel, R., Matsumoto-Oda, A., Crofoot, M. C., Gomez, S., Börger, L., & Rotics, S. (2024). How to treat mixed behavior segments in supervised machine learning of behavioural modes from inertial measurement data. *Movement Ecology*, *12*(1), 44. https://doi.org/10.1186/s40462-024-00485-7

Resheff, Y. S., Rotics, S., Harel, R., Spiegel, O., & Nathan, R. (2014). AcceleRater: A web application for supervised learning of behavioral modes from acceleration measurements. *Movement Ecology*, *2*(1), 27. https://doi.org/10.1186/s40462-014-0027-0

Rodríguez-Pérez, J., Bravo-Chaparro, E., Fernández-García, M., González, J. C., Báguena, G., González-Quirós, P., López-Bao, J. V., & Mateo-Tomás, P. (in press). Integrating GPS-tracked sentinel species into the fight against wildlife poisoning. *Conservation Letters*. https://doi.org/10.1111/conl.13086

Sergio, F., Tanferna, A., Blas, J., Blanco, G., & Hiraldo, F. (2019). Reliable methods for identifying animal deaths in GPS- and satellite-tracking data: Review, testing, and calibration. *Journal of Applied Ecology*, *56*(3), 562–572. https://doi.org/10.1111/1365-2664.13294

Serratosa, J., Oppel, S., Rotics, S., Santangeli, A., Butchart, S. H. M., Cano-Alonso, L. S., Tellería, J. L., Kemp, R., Nicholas, A., Kalvāns, A., Galarza, A., Franco, A. M. A., Andreotti, A., Kirschel, A. N. G., Ngari, A., Soutullo, A., Bermejo-Bermejo, A., Botha, A. J., Ferri, A., … Jones, V. R. (2024). Tracking data highlight the importance of human-induced mortality for large migratory birds at a flyway scale. *Biological Conservation*, *293*, 110525. https://doi.org/10.1016/j.biocon.2024.110525

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., Newman, C., & Macdonald, D. W. (2008). Identification of animal movement patterns using tri-axial accelerometry. *Endangered Species Research*, *10*, 47–60. https://doi.org/10.3354/esr00084

Sheppard, J. K., McGann, A., Lanzone, M., & Swaisgood, R. R. (2015). An autonomous GPS geofence alert system to curtail avian fatalities at wind farms. *Animal Biotelemetry*, *3*(1), 43. https://doi.org/10.1186/s40317-015-0087-y

Slabe, V. A., Anderson, J. T., Millsap, B. A., Cooper, J. L., Harmata, A. R., Restani, M., Crandall, R. H., Bodenstein, B., Bloom, P. H., Booms, T., Buchweitz, J., Culver, R., Dickerson, K., Domenech, R., Dominguez-Villegas, E., Driscoll, D., Smith, B. W., Lockhart, M. J., McRuer, D., … Katzner, T. E. (2022). Demographic implications of lead poisoning for eagles across North America. *Science*, *375*(6582), 779–782. https://doi.org/10.1126/science.abj3068

Smith, J. E., & Pinter-Wollman, N. (2021). Observing the unwatchable: Integrating automated sensing, naturalistic observations and animal social network analysis in the age of big data. *Journal of Animal Ecology*, *90*(1), 62–75. https://doi.org/10.1111/1365-2656.13362

Spiegel, O., Harel, R., Centeno-Cuadros, A., Hatzofe, O., Getz, W. M., & Nathan, R. (2015). Moving beyond Curve Fitting: Using Complementary Data to Assess Alternative Explanations for Long Movements of Three Vulture Species. *The American Naturalist*, *185*(2), E44–E54. https://doi.org/10.1086/679314

Spiegel, O., Harel, R., Getz, W. M., & Nathan, R. (2013). Mixed strategies of griffon vultures' (Gyps fulvus) response to food deprivation lead to a hump-shaped movement pattern. *Movement Ecology*, *1*(1), 5. https://doi.org/10.1186/2051-3933-1-5

Tuia, D., Kellenberger, B., Beery, S., Costelloe, B. R., Zuffi, S., Risse, B., Mathis, A., Mathis, M. W., van Langevelde, F., Burghardt, T., Kays, R., Klinck, H., Wikelski, M., Couzin, I. D., van Horn, G., Crofoot, M. C., Stewart, C. V., & Berger-Wolf, T. (2022). Perspectives in machine learning for wildlife conservation. *Nature Communications*, *13*(1), 792. https://doi.org/10.1038/s41467-022-27980-y

Valletta, J. J., Torney, C., Kings, M., Thornton, A., & Madden, J. (2017). Applications of machine learning in animal behaviour studies. *Animal Behaviour*, *124*, 203–220. https://doi.org/10.1016/j.anbehav.2016.12.005

van Eeden, L. M., Crowther, M. S., Dickman, C. R., Macdonald, D. W., Ripple, W. J., Ritchie, E. G., & Newsome, T. M. (2018). Managing conflict between large carnivores and livestock. *Conservation Biology*, *32*(1), 26–34. https://doi.org/10.1111/cobi.12959

Venables, W., & Ripley, B. (2002). *Modern Applied Statistics with S* (4th Edition). Springer.

Venter, O., Sanderson, E. W., Magrach, A., Allan, J. R., Beher, J., Jones, K. R., Possingham, H. P., Laurance, W. F., Wood, P., Fekete, B. M., Levy, M. A., & Watson, J. E. M. (2016). Sixteen years of change in the global terrestrial human footprint and implications for biodiversity conservation. *Nature Communications*, *7*(1), 12558. https://doi.org/10.1038/ncomms12558

Wall, J., Wittemyer, G., Klinkenberg, B., & Douglas-Hamilton, I. (2014). Novel opportunities for wildlife conservation and research with real-time monitoring. *Ecological Applications*, *24*(4), 593–601. https://doi.org/10.1890/13-1971.1

Wand, M. (2024). *KernSmooth: Functions for Kernel Smoothing Supporting Wand & Jones (1995).* [Computer software].

Wang, G. (2019). Machine learning for inferring animal behavior from location and movement data. *Ecological Informatics*, *49*, 69–76. https://doi.org/10.1016/j.ecoinf.2018.12.002

Williams, H. J., Shepard, E. L. C., Duriez, O., & Lambertucci, S. A. (2015). Can accelerometry be used to distinguish between flight types in soaring birds? *Animal Biotelemetry*, *3*(1), 45. https://doi.org/10.1186/s40317-015-0077-0

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., Griffiths, R. C., Kay, W. P., Morales, J. M., Potts, J. R., Rogerson, K. F., Rutz, C., Spelt, A., Trevail, A. M., Wilson, R. P., & Börger, L. (2020). Optimizing the use of biologgers for movement ecology research. *Journal of Animal Ecology*, *89*(1), 186–206. https://doi.org/10.1111/1365-2656.13094

Wright, M. N., & Ziegler, A. (2017). ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R. *Journal of Statistical Software*, *77*, 1–17. https://doi.org/10.18637/jss.v077.i01

Yu, H., Deng, J., Nathan, R., Kröschel, M., Pekarsky, S., Li, G., & Klaassen, M. (2021). An evaluation of machine learning classifiers for next-generation, continuous-ethogram smart trackers. *Movement Ecology*, *9*(1), 15. https://doi.org/10.1186/s40462-021-00245-x

Supporting Information

Data S1 - Transmitter schedule of free-roaming griffon vultures

The GPS-ACC transmitters were programmed with different schedules, depending on the analysis:

Conducting observations to build an ACC training dataset

To build the ACC training dataset, bouts of 5 sec at 20Hz were recorded for free-roaming griffons every 10 min, if the transmitter’s battery was more than 50% charged, or every 60 min if the transmitter’s battery was below 50% charge. The GPS positions of free-roaming vultures were recorded every 10 min if the transmitter’s battery was above 75%, every 20 min if the battery was between 50-75%, or every 60 min if the battery was below 50% charge.

Using the confidence score to validate ‘Feeding’ predictions of free-roaming vultures

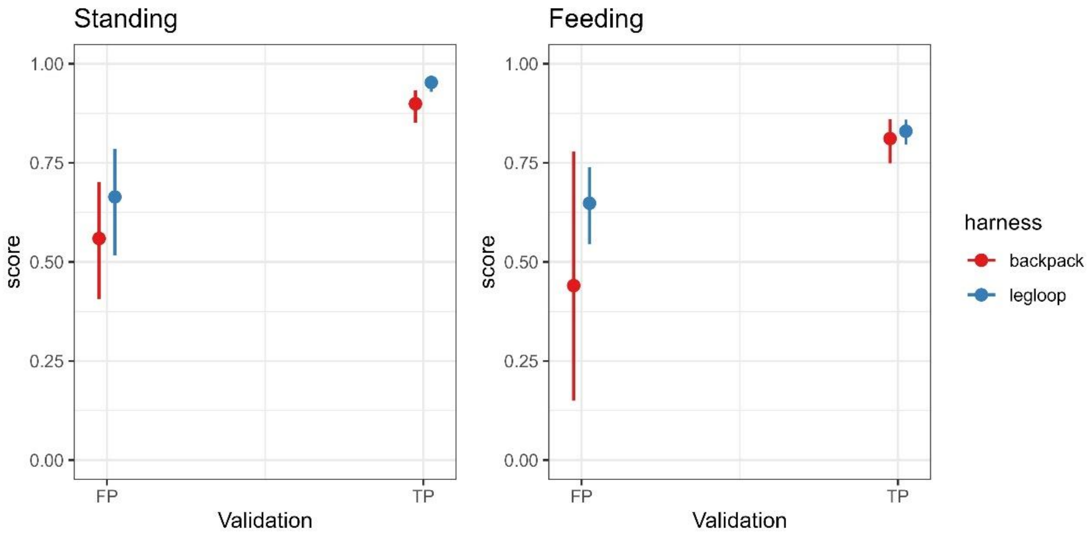
To use the confidence score to validate the feeding predictions of free-roaming vultures, the GPS positions were recorded every 10 min to 60 min (depending on the battery charge). The transmitter was programmed to collect 5 sec bouts of accelerometer data at 20Hz every 10 min if the battery was over 50% charge.

Case study: identification of vulture feeding hotspots to prevent poisoning

To build a map of feeding hotspots, the transmitters were programmed to record GPS positions every 10 min if the battery was over 50%, and every 60 min, if the battery was below 50% charge. Accelerometer data were recorded every 5 min if the battery was over 50% charge.

Data S2 – Details of the model building sequence

Using the ‘training subset’, we built a random forest model (number of trees = 1000, number of statistical features per tree = 10, and minimum number of observations per leaf = 10). We evaluated the performance of our model using the ‘testing’ subset. We built a confusion matrix and calculated three performance metrics, for the full model and for each behavior: (i) accuracy (sum of true-positives and true-negatives divided by all predictions); (ii) precision (true-positives divided by the sum of true-positives and false-positives); and (iii) recall (true-positives divided by the sum of true-positives and false-negatives). The equations for each metric can be found in Table 1. Because only 67% of the data was used to train the algorithm, these metrics are likely an underestimation of the real performance of the final algorithm.



Supplementary Figure S1 - Comparison of the confidence scores of true-positive (TP) and false-positive (FP) 'Standing' and 'Feeding' classifications, recorded with a backpack (red) and with a leg-loop (blue) harness. Each panel shows the model predictions and the 95% confidence intervals for the effect of the harness on the confidence scores of true positive and false positive behavioral classifications.

Leaflet | © OpenStreetMap, ODbL, Tiles © Esri — Source: Esri, i-cubed, USDA, USGS, AEX, GeoEye, Getmapping, Aerogrid, IGN, IGP, UPR-EGP, and the GIS User Community

Supplementary Figure S2 – Using GPS-derived metrics to identify feeding hotspots, without using acceleration-based behavioral classification. The red points show vulture locations with GPS ground speed equal to or slower than 4m/s, outside the roosting areas and outside feeding stations for vultures, during November 2022 (n = 1,938 locations). The polygons indicate the density of vulture locations, showing potential areas where vultures could be foraging and would be at greater risk of poisoning. The polygon colors indicate the density of vulture locations, with blue areas having lower density and red areas higher density. Comparison with Figure 6 (based on accelerometer-classified ‘Feeding’ behavior) shows that incorporating behavioral data produces spatially distinct and more accurate hotspot identification.

Supplementary Table S1 – Full list of statistical features used to summarize each acceleration bout. Note that X, Y and Z refer to the orthogonal acceleration axes (sway – X; surge – Y; heave - Z), while *x* and *y* refer to the variables that should be included in the R functions.

Feature	**R function**
Mean (X, Y, Z)	mean(*x*)
Maximum (X, Y, Z)	max(*x*)
Minimum (X, Y, Z)	min(*x*)
Range (X, Y, Z)	max(*x*) – min(*x*)
Standard deviation (X, Y, Z)	sd(*x*)
Skewness (X, Y, Z)	moments::skewness(*x*)
Kurtosis (X, Y, Z)	moments::kurtosis(*x*)
Euclidian norm (X, Y, Z)	sqrt(sum(*x*^2))
25% quantile (X, Y, Z)	quantile(*x*, probs = 0.25)
50% quantile (X, Y, Z)	quantile(*x*, probs = 0.50)
75% quantile (X, Y, Z)	quantile(*x*, probs = 0.75)
Covariance (X-Y, X-Z, Y-Z)	cov(*x*, *y*)
Mean difference (X-Y, X-Z, Y-Z)	mean(*x* – *y*)
Standard deviation of the difference (X-Y, X-Z, YZ)	sd(*x* – *y*)
Mean amplitude (X, Y, Z)	mean_amplitude(*x*) – user defined function provided in the R code

Supplementary Table S2 – Results of the random forest model to classify vulture behavior based on accelerometer data. The random forest model was trained on 67% of the total number of observations and its performance was tested on 33% of the data. The true-positives and false-positives are the result of the performance test (for a full confusion matrix see Figure 3 of the manuscript). For each prediction, the model calculates a confidence score, i.e., the proportion of trees that agree on the highest scoring prediction. Here we show the mean confidence score values for each predicted behavior category.

Behavior	Total number of observations (number of vultures)	True positives	False positives	Precision	Recall	Mean confidence score of true-positives (sd)
Feeding	587 (10)	189	17	0.87	0.92	0.82 (0.15)
Lying	364 (5)	104	5	0.94	0.95	0.98 (0.06)
Standing	3488 (11)	1142	20	0.98	0.98	0.98 (0.08)
Ground	67 (8)	4	23	0.57	0.15	0.41 (0.03)
Flapping	122 (18)	42	2	0.98	0.95	0.93 (0.12)
Soaring	1155 (9)	375	4	0.99	0.99	0.99 (0.08)

Supplementary Table S3 – Comparison of the confidence scores of true-positive and false-positive behavioral classifications. Results of the GLMM beta binomial model, with confidence score (0-1) as the response variable, and correctness (true-positive / false-positive), behavioral class (Feeding, Flapping, Ground, Lying, Soaring and Standing) and their interaction as predictors. The model had device ID as random effect. A p-value <0.05 indicates a statistical significant relationship.

Variable	Estimate	Std. Error	Z value	p-value
Intercept	0.470	0.215	2.191	0.028
Validation – True positive	0.876	0.195	4.492	<0.001
Behavior – Flapping	-0.775	0.578	-1.343	0.179
Behavior – Ground	0.173	0.248	0.698	0.485
Behavior – Lying	-0.133	0.403	-0.330	0.741
Behavior – Soaring	0.902	0.456	1.979	0.048
Behavior – Standing	-0.020	0.254	-0.079	0.937
Valid. True Positive x Beh. Flapping	1.761	0.586	3.007	0.003
Valid. True Positive x Beh. Ground	-1.749	0.445	-3.926	<0.001
Valid. True Positive x Beh. Lying	1.807	0.416	4.342	<0.001
Valid. True Positive x Beh. Soaring	1.017	0.442	2.304	0.021
Valid. True Positive x Beh. Standing	1.341	0.260	5.162	<0.001

Random effects	Variance	Std. Dev.
Device id	0.0995	0.3155

Supplementary Table S4 – Comparison of the confidence scores of true-positive and false-positive 'Standing' and 'Feeding' classifications, recorded with a backpack and with a leg-loop harness. Results of the GLMM model beta binomial model, with confidence score (0-1) as the response variable, and the correctness (true-positive / false-positive), the type of harness (backpack / leg-loop) and their interaction as predictors. The model had device ID as random effect. A p-value <0.05 indicates a statistical significant relationship.

Behavior	**Variable**	**Estimate**	**Std. Error**	**Z value**	**p-value**	**Random effect variance**	**Random effect std. dev.**
Standing	**Intercept**	0.238	0.315	0.754	0.451	-	-
	Correctness – True positive	1.949	0.250	7.808	<0.001	-	-
	Harness – Leg-loop	0.444	0.451	0.985	0.325	-	-
	Correctness (True positive) x Harness (Leg-loop)	0.374	0.349	1.074	0.283	-	-
	Device id	-	-	-	-	0.185	0.43
Feeding	**Intercept**	-0.240	0.764	-0.314	0.753	-	-
	Correctness – True positive	1.698	0.769	2.207	0.0273	-	-
	Harness – Leg-loop	0.851	0.796	1.069	0.285	-	-
	Correctness (True positive) x Harness (Leg-loop)	-0.725	0.798	-0.908	0.364	-	-
	Device id	-	-	-	-	0.022	0.149

Supplementary Table S5 - Confusion matrix of the random forest assessing the influence of the harness type on the performance of the algorithm. This random forest was trained solely on leg-loop data and tested on backpack data.

Predicted / Observed	**Feeding (n = 104)**	**Ground (n = 47)**	**Standing (n = 563**
Feeding	104	43	52
Ground	0	0	0
Standing	0	4	511

Precision	0.52	NA	0.99
Recall	1.00	0	0.91

Supplementary Table S6 – Comparison of the confidence scores of feeding bouts likely to be true-positives or false-positives. Results of the GLMM model beta binomial model, with confidence score (0-1) as the response variable, and correctness (likely to be true-positive – inside feeding stations; likely to be false-positives – on cliffs or in flight) as predictor. The model had device ID as random effect. A p-value <0.05 indicates a statistical significant relationship.

Variable	Estimate	Std. Error	Z value	p-value
Intercept	1.101	0.078	14.140	<0.001
Correctness – Low probability of true positive	-0.805	0.167	-4.827	<0.001

Random effects	Variance	Std. Dev.
Device id	<0.001	<0.001

Data S3 – Temporal correlation between behaviors

To test if there is temporal correlation between behaviors, we used the dataset of 51 free-roaming vultures collected during November 2022, consisting of 197,641 behaviors with confidence scores over 0.5.

We first calculated the overall occurrence rate of each behavior (*OR*, Supplementary Figure S3A). Then, for each individual on each day, we quantified how often a behavior was followed by another (Behavior Transition Rate, *BTR*, Supplementary Figure S3B). For example, how often ‘Feeding’ was followed by ‘Standing’ (Feeding-to-Standing BTR), by ‘Feeding’, by ‘Ground’, etc. We excluded transitions between behaviors that were separated by more than 10 minutes.

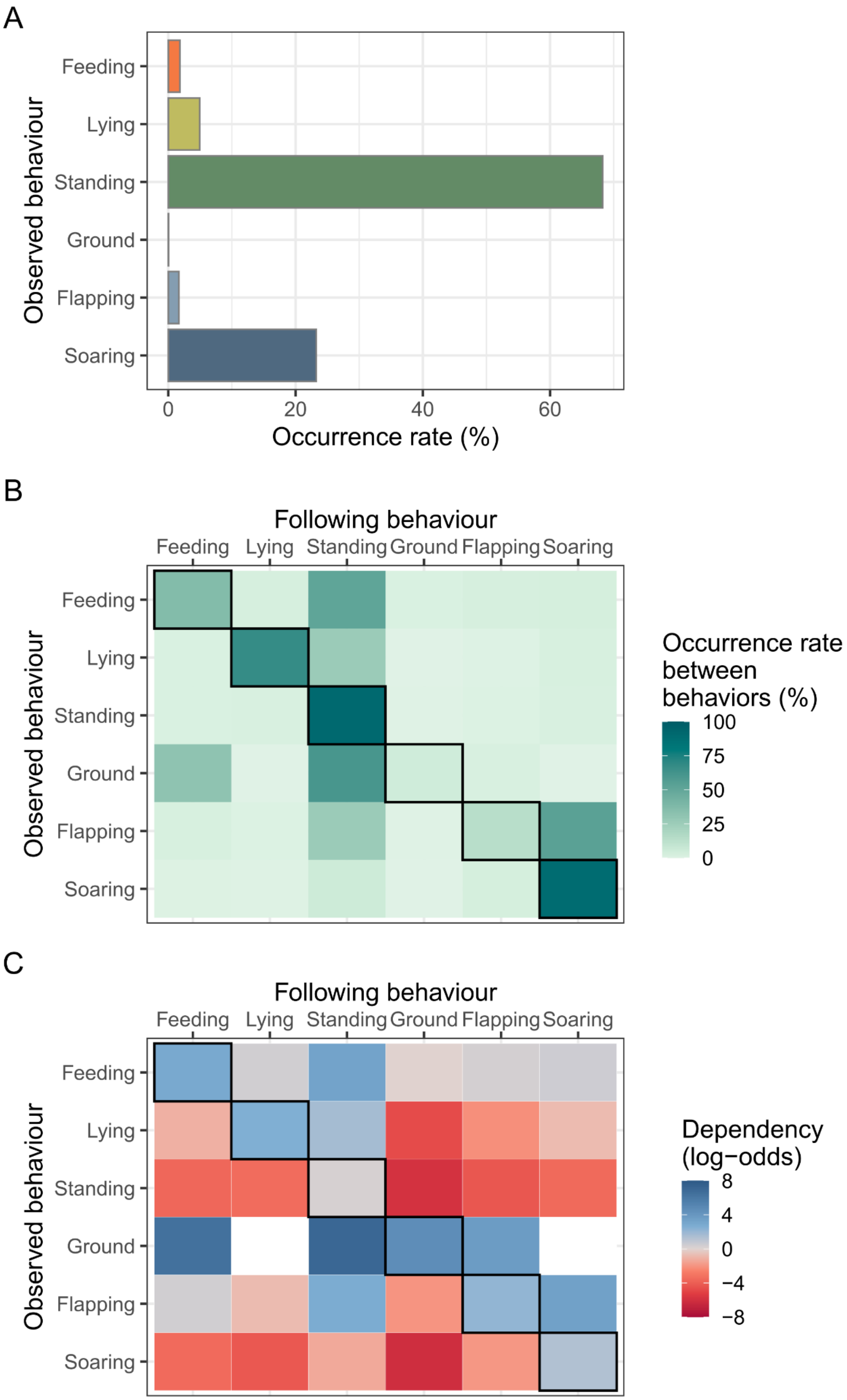
Frequent behaviors may inflate *BTR* values simply due to their high occurrence (e.g., ‘Feeding’ followed by ‘Standing’ could be common because ‘Standing’ is generally frequent in the dataset due to the time vultures spend standing). We corrected for this by computing the dependency between behaviors as the log-odds of the ratio between *BTR* and the overall *OR*. Positive dependency values indicate strong temporal associations (e.g., Feeding is very often followed by Standing), negative dependency values indicate avoidance or rare transitions (e.g., Soaring is rarely followed by Feeding), and values near zero indicate weak or random associations (Supplementary Figure S3C).

$$OR = \frac{number\ of\ times\ a\ behavior\ occurred}{total\ number\ of\ behaviors} \times 100$$

$$ORB = \frac{number\ of\ times\ behavior\ A\ is\ followed\ by\ behavior\ B}{total\ number\ of\ behavior\ A} \times 100$$

$$dependency = \ln\left(\frac{ORB}{OR}\right)$$

Despite the relatively low temporal resolution of our data (5 to 10min, depending on the dataset, i.e., longer than many of the behavior duration), we still detected temporal correlation between some behaviors. After adjusting for overall behavior frequency, we found that ‘Feeding’ was frequently followed by ‘Feeding’ or ‘Standing’, but not by ‘Ground’. In contrast, ‘Ground’ was often followed by ‘Feeding’, ‘Standing’, ‘Ground’, and ‘Flapping’ (the latter likely due to the running behaviors while taking-off). Yet, we note that this behavior was rare in the dataset (0.05% occurrence rate, n = 95), so these dependencies may reflect the limited sample size.


A
Observed behaviour
Feeding
Lying
Standing
Ground
Flapping
Soaring
0
20
40
60
Occurrence rate (%)
B
Following behaviour
Feeding
Lying
Standing
Ground
Flapping
Soaring
Observed behaviour
Feeding
Lying
Standing
Ground
Flapping
Soaring
Occurrence rate between behaviors (%)
100
75
50
25
0
C
Following behaviour
Feeding
Lying
Standing
Ground
Flapping
Soaring
Observed behaviour
Feeding
Lying
Standing
Ground
Flapping
Soaring
Dependency (log−odds)
8
4
0
−4
−8

Supplementary Figure S3 – Temporal correlation between consecutive behaviors. Figure A shows the total occurrence rate of all behaviors (*OR*); Figure B shows percentage of times that a behavior is followed by another (*BTR*); Figure C shows the dependency between behaviors, calculated as the log-odds of the ratio between occurrence rate between behaviors (*BTR*) and the total occurrence rate of each behavior (*OR*). In C, positive dependencies are depicted by shades of blue, negative dependencies are depicted by shades of red, and white tiles depict undefined values (log of 0). In B and C, darker tones represent stronger relationships, and the black outlines in the diagonal highlight the self-dependency (e.g., Feeding-Feeding dependency).