1   # Using accelerometer-based behavioral classification to
2   # enhance scavenger conservation

3

4   Gideon Vaadia*[1], Marta Acácio*[1,2,3,4], Tal Agassi[5], Nili Anglister[1], Yigal Miller[6],

5   Ohad Hatzofe[6], Patricia Mateo-Tomás[7], Jorge Rodríguez-Pérez[7], María

6   Fernández-García[7]; Pedro P. Olea[8,9], Ignacio Otero[10], Noa Pinter-Wollman[11],

7   Moni Shahar[5] & Orr Spiegel[1]

8

9   * Equal contributions

10   Corresponding author

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

12

13   [1] School of Zoology, Faculty of Life Sciences, Tel Aviv University, Tel Aviv, Israel

14   [2] CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, InBIO
15   Laboratório Associado, Campus de Vairão, University of Porto, 4485-661 Vairão,
16   Portugal

17   [3] CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, InBIO
18   Laboratório Associado, Instituto Superior de Agronomia, University of Lisboa, 1349-017
19   Lisboa, Portugal

20   [4] BIOPOLIS Program in Genomics, Biodiversity and Land Planning, CIBIO, Campus de
21   Vairão, 4485-661 Vairão, Portugal

22   [5] Center of AI and Data Science, Tel Aviv University (TAD), Tel Aviv, Israel

23   [6] Israeli Nature and Parks Authority, Am Ve'Olamo 3, Jerusalem, Israel

24   [7] Biodiversity Research Institute (CSIC – University of Oviedo – Principality of Asturias),
25   Mieres, Spain

26   [8] Terrestrial Ecology Group (TEG_UAM), Departamento de Ecología, Facultad de
27   Ciencias, Universidad Autónoma de Madrid (UAM), Madrid, Spain

28   [9] Centro de Investigación en Biodiversidad y Cambio Global (CIBC-UAM), Universidad
29   Autónoma de Madrid (UAM), Madrid, Spain

30   [10] GREFA (Grupo de Rehabilitación de la Fauna Autóctona y su Hábitat), Majadahonda,
31   Spain

32   [11] Department of Ecology and Evolutionary Biology, University of California Los Angeles,
33   Los Angeles, California, USA

Abstract

35   1. Human activities are endangering animal species globally and implementing

36   effective conservation strategies requires understanding animal behavior and

37   ecology. Advancements in GPS tracking technology, accelerometry, and

38   machine learning algorithms are allowing the *in-situ* study of animal movement

39   and behavior remotely. However, the challenge of building supervised machine

40   learning algorithms and collecting the large datasets required to train them is

41   hampering the widespread use of these tools. Additionally, the reliability of these

42   models in classifying unobserved behaviors is rarely validated, resulting in

43   possible classification errors.

44   2. We built a supervised accelerometer-based behavioral classification model for

45   griffon vultures (*Gyps fulvus*). Similarly to most other avian scavenger

46   populations worldwide, griffons are critically endangered in Israel and

47   neighboring countries, mostly due to feeding on poisoned carcasses. Thus,

48   identifying this scavenger's feeding behavior and foraging areas is crucial for their

49   conservation.

50   3. We trained a Random Forest model on acceleration data of 14 captive and 17

51   free-roaming griffons. We classified 5783 behavioral observations into 6 classes:

52   feeding, lying, standing, other ground behaviors, flapping and soaring flight. The

53   model performed well (0.96 accuracy, 0.89 precision and 0.82 recall) and,

54   importantly, feeding behaviors were accurately classified (0.87 precision, 0.92

55   recall). We calculated an observation-specific confidence score and

56   demonstrated its effectiveness in identifying true- and false-positive

57   classifications, in both captive and free-roaming individuals. Finally, we used our

58   model to reliably identify feeding hotspots, where vultures can be at higher risk

59   of poisoning.

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

66 ## Resumen

67   1. Las actividades humanas suponen una amenaza para la fauna a escala global e
68   implementar estrategias de conservación efectivas al respecto requiere entender
69   el comportamiento y la ecología animal. Avances en tecnologías como el
70   seguimiento GPS, los acelerómetros y los algoritmos de aprendizaje automático
71   facilitan el estudio remoto in situ de los movimientos y comportamientos
72   animales. Sin embargo, elaborar algoritmos supervisados de aprendizaje
73   automático y recopilar la gran cantidad de datos necesarios para entrenarlos son
74   retos que dificultan el uso generalizado de estas herramientas. Además, la
75   fiabilidad de estos modelos al clasificar comportamientos no observados no
76   suele validarse, lo que pueda dar lugar a errores de clasificación.

77   2. Elaboramos un modelo supervisado de clasificación de comportamientos de
78   buitre leonado (Gyps fulvus) con datos de acelerómetro. En consonancia con la
79   mayoría de las poblaciones de carroñeros en el mundo, los buitres están
80   críticamente amenazados en Israel y países vecinos, principalmente por
81   consumir carroñas envenenadas. Por tanto, identificar comportamientos y áreas
82   de alimentación de estos carroñeros es clave para su conservación.

83   3. Entrenamos un modelo "Random Forest" con datos de 14 buitres en cautividad
84   y 17 en libertad, clasificando 5.783 comportamientos observados en 6 clases:
85   alimentación, tumbado, posado, otros comportamientos en suelo, aleteo y
86   planeo. El modelo funcionó bien (exactitud: 0,96, precisión: 0,89 y sensibilidad:

87   0,82), identificando con exactitud los comportamientos de alimentación

88   (precisión: 0,87 y sensibilidad: 0,92). Calculamos un índice de confianza

89   específico de cada observación, que discriminó con solvencia verdaderos y

90   falsos positivos tanto en buitres en cautividad como en libertad. Finalmente

91   usamos el modelo para identificar zonas de alimentación, dónde el riesgo de

92   envenenamiento de los buitres pudiera incrementarse.

93   4. Síntesis y aplicaciones: Proporcionamos una herramienta útil para identificar

94   zonas de alimentación de buitres, apoyando así tareas de gestión para prevenir

95   su envenenamiento. Junto con el seguimiento casi en tiempo real, nuestro

96   modelo puede ayudar en la lucha contra el envenenamiento de fauna carroñera.

97   Los datos de entrenamiento, el modelo y los códigos, junto con un marco

98   conceptual, están disponibles en una plataforma de fácil uso para fomentar su

99   utilización por ecólogos y profesionales de la conservación.

100

101   **Keywords (up to 8):**

102   Accelerometer, Behavior classification, Random Forest, Griffon Vulture, Poisoning,

103   Conservation, Biotelemetry, Supervised machine learning.

104

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

132 varying resolution (i.e., frequency, in Hz) and intervals (for example, recording for 5

133 seconds at 20Hz, every 10 minutes). Different acceleration signatures enable the

134 measurement of movement-related energy expenditure (Gleiss et al., 2011; Halsey et

135 al., 2009) and can be used to distinguish among different behaviors (Shepard et al.,

136 2008), for example, for estimating flight duration in small migratory passerines (Bäckman

137 et al., 2017).

138   Machine learning algorithms are used to classify raw acceleration bouts into

139 different behavioral classes (Nathan et al., 2012; Resheff et al., 2014; Valletta et al.,

140 2017; Wang, 2019; Yu et al., 2021). These algorithms can operate in an unsupervised

141 manner, identifying similarities in acceleration data to produce unlabeled clusters of

142 similar measurements that subsequently need to be manually classified into specific

143 behaviors (Chimienti et al., 2016; Wang, 2019). Alternatively, supervised learning

144 involves training an algorithm with a dataset in which each behavior is labeled, allowing

145 the algorithm to 'learn' the distinctive acceleration patterns of different behaviors (Nathan

146 et al., 2012; Wang, 2019; Yu et al., 2021). However, depending on the level of detail

147 required and on how distinctive the behaviors are, assembling a training dataset can be

148 laborious, as it typically requires direct observations of animals in the wild or in captivity,

149 synchronized with the ACC measurements (Campbell et al., 2013; Dickinson et al.,

150 2021). Despite these difficulties, supervised machine learning algorithms have been

151 successfully used to classify behaviors across diverse animal groups, including baboons

152 (Fehlmann et al., 2023), large pelagic fish (Clarke et al., 2021), sea turtles (Jeantet et

153 al., 2020), condors, and other vultures (Rast et al., 2024; Spiegel et al., 2013; Williams

154 et al., 2015). Commonly used algorithms include artificial neural networks, extreme

155 gradient boosting, and random forests (Resheff et al., 2014; Yu et al., 2021). Random

156 forests have the advantage of being able to model complex interactions between the,

157 often correlated, predictor variables, therefore not requiring the pre-processing and

158 filtering of variables (Cutler et al., 2007), and simplifying behavioral classification.

159   Once a machine learning model is trained, it can classify new, unseen

160   accelerometer data into the trained behavioral classes, invariably resulting in some

161   classification errors (Glass et al., 2020; Jeantet et al., 2020). Errors emerge from a few,

162   non-mutually exclusive, processes. First, acceleration bouts, particularly long ones, may

163   include transitions among behaviors resulting in a mixture of different acceleration

164   signatures (Resheff et al., 2024). Second, rare behaviors may be underrepresented or

165   missing from the limited training dataset (e.g., seasonal and rare behaviors such as

166   copulation). Third, the behavioral repertoire of some individuals may be broader than

167   what the algorithm is trained for. Because some behaviors might be difficult to observe

168   in captivity (e.g., flight behaviors, Williams et al., 2015), this last error is particularly

169   relevant for algorithms trained on captive individuals that are used to predict the

170   behaviors of wild animals (Dickinson et al., 2021). Still, the models must choose the best

171   fitting behavioral class among the available options, even if none provides a particularly

172   good fit. These errors demand a mechanism to verify the accuracy of each behavioral

173   classification, allowing the model to distinguish between true-positive and false-positive

174   classifications (Bidder et al., 2014; Glass et al., 2020). While some studies offer guidance

175   on how to best use and analyze large acceleration datasets (e.g., Leos-Barajas et al.,

176   2017; Resheff et al., 2014; Williams, Taylor, et al., 2020), the complexity of these tools

177   remains a barrier for non-experts, hindering their use in conservation science and

178   practice.

179   Here we develop an accelerometer-based behavioral classification tool and

180   validate its real-world application in ecology and conservation, using griffon vultures

181   (*Gyps fulvus*) as a case study. As obligate scavengers, vultures support key ecosystem

182   functions by consuming carcasses and recycling nutrients (Buechley & Şekercioğlu,

183   2016). Yet, around the world, 70% of vulture species are in danger of extinction, with

184   poisoning driven by consuming carcasses containing toxic substances being one of the

185   leading causes for population declines (Ives et al., 2022; Ogada et al., 2012; Plaza et al.,

186   2019). Poisoning can be either intentional or unintentional. For instance, poachers may
187   lace carcasses with poison to prevent these raptors from alerting environmental
188   authorities of poached wildlife (Mateo-Tomás & López-Bao, 2020; Ogada et al., 2016),
189   and farmers may do so for combating pests and mammalian carnivores. Anti-
190   inflammatory drugs used to treat cattle are also lethal to vultures, leading to poisoning at
191   these carcasses (López-Bao & Mateo-Tomás, 2022; Ogada et al., 2012; Plaza et al.,
192   2019). Several hundred vultures may quickly gather to eat at a single carcass, increasing
193   their vulnerability to mass poisoning events (McNutt & Bradley, 2014). Other scavenger
194   species also feed on carcasses (Olea et al., 2019), exposing them to similar risks of
195   poisoning (Katzner et al., 2024; López-Bao & Mateo-Tomás, 2022). Early detection of
196   carcasses might facilitate their proper management to, for example, prevent vultures and
197   other wild species from feeding on carrion with toxic substances by removing these from
198   the environment. Moreover, when poisoning events are promptly detected, vultures and
199   other animals can undergo medical intervention (Acácio et al., 2023; Anglister et al.,
200   2023). Considering the vulture's large roaming areas (Spiegel et al., 2015), tracking
201   technology and behavioral classification are essential tools to identify vultures' feeding
202   events. Vultures can therefore act as sentinels, facilitating carcass detection, and
203   maximizing prompt intervention efforts to reduce detrimental effects associated with
204   consumption of contaminated carrion. For instance, by reducing the number of fatalities
205   at a poisoning event to avoid long-term effects on species' populations (Acácio et al.,
206   2023; Slabe et al., 2022).

207   In this study, our goals are to 1) develop an ACC-based behavioral classification
208   algorithm, which, together with the training dataset and a conceptual framework of the
209   methodological workflow, is made freely available to conservationists and ecologists; 2)
210   validate the algorithm's classifications, by comparing the confidence scores of true-
211   positive and false-positive classifications, using both the training dataset and data from
212   free-roaming vultures; and 3) apply our novel algorithm to real-life scenarios with

213 important conservation implications – i.e., rapid carcass detection to prevent vulture

214 poisoning. Ultimately, our goal is to combine technological advancements in GPS and

215 accelerometry to improve wildlife conservation efforts, and to develop a tool that is easily

216 transferable to other systems.

217

Materials and Methods

219 *Study system*

220 The study took place in Israel, where griffon vultures are critically endangered (Mayrose

221 et al., 2017). A historical population of thousands of griffons is currently declining; three

222 decades ago, there were only 400 griffons in this population, and fewer than 200

223 individuals remain today (Hatzofe, 2020). Pesticide poisoning from consuming laced

224 carcasses is the leading cause of griffon mortality, accounting for 45% of documented

225 deaths between 2010-2021 in this region (Anglister et al., 2023). Lead poisoning and

226 ingestion of animals treated with anti-inflammatory drugs each contribute to 6% of

227 mortality events, posing additional threats to this population (Anglister et al., 2023). To

228 prevent the local extinction of this species, the Israel Nature and Parks Authority (INPA)

229 runs an intricate management program, including the provisioning of contaminant-free

230 food at supplementary feeding stations (Spiegel et al., 2013, 2015), the release of

231 captive-bred and translocated griffons (Efrat et al., 2020), and individually tracking

232 vultures using GPS-Accelerometer transmitters, to identify poisoning events and other

233 threats. When wild carcasses are detected in a random location within areas of known

234 pastoral activity and poisoning history, or when vultures exhibit minimal movement,

235 suggesting they are unwell, rangers are sent to the field to remove the carcasses and/or

236 transport affected individuals to a wildlife hospital, underscoring the critical role of GPS-

237 tracking data for the conservation of this population (Acácio et al., 2023).

238   Every year, approximately 100 free-roaming griffons are captured by the INPA

239   using a cage trap, to identify the individuals with metal and color rings and with patagial

240   tags. In these trapping events, a few individuals are fitted with a GPS-ACC transmitter

241   (Ornitela OT-50) using a Teflon harness in a leg-loop configuration. The transmitters are

242   equipped with solar panels that recharge the batteries, and transmit the collected data

243   via the GSM network, eliminating the need for recapturing individuals to retrieve

244   information. The current study did not require additional ethical approval, since it uses

245   tracking data collected from these GPS-ACC tags and no captures of free-roaming

246   vultures were conducted specifically for this study. The capture and tagging procedures

247   were approved by ethics committee of the Israel Nature and Parks Authority (permit

248   number 42166).

249

250   *Conducting observations to build an ACC training dataset*

251   Between January 2020 and February 2022, we deployed GPS-ACC transmitters on 31

252   griffon vultures, 14 captive vultures and 17 free-roaming individuals. The captive vultures

253   were housed in in 4 breeding programs, rehabilitation or wildlife facilities, in Israel and in

254   Spain: Ramat Hanadiv (Israel, n = 4), Hai-Bar Carmel (Israel, n = 4), Cabárceno Wildlife

255   Park (Spain, n = 3) and GREFA wildlife hospital (Spain, n = 3). In each cage, there were

256   6 to 12 vultures. Additional behavioral data was collected in Israel for 17 free-roaming

257   griffon vultures. One individual dropped his transmitter and was deployed with another

258   device (thus there were 31 individuals but 32 transmitters). In Israel, the transmitters

259   were deployed using a leg-loop harness and in Spain the loggers were deployed using

260   a backpack harness.

261   The transmitters were programmed to collect GPS and ACC at independent

262   schedules, and differently for captive and free-roaming griffons. The transmitters of

263   captive griffons were programmed to collect tri-axial accelerometer data at 20Hz almost

264 continuously (10-min-long periods, with a 1 sec interval in between). These 10min bouts

265 were parsed into 5 sec bouts to match the free-roaming dataset. Bouts of 5 sec at 20Hz

266 were recorded for free-roaming griffons every 10 min, depending on the transmitter's

267 battery charge (see Supplementary Material for details).

268   To classify each 5 sec ACC bout as a particular behavior, we conducted direct

269 observations and video recordings of the tagged griffons, both in captivity and in the wild.

270 In total, we performed observations for 79 days. Direct observations of captive and wild

271 griffons were conducted with a spotting scope (Swarovski ATX spotting scope 85mm),

272 ensuring a sufficient distance to not disturb the vulture's natural behavior. The video

273 recordings were captured using a camera mounted on a wall support in Spain, and with

274 nest cameras at the captive breeding facilities, or live streaming nest cameras at wild

275 nests in Israel (BirdLife Israel, 2022). The direct observations of wild vultures were

276 performed at roosting sites, at approximately 250 of the individuals.

277   We recorded six ecologically important behavioral classes: 'Standing' - vulture is

278 resting upright (could be roosting, and may include minor preening and changes in body

279 posture); 'Lying' - vulture is lying parallel to the ground, either resting or incubating;

280 'Feeding' – vulture is either directly eating from a carcass, or engaged in intense social

281 interactions next to the carcass (e.g., fighting or posturing towards other vultures before

282 eating); 'Ground' - includes all other active ground behaviors which are not directly

283 related to feeding or resting (e.g., walking, running, hopping, etc.); 'Flapping' - active

284 flight with wingbeats; and 'Soaring' - passive flight (e.g., thermal soaring, gliding, etc.).

285 Because long flights do not occur in captivity, we used GPS-ACC data from 17 free-

286 roaming griffons in southern Israel to classify 'Soaring' (passive) and 'Flapping' (active)

287 flight behaviors. We identified segments of continuous flight using the GPS ground speed

288 (ground speed >4m/sec) and plotted the acceleration measurements taken during these

289 flights. The acceleration signatures of soaring and passive flights are so distinctive

290 (Figure 1B,C, Williams et al., 2015) that there was no need to ground-truth these

291 behaviors with visual sightings (which would be challenging, considering their large
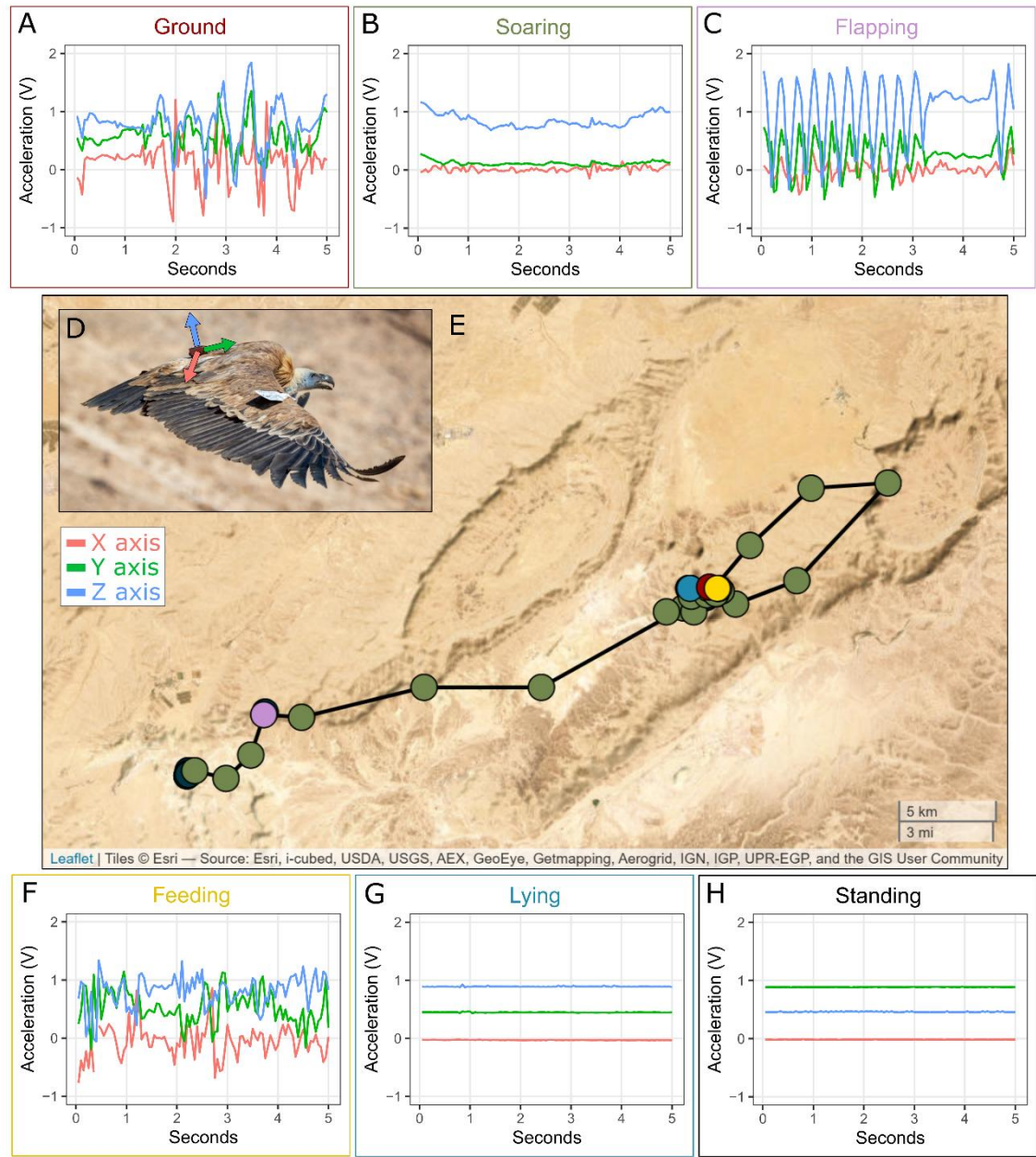
292 roaming areas).

293

294

295 **Figure 1 - Examples of accelerometer-based classification of griffon vulture's**

296 **behaviors.** Acceleration measurements of bouts classified as: (A) 'Ground, (B) 'Soaring'

297 flight, (C), 'Flapping' flight, (F) 'Feeding', (G) 'Lying' and (H) 'Standing'. The acceleration

298 data was collected at 20Hz during 5 seconds for three orthogonal axes (D): sway – X

299   (red), surge – Y (green), and heave - Z (blue). (E) GPS-tracking of a griffon vulture over

300   one day. The colors of the GPS locations match the behaviors recorded on that location:

301   'Ground' – red; 'Soaring' – green; 'Flapping' – violet; 'Feeding' – yellow; 'Lying' – light

302   blue; 'Standing' – black. This illustrates the large daily movements of griffon vultures,

303   emphasizing the logistical challenges associated with surveillance in the desert study

304   area. Photo credit: Yacov Ben Bunan.

305

306   _Pre-processing the ACC data and model training_

307   Before deployment on the griffons, 50 transmitters were calibrated on a leveled surface,

308   in all six possible perpendicular orientations. This calibration allowed us to obtain a

309   transmitter-specific instrument error for translating raw acceleration data (in mV) into

310   acceleration units (m/sec^2). For 14 transmitters (out of 32) without specific error values,

311   we used the average error across the measured transmitters (n = 50). The calibration

312   values used are publicly available on GitHub.

313   We identified the start and end of each accelerometer bout and excluded from

314   the ACC behavioral dataset all bouts shorter than 5 sec, as well as all bouts that matched

315   more than one behavioral class during the 5 sec period. Each acceleration bout was

316   summarized into 47 statistical features commonly used in other studies using machine

317   learning algorithms to perform behavioral classifications of ACC data (e.g., Nathan et al.,

318   2012; Yu et al., 2021). For a full list of features, see Supplementary Table S1. All

319   analyses were performed in R (R Core Team, 2023).

320   Using the R packages _ranger_ (Wright & Ziegler, 2017) and _parsnip_ (Kuhn &

321   Vaughan, 2024), we built a random forest model to classify behaviors using the

322   annotated acceleration bouts. We started by splitting this dataset into 'training' (67%)

323   and 'testing' (33%) subsets, an _ad hoc_ measure commonly found in other machine-

324   learning applications (e.g., (Jeantet et al., 2020). Using the 'training subset', we built a

325   random forest model and we evaluated the performance of our model using the 'testing'

326   subset. We built a confusion matrix and calculated three performance metrics, for the full

327   model and for each behavior: (*i*) accuracy; (*ii*) precision; and (*iii*) recall. The equations

328   and descriptions for each metric can be found in Table 1. For example, a model may

329   have 0.90 accuracy (i.e., 90% of all behaviors were predicted correctly), 0.85 precision

330   for a specific behavior (e.g., 85% of all 'Feeding' predictions were indeed 'Feeding'

331   observations and 15% were a different behavior and wrongly identified as 'Feeding'),

332   and 0.80 recall of a specific behavior (e.g., 80% of 'Feeding' observations were correctly

333   predicted as 'Feeding' and 20% were wrongly classified as another behavior).

334   After training and evaluating the performance of the algorithm with the split

335   annotated dataset, we built the final random forest model using the full dataset for

336   training, likely improving the performance of the algorithm. This full algorithm was then

337   used to classify unobserved accelerometer bouts to identify feeding in free roaming

338   vultures (see below).

339   More details of the model building sequence can be found in Supplementary

340   Material 2, and a full description of the model building process can be found in Figure 2.

341   All the training data and the code necessary to train and build the algorithm are publicly

342   available on Zenodo (Acácio et al., 2025) and GitHub. The repository includes a tutorial

343   suitable for two types of users: those who may wish to apply our (already-trained) model

344   to their own data (e.g., researchers and conservationists working on similar vulture/raptor

345   species), and those wishing to use our pipeline for training and building their own model

346   (e.g., researchers and conservationists working on other species, or with different

347   sampling protocols). With these tutorials, our main goal is to bridge the gap between

348   researchers and practitioners.

349

350 **Table 1** – Performance metrics used to evaluate the random forest model performance,

351 considering the true-positive (TP), true-negative (TN), false-positive (FP) and false-

352 negative (FN) predictions.

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

353

354

355 *Calculating confidence scores to validate model predictions*

356 Using the training dataset, we calculated a confidence score for each behavioral

357 classification (i.e., for every bout). This confidence score is the level of consensus

358 amongst the different decision trees within the random forest (i.e., the proportion of trees

359 that agree on the highest scoring prediction). For example, if the model classifies a given

360 bout as 'Feeding' with a confidence score of 0.7, then 70% of the trees agreed on that

361 classification. To determine the validity of this score as an indicator of the behavioral

362 classification's reliability, we compared the scores of correctly identified behaviors (true-

363 positives) and of incorrectly identified behaviors (false-positives) in the testing subset.

364 We then used a generalized linear mixed model (GLMM) with an ordered beta

365 distribution and a logit link to compare scores of the two groups. The confidence score

366 (range 0 to 1) was the response variable, and the explanatory variables were the Boolean

367 correctness of the model prediction (categorical; true-positive or false-positive), the

368 predicted behavior (categorical), and their interaction. The model included device ID as

369 a random intercept. The GLMM was built using *glmmTMB* R package (Brooks et al.,

15

370   2017), and the fit of the model and residuals were evaluated using *DHARMa* R package

371   (Hartig, 2022).

372   To understand the effect of the harness configuration (backpack or leg-loop) on

373   the confidence scores, we compared the confidence scores of true-positive and false-

374   positive classifications of behaviors recorded with the two different harnesses. We

375   performed this comparison for the two behaviors with most observations: standing and

376   feeding. We built two separate GLMMs for each behavior. Each GLMM included the

377   confidence score as a response variable, and the Boolean correctness of the model

378   prediction (true-positive or false-positive), the predicted behavior, and their interaction,

379   as explanatory variables. We also included device ID as a random intercept. To further

380   explore the influence of harness type on the behavioral classification, we trained a new

381   random forest model using only the leg-loop data (n = 3428) and tested it on the

382   backpack dataset (n = 714). This model was trained on a subset of bouts that included

383   only the three behavioral categories present in both datasets ('Standing', 'Ground' and
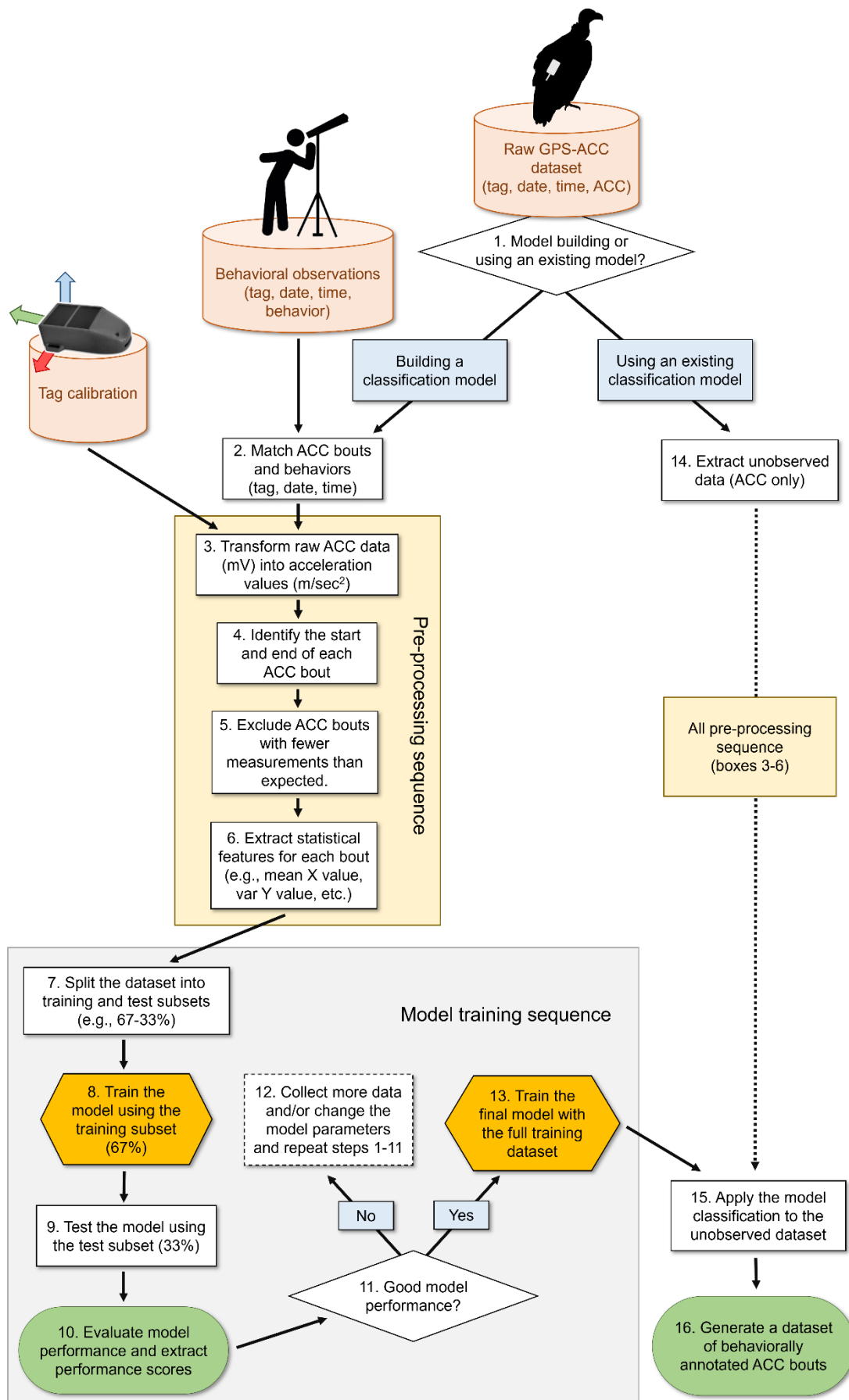
384   'Feeding').

385

386

387 **Figure 2 - Methodological workflow outlining the process of identifying behaviors**

388 **of wild animals based on accelerometer data.** All the training data, code and tutorials

389 necessary to follow this workflow are available on GitHub

390 (www.github.com/Orrslab/ACC_behavior_classification) and at Acácio et al., 2025.

391

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

393 To assess the reliability of our algorithms at classifying unobserved data, we validated

394 predicted 'Feeding' behaviors of free-roaming vultures. We focused on this behavior due

395 to its importance for identifying poisoning events, the main cause of vulture mortality in

396 our study area (Anglister et al., 2023). This validation is important because even a highly

397 accurate model introduces classification errors. For example, considering a transmitter

398 collecting 72 bouts a day, with 10 of those classified as 'Feeding', and a model with 90%

399 precision for 'Feeding'. Over the course of one week, the device would collect 504 bouts,

400 70 of which classified as 'Feeding'. Considering the model's precision, 7 of these

401 'Feeding' classifications would be false-positives which, extrapolating for a population of

402 50vultures, this would correspond to approximately 350 false-positive feeding predictions

403 per week.

404　　We combined information about the location of supplementary feeding stations,

405 satellite imagery, and GPS positions from griffon-borne transmitters to assess the

406 likelihood that the unobserved vultures' ACC-predicted 'Feeding' behavior represents a

407 true feeding event. Between November-December 2020, we collected GPS and

408 accelerometer data from 7 tagged free-roaming griffons in southern Israel (transmitter

409 schedule described in the Supplementary Material). These individuals were selected

410 because they provided consistent high-resolution data throughout this two-month period,

411 making them suitable for the fine-scale analysis of feeding behavior. We matched a GPS

412 position to an accelerometer bout if they were recorded within 5 min of each other. We

413   designated four situations with decreasing probability of representing real feeding events

414   based on the GPS location and GPS ground speed: 'Station' – if the ACC identified a

415   feeding event within 250m from a supplementary feeding station, it likely represents a

416   true feeding event (likely true-positive predictions); 'Open area' – if the ACC identified a

417   feeding event at a GPS position that is at an open landscape (but not on a cliff), where

418   naturally-occurring food is sometimes available (wildlife or livestock carcasses), it may

419   represent a true feeding event ; 'Cliff' – if the ACC identified a feeding event on cliff faces,

420   where food is largely absent but where vultures spend a large proportion of their time

421   roosting, it is not likely to be a true feeding event; 'Flight' – if the GPS ground speed was

422   >4m/sec the vulture was probably flying, and therefore it is likely a false feeding event. .

423   The 250 m radius around the feeding station accounted for the vultures' behavior of

424   standing nearby and overlooking the station before feeding. Given the potential 5-minute

425   offset between GPS fixes and accelerometer bouts, this buffer allows for the possibility

426   that a griffon could move to the station and begin feeding within that time window. All

427   points were mapped on satellite images of the study area and were visually examined

428   after this classification to confirm the assignment to each situation (for example, to

429   confirm vulture presence on a cliff, in an open area or near a feeding station).

430   To determine if the confidence score of the classification can be used to identify

431   false-positives in free-roaming griffons, we compared the algorithm's confidence scores

432   of 'Feeding' predictions at 'Stations' (i.e., high probability of true-positives) with 'Feeding'

433   predictions at 'Cliffs' or during 'Flight' (i.e., high probability of false-positives). We omitted

434   the 'Open area' situation since it could represent a mixture of feeding and non-feeding

435   behaviors and were therefore less conclusive for this comparison. We used a GLMM,

436   with a beta distribution and a logit link, in which the response variable was the algorithm's

437   confidence score, and the explanatory variable was the classification accuracy according

438   to the GPS location (likely true-positive or likely false-positive). We included device ID

439   as a random intercept.

440

Case study: identification of vulture feeding hotspots to prevent poisoning

442 To demonstrate the applicability of the ACC algorithm to a real-world conservation

443 problem, we used it to identify griffons' feeding hotspots outside supplementary feeding

444 stations (i.e., places where safe carcasses are provided to vultures). Considering the

445 high risk of carcasses outside feeding stations being contaminated with toxic substances

446 for vultures (e.g., pesticides or NSAIDs, Anglister et al., 2023), their rapid detection and

447 removal from the field is a priority for wildlife authorities in Israel (Acácio et al., 2023).

448 Accordingly, mapping those areas where vultures are feeding on potentially

449 contaminated carcasses may guide management actions.

450 In November 2022, we collected one month of GPS and accelerometer data for

451 51 free-roaming griffons in Southern Israel, aiming to identify the locations of feeding

452 events that occurred outside supplementary feeding stations (events that present higher

453 risk of poisoning). After applying the random forest algorithm to this dataset, we matched

454 the accelerometer 'Feeding' bouts with a GPS location using three criteria. First, if they

455 were collected within 5 min of each other, and if the GPS ground speed was below

456 4m/sec (indicating the bird was not flying). Second, if no GPS position matched these

457 criteria, we matched ACC bouts with GPS locations if they were collected within 11 min

458 of each other (while maintaining the ground speed criteria), to account for a possible

459 delay in the time to acquire a position by the GPS. If no GPS position matched these

460 criteria, the 'Feeding' bout was discarded from further analysis, because we could not

461 infer where the feeding event took place.

462 Using the results of the previous analyses, where we assessed if the confidence

463 score could be used to minimize the number of false-positives, we excluded bouts with

464 confidence scores below 0.5. This conservative threshold was chosen to avoid

465 eliminating true-positives, as failing to detect feeding areas posed a greater risk for

20

466   griffon conservation than including false-positive observations. However, this threshold

467   is system- and data-specific and is expected to be different for other species and

468   systems. We also excluded 'Feeding' bouts that occurred within supplementary feeding

469   stations, at known roost sites (the latter likely represent false-positives), and outside the

470   study area (southern Israel and Jordan). With the remaining locations (n = 264), we

471   created a 2D kernel (grid size = 1000, bandwidth = bandwidth.nrd function from MASS

472   R package, Venables & Ripley, 2002), portraying the density of locations, using *bkde2D*

473   function of *KernSmooth* R package (Wand, 2024). On this density map, we overlaid

474   the information of known carcasses independently identified in the field, outside feeding

475   stations, during this same time period (n = 5). The carcasses were located by local

476   rangers, either through reports from farmers, chance encounters during field patrols, or

477   via an alert system that flags unusual landings of tracked raptors based on GPS data.

478   This system uses location data from several species to identify potential poisoning

479   events, independent of accelerometer-based behavioral classifications used in this

480   study.

481   Finally, to assess if the behavioral classification impacted the designation of

482   feeding hotspots compared to a mapping based on GPS metrics alone, we built an

483   additional density map using all GPS locations from the same dataset, without filtering

484   for "feeding-only" locations. We excluded locations that occurred within supplementary

485   feeding stations, at known roost sites, or where the ground speed exceeded >4m/sec.

486

487   ## Results

488   *Behavioral classification*

489   We collected 5783 behavioral observations for 14 captive and 17 free-roaming griffon

490   vultures (a total of 31 individuals), during 57 days (18 days for captive individuals and 39

491   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).

Observed behavior

	Feeding	Lying	Standing	Ground	Flapping	Soaring
Feeding	189	1	13	13	2	0
Lying	0	104	5	1	0	1
Standing	15	2	1142	8	0	3
Ground	2	0	1	4	0	0
Flapping	0	0	0	1	42	0
Soaring	0	2	1	0	0	375

Precision
1.00
0.75
0.50

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

22

508 of false-positives. For example, more 'Feeding' bouts were wrongly classified as

509 'Standing' than as 'Ground', and none of the 'Feeding' bouts were wrongly classified as

510 'Soaring', 'Flapping', or 'Lying'.

511

512 *Performance of the confidence score in validating model predictions*

513 Overall, the confidence scores of correctly-identified behaviors (true-positives) were

514 significantly higher than the scores of inaccurately-identified behaviors (false-positives;

515 GLMM: model estimate ± se = 0.876 ± 0.195, p-value < 0.001, Figure 4, Supplementary

516 Table S3). 'Ground' behaviors, which had the lowest number of observations (n=67),

517 were the exception, with significantly higher confidence scores of false-positives

518 compared to true-positives (Figure 4). 'Ground' true-positive behaviors also had the

519 lowest confidence scores (mean confidence score ± sd = 0.41 ± 0.03). 'Soaring', 'Lying',

520 'Standing' and 'Flapping' were the behaviors with highest true-positive confidence scores

521 (mean confidence score ± sd; 'Soaring' = 0.99 ± 0.08, 'Lying' = 0.98 ± 0.06, 'Standing' =

522 0.98 ± 0.08, 'Flapping' = 0.93 ± 0.12). 'Feeding' had on average a high confidence score

523 but also a large variation (mean confidence score ± sd = 0.82 ± 0.15, Figure 4,

524 Supplementary Table S2, Supplementary Table S3).

525 There were no significant differences between the confidence scores of

526 behaviors recorded with backpack or leg-loop harnesses for the two tested behaviors:

527 Standing' (GLMM: estimate ± se = 0.374 ± 0.283, p-value = 0.283) and 'Feeding' (GLMM:

528 estimate ± se = -0.725 ± 0.798, p-value = 0.364, Supplementary Figure S1 and

529 Supplementary Table S4). The model trained on leg-loop data and tested on backpack

530 data had high overall accuracy (accuracy = 0.86) and performed well at classifying

531 'Standing' behaviors ('Standing' precision = 0.99; recall = 0.91). All true 'Feeding'

532 behaviors were correctly identified as such ('Feeding' recall = 1). However, most

533 'Ground' behaviors were misclassified as 'Feeding', which reduced the precision of the

534   'Feeding' category ('Feeding' precision = 0.52). Similarly to the full model, 'Ground'

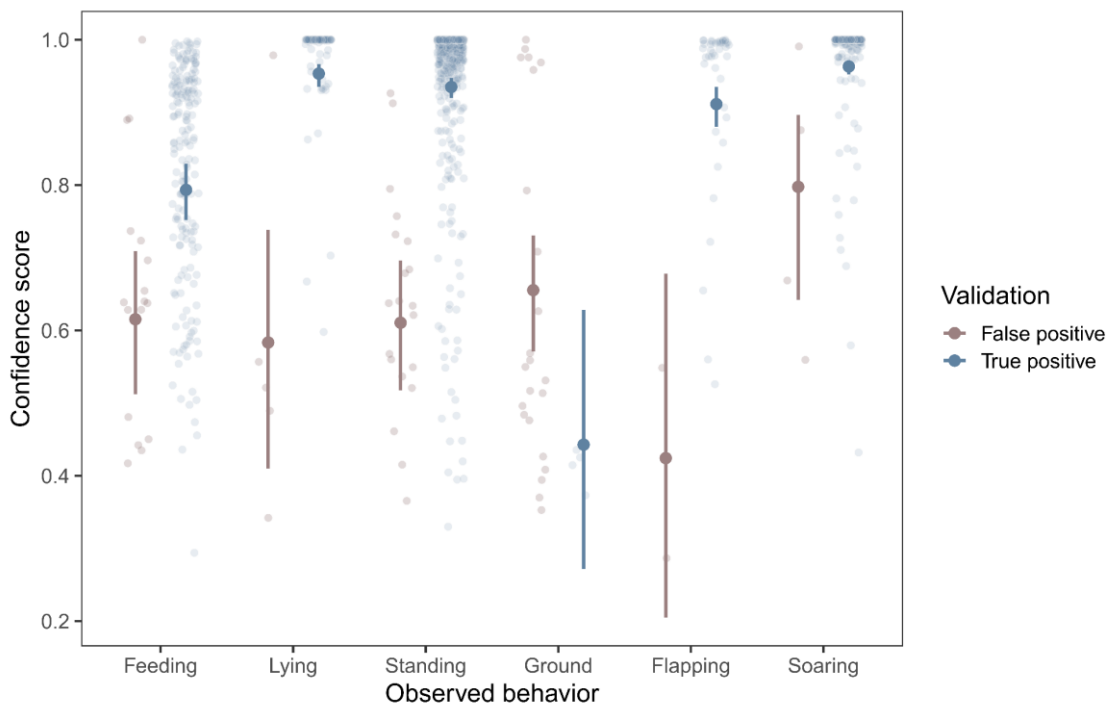535   behaviors had the poorest performance (Supplementary Table S5).

536

537

538   **Figure 4 – Model predicted confidence scores of true-positive (blue) and false-**

539   **positive (dark red) behavioral classifications.** The darker points and error bars

540   indicate the model predicted confidence scores and 95% confidence intervals for true-

541   positive and false-positive behavioral classifications. The lighter points show the raw

542   data.

543

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

545   We used the GPS locations to validate 175 'Feeding' bouts from 7 free-roaming vultures

546   in Southern Israel. Overall, 126 'Feeding' bouts (72%) occurred within a supplementary

547   feeding station ('Station' – likely true-positives), 22 bouts (13%) were located on 'Open

548   areas' (likely a mix of true- and false-positives), 20 bouts (11%) were on 'Cliffs' (likely

549   false-positives), and 7 bouts (4%) were in 'Flight' (likely false-positives, Figure 5). The

550 relatively high proportion of feeding bouts identified on cliffs likely reflects the

551 considerable amount of time griffons spend in these areas. Overall, of all 'Feeding' bouts

552 identified by the algorithm, 72-85% (all 'Station' bouts + at least part of the 'Open area'

553 bouts) were likely real feeding events. Furthermore, after removing 'Cliff' and 'Flight'

554 bouts (easily identifiable using only the GPS location, satellite imagery, and ground

555 speed), 85-100% of the 'Feeding' predictions (all 'Station' bouts + at least part of the

556 'Open area' bouts) were indeed likely feeding events.

557 Importantly, the confidence scores of 'Feeding' bouts likely to be true-positives

558 were higher (mean ± sd: 'Station' = 0.75 ± 0.16) than the scores of bouts likely to be

559 false-positives ('Cliff' and 'Flight' = 0.56 ± 0.19). This comparison was statistically

560 significant (GLMM: estimate ± se = -0.805 ± 0.167, p-value <0.001, Supplementary Table

561 S5). When considering solely the 'Feeding' bouts with a confidence score over 0.5, 114

562 bouts (88.4%) occurred within a supplementary feeding station and were likely true-

563 positives. This threshold maximizes the number of true-positive predictions, while

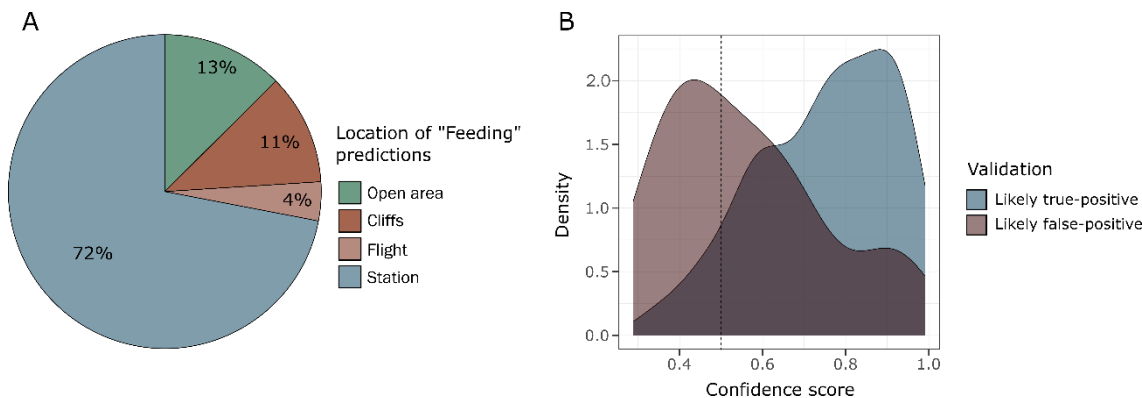564 minimizing the number of false-positives (Figure 5).

565

566

567 **Figure 5 - Validation of 'Feeding' behaviors using data from free-roaming griffons.**

568 A - Percentage of 'Feeding' predictions (n = 175) located within a supplementary feeding

569 station ('Station' - blue), on open landscape ('Open area' - green), on cliffs ('Cliffs' - dark

570 brown) or in flight ('Flight' - light brown). B - Distribution of the confidence scores of

571 'Feeding' bouts likely to be true-positives (located within a feeding station, in blue) and

572 likely to be false-positives (located on cliffs or in flight, in dark red). The dashed line

573 indicates the confidence threshold of 0.5, a conservative threshold that reduces the

574 number of false-positives, while including nearly all true-positive predictions.

575

576 *Mapping vulture's feeding hotspots to facilitate poisoning identification*

577 In November 2022, we collected 4595 'Feeding' bouts of 51 griffon vultures in our study

578 area. After sequentially removing the bouts without a GPS location (n = 586), bouts inside

579 feeding stations (n = 2534), outside Southern Israel and Jordan (n = 157), bouts located

580 in known roosts (n = 979), and bouts with a confidence score below 0.5 (n = 60, Figure

581 5), we retained 264 bouts of 31 vultures, that allowed us to map their feeding hotspots.

582 We built a KDE with the remaining 264 'Feeding' bouts and detected a hotspot of

583 feeding events in the Judean desert. This hotspot is consistent with the INPA ranger's

584 reports for this same period, where 4 out of 5 reported carcasses outside feeding stations

585 were within the KDE (Figure 6). This density map differed substantially from the one

586 based solely on GPS-derived metrics (Supplementary Figure S2), containing 1,938

587 potential feeding locations. Notably, the GPS-only KDE failed to identify a key feeding

588 hotspot in Jordan. This confirms that the algorithm can be used to identify areas with

589 high probability of vultures' feeding on potentially contaminated carcasses and highlights

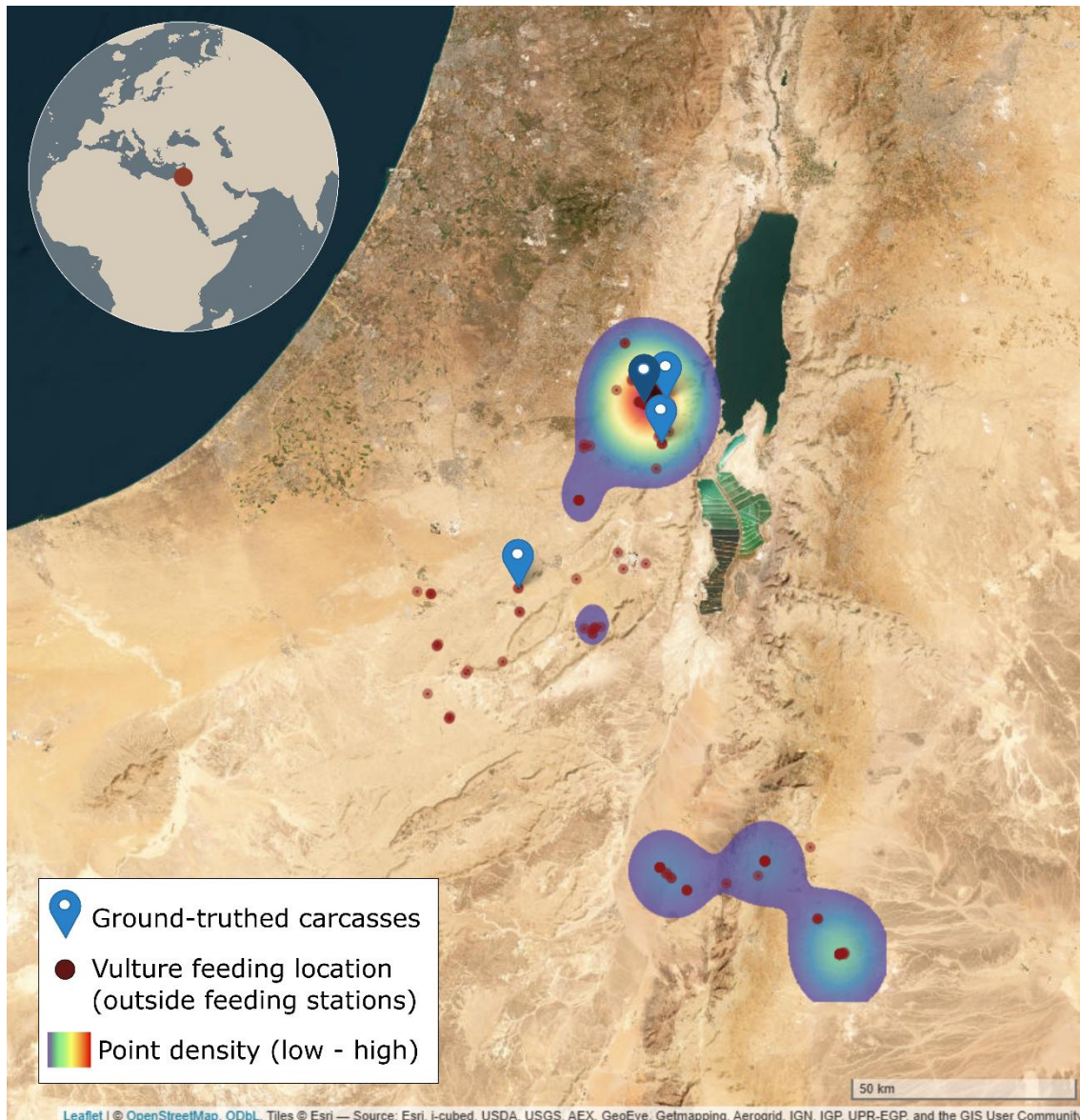590 the added value of the behavioral classification.

591

592

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

601 areas having lower density and red areas higher density. The inset shows the location

602 of the study area in the world.

603

604 Discussion

605 Recent advancements in tracking technology and analytical tools are enhancing our

606 understanding of animal ecology and behavior and improving its applications for

607 biodiversity conservation (Tuia et al., 2022; Williams et al., 2020). In this study, we add

608 to this body of literature by developing a machine learning algorithm to classify griffon

609 vultures' behaviors, thoroughly validating the behavioral classifications, and using them

610 to inform conservation efforts, namely carcass detection to prevent poisoning. Our model

611 accurately predicted griffons' behaviors, allowing for the identification of potential feeding

612 events outside feeding stations and the mapping of feeding hotspots where vultures and

613 other scavengers can engage in risky behaviors such as the consumption of poisoned

614 carcasses (Peters et al., 2023). These maps may become fundamental tools for

615 monitoring-effort prioritization and for optimizing on-the-ground actions for the

616 conservation of vultures and other scavengers (e.g., the detection of poisoning events,

617 Rast et al., 2024). Another major contribution of this study is the use and validation of

618 the algorithm's confidence in each behavioral classification, showing the utility of this

619 approach for other behaviors and contexts. Assessing the degree of confidence in this

620 manner is rarely done in ecological studies (Bidder et al., 2014), but we highlight how

621 this approach may minimize misclassifications (e.g., false-positives) especially when

622 resources for ground-truthing are limited. Our algorithm and training dataset are made

623 accessible to other researchers and conservationists studying vultures and similar

624 species. Moreover, they can be easily adapted to classify the behaviors of other species

625 in diverse study systems. Importantly, to further promote this usability, we provide a

626 methodological workflow to guide potential users in the process of identifying behaviors

627 of wild animals based on accelerometer data.

628

629 <u>Accelerometer-based behavioral classification as a tool for vulture conservation</u>

630 With our behavioral classification model, we were able to identify vulture feeding hotspots

631 in Southern Israel. Indeed, the areas where our tracked vultures displayed 'Feeding'

632 behaviors matched the locations of known 'wild' carcasses (i.e., outside feeding stations)

633 during this same period. This case study used only a single month worth of high-

634 resolution data embedded within a long-term lower resolution tracking effort (Acácio et

635 al., 2024; Spiegel et al., 2013), but it exemplifies how GPS and accelerometer data can

636 be used to direct conservation efforts. The use of GPS tracking has been instrumental

637 for vulture conservation in Israel (Spiegel et al., 2013), particularly for the detection of

638 poisoning events (Acácio et al., 2023; Anglister et al., 2023). The local government

639 environmental agency, INPA, developed a near-real time alert system that warns rangers

640 whenever a vulture lands at a suspicious area and when vultures are either moving very

641 little or are suspected to be dead (Nemtzov et al., 2021). Rangers then actively respond

642 to these alerts by inspecting the area and removing the carcasses; therefore, reducing

643 the number of false alarms is important – both to reduce costs and workload, as well

644 avoiding erosion of rangers' responsiveness.

645 A similar near real-time alert system, using GPS data, has also been used for the

646 monitoring of African elephants (*Loxodonta Africana,* Wall et al., 2014) and to track

647 California condors (*Gymnogyps californianus*) in the vicinity of wind farms (Sheppard et

648 al., 2015). It has also been suggested as an anti-poaching tool to prevent the extinction

649 of large mammals (O'Donoghue & Rutz, 2016). We propose that all these systems could

650 be improved by using accelerometer data to remotely identify animal behavior and risky

651 events sooner and more reliably. Indeed, our results show that relying solely on GPS-

652   derived filters to identify potential feeding hotspots failed to identify a key feeding hotspot

653   in Jordan and produced over seven times more data points, many of which were likely

654   false positives. Such an overload of low-quality alerts could lead to reduced

655   responsiveness by the rangers, ultimately undermining conservation efforts on the

656   ground.

657   Combining maps of feeding hotspots (either fixed or season-specific ones) with

658   similar near real-time alert systems may be crucial for vulture management and

659   conservation. For example, the feeding areas that griffons use systematically throughout

660   the year should be prioritized in terms of surveillance and sanitation efforts, to prevent

661   vultures (and other scavengers) from accessing carcasses contaminated with toxic

662   substances. Additional management actions could be implemented, such as establishing

663   new supplementary feeding stations in these areas, or increasing carcass supply at

664   existing stations, either all year-round or during particular seasons, to match potential

665   seasonal changes in vulture's activity areas. Additionally, the hotspots could be used to

666   implement geofences where data collection and transmissions would be at higher

667   frequency. This increased resolution may be critical in poisoning events, where the actual

668   feeding may be quite fast (sometimes consuming a carcass within minutes) and vultures

669   may perish quickly, depending on the type and amount of toxic substance ingested.

670   Then, information regarding the griffon's location and behavior is obtained and

671   communicated faster: when a griffon lands in these areas and only if it feeds there (as

672   indicated by the ACC classification), an alert should be sent to the rangers for immediate

673   carcass inspection. While the system should also trigger alerts for any feeding events

674   detected outside feeding stations (to allow for carcass inspection and potential removal

675   to reduce the risk of poisoning), identifying risky hotspots can help optimize resource

676   allocation and prioritize conservation actions in high-risk areas.

677   Around the world, an increasing number of individuals of multiple vulture species

678   are being tracked with GPS-Accelerometer devices, showing that they roam

679   exceptionally large areas on their daily movements (Kane et al., 2022; Spiegel et al.,

680   2015). Considering that about 70% of vulture species are endangered (Ives et al., 2022;

681   Ogada et al., 2012; Plaza et al., 2019), surveillance systems that combine GPS tracking

682   with accelerometry may be a useful tool to improve management actions in their large

683   roaming areas to combat major threats such poisoning. For instance, such ACC-based

684   systems will enhance existing applications of GPS-tagged vultures to inform on-ground

685   actions against illegal wildlife persecution (Rast et al., 2024; Rodríguez-Pérez et al., in

686   press) or to improve regulations for carrion disposal to feed vultures and other wild

687   species (Mateo-Tomás et al., 2023). Future studies could apply our methodology to

688   publicly available GPS-ACC datasets to identify high-risk areas for vultures and guide

689   targeted conservation interventions at a larger geographic scale.

690   Our thoroughly validated training dataset can also be used to classify the

691   behavior of griffons in other populations, as well as other old and new-world vulture

692   species, particularly in Africa and Asia, where vulture populations continue to decrease.

693   For example, our algorithm and training dataset could be used to predict the behaviors

694   of endangered *Gyps* species in Africa and Asia (e.g., *Gyps africanus*, *Gyps coprotheres*,

695   *Gyps bengalensis*, amongst others), or even other vulture species (e.g., *Torgos* and

696   *Trigonoceps* species), considering their morphological and behavioral similarities with

697   the griffon vulture. The use of surrogate species to identify accelerometer-based

698   behaviors has been examined in other systems, with a variety of results. For example,

699   the behaviors of domestic dogs were good predictors of the behaviors of dingoes and

700   cheetahs (Campbell et al., 2013), but the behavior of domestic caprids did not predict

701   well the behavior of their wild counterparts (Dickinson et al., 2021). Therefore, we

702   recommend caution when using our trained model to classify the behavior of other

703   vulture species. In addition, our algorithm and modelling pipeline can be easily adapted

704   for other, not related, animal species, as long as researchers provide their own training

705   dataset for their study species.

706

<u>Validating the accuracy of predictions of unobserved behaviors</u>

708   Tri-axial accelerometers and classification algorithms have increasingly been used to

709   obtain fine-scale behavior of wild animals (Nathan et al., 2012; Resheff et al., 2014;

710   Wang, 2019; Yu et al., 2021). However, after training and testing the model on a validated

711   dataset, the model must classify unobserved and, sometimes, unknown behaviors. In

712   this case, the model then matches the unknown behavior with the best fitting known

713   acceleration signature, resulting in misclassifications (Glass et al., 2020). Most

714   ecological studies fail to acknowledge this limitation and do not provide a metric of how

715   likely a particular classification is to be true (Glass et al., 2020). Here we tackle this

716   methodological gap and calculate a confidence score, which allows us to distinguish

717   between true-positive and false-positive classifications. Our approach is computationally

718   simple to implement and does not require running more complex classification models.

719   In addition to the confidence scores, we used biologically relevant information to validate

720   observations classified as 'Feeding'. For this subset of data, 15% of the observations

721   were likely misclassifications because they occurred on cliffs (where there is no food in

722   our case) or in flight. Filtering out observations based on easily accessible metrics (here,

723   the topography, knowledge of the behavior of the species, spatial position, and GPS

724   ground speed) as well as any observations with a confidence score below a relevant

725   threshold, increases the accuracy of behavioral classifications.

726   Selecting filtering thresholds is always a balance between two types of errors.

727   Here, we considered a conservative threshold of 0.5 to distinguish true-positive and

728   false-positive 'Feeding' predictions of free-roaming griffons, at the risk of including some

729   false-positive predictions in our dataset (Type I error). However, in this case, the risk of

730   not including part of the true-positive predictions (Type II error) is higher than including

731   some false-positives; not including all true-positives could mean that some feeding

732   hotspots would not be identified, potentially compromising sanitary management and

733   overlooking potential feeding and poisoning events. We encourage other researchers to

734   use a similar approach whenever possible, combining confidence scores with ground-

735   truthing information, to improve the accuracy of their conclusions.

736   In addition, we note that different behaviors, or even the same behavior in

737   different species or with different accelerometer devices, may have different confidence

738   scores distributions. For example, with our dataset a threshold of 0.9 could have been

739   suitable for distinguishing flapping and soaring flight behaviors, to study, for instance,

740   flight biomechanics. Therefore, the threshold of confidence should consider the

741   underlying distribution of confidence scores for the behaviors in mind and should be

742   defined according to this and the study objectives, balancing the risks of data loss with

743   the costs of including false-positives in the dataset.

744   Finally, quantifying temporal correlations between behaviors could also help

745   improve model performance or assist in post-processing filtering of the classifications

746   (Supplementary Material 3, Supplementary Figure S3). For instance, it is possible to

747   combine a correlation matrix of the behaviors with the confidence scores. In our dataset,

748   'Feeding' is often followed by other 'Feeding' behaviors (Supplementary Figure S3C).

749   Thus, if a high-confidence 'Feeding' behavior is followed by a low-confidence 'Feeding'

750   prediction, the strong positive dependency between these two behaviors could support

751   treating the second 'Feeding' as likely true-positive. Future studies could also implement

752   more complex models that allow for the incorporation of the correlation matrix within the

753   model.

754

755   Challenges and considerations of accelerometer-based behavioral classification

756   Different tag placement and different attachment methods can greatly influence

757   accelerometer signatures, and consequently the behavioral classifications (Garde et al.,

758   2022). Nevertheless, our results show that our algorithm is reliable for more than one

759   attachment method, further increasing its usefulness. These non-significant differences

760   may result from the similarity in logger placement between the two attachment types

761   (about 3cm difference), as well as from the limited spinal flexibility of griffon vultures. Still,

762   the large confidence intervals in this comparison (due to the small sample size for

763   backpack harnesses), as well as difficulty of our algorithm in classifying 'Ground'

764   behaviors recorded with a backpack harness shows that this topic deserves further

765   investigation with larger sample sizes and with other species.

766   After building the random forest, it is crucial to validate its accuracy, precision,

767   and recall. In our study, the overall model was highly accurate, yet behaviors differed in

768   their precisions. As expected, 'Ground' was consistently the behavior with the poorest

769   predictions across all our validations. This behavioral category included several, quite

770   distinct, ground behaviors (like walking, running, hopping, etc.), in an attempt to account

771   for all possible behaviors a griffon may display and to minimize misclassifications at the

772   cost of losing accuracy (Glass et al., 2020). 'Feeding', on the other hand, was accurately

773   classified by our model, however despite a relatively large number of bouts in the training

774   dataset (n = 587), the confidence scores of this behavior had large variance. A possible

775   reasoning is that griffon's feeding behavior is highly complex and may include rapid shifts

776   between fighting, posturing (spreading the wings), as well as eating per se (Bosè &

777   Sarrazin, 2007) – all inseparable within a 5s timeframe. Including so many different

778   postures in a single behavioral category results in high variation of confidence scores.

779   In addition, the number of conspecifics within a feeding event may further

780   influence the behaviors that individuals display while foraging (Bosè et al., 2012),

781   increasing within-individual variability for both wild and captive vultures. To mitigate the

782   effect of within-individual variability in our training dataset, we ensured that multiple

783   captive individuals were feeding at the same carcass, to replicate the wild feeding

784   conditions. Finally, it is likely that individuals differ in their behavior while foraging (e.g.,

785   dominant vs subordinates, Bosè et al., 2012; Bosè & Sarrazin, 2007), emphasizing the

786   need to improve behavioral classification models and account for individual differences

787   in behavior (Kirchner et al., 2023). In general, we suggest that future models can improve

788   accuracy and precision by further splitting our six classes into subclasses that reflect

789   more homogenous elementary behaviors (e.g., pecking, tearing meat apart, fighting).

790   Merging 'Ground' and 'Feeding' categories could also potentially improve the model's

791   accuracy; however, since these two behaviors are not necessarily linked, this would

792   come at the cost of decreased resolution in detecting true feeding events, compromising

793   management and conservation applications. For specific applications focused solely on

794   identifying feeding activity, an alternative approach could involve merging all feeding and

795   all non-feeding behaviors in a binary classification. This could simplify the interpretation,

796   and we suggest that this option is worth exploring in a future study

797   Despite the potential of accelerometer-based behavioral classification, collecting

798   such large volumes of data can be costly, both in terms of data transmission and storage,

799   as well as in terms of device memory and battery (Hounslow et al., 2019). Short sampling

800   intervals (2-3 secs) at high resolution may reduce the probability of having multiple

801   behaviors within a single bout, but may drain batteries faster, which can result in

802   incomplete sampling designs and lower the device's lifespan. Integrating low-frequency

803   accelerometry with additional sensors (e.g., time-depth recorders for marine species,

804   Jeantet et al., 2020), may still effectively allow the study of animal behavior without

805   significantly increasing costs or reducing device longevity (Hounslow et al., 2019). In

806   addition, analyzing such large volumes of data can also be challenging, so we emphasize

807   the need for collaboration in between fields of knowledge, with ecologists and data

808   scientists working together for the conservation of biodiversity (Tuia et al., 2022). Lastly,

809   as human activities are increasingly impacting the planet and driving species towards

810   extinction, it is critical to harness technological advances for effective conservation and

811   to safeguard the future of our planet's species and ecosystems.

812

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.

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.

36

850

Data accessibility

852   All the training data, the code necessary to train and build the algorithm and a tutorial

853   are publicly on Zenodo Repository https://doi.org/10.5281/zenodo.16152589 (Acácio et

854   al., 2025) and on GitHub (www.github.com/Orrslab/ACC_behavior_classification).

855

Author contributions

857   Gideon Vaadia, Orr Spiegel, Moni Shahar and Marta Acácio conceived the ideas and

858   designed the methodology. Gideon Vaadia, Orr Spiegel, Nili Anglister, Yigal Miller,

859   Ohad Hatzofe collected the data. Patricia Mateo-Tomás and María Fernández-García

860   tagged the Spanish vultures, whose behaviors were recorded by Patricia Mateo-

861   Tomás, Jorge Rodríguez-Pérez, María Fernandéz-García, Ignacio Otero and Pedro P.

862   Olea, and processed by Patricia Mateo-Tomás and Jorge Rodríguez-Pérez. Gideon

863   Vaadia, Tal Agassi and Marta Acácio analyzed the data. Marta Acácio and Gideon

864   Vaadia wrote the first version of the manuscript, with critical input from Noa Pinter-

865   Wollman and Orr Spiegel. All authors contributed critically to drafts of the manuscript

866   and gave approval for publication.

867

868

869 References

870 Acácio, M., Anglister, N., Vaadia, G., Harel, R., Nathan, R., Hatzofe, O., & Spiegel, O.
871   (2023). A lifetime track of a griffon vulture: The moving story of Rehovot (Y64).
872   *Ecology*, *104*(4), e3985. https://doi.org/10.1002/ecy.3985
873 Acácio, M., Gahm, K., Anglister, N., Vaadia, G., Hatzofe, O., Harel, R., Efrat, R.,
874   Nathan, R., Pinter-Wollman, N., & Spiegel, O. (2024). Behavioral plasticity
875   shapes population aging patterns in a long-lived avian scavenger. *Proceedings*
876   *of the National Academy of Sciences*, *121*(35), e2407298121.
877   https://doi.org/10.1073/pnas.2407298121
878 Acácio, M., Vaadia, G., Agassi, T., Anglister, N., Miller, Y., Hatzofe, O., Mateo-Tomás,
879   P., Rodríguez Pérez, J., Fernández-García, M., Olea, P. P., Otero, I., Pinter-
880   Wollman, N., Shahar, M., & Spiegel, O. (2025). *Data for: "Using accelerometer-*
881   *based behavioral classification to enhance scavenger conservation"* [Dataset].
882   Zenodo. https://doi.org/10.5281/zenodo.16152589
883 Anglister, N., Gonen-Shalom, S., Shlanger, P., Blotnick-Rubin, E., Rosenzweig, A.,
884   Horowitz, I., Hatzofe, O., King, R., Anglister, L., & Spiegel, O. (2023). Plasma
885   cholinesterase activity: A benchmark for rapid detection of pesticide poisoning
886   in an avian scavenger. *Science of The Total Environment*, *877*, 162903.
887   https://doi.org/10.1016/j.scitotenv.2023.162903
888 Bäckman, J., Andersson, A., Pedersen, L., Sjöberg, S., Tøttrup, A. P., & Alerstam, T.
889   (2017). Actogram analysis of free-flying migratory birds: New perspectives
890   based on acceleration logging. *Journal of Comparative Physiology A*, *203*(6),
891   543–564. https://doi.org/10.1007/s00359-017-1165-9
892 Bidder, O. R., Campbell, H. A., Gómez-Laich, A., Urgé, P., Walker, J., Cai, Y., Gao, L.,
893   Quintana, F., & Wilson, R. P. (2014). Love Thy Neighbour: Automatic Animal
894   Behavioural Classification of Acceleration Data Using the K-Nearest Neighbour
895   Algorithm. *PLOS ONE*, *9*(2), e88609.
896   https://doi.org/10.1371/journal.pone.0088609
897 BirdLife Israel. (2022). *BirdLife Israel live nest cameras* [Computer software].
898   https://www.birds.org.il/en
899 Bosè, M., Duriez, O., & Sarrazin, F. (2012). Intra-specific competition in foraging
900   Griffon Vultures Gyps fulvus: 1. Dynamics of group feeding. *Bird Study*, *59*(2),
901   182–192. https://doi.org/10.1080/00063657.2012.658639
902 Bosè, M., & Sarrazin, F. (2007). Competitive behaviour and feeding rate in a
903   reintroduced population of Griffon Vultures Gyps fulvus. *Ibis*, *149*(3), 490–501.
904   https://doi.org/10.1111/j.1474-919X.2007.00674.x
905 Brooks, M., Kristensen, K., van Benthem, K., Magnusson, A., Berg, C., Nielsen, A.,
906   Skaug, H., Maechler, M., & Bolker, B. (2017). glmmTMB Balances Speed and
907   Flexibility Among Packages for Zero-inflated Generalized Linear Mixed
908   Modeling. *The R Journal*, *9*(2), 378–400. https://doi.org/doi:10.32614/RJ-2017-
909   066
910 Brown, D. D., Kays, R., Wikelski, M., Wilson, R., & Klimley, A. P. (2013). Observing the
911   unwatchable through acceleration logging of animal behavior. *Animal*
912   *Biotelemetry*, *1*(1), 20. https://doi.org/10.1186/2050-3385-1-20
913 Buechley, E. R., & Şekercioğlu, Ç. H. (2016). The avian scavenger crisis: Looming
914   extinctions, trophic cascades, and loss of critical ecosystem functions.
915   *Biological Conservation*, *198*, 220–228.
916   https://doi.org/10.1016/j.biocon.2016.04.001
917 Campbell, H. A., Gao, L., Bidder, O. R., Hunter, J., & Franklin, C. E. (2013). Creating a
918   behavioural classification module for acceleration data: Using a captive
919   surrogate for difficult to observe species. *Journal of Experimental Biology*,
920   *216*(24), 4501–4506. https://doi.org/10.1242/jeb.089805

921 Chimienti, M., Cornulier, T., Owen, E., Bolton, M., Davies, I. M., Travis, J. M. J., &
922   Scott, B. E. (2016). The use of an unsupervised learning approach for
923   characterizing latent behaviors in accelerometer data. *Ecology and Evolution*,
924   *6*(3), 727–741. https://doi.org/10.1002/ece3.1914
925 Clarke, T. M., Whitmarsh, S. K., Hounslow, J. L., Gleiss, A. C., Payne, N. L., &
926   Huveneers, C. (2021). Using tri-axial accelerometer loggers to identify
927   spawning behaviours of large pelagic fish. *Movement Ecology*, *9*(1), 26.
928   https://doi.org/10.1186/s40462-021-00248-8
929 Cutler, D. R., Edwards Jr., T. C., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J., &
930   Lawler, J. J. (2007). Random Forests for Classification in Ecology. *Ecology*,
931   *88*(11), 2783–2792. https://doi.org/10.1890/07-0539.1
932 Dickinson, E. R., Twining, J. P., Wilson, R., Stephens, P. A., Westander, J., Marks, N.,
933   & Scantlebury, D. M. (2021). Limitations of using surrogates for behaviour
934   classification of accelerometer data: Refining methods using random forest
935   models in Caprids. *Movement Ecology*, *9*(1), 28.
936   https://doi.org/10.1186/s40462-021-00265-7
937 Efrat, R., Hatzofe, O., Miller, Y., & Berger-Tal, O. (2020). Determinants of survival in
938   captive-bred Griffon Vultures Gyps fulvus after their release to the wild.
939   *Conservation Science and Practice*, *2*(12), e308.
940   https://doi.org/10.1111/csp2.308
941 Fehlmann, G., O'Riain, M. J., Kerr-Smith, C., Hailes, S., Holton, M., Hopkins, P., &
942   King, A. J. (2023). Using behavioral studies to adapt management decisions
943   and reduce negative interactions between humans and baboons in Cape Town,
944   South Africa. *Conservation Science and Practice*, *5*(7), e12948.
945   https://doi.org/10.1111/csp2.12948
946 Garde, B., Wilson, R. P., Fell, A., Cole, N., Tatayah, V., Holton, M. D., Rose, K. A. R.,
947   Metcalfe, R. S., Robotka, H., Wikelski, M., Tremblay, F., Whelan, S., Elliott, K.
948   H., & Shepard, E. L. C. (2022). Ecological inference using data from
949   accelerometers needs careful protocols. *Methods in Ecology and Evolution*,
950   *13*(4), 813–825. https://doi.org/10.1111/2041-210X.13804
951 Glass, T. W., Breed, G. A., Robards, M. D., Williams, C. T., & Kielland, K. (2020).
952   Accounting for unknown behaviors of free-living animals in accelerometer-
953   based classification models: Demonstration on a wide-ranging mesopredator.
954   *Ecological Informatics*, *60*, 101152. https://doi.org/10.1016/j.ecoinf.2020.101152
955 Gleiss, A. C., Wilson, R. P., & Shepard, E. L. C. (2011). Making overall dynamic body
956   acceleration work: On the theory of acceleration as a proxy for energy
957   expenditure. *Methods in Ecology and Evolution*, *2*(1), 23–33.
958   https://doi.org/10.1111/j.2041-210X.2010.00057.x
959 Halsey, L. G., Green, J. A., Wilson, R. P., & Frappell, P. B. (2009). Accelerometry to
960   Estimate Energy Expenditure during Activity: Best Practice with Data Loggers.
961   *Physiological and Biochemical Zoology*, *82*(4), 396–404.
962   https://doi.org/10.1086/589815
963 Hartig, F. (2022). *_DHARMa: Residual Diagnostics for Hierarchical (Multi-Level /*
964   *Mixed) Regression Models_* (Version 0.4.6) [R Package]. https://CRAN.R-
965   project.org/package=DHARMa
966 Hatzofe, O. (2020). *Summary of the griffon vulture counts in Israel* (pp. 1–5) [Internal
967   report]. Israel Nature and Parks Authority.
968 Hays, G. C., Bailey, H., Bograd, S. J., Bowen, W. D., Campagna, C., Carmichael, R.
969   H., Casale, P., Chiaradia, A., Costa, D. P., Cuevas, E., Bruyn, P. J. N. de, Dias,
970   M. P., Duarte, C. M., Dunn, D. C., Dutton, P. H., Esteban, N., Friedlaender, A.,
971   Goetz, K. T., Godley, B. J., … Sequeira, A. M. M. (2019). Translating Marine
972   Animal Tracking Data into Conservation Policy and Management. *Trends in*
973   *Ecology & Evolution*, *34*(5), 459–473. https://doi.org/10.1016/j.tree.2019.01.009
974 Hounslow, J. L., Brewster, L. R., Lear, K. O., Guttridge, T. L., Daly, R., Whitney, N. M.,
975   & 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.,

1031   Campos-Candela, A., Cooke, S. J., Gjelland, K. Ø., Gupte, P. R., Harel, R.,
1032   Hellström, G., Jeltsch, F., … Jarić, I. (2022). Big-data approaches lead to an
1033   increased understanding of the ecology of animal movement. *Science*,
1034   *375*(6582), eabg1780. https://doi.org/10.1126/science.abg1780

1035 Nathan, R., Spiegel, O., Fortmann-Roe, S., Harel, R., Wikelski, M., & Getz, W. M.
1036   (2012). Using tri-axial acceleration data to identify behavioral modes of free-
1037   ranging animals: General concepts and tools illustrated for griffon vultures.
1038   *Journal of Experimental Biology*, *215*(6), 986–996.
1039   https://doi.org/10.1242/jeb.058602

1040 Nemtzov, S., Hatzofe, O., Steinitz, O., & Vine, G. (2021). *An Innovative Automatic*
1041   *Location-Based Real-Time Alert System to Prevent Wildlife Poisoning Using*
1042   *GPS-Tagged Griffon Vultures has Led to Better Conservation of Endangered*
1043   *Species in Protected Areas* [Computer software]. https://panorama.solutions/en/
1044   solution/innovative-automatic-location-based-real-time-alertsystem-prevent-
1045   wildlife-poisoning-using

1046 O'Donoghue, P., & Rutz, C. (2016). Real-time anti-poaching tags could help prevent
1047   imminent species extinctions. *Journal of Applied Ecology*, *53*(1), 5–10.
1048   https://doi.org/10.1111/1365-2664.12452

1049 Ogada, D., Keesing, F., & Virani, M. Z. (2012). Dropping dead: Causes and
1050   consequences of vulture population declines worldwide. *Annals of the New*
1051   *York Academy of Sciences*, *1249*(1), 57–71. https://doi.org/10.1111/j.1749-
1052   6632.2011.06293.x

1053 Ogada, D., Shaw, P., Beyers, R. L., Buij, R., Murn, C., Thiollay, J. M., Beale, C. M.,
1054   Holdo, R. M., Pomeroy, D., Baker, N., Krüger, S. C., Botha, A., Virani, M. Z.,
1055   Monadjem, A., & Sinclair, A. R. E. (2016). Another Continental Vulture Crisis:
1056   Africa's Vultures Collapsing toward Extinction. *Conservation Letters*, *9*(2), 89–
1057   97. https://doi.org/10.1111/conl.12182

1058 Olea, P. P., & Mateo-Tomás, P. (2014). Living in risky landscapes: Delineating
1059   management units in multithreat environments for effective species
1060   conservation. *Journal of Applied Ecology*, *51*(1), 42–52.
1061   https://doi.org/10.1111/1365-2664.12176

1062 Olea, P. P., Mateo-Tomás, P., & Sánchez-Zapata, J. A. (Eds.). (2019). *Carrion Ecology*
1063   *and Management* (Vol. 2). Springer International Publishing.
1064   https://doi.org/10.1007/978-3-030-16501-7

1065 Peters, N., Kendall, C. J., Davies, J. G., Bracebridge, C., Nicholas, A., Mgumba, M. P.,
1066   & Beale, C. M. (2023). Identifying priority locations to protect a wide-ranging
1067   endangered species. *Biological Conservation*, *277*, 109828.
1068   https://doi.org/10.1016/j.biocon.2022.109828

1069 Plaza, P. I., Martínez-López, E., & Lambertucci, S. A. (2019). The perfect threat:
1070   Pesticides and vultures. *Science of The Total Environment*, *687*, 1207–1218.
1071   https://doi.org/10.1016/j.scitotenv.2019.06.160

1072 R Core Team. (2023). *R: A language and environment for statistical computing*
1073   [Computer software]. R Foundation for Statistical Computing. https://www.R-
1074   project.org/

1075 Rast, W., Portas, R., Shatumbu, G. I., Berger, A., Cloete, C., Curk, T., Götz, T.,
1076   Aschenborn, O., & Melzheimer, J. (2024). Death detector: Using vultures as
1077   sentinels to detect carcasses by combining bio-logging and machine learning.
1078   *Journal of Applied Ecology*, *61*(12), 2936–2945. https://doi.org/10.1111/1365-
1079   2664.14810

1080 Resheff, Y. S., Bensch, H. M., Zöttl, M., Harel, R., Matsumoto-Oda, A., Crofoot, M. C.,
1081   Gomez, S., Börger, L., & Rotics, S. (2024). How to treat mixed behavior
1082   segments in supervised machine learning of behavioural modes from inertial
1083   measurement data. *Movement Ecology, 12*(1), 44.
1084   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

1140   carnivores and livestock. *Conservation Biology*, *32*(1), 26–34.
1141   https://doi.org/10.1111/cobi.12959
1142 Venables, W., & Ripley, B. (2002). *Modern Applied Statistics with S* (4th Edition).
1143   Springer.
1144 Venter, O., Sanderson, E. W., Magrach, A., Allan, J. R., Beher, J., Jones, K. R.,
1145   Possingham, H. P., Laurance, W. F., Wood, P., Fekete, B. M., Levy, M. A., &
1146   Watson, J. E. M. (2016). Sixteen years of change in the global terrestrial human
1147   footprint and implications for biodiversity conservation. *Nature Communications*,
1148   *7*(1), 12558. https://doi.org/10.1038/ncomms12558
1149 Wall, J., Wittemyer, G., Klinkenberg, B., & Douglas-Hamilton, I. (2014). Novel
1150   opportunities for wildlife conservation and research with real-time monitoring.
1151   *Ecological Applications*, *24*(4), 593–601. https://doi.org/10.1890/13-1971.1
1152 Wand, M. (2024). *KernSmooth: Functions for Kernel Smoothing Supporting Wand &*
1153   *Jones (1995).* [Computer software].
1154 Wang, G. (2019). Machine learning for inferring animal behavior from location and
1155   movement data. *Ecological Informatics*, *49*, 69–76.
1156   https://doi.org/10.1016/j.ecoinf.2018.12.002
1157 Williams, H. J., Shepard, E. L. C., Duriez, O., & Lambertucci, S. A. (2015). Can
1158   accelerometry be used to distinguish between flight types in soaring birds?
1159   *Animal Biotelemetry*, *3*(1), 45. https://doi.org/10.1186/s40317-015-0077-0
1160 Williams, H. J., Taylor, L. A., Benhamou, S., Bijleveld, A. I., Clay, T. A., de Grissac, S.,
1161   Demšar, U., English, H. M., Franconi, N., Gómez-Laich, A., Griffiths, R. C., Kay,
1162   W. P., Morales, J. M., Potts, J. R., Rogerson, K. F., Rutz, C., Spelt, A., Trevail,
1163   A. M., Wilson, R. P., & Börger, L. (2020). Optimizing the use of biologgers for
1164   movement ecology research. *Journal of Animal Ecology*, *89*(1), 186–206.
1165   https://doi.org/10.1111/1365-2656.13094
1166 Wright, M. N., & Ziegler, A. (2017). ranger: A Fast Implementation of Random Forests
1167   for High Dimensional Data in C++ and R. *Journal of Statistical Software*, *77*, 1–
1168   17. https://doi.org/10.18637/jss.v077.i01
1169 Yu, H., Deng, J., Nathan, R., Kröschel, M., Pekarsky, S., Li, G., & Klaassen, M. (2021).
1170   An evaluation of machine learning classifiers for next-generation, continuous-
1171   ethogram smart trackers. *Movement Ecology*, *9*(1), 15.
1172   https://doi.org/10.1186/s40462-021-00245-x
1173
1174

Supporting Information

Data S1 - Transmitter schedule of free-roaming griffon vultures

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

<u>Conducting observations to build an ACC training dataset</u>

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.

<u>Using the confidence score to validate 'Feeding' predictions of free-roaming vultures</u>

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.

<u>Case study: identification of vulture feeding hotspots to prevent poisoning</u>

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.

1200 **Data S2 – Details of the model building sequence**

1201 Using the 'training subset', we built a random forest model (number of trees = 1000,

1202 number of statistical features per tree = 10, and minimum number of observations per

1203 leaf = 10). We evaluated the performance of our model using the 'testing' subset. We

1204 built a confusion matrix and calculated three performance metrics, for the full model and

1205 for each behavior: (i) accuracy (sum of true-positives and true-negatives divided by all

1206 predictions); (ii) precision (true-positives divided by the sum of true-positives and false-

1207 positives); and (iii) recall (true-positives divided by the sum of true-positives and false-

1208 negatives). The equations for each metric can be found in Table 1. Because only 67%

1209 of the data was used to train the algorithm, these metrics are likely an underestimation

1210 of the real performance of the final algorithm.

1211

1212

1213
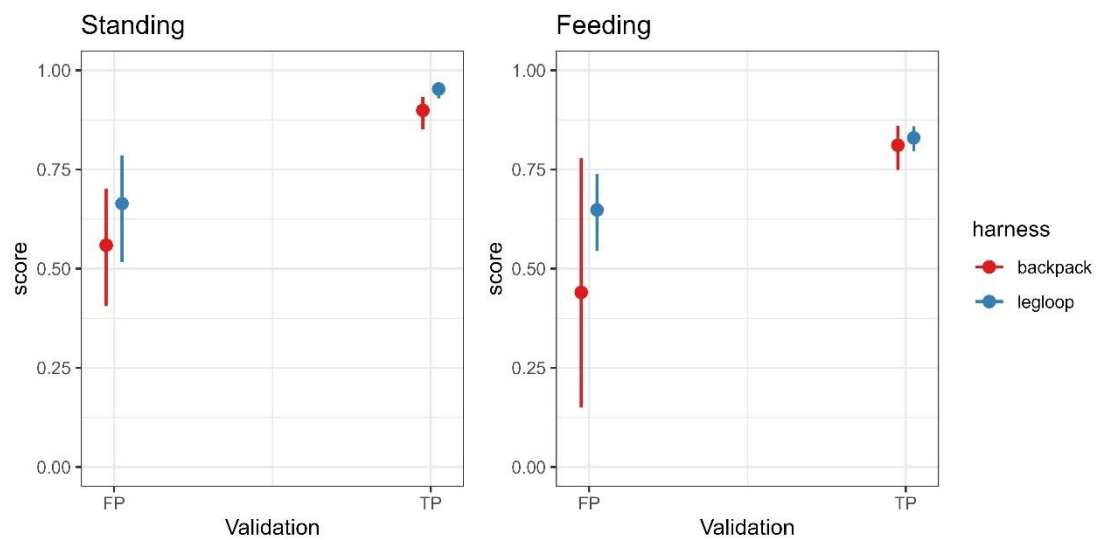
1214

1215 **Supplementary Figure S1 - Comparison of the confidence scores of true-positive**
1216 **(TP) and false-positive (FP) 'Standing' and 'Feeding' classifications, recorded**
1217 **with a backpack (red) and with a leg-loop (blue) harness.** Each panel shows the
1218 model predictions and the 95% confidence intervals for the effect of the harness on the
1219 confidence scores of true positive and false positive behavioral classifications.

1220

1221

1222

1223 **Supplementary Figure S2 – Using GPS-derived metrics to identify feeding**
1224 **hotspots, without using acceleration-based behavioral classification.** The red
1225 points show vulture locations with GPS ground speed equal to or slower than 4m/s,
1226 outside the roosting areas and outside feeding stations for vultures, during November
1227 2022 (n = 1,938 locations). The polygons indicate the density of vulture locations,
1228 showing potential areas where vultures could be foraging and would be at greater risk
1229 of poisoning. The polygon colors indicate the density of vulture locations, with blue
1230 areas having lower density and red areas higher density. Comparison with Figure 6
1231 (based on accelerometer-classified 'Feeding' behavior) shows that incorporating
1232 behavioral data produces spatially distinct and more accurate hotspot identification.

1233 **Supplementary Table S1 – Full list of statistical features used to summarize each**
1234 **acceleration bout.** Note that X, Y and Z refer to the orthogonal acceleration axes
1235 (sway – X; surge – Y; heave - Z), while x and y refer to the variables that should be
1236 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

1237

1238

1239 **Supplementary Table S2 – Results of the random forest model to classify vulture**
1240 **behavior based on accelerometer data.** The random forest model was trained on
1241 67% of the total number of observations and its performance was tested on 33% of the
1242 data. The true-positives and false-positives are the result of the performance test (for a
1243 full confusion matrix see Figure 3 of the manuscript). For each prediction, the model
1244 calculates a confidence score, i.e., the proportion of trees that agree on the highest
1245 scoring prediction. Here we show the mean confidence score values for each predicted
1246 behavior category.

1247

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)

1248

1249

1250 **Supplementary Table S3 – Comparison of the confidence scores of true-positive**
1251 **and false-positive behavioral classifications.** Results of the GLMM beta binomial
1252 model, with confidence score (0-1) as the response variable, and correctness (true-
1253 positive / false-positive), behavioral class (Feeding, Flapping, Ground, Lying, Soaring
1254 and Standing) and their interaction as predictors. The model had device ID as random
1255 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

1256

Random effects	Variance	Std. Dev.
Device id	0.0995	0.3155

1257

1258

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

1269 **Supplementary Table S5 - Confusion matrix of the random forest assessing the**
1270 **influence of the harness type on the performance of the algorithm.** This random
1271 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

1272

1273 **Supplementary Table S6 – Comparison of the confidence scores of feeding**
1274 **bouts likely to be true-positives or false-positives.** Results of the GLMM model
1275 beta binomial model, with confidence score (0-1) as the response variable, and
1276 correctness (likely to be true-positive – inside feeding stations; likely to be false-
1277 positives – on cliffs or in flight) as predictor. The model had device ID as random effect.
1278 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

1279

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

1280

1281

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.
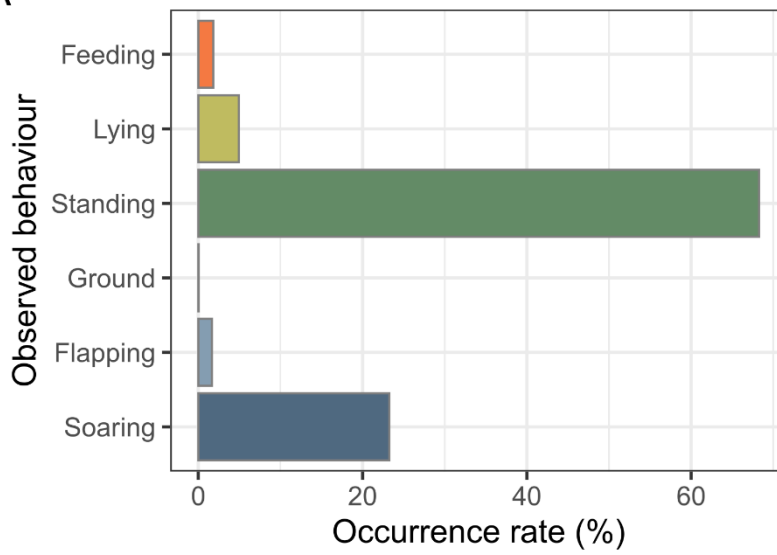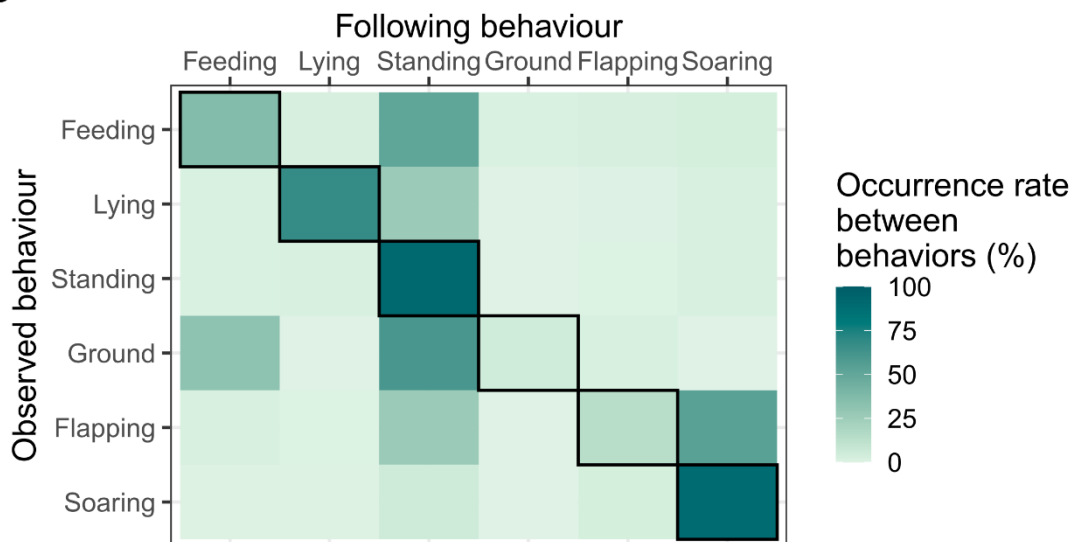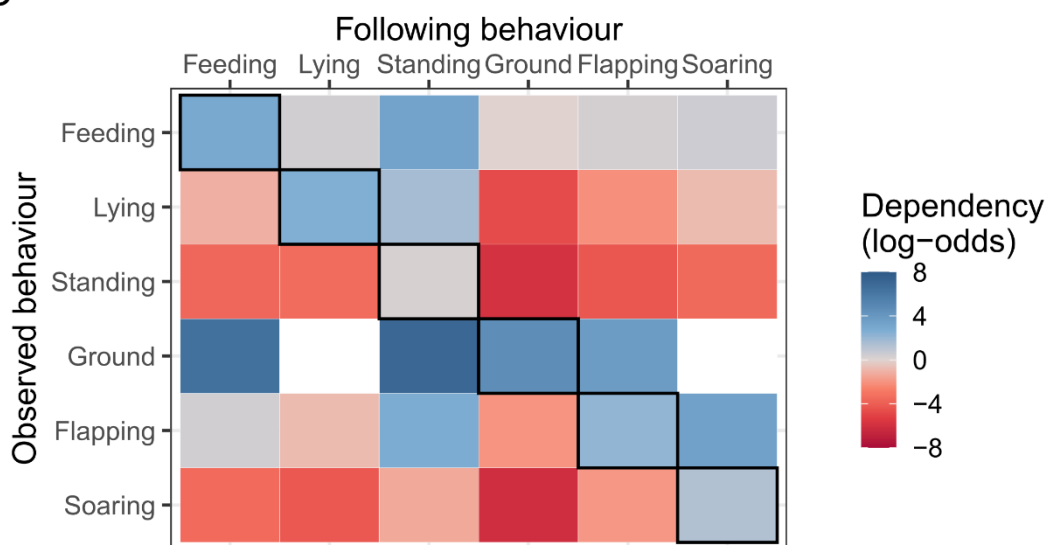
1318 **Supplementary Figure S3 – Temporal correlation between consecutive**
1319 **behaviors.** Figure A shows the total occurrence rate of all behaviors (*OR*); Figure B
1320 shows percentage of times that a behavior is followed by another (*BTR*); Figure C
1321 shows the dependency between behaviors, calculated as the log-odds of the ratio
1322 between occurrence rate between behaviors (*BTR*) and the total occurrence rate of
1323 each behavior (*OR*). In C, positive dependencies are depicted by shades of blue,
1324 negative dependencies are depicted by shades of red, and white tiles depict undefined
1325 values (log of 0). In B and C, darker tones represent stronger relationships, and the
1326 black outlines in the diagonal highlight the self-dependency (e.g., Feeding-Feeding
1327 dependency).

1328