



## Classifying the activity states of small vertebrates using automated VHF telemetry

Jannis Gottwald, Raphaël Royauté, Marcel Becker, Tobias Geitz, Jonas Höchst, Patrick Lampe, Lea Leister, Kim Lindner, Julia Maier, Sascha Rösner, et al.

### ► To cite this version:

Jannis Gottwald, Raphaël Royauté, Marcel Becker, Tobias Geitz, Jonas Höchst, et al.. Classifying the activity states of small vertebrates using automated VHF telemetry. *Methods in Ecology and Evolution*, 2023, 14 (1), pp.252-264. 10.1111/2041-210X.14037 . hal-03952198

**HAL Id: hal-03952198**

**<https://hal.science/hal-03952198>**

Submitted on 6 Mar 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

## RESEARCH ARTICLE

## Classifying the activity states of small vertebrates using automated VHF telemetry

Jannis Gottwald<sup>1</sup>  | Raphaël Royauté<sup>2</sup>  | Marcel Becker<sup>3</sup> | Tobias Geitz<sup>3</sup> | Jonas Höchst<sup>4</sup> | Patrick Lampe<sup>4</sup> | Lea Leister<sup>1</sup> | Kim Lindner<sup>3</sup> | Julia Maier<sup>1</sup> | Sascha Rösner<sup>3</sup> | Dana G. Schabo<sup>3</sup>  | Bernd Freisleben<sup>4</sup> | Roland Brandl<sup>3</sup> | Thomas Müller<sup>2,5</sup> | Nina Farwig<sup>3</sup> | Thomas Nauss<sup>1</sup>

<sup>1</sup>Department of Geography, Philipps-Universität Marburg, Marburg, Germany

<sup>2</sup>Senckenberg Biodiversity and Climate Research Centre, Senckenberg Gesellschaft für Naturforschung, Frankfurt (Main), Germany

<sup>3</sup>Department of Biology, Philipps-Universität Marburg, Marburg, Germany

<sup>4</sup>Department of Mathematics and Computer Science, Philipps-Universität Marburg, Marburg, Germany

<sup>5</sup>Department of Biological Sciences, Goethe University Frankfurt, Frankfurt (Main), Germany

## Correspondence

Jannis Gottwald

Email: [jannis.gottwald@geo.uni-marburg.de](mailto:jannis.gottwald@geo.uni-marburg.de)

## Funding information

Hessisches Ministerium für Wissenschaft und Kunst

Handling Editor: Thomas White

## Abstract

1. The most basic behavioural states of animals can be described as active or passive. While high-resolution observations of activity patterns can provide insights into the ecology of animal species, few methods are able to measure the activity of individuals of small taxa in their natural environment. We present a novel approach in which a combination of automatic radiotracking and machine learning is used to distinguish between active and passive behaviour in small vertebrates fitted with lightweight transmitters (<0.4 g).
2. We used a dataset containing >3 million signals from very-high-frequency (VHF) telemetry from two forest-dwelling bat species (*Myotis bechsteinii* [ $n = 52$ ] and *Nyctalus leisleri* [ $n = 20$ ]) to train and test a random forest model in assigning either active or passive behaviour to VHF-tagged individuals. The generalisability of the model was demonstrated by recording and classifying the behaviour of tagged birds and by simulating the effect of different activity levels with the help of humans carrying transmitters. The model successfully classified the activity states of bats as well as those of birds and humans, although the latter were not included in model training (F1 0.96–0.98).
3. We provide an ecological case-study demonstrating the potential of this automated monitoring tool. We used the trained models to compare differences in the daily activity patterns of two bat species. The analysis showed a pronounced bimodal activity distribution of *N. leisleri* over the course of the night while the night-time activity of *M. bechsteinii* was relatively constant. These results show that subtle differences in the timing of species' activity can be distinguished using our method.
4. Our approach can classify VHF-signal patterns into fundamental behavioural states with high precision and is applicable to different terrestrial and flying vertebrates. To encourage the broader use of our radiotracking method, we provide

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2022 The Authors. *Methods in Ecology and Evolution* published by John Wiley & Sons Ltd on behalf of British Ecological Society.

the trained random forest models together with an R package that includes all necessary data processing functionalities. In combination with state-of-the-art open-source automated radiotracking, this toolset can be used by the scientific community to investigate the activity patterns of small vertebrates with high temporal resolution, even in dense vegetation.

#### KEYWORDS

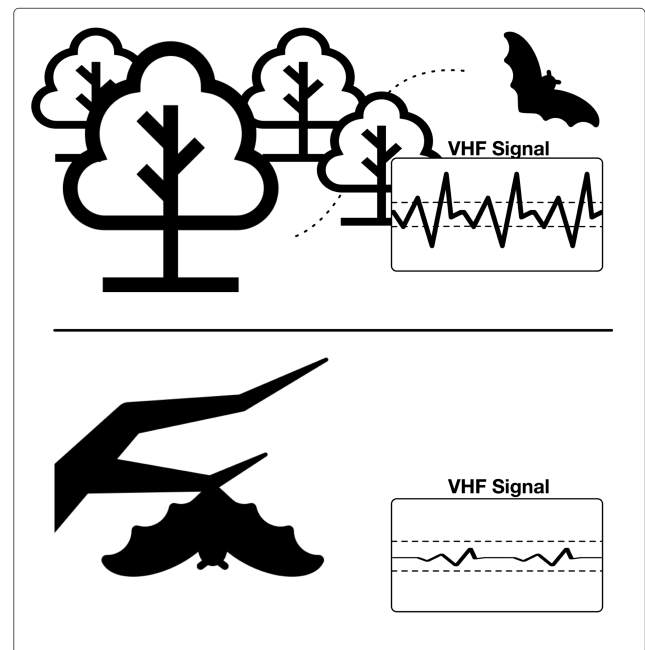
automated radiotelemetry system, bats, behaviour, birds, generalised additive models, machine learning, *Myotis bechsteinii*, *Nyctalus leisleri*, random forest, small animals, tRackIT

## 1 | INTRODUCTION

The behaviour of an animal can be fundamentally divided into active and passive states (Halle & Stenseth, 2000), with the former requiring a much higher energy expenditure (Rowcliffe et al., 2014). Quantifying the distribution of activity periods throughout the day provides important insights into species' responses to their environment, foraging strategies, bioenergetics and adaptations (Torney et al., 2021). Moreover, temporal segregation of species that share the same niche is one recognised mechanism that can facilitate species coexistence (Nakabayashi et al., 2021).

Detailed analysis of individual activity patterns requires high-resolution observations (Nathan et al., 2022), which are often difficult to obtain. The observer's presence may influence animal behaviour and thus bias conclusions (Crofoot et al., 2010) and continuous observation of elusive or highly mobile species in habitats with dense vegetation is close to impossible (Maffei et al., 2005). Information on medium to large-sized species can be obtained using camera traps, GPS transmitters and accelerometers (Kays et al., 2015), as demonstrated by investigations of dynamic habitat and resource use (Wyckoff et al., 2018), behaviour (Freeman et al., 2010) and migration and dispersal (Walton et al., 2018). However, these devices are of limited use for small animals (<100g), due to low detection probabilities, the trade-off between transmitter size and weight, battery life, and data collection intensity (Hallworth & Marra, 2015; Hammond et al., 2016; Wikelski et al., 2007). Newer technical solutions such as the ATLAS system (Advanced Tracking and Localisation of Animals in real-life Systems; Nathan et al., 2022) or the Wildlife Biologging Network (WBN; Ripperger et al., 2020) allow the tracking of small animals with high temporal and spatial resolution, but the required installation effort and costs remain high.

Very-high-frequency (VHF) telemetry has been used in wildlife tracking since the 1960s (Cochran et al., 1965), with the ongoing miniaturisation of VHF transmitters (<0.2 g) allowing the tracking of small taxa (body mass <5 g), ranging from large insects to small vertebrates (Naef-Daenzer et al., 2005). Some studies take advantage of the fact that even small movements of tagged animals result in discernible variations in the strength of the received signal (Kjos & Cochran, 1970) that reflect changes in the angle and distance between the transmitter and receiver (Figure 1). However, collecting reasonable amounts of data on activity bouts using manual radiotelemetry requires an enormous amount of fieldwork, which implies a high level of wildlife



**FIGURE 1** Principle of activity-recognition-based on very-high-frequency (VHF) signal patterns. Top: flying bat; bottom: resting bat. The amplitude and variation of the signal strength over time increase when the tagged individual is moving.

disturbance (Kenward, 2000), and the risk of missing critical events in the life of the tagged individuals is high (Lambert et al., 2009).

Kays et al. (2011) proposed a method for automatically classifying active and passive behaviour based on a threshold in the difference in the signal strength of successive VHF-signals recorded by a system commercially available at the time. Schofield et al. (2018) applied a similar system to investigate the activity pattern of 241 individuals out of three migrating songbird species during stopover. They found that a threshold of 2.5 dBm optimally separates active from passive behaviour. However, the applied systems could only track one tag at a time, resulting in a low temporal resolution (i.e. a few seconds of observations approximately every 10 min due to switching through frequency channels). High-throughput tracking systems (<10-s data interval, many individuals at a time) are now widely available and enable ground-breaking research in animal behaviour, evolution, and ecology (Nathan et al., 2022). In recent years, with the ongoing development of low-cost open-source solutions, automatic VHF radiotracking now enables such high-resolution

capacities. Systems such as the Motus Wildlife Tracking System (Taylor et al., 2017) or the tRackIT-system (Höchst et al., 2021) allow the tracking of many individuals simultaneously and with a very high temporal resolution (seconds) over the complete tagging period. Continuous, high-resolution recording of the VHF-signals makes the entire signal pattern available for subsequent data analysis.

In this work, we build on the methodology of Kays et al. (2011) by calibrating a machine learning (ML) model based on millions of data points representing the behaviours of multiple tagged individuals of two temperate bat species (*Myotis bechsteinii*, *Nyctalus leisleri*). Many ML algorithms are optimised for the recognition of complex patterns in a dataset and may be more robust against factors that influence signal propagation, such as changes in temperature and humidity, physical contact with conspecifics and/or multipath signal propagation (Alade, 2013) than a rule-based approach relying on a single separation value. ML approaches may therefore provide substantial improvements in the accuracy of individual activity states classification compared to threshold-based approaches.

Although deep learning methods have been successfully applied to several ecological problems where large amounts of data are available (Christin et al., 2019), we chose a random forest model due to the following reasons: (a) developing a (supervised) deep learning method requires considerable effort for selecting an appropriate neural network architecture, choosing an appropriate framework to implement the neural network, training, validating, testing and refining the neural network (Christin et al., 2019), (b) our classification tasks resolve to a simple binary classification of active/passive states based on tabular data. In this setting, tree ensemble methods such as random forests seem to have clear advantages—they are less computationally intensive, easy to implement, robust and at least as performant as deep learning (Shwartz-Ziv & Armon, 2022) and (c) in a large study comparing 179 classifiers applied to the 121 classification data sets of the UCI repository, random forests are the best classifiers in over 90% of the cases (Fernández-Delgado et al., 2014).

Our random forest model was used in conjunction with recent developments in automated radiotelemetry (Gottwald et al., 2019; Höchst et al., 2021) to develop a toolset that allows researchers to record the activity patterns of even very small species (body mass <5 g) in their natural habitat and with high resolution. The method was tested by applying it to independent data from bats, humans and birds recorded in a densely vegetated and hilly area and then comparing the results with those obtained by the threshold-based approach of Kays et al. (2011) using the separation value suggested by Schofield et al. (2018).

In our method, activity states are recognised with high temporal resolution (<10 s) and high accuracy. In the following, we provide detailed information on the application of the random forest model and its validation using data on the behaviour of tagged bat and bird individuals generated with an open-source multi-sensor tool (Gottwald et al., 2021). In a case study, we demonstrate the use of the approach to detect differences in activity patterns between those of the two forest dwelling bat species *Myotis bechsteinii* and *N. leisleri*.

In the next sections, we detail our process for developing and validating a random forest model to classify VHF-signals based on

activity data gathered on bats, birds and humans. We showcase the possible insights in wildlife monitoring that our approach may bring by providing an ecological case study focusing on the comparison of activity patterns between two bat species. Detailed information on data processing and analysis is provided, along with an R package, example scripts and data in a hope to promote broad application in wildlife monitoring and ecology.

## 2 | FIELD METHODS

### 2.1 | Study area

The study was conducted in the Marburg Open Forest (MOF), Hesse, Germany, a densely vegetated mixed forest of 200 ha, dominated by European beech (*Fagus sylvatica*) with some clearings and a relatively strong relief for low mountain ranges (lowest elevation ~200 m, highest ~400 m; Figure 2). The forest is home to 13 species of bats and 43 species of birds.

### 2.2 | Tagging of bats and birds

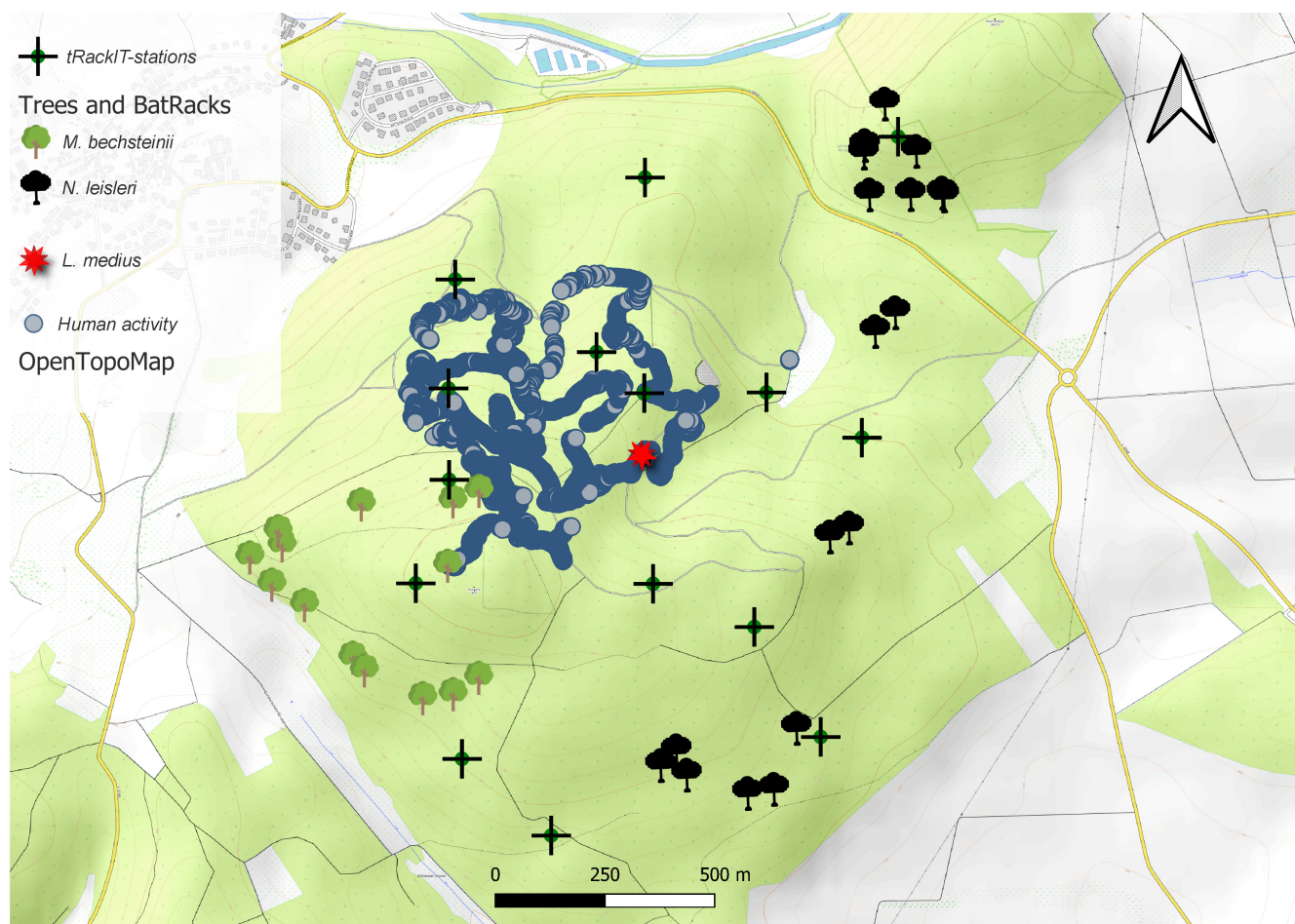
Every year, we caught and then tagged bats and birds with customised VHF-transmitters of different sizes and weights (V3+, Dessau Telemetrie-Service; 0.3 g 1 g). Tag weights were always <4% of the body mass of the tagged individual (see S2 for technical details, methods and permits). For the ecological case study on bats, we captured and tagged 91 bat individuals from two focus species (66 *M. bechsteinii* (~12 g) and 25 *N. leisleri* (~18 g)). For the evaluation of our approach (see 'transferability to small diurnal flying vertebrates' section) we used data of 19 bird individuals tagged in another study conducted in parallel (one *Leipicus medius*, three *Cyanistes caeruleus*, three *Erithacus rubecula*, three *Garrulus glandarius*, three *Parus major*, three *Sylvia atricapilla* and three *Turdus merula*). The frequency separation between transmitters used simultaneously was at least 3 kHz. For the handling and tagging of the bats and birds, a licence was issued by the Nature Conservancy Department of Central Hessen ('Obere Naturschutzbehörde Mittelhessen, Regierungspräsidium Gießen', v54-19c 2015 h01; v54-19c 2015 h01 MR 20/15 Nr. G 10/2019).

### 2.3 | Radiotracking

From 2018 to 2021, we operated a network of 15 custom-designed automatic radiotracking stations (henceforth 'tRackIT-stations'; (Gottwald et al., 2019; Höchst et al., 2021)) distributed over the MOF (Figure 2). The stations recorded signal frequency, duration, and strength as well as the timestamp of the signal of all individuals tagged at a given time simultaneously and automatically.

Each tRackIT-station consisted of four directional antennas with moderate directivity (HB9CV-antenna). While this antenna design reduces the reception range to <1000 m in hilly and vegetated





**FIGURE 2** The Marburg Open Forest in Hesse, Germany. The map shows the locations of the tRackIT-stations (Gottwald et al., 2019; Höchst et al., 2021), the roost trees of bats (*M. bechsteinii*, *N. leisleri*) observed by BatRack multi-sensor stations (Gottwald et al., 2021), the breeding site of a woodpecker (*Leipicus medius*) and the GPS track (shown in blue) of the activity simulation used to test the transferability of the classification method to birds and humans. (Map data from OpenStreetMap)

terrain, it guarantees overlapping radiation patterns of neighbouring antennas which was necessary for bearing calculation and subsequent triangulation as described in Gottwald et al. (2019). However, tracking of positions of tagged individuals is not part of this study. The towers of the stations had a height of approximately 8 m, and antennas were oriented north, east, south and west. We permanently monitored a frequency range of 150.000–150.300 MHz.

From 2018 to 2020, we used the pair 4.3 software developed by the open-source project radiotracking.eu (Gottwald et al., 2019) but switched to the tRackIT operating system (<https://github.com/Nature40/tRackIT-OS>) in 2021 due to high amounts of noise and frequent software failures that often went unnoticed (Höchst et al., 2021). The tRackIT-system enables live transmission of parameters to assess the health of the stations as well as transmission, processing and visualisation of VHF-signals. The former greatly reduces maintenance time and the latter enables tracking of activity, positions and body temperature in near real time. For a detailed description of the hardware and software, please see Gottwald et al. (2019) and Höchst et al. (2021).

The VHF data were filtered by tag frequency  $\pm 3$  kHz and signal duration  $\pm 5$  milliseconds according to settings given by the

manufacturer. For the data recorded with the radiotracking.eu software, we had to visually assess the success of the filtering procedure and in some cases remove recordings below a station- and frequency-specific threshold in dBW due to high amounts of electromagnetic noise.

In total, we used data from 72 individuals (*M. bechsteinii*:  $N_{ID} = 52$ ,  $N_{Obs} = 577,977$ ; *N. leisleri*:  $N_{ID} = 20$ ,  $N_{Obs} = 204,443$ ) monitored for an average of 19 days (according to battery power) to distinguish active from passive states.

### 3 | A RANDOM FOREST MODEL TO CLASSIFY ACTIVITY STATES BASED ON AUTOMATICALLY RECORDED VERY-HIGH-FREQUENCY SIGNALS

#### 3.1 | Groundtruth

We used the patterns in the strength of the recorded VHF-signals together with a supervised ML algorithm to classify the activity of the tagged individuals. Supervised ML requires training and test data

for implementation. We monitored 23 out of the 72 tagged bat individuals (6 *N. leisleri* and 17 *M. bechsteinii*) using a multi-sensor tool (Gottwald et al., 2021) to supply the random forest model with periods of known activity and inactivity.

First, the roost trees of tagged bats were located via manual radiotelemetry between 9 June 2020 and 26 July 2020 and between 10 May 2021 and 18 August 2021. We then set up custom-made video recorder ('BatRack') units to automatically record videos of tagged individuals (Gottwald et al., 2021; <https://nature40.github.io/BatRack/> (vid. 2)). BatRacks consist of a VHF-antenna and an infrared video unit connected to a Raspberry Pi single board computer. We installed the cameras with a focus on the roost entrance and its surrounding area (40-m radius), which allowed the motion of tagged individuals to be captured on the video tracks. The infrared camera unit was automatically triggered by the VHF-signal of the bat transmitters and started recording if the VHF-signal strength exceeded a threshold of  $-60$  dBW, i.e., when a tagged bat flew close to the roosting tree and the BatRack system.

We manually reviewed the video tracks recorded by BatRack units in conjunction with the VHF-signal and classified the observed behavioural sequence into the categories swarming, passing, entering or emerging from the roost. Sequences that showed swarming, passing or emerging were classified as active, and the time between entering and emerging from the roost as inactive. In addition to the sequences recorded on video, we classified periods of time as active if an individual was recorded in short time intervals on widely separated VHF-receivers (tRackIT-stations and BatRacks). From the three (2020) to nine (2021) BatRacks set in front of a total of 30 roosting trees of 6 *N. leisleri* and 17 *M. bechsteinii* individuals (Figure 2), 723 h of behaviour were recorded. For these periods of known activity type, we assigned a passive or active label to the VHF data recorded by one or more of the 15 tRackIT-stations.

### 3.2 | Predictor variables

We calculated 29 predictor variables thought to capture the patterns in the signal strengths over time by applying rolling windows of  $\pm 10$  data entries, corresponding to an approximate time window of 20 s, to each observation of the classified VHF-data recorded by the tRackIT-stations. We chose the window size to capture the dominant signal strength pattern without smoothing out even short changes of the activity state. To prevent averaging over longer periods, the dataset was split into 5 min bins per station before applying the rolling windows. For each bin, we selected the receiver with the most data entries, that is, the best data coverage. We only evaluated

bins with at least 60 observations, that is, three times the window size. This procedure ensures that only stations and receivers with relatively good reception are considered for classification.

To smooth out noise or potentially distracting fluctuations in the signal, we calculated a Hampel filter, in which data points that differ from the window median by more than three standard deviations are replaced by the median (Hampel, 1974). We also applied a mean and a max filter on the raw data of the main receiver whereby the respective data point was replaced with the mean or max of the rolling window. Next, we calculated the variance, standard deviation, kurtosis, skewness and sum of squares for both the raw and the smoothed data, to capture the variability and shape of the data distribution within the window.

Only one antenna is necessary to classify VHF-signals into active versus passive states. However, agreement between receivers of the same station provides additional information and can improve the reliability of the classification. This is especially likely if the individual is relatively close to the station ( $<400$  m in our scenario). When data were available from more than one receiver at the same station, we calculated the variance of signal strength between the receiver with the most and the receiver with the second most observations, together with the correlation coefficient and the covariance of signal strength in a rolling window of  $\pm 10$  data entries. All variables are described in Supplement S1.

### 3.3 | Training and test data

To give equal weight to each class and to avoid overoptimistic accuracy metrics caused by a comparably well-detected majority class, we balanced the groundtruth dataset by randomly down-sampling the activity class with the most data to the amount of data contained by the class with the least data. We then split these balanced data sets into 50% training data and 50% test data for data originating from one receiver. We used the same procedure for data derived from the signals of two receivers, resulting in two training and two test datasets. From a total of 3,243,753 VHF-signals, we assigned 249,796 signals to train the two-receiver model and 588,880 signals to train the one-receiver model (Table 1).

### 3.4 | Model tuning

We used a random forest model as our classification method because it tends to outperform other classifiers, as shown in an extensive comparative study (Fernández-Delgado et al., 2014). This model type is

**TABLE 1** Characteristics of the test and training data obtained from 723 h of video observation on 23 tagged individuals

Setup	Active data	Passive data	Total data points	Balanced active (train/test)	Balanced passive (train/test)
One receiver	588,880	2,654,873	3,243,753	294,440	294,440
Two receivers	249,796	1,469,674	1,719,470	124,898	124,898

also robust against multicollinearity in predictor variables, especially when used with feature selection procedures (Gregorutti et al., 2017), as used in our approach. Since not all variables are equally important to the model and some may even be misleading, we used 50% of the data recorded by either one or two receivers to perform a forward feature selection as implemented in the CAST package (Meyer et al., 2018). This resulted in two random forest models, for data collected by one receiver and two receivers, respectively.

### 3.5 | Groundtruth for controlled walks with human subjects

We conducted a series of 61 controlled walks with a human volunteer to test the reliability of the trained models when applied to various activity patterns and tag positions. This was achieved by moving the VHF transmitters at two different heights, 15 cm above the ground at the ankle and 4 m above the ground, on a pole, around the tRackIT-stations. We simulated inactive states standing still and movements on a small spatial scale were simulated by walking and hopping back and forth over an area of about 1 m<sup>2</sup>. We simulated movements at a medium spatial scale by walking within areas of 40 m<sup>2</sup>, and multiple back and forth displacements and displacements of at least 200 m were used to simulate large-scale movements. We performed each movement type for 3–10 min at different positions within the north-western part of the study area, which is characterised by a diverse topography and complete forest coverage (Figure 2). We recorded the beginning and end times of each sequences and all signals simultaneously recorded by one or more of the 15 tRackIT-stations and then manually assigned the known activity type (active or passive). The human activity dataset consisted of 32,175 data points (26,133 active, 6042 inactive).

### 3.6 | Model validation

We applied the trained random forest models to the 50% of the data withheld for testing to evaluate their performance in classifying bat activity. The same trained models were applied to the human activity datasets. In a first step, we calculated the sensitivity ( $\text{true positives} / (\text{true positives} + \text{false negatives})$ ) and specificity ( $\text{true negatives} / (\text{true negatives} + \text{false positives})$ ) based on a comparison of the observed data with the activity class attributed by the random forest models for both datasets. Additionally, we calculated the F1 metric as the harmonic mean of the precision ( $\text{true positives} / (\text{true positives} + \text{false positives})$ ) and sensitivity, the ROC-AUC and the Kappa index, which takes the probability distribution of each class into account. Values

vary between 0 and 1 (<0 and <1 for Kappa), with values close to 1 indicating that the model shows an almost perfect agreement (Chinchor, 1992; Landis & Koch, 1977).

## 4 | RESULTS OF MODEL VALIDATION

The trained random forest models performed equally well, with F1-scores of at least 0.96 and sensitivities and specificities no less than 0.95, when applied to the validation data of bats and the human activity-simulation (Table 2). Whether the tag was positioned 15 cm or 4 m above the ground during the human activity simulation had no impact on the classification accuracy. The four activity levels simulated by a human were detected similarly well, with sensitivities between 0.95 and 0.97.

### 4.1 | Comparison with a threshold-based approach

We compared the results of the ML-based approach with those of a threshold-based approach by calculating the difference in the signal strength between successive signals for the test datasets of bats and humans (for methods and results on the bird data see 'transferability to small diurnal flying vertebrates' section). We applied a threshold of 2.5 dB, which was deemed appropriate to optimally separate active and passive behaviours in previous studies (Schofield et al., 2018). In addition, we used the `optimise-function` of the R-package `STATS` (R Core Team, 2021) to identify the value of the signal strength difference that separated the training dataset into active and passive with the highest accuracy (i.e. 1.08 dB) and applied it to the test datasets. We calculated the same metrics as described above, except for ROC-AUC, which requires probabilities for each classification.

Regardless of the method used, all F1 values were <0.9 (2.5 dB threshold, bats: 0.63, humans: 0.74; 1.08 dB threshold, bats: 0.74, humans: 0.88) and Kappa values <0.60 (2.5 dB threshold, bats: 0.37, humans: 0.3; 1.08 dB threshold, bats: 0.46, humans: 0.53), which correspond to a moderate to fair agreement (Landis & Koch, 1977). These values remain well below those obtained from our ML models however.

### 4.2 | Ecological case study: Comparison of activity patterns in two forest bat species

In the following, we present an ecological case study to highlight the advantages of the fine-scale classification of activity states at a

TABLE 2 Performance metrics of the test datasets classified by the trained random forest model

Dataset	n passive	n active	F1	ROC-AUC	Sensitivity	Specificity	Precision	Kappa
Bats 1 receiver	294,172	294,172	0.96	0.99	0.96	0.97	0.97	0.93
Bats 2 receivers	110,273	110,273	0.98	1.0	0.98	0.98	0.98	0.95
Human activity	6150	26,504	0.98	1.0	0.97	0.95	0.99	0.90

1-min rate for two species monitored over four consecutive years. Both *M. bechsteinii* and *N. leisleri* are protected species (Habitats Directive 92/43/EEC) endemic to Eurasian forests, but they differ substantially in their foraging behaviour. *Nyctalus leisleri* feeds on ephemeral insects that occur in large numbers, but only for short periods at dusk and dawn (Beck, 1995; Rydell et al., 1996) while *M. bechsteinii* partially collects its prey from the vegetation (Dietz & Pir, 2011) and is thus generally less dependent on the timing of insect flight activity (Rydell et al., 1996).

We focused on the following questions: (1) Do *M. bechsteinii* and *N. leisleri* differ in their overall probability of activity? (2) Do *M. bechsteinii* and *N. leisleri* differ in their timing of activity over the course of their circadian rhythms? To answer these questions, we compared the timing of the onset and end of activity periods, the timing of maximum activity and the overall duration of night-time activity bouts using the data processed with the random forest model.

### 4.3 | Statistical analyses

All analyses were conducted with R v. 4.1.2 (R Core Team, 2021), using the *mgcv* package for additive models (Wood, 2011).

We used hierarchical generalised additive models (HGAMs) to compare differences in the overnight activity patterns of *M. bechsteinii* and *N. leisleri*. These classes of models can be applied to estimate nonlinear relations between responses while allowing for a variety of error terms and random effect specifications (Pedersen et al., 2019). In this study, we modelled activity over the course of the 24-h cycle as shown in Equation 1:

$$P(\text{activity})_i = f(\text{time})_i + \zeta_{\text{ID}} + \zeta_{\text{DATE}}, \quad (1)$$

where the probability of activity for observation *i* is modelled as a binomial variable (0: inactive, 1: active) as a function of the time of day (centered around sunset to account for seasonal shifts in daylight). We used a circular cubic spline with 120 equally spaced knots to constrain the beginning and end of the 24-h cycle so that they matched. Individual identity and date were added as random effects to account for individual, seasonal and yearly effects. Given the volume of data (>700,000 observations), all models were fitted through the *bam()* function for faster model estimation.

Given the short timespan between observations, our models had highly autocorrelated residuals ( $\rho > 0.50$ ). While there are no strict guidelines for accounting for autocorrelation with binomial data in HGAMs, the residual autocorrelation was not influenced by the choice of the error family specified (gaussian vs. binomial). We, therefore, set the autocorrelation manually at a value equal to that of the first lag ( $\rho = 0.57$ ) using the *start\_value\_rho()* from the *ITSADUG* package (van Rij et al., 2020). Next, we refitted with the estimated autocorrelation value with an AR1 structure. This procedure successfully accounted for autocorrelation, as evidenced by the decrease in the median autocorrelation to  $-0.13$  in the refitted model.

Visual inspection of the autocorrelation confirmed that  $\rho$  remained  $< |0.15|$  at all lags.

We compared the activity patterns of the two species by contrasting the Akaike information criterion (AIC) values for a model in which species did not vary in their daily activity patterns (Model 0) against one in which the effect of time of day varied between species (Model 1, using the 'by = species' argument to specify a time  $\times$  species interaction). To visualise the fine-scale difference in activity patterns between *M. bechsteinii* and *N. leisleri*, we calculated the difference in spline functions,  $\Delta f(\text{time})$ . This more precisely revealed the period of the day when the two species were most likely to differ in their probability of activity (negative value:  $P(\text{activity})_{\text{Bechstein}} < P(\text{activity})_{\text{Leisler}}$ , positive value:  $P(\text{activity})_{\text{Bechstein}} > P(\text{activity})_{\text{Leisler}}$ ).

We further characterised the activity patterns of the two bat species by calculating the following metrics based on the predicted values for Model 1:

- Onset and end of activity periods, defined as the first and last time of day when the probability of activity was larger than chance (i.e.  $p(\text{activity}) > 0.5$ ).
- Time of peak activity, calculated as the time of the day when the probability of activity, was maximal.
- Activity duration, defined as the duration of the activity period during a 24 h period weighted by the average probability of being active (in hours). This metric was calculated as the area under the curve between the onset and end of the activity period.

### 4.4 | Species comparisons of circadian activity

*Nyctalus leisleri* and *M. bechsteinii* showed pronounced differences in the shapes of their activity curves and these species differences were also supported by AIC model selection ( $\Delta\text{AIC} = 15,092$ , Table 3; Figure 3). While both species appeared to synchronise their onset of activity with sunset, *N. leisleri* was active an average of 19 min earlier than *M. bechsteinii*. *N. leisleri* also reached peak activity earlier, but its activity markedly declined as soon as *M. bechsteinii* became highly active.

The latter species was highly active throughout most of the night, as indicated by a significantly higher activity duration (area under the curve when  $p(\text{activity}) > 0.5$  [95% CI]; *M. bechsteinii*: 4.70 [4.56; 4.83]; *N. leisleri*: 3.42 [3.23; 3.62]). However, *M. bechsteinii* reached the end of its activity period an average of 12 min sooner than *N. leisleri* (Table 4).

### 4.5 | Transferability of the models to diurnal flying vertebrates (birds)

Our previous section shows that the tRackIT-system can provide important insights into ecological differences between bat species with which the model was trained on. We now focus on the broader application of this method to other flying vertebrates.



**TABLE 3** Model coefficients ( $\beta$ ) and standard errors (SE) and test statistics ( $z$ ,  $p$ ) for the linear portion of the additive models (intercept) along with smoothed parameters for nonlinear terms (edf: Estimated degrees of freedom, chi-squared and  $p$ -values)

Model 0 (AIC = 263,401; R <sup>2</sup> = 0.45)							Model 1 (AIC = 248,309; R <sup>2</sup> = 0.47)								
Linear terms			β	SE	z	p	Linear terms			β	SE	z	p		
Intercept			-1.74	0.13	-12.98	<2 × 10 <sup>16</sup>	Intercept			-1.74	0.13	-13.29	<2 × 10 <sup>16</sup>		
Smoothed terms			edf	df	χ <sup>2</sup>	p	% Variance	Smoothed terms			edf	df	χ <sup>2</sup>	p	% Variance
Time			51.84	118	4,061,287	<2 × 10 <sup>16</sup>	17.42	Time:Leisler			45.82	118	188,033	<2 × 10 <sup>16</sup>	8.83
								Time:Bechstein			47.23	118	3,571,852	<2 × 10 <sup>16</sup>	22.46
Random effects								Random effects							
ID			67.07	71	1,976,322	<2 × 10 <sup>16</sup>	10.18	ID			66.32	71	1,824,189	<2 × 10 <sup>16</sup>	4.78
DATE			280.26	306	3,084,452	<2 × 10 <sup>16</sup>	17.44	DATE			282.01	306	2,421,809	<2 × 10 <sup>16</sup>	10.73

To test the reliability of the model on birds, we attached a transmitter to the back of a middle spotted woodpecker (*L. medius*), and placed a daylight variant of the BatRack ('BirdRack') in front of its nesting tree for 4 consecutive days. The tree was located on a steep and completely forested slope of a small valley (Figure 2). A typical recorded sequence consisted of flying, hopping up the stem, and a very short feeding sequence during which the bird remained motionless at the entrance of its breeding cavity. Since the feeding sequence was usually shorter than three consecutive VHF-signals (~2.5 s), we classified all recorded signals within such a sequence as active. To generate sufficient inactive sequences, 2200 random data points were sampled from signals recorded by tRackIT-stations each night between 0:00h and 2:00h, while the woodpecker was presumably asleep, over four consecutive nights. The dataset of the woodpecker, based on the 75 observed activity sequences, consisted of 17,541 data points (8741 active, 8800 inactive).

We applied the two random forest models to all recordings of the tagged woodpecker and calculated the same performance metrics as for bats and human activity for the sequences of known activity. We used the entire woodpecker data set as well as the activity classifications of six additional bird species, each represented by three individuals, to assess the transferability of the model to birds of different size and movement habits. Since there are no actual observations for the latter and only partial observations for the woodpecker, we visually compared the classified activity of the woodpecker to patterns expected for diurnal vertebrates. Then, we calculated activity probability in relation to the time after sunset for three individuals from each of six small to medium-sized bird species (Table 5) using methods comparable to those of the ecological field study for bats.

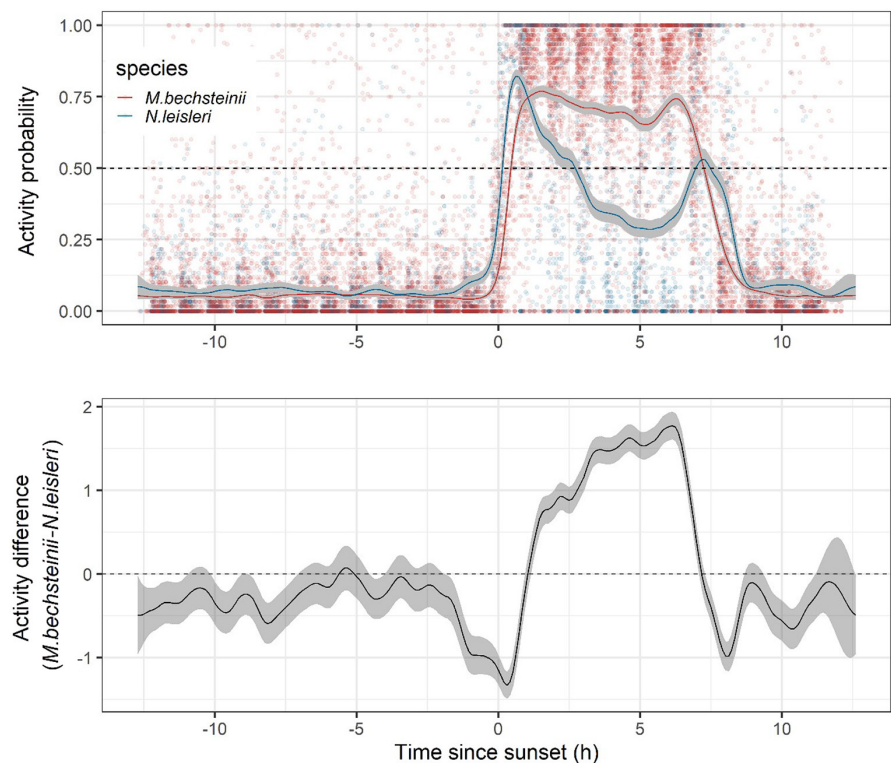
Performance metrics for the sequences of known activity type of the woodpecker were in line with those for bats and human activity (F1 = 0.97; ROC-AUC = 1, Sensitivity = 0.95; Specificity = 1; Precision = 1, Kappa = 0.94). Note that a threshold-based approach also behaved poorly on this dataset (2.5 dB threshold, F1 = 0.62, Kappa = 0.38; 1.08 threshold, F1 = 0.79, Kappa = 0.58).

Visual assessment of the active/passive sequences for the woodpecker showed typical patterns of high activity during the day, starting around sunrise (05:12) and ending around sunset (21:30; Figure 4). The activity probability in relation to time after sunrise of the six additional bird species also correspond to the expected patterns for diurnal birds (Figure 4). Even though no actual observations were available, these patterns suggest a successful classification of the activity of different bird species.

## 5 | DISCUSSION

Using a large dataset consisting of the observed behaviour of tagged bat individuals, we trained two random forest models to classify novel data from the same species into fundamental behaviour, and with high precision and high temporal resolution (~1 s interval). Our approach outperformed previous methods based on a

**FIGURE 3** *Nyctalus leisleri* was consistently active sooner than *Myotis bechsteinii*, but the latter species had longer periods of continuous activity. Top panel: The points represent the activity probability calculated over 1-h intervals, and the solid lines the predicted values from the best hierarchical generalised additive model (HGAM) model. The dashed line indicates the times when the population was equally likely to be detected as active or passive. Bottom panel: Difference in the activity probability calculated from the best HGAM model. Positive values indicate a larger activity probability for *M. bechsteinii* than for *N. leisleri*.



**TABLE 4** Activity metrics of *Nyctalus leisleri* and *Myotis bechsteinii*. Wake-up and sleep times were calculated as the first and last time when the probability of activity was  $>0.5$ . All times are presented as hours since sunset. The time of peak activity represents the time of day when the probability of activity was maximal. Activity duration was calculated as the area under the curve between wake-up and sleep times

Metric	<i>N. leisleri</i>	<i>M. bechsteinii</i>
Activity onset (h)	00:12	00:31
Time of peak activity (h)	00:37	01:33
Activity end (h)	07:23	07:11
Peak P(activity)	0.82 [0.80; 0.84]	0.70 [0.75; 0.79]
Activity duration (h)	3.42 [3.23; 3.62]	4.70 [4.56; 4.83]

**TABLE 5** Number of activity observation (1-min resolution) per bird species

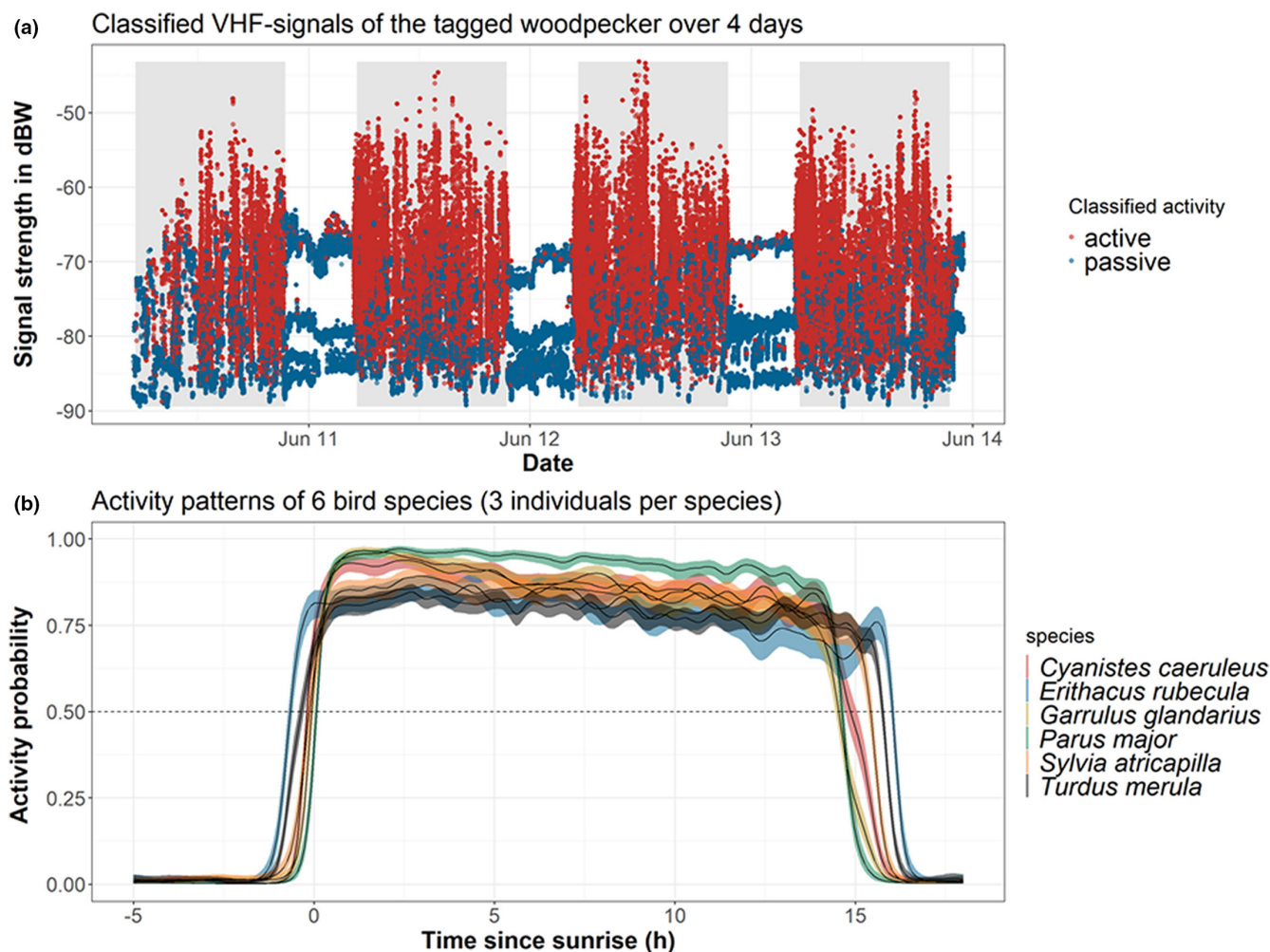
Species	Number of observations
<i>Cyanistes caeruleus</i>	55,735
<i>Erithacus rubecula</i>	54,458
<i>Garrulus glandarius</i>	58,088
<i>Parus major</i>	84,170
<i>Sylvia atricapilla</i>	59,504
<i>Turdus merula</i>	129,521

threshold-based approach even when using a value calibrated with a large groundtruth dataset. Although not inadequate, the threshold-based approach had generally lower and more variable performance

metrics compared with our ML model. We also achieved similar precision when applying the ML models to groundtruth data from other species (woodpecker and controlled human walks). The activity probability estimates of 18 additional bird individuals out of six species also matched expected activity patterns for diurnal vertebrates. This strongly suggests that our method generalises well and could be applied to a variety of vertebrates with similar accuracy (e.g. down to a body mass of 4 g with 0.2 g transmitters; Naef-Daenzer et al., 2005).

In the ecological case study, we demonstrate that our approach enables the detection of even subtle differences in the timing of activity according to a species' ecological preferences (differences in activity onset of  $<20$  min). Specifically, we were able to show distinct activity patterns for these two species, characterised by a slight shift in their timing of activity and a significantly lower activity of *N. leisleri* during the night. Given that these species have evolved to occupy different ecological niches, these patterns are much more likely due to a synchronisation of activity peaks with prey abundance rather than to an avoidance of competition (Ruczyński et al., 2017). *N. leisleri*, like other aerial hawking bats, has likely evolved to exploit insect emergence at dusk and dawn, thus avoiding the greater predation risk that may occur at higher light levels (Rydell et al., 1996). By contrast, *M. bechsteinii* and other gleaning bats are less constrained to flying insects as a food source such that an onset of activity comparable to that of *N. leisleri* would not bring substantial additional benefit.

Our findings are generally in line with previous observations of the activity patterns of *N. leisleri* (Ruczyński et al., 2017; Shiel et al., 1999). No comparable studies exist for *M. bechsteinii*, but in acoustic studies with results reported at the genus level all-night



**FIGURE 4** (a) Signal strength [dBW] from a woodpecker tagged over four consecutive days and nights and the corresponding classification of the bird's activity into active ( $N = 146,962$ ) and passive states ( $N = 303,802$ ). (b) Probability of activity in relation to time since sunrise of six bird species calculated from activity classifications of three individuals per species (see ecological case study for methods). Periods of high activity are consistent with the diurnal activity patterns expected for these species.

activity was determined for *Myotis* (Perks & Goodenough, 2020). However, our study is the first to investigate the overlap of these two species within the same study area. Our approach also allows to detect changes in activity probability according to the reproductive status of individuals and indicates that these shifts are species specific (see Supplementary Material S4).

The high precision and high temporal resolution of our approach together with the easy accessibility of the developed methods may open new research avenues on the variations in the activity patterns among and within species in their response to the environment.

A more in-depth analysis of activity bouts as a function of abiotic factors or the detection of changes in patterns indicating, for example, breeding has not been conducted here, but such studies are likely to be feasible. Whether specific behaviours can be recognised (i.e., foraging, parental care, grooming), as is possible with accelerometers, also remains to be determined. The fact that the variance in the signal pattern depends less on the intensity of the movement than on the signal path remains an issue, however. While the amplitude of the measurements from accelerometers can be

directly related to different behaviour classes (Kays et al., 2015), the amplitude of stationary recorded VHF-signals also changes due to the distance to the radio tracking station. The spatial context of the receiving stations as well as the localisation algorithms presented in Gottwald et al., 2019 could provide additional information, such as distance to the station and direction of movement. However, for localisations, at least two radiotracking stations are necessary and the spatial accuracy and reliability of position tracking when operating in cluttered environments such as forests is still under investigation.

The tRackIT-system can currently record up to 90 individuals at a time within the same spatial context, but technology that allows for higher numbers is under development. Given the relatively low costs of the transmitters (~130 €) and tRackIT-stations (~1500 €), the monitoring of an entire community of small vertebrates at high temporal resolution becomes possible with this system. For instance, a study investigating the activity states of an entire temperate forest bird community is currently conducted in the Marburg Open Forest. The tRackIT-system now allows activity classification in real-time, which

opens several exciting research avenues. For example, it is now being used to narrow down the time of death of chicks in meadow-breeding birds to subsequently reduce error bars in nest survival models (<https://www.audi-umweltstiftung.de/umweltstiftung/de/projects/greenovation/telemetry-technology.html>). Personal observations during the bird-breeding season also showed clear shifts in the frequency and regularity of activity periods during the transition from the nonbreeding to the breeding season (J. Gottwald). Future applications may also help automatically determine the (species-specific) onset of the breeding season in songbird communities.

Over the 4 years of the study, we collected data with two different software designs (radiotracking.eu and tRackITOS) that show significant differences in data quality. We also covered a range of suboptimal recording conditions caused by topography and vegetation, which leads us to the assumption that the approach presented here is not exclusively applicable to data recorded with tRackIT-stations. Other open-source systems such as Motus (sensorgnome; Taylor et al., 2017), but also commercial systems such as the Lotek SRX/DX series receivers (Taylor et al., 2017), record comparable data that may be used with the functionalities and models presented here. However, this was not tested as part of this study.

The scientific insights that can be expected from automatic radiotracking based activity studies have the potential to deepen our understanding of the ecology and behaviour of small animal species in unprecedented ways (Nathan et al., 2022). With the recent advances in open-source automatic radiotracking (Gottwald et al., 2019; Höchst et al., 2021; Taylor et al., 2017) together with the trained models and data-processing functionalities of the tRackIT R-package, the scientific community is now equipped with an accessible toolset that allows the activity patterns of small animals to be analysed and classified at high temporal resolution.

## AUTHOR CONTRIBUTIONS

Jannis Gottwald planned the study and organised the field work for collecting groundtruth and bat activity data. Jannis Gottwald, Lea Leister, Julia Maier, Tobias Geitz (bats and humans), Dana G. Schabo, Sascha Rösner, Kim Lindner, Marcel Becker (woodpecker), Patrick Lampe, Jannis Gottwald and Jonas Höchst (hardware and software development and implementation) participated in the fieldwork effort required for locating, tagging and monitoring the bat and bird populations. Bernd Freisleben, Roland Brandl, Thomas Müller, Nina Farwig and Thomas Nauss supervised the study. Jannis Gottwald developed the associated R-Package and defined the research questions in association with Dana G. Schabo, Sascha Rösner and, Raphaël Royauté. Jannis Gottwald applied and then validated the ML model. Raphaël Royauté was responsible for the statistical analysis of the activity patterns of bats. Jannis Gottwald and Raphaël Royauté jointly wrote the first draft of the manuscript. All authors contributed to subsequent versions.

## ACKNOWLEDGEMENTS

The research was funded by the Hessen State Ministry for Higher Education, Research and the Arts, Germany, as part of the LOEWE

priority project Nature 4.0—Sensing Biodiversity (<https://uni-marburg.de/natur40>). We thank all Master and Bachelor students who supported the fieldwork. Special thanks to C. Kodges and F. Strehmann for their tireless work in catching and tagging the birds and bats. Open Access funding enabled and organized by Projekt DEAL.

## CONFLICT OF INTEREST

All authors declare that they have no conflict of interest.

## PEER REVIEW

The peer review history for this article is available at <https://publons.com/publon/10.1111/2041-210X.14037>.

## DATA AVAILABILITY STATEMENT

To ensure the complete reproducibility of our research, all data and the code used are stored in a data collection at data\_UMR, the research data repository of Philipps-Universität Marburg (<https://data.uni-marburg.de/handle/dataumr/147>). Detailed information on data processing and analysis is provided, along with an R package (<https://github.com/Nature40/tRackIT>; <https://doi.org/10.5281/zenodo.7310599>). The trained models are available at: <https://doi.org/10.17192/fdr/79>. The workflow for model tuning, evaluation and comparison with a threshold-based approach is available at: [https://nature40.github.io/tRackIT\\_activity\\_classification\\_model\\_tuning\\_and\\_evaluation/](https://nature40.github.io/tRackIT_activity_classification_model_tuning_and_evaluation/). Example data processing routines from raw vhf-signals to activity classifications using a small dataset (<https://doi.org/10.17192/fdr/104>) can be found at the package GitHub page: <https://nature40.github.io/tRackIT/>. Reproducible scripts for the ecological case study are available at: [https://nature40.github.io/tRackIT\\_activity\\_ecological\\_case\\_study/](https://nature40.github.io/tRackIT_activity_ecological_case_study/). R-markdown versions of all scripts can be found in the rmd folder of the tRackIT R-Package.

## ORCID

Jannis Gottwald  <https://orcid.org/0000-0001-7763-1415>  
 Raphaël Royauté  <https://orcid.org/0000-0002-5837-633X>  
 Dana G. Schabo  <https://orcid.org/0000-0002-2038-0608>

## REFERENCES

- Alade, M. O. (2013). Investigation of the effect of ground and air temperature on very high frequency radio signals. *Journal of Information Engineering and Applications*, 3, 16–21.
- Beck, A. (1995). Fecal analyses of European bat species. *Myotis*, 32–33, 109–119.
- Chinchor, N. (1992). MUC-4 evaluation metrics in *Proceedings of the Fourth Message Understanding Conference* 22–29. Morgan Kaufmann.
- Christin, S., Hervet, É., & Lecomte, N. (2019). Applications for deep learning in ecology. *Methods in Ecology and Evolution*, 10(10), 1632–1644.
- Cochran, W. W., Warner, D. W., Tester, J. R., & Kuechle, V. B. (1965). Automatic radio-tracking system for monitoring animal movements. *Bioscience*, 15(2), 98–100.
- Crofoot, M. C., Lambert, T. D., Kays, R., & Wikelski, M. C. (2010). Does watching a monkey change its behaviour? Quantifying observer effects in habituated wild primates using automated radiotelemetry. *Animal Behaviour*, 80(3), 475–480.
- Dietz, M., & Pir, J. B. (2011). *Distribution, ecology and habitat selection by Bechstein's Bat (Myotis Bechsteinii) in Luxembourg*. Laurenti.



- Fernández-Delgado, M., Cernadas, E., Barro, S., & Amorim, D. (2014). Do we need hundreds of classifiers to solve real world classification problems? *Journal of Machine Learning Research: JMLR*, 15(1), 3133–3181.
- Freeman, R., Dennis, T., Landers, T., Thompson, D., Bell, E., Walker, M., & Guilford, T. (2010). Black petrels (*Procellaria parkinsoni*) patrol the ocean shelf-break: GPS tracking of a vulnerable procellariiform seabird. *PLoS ONE*, 5(2), e9236.
- Gottwald, J., Lampe, P., Höchst, J., Friess, N., Maier, J., Leister, L., Neumann, B., Richter, T., Freisleben, B., & Nauss, T. (2021). BatTrack: An open-source multi-sensor device for wildlife research. *Methods in Ecology and Evolution*, 12(10), 1867–1874.
- Gottwald, J., Zeidler, R., Friess, N., Ludwig, M., Reudenbach, C., & Nauss, T. (2019). Introduction of an automatic and open-source radio-tracking system for small animals. *Methods in Ecology and Evolution*, 10(12), 2163–2172.
- Gregorutti, B., Michel, B., & Saint-Pierre, P. (2017). Correlation and variable importance in random forests. *Statistics and Computing*, 27(3), 659–678.
- Halle, S., & Stenseth, N. C. (2000). *Activity patterns in small mammals: An ecological approach*. Springer Berlin Heidelberg.
- Hallworth, M. T., & Marra, P. P. (2015). Miniaturized GPS tags identify non-breeding territories of a small breeding migratory songbird. *Scientific Reports*, 5(1), 1–6. <https://doi.org/10.1038/srep11069>
- Hammond, T. T., Springthorpe, D., Walsh, R. E., & Berg-Kirkpatrick, T. (2016). Using accelerometers to remotely and automatically characterize behavior in small animals. *The Journal of Experimental Biology*, 219(Pt 11), 1618–1624.
- Hampel, F. R. (1974). The influence curve and its role in robust estimation. *Journal of the American Statistical Association*, 69(346), 383–393.
- Höchst, J., Gottwald, J., Lampe, P., Zobel, J., Nauss, T., Steinmetz, R., & Freisleben, B. (2021). tRackIT OS: Open-source software for reliable VHF wildlife tracking. In *INFORMATIK 2021* (pp. 425–442). Gesellschaft für Informatik.
- Kays, R., Crofoot, M. C., Jetz, W., & Wikelski, M. (2015). Terrestrial animal tracking as an eye on life and planet. *Science*, 348(6240), aaa2478.
- Kays, R., Tilak, S., Crofoot, M., Fountain, T., Obando, D., Ortega, A., Kuemmeth, F., Mandel, J., Swenson, G., Lambert, T., Hirsch, B., & Wikelski, M. (2011). Tracking animal location and activity with an automated radio telemetry system in a tropical rainforest. *Computer Journal*, 54(12), 1931–1948.
- Kenward, R. E. (2000). *A manual for wildlife radio tagging*. Academic Press.
- Kjos, C. G., & Cochran, W. W. (1970). Activity of migrant thrushes as determined by radio-telemetry. *The Wilson Bulletin*, 82(2), 225–226.
- Lambert, T. D., Kays, R. W., Jansen, P. A., Aliaga-Rossel, E., & Wikelski, M. (2009). Nocturnal activity by the primarily diurnal central American agouti (*Dasyprocta punctata*) in relation to environmental conditions, resource abundance and predation risk. *Journal of Tropical Ecology*, 25(2), 211–215. <https://doi.org/10.1017/s0266467408005804>
- Landis, J. R., & Koch, G. G. (1977). An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics*, 33(2), 363–374.
- Maffei, L., Noss, A. J., Cuéllar, E., & Rumiz, D. I. (2005). Ocelot (*Felis pardalis*) population densities, activity, and ranging behaviour in the dry forests of eastern Bolivia: Data from camera trapping. *Journal of Tropical Ecology*, 21(3), 349–353.
- Meyer, H., Reudenbach, C., Hengl, T., Katurji, M., & Nauss, T. (2018). Improving performance of spatio-temporal machine learning models using forward feature selection and target-oriented validation. *Environmental Modelling & Software*, 101, 1–9.
- Naef-Daenzer, B., Früh, D., Stalder, M., Wetli, P., & Weise, E. (2005). Miniaturization (0.2 g) and evaluation of attachment techniques of telemetry transmitters. *The Journal of Experimental Biology*, 208(Pt 21), 4063–4068.
- Nakabayashi, M., Kanamori, T., Matsukawa, A., Tangah, J., Tuuga, A., Malim, P. T., Bernard, H., Ahmad, A. H., Matsuda, I., & Hanya, G. (2021). Temporal activity patterns suggesting niche partitioning of sympatric carnivores in Borneo, Malaysia. *Scientific Reports*, 11(1), 19819.
- Nathan, R., Monk, C. T., Arlinghaus, R., Adam, T., Alós, J., Assaf, M., Baktoft, H., Beardsworth, C. E., Bertram, M. G., Bijleveld, A. I., Brodin, T., Brooks, J. L., Campos-Candela, A., Cooke, S. J., Gjelland, K. Ø., Gupte, P. R., Harel, R., Hellström, G., Jeltsch, F., ... Jarić, I. (2022). Big-data approaches lead to an increased understanding of the ecology of animal movement. *Science*, 375(6582), eabg1780.
- Pedersen, E. J., Miller, D. L., Simpson, G. L., & Ross, N. (2019). Hierarchical generalized additive models in ecology: An introduction with mgcv. *PeerJ*, 7, e6876.
- Perks, S. J., & Goodenough, A. E. (2020). Abiotic and spatiotemporal factors affect activity of European bat species and have implications for detectability for acoustic surveys. *Wildlife Biology*, 2020(2), 1–8. <https://doi.org/10.2981/wlb.00659>
- R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing <https://www.R-project.org/>
- Ripperger, S. P., Carter, G. G., Page, R. A., Duda, N., Koelpin, A., Weigel, R., Hartmann, M., Nowak, T., Thielecke, J., Schadhauser, M., Robert, J., Herbst, S., Meyer-Wegener, K., Wagemann, P., Schröder-Preikschat, W., Cassens, B., Kapitza, R., Dressler, F., & Mayer, F. (2020). Thinking small: Next-generation sensor networks close the size gap in vertebrate biologging. *PLoS Biology*, 18(4), e3000655.
- Rowcliffe, J. M., Kays, R., Kranstauber, B., Carbone, C., & Jansen, P. A. (2014). Quantifying levels of animal activity using camera trap data. *Methods in Ecology and Evolution/British Ecological Society*, 5(11), 1170–1179.
- Ruczyński, I., Zahorowicz, P., Borowik, T., & Hałat, Z. (2017). Activity patterns of two syntopic and closely related aerial-hawking bat species during breeding season in Białowieża Primateval Forest. *Mammal Research*, 62(1), 65–73.
- Rydell, J., Entwistle, A., & Racey, P. A. (1996). Timing of foraging flights of three species of bats in relation to insect activity and predation risk. *Oikos*, 76(2), 243–252.
- Schofield, L. N., Deppe, J. L., Zenzal, T. J., Ward, M. P., Diehl, R. H., Bolus, R. T., & Moore, F. R. (2018). Using automated radio telemetry to quantify activity patterns of songbirds during stopover. *The Auk*, 135, 949–963.
- Shiel, C. B., Shiel, R. E., & Fairley, J. S. (1999). Seasonal changes in the foraging behaviour of Leisler's Bats (*Nyctalus Leisleri*) in Ireland as revealed by radio-telemetry. *Journal of Zoology*, 249(3), 347–358.
- Shwartz-Ziv, R., & Armon, A. (2022). Tabular data: Deep learning is not all you need. *An International Journal on Information Fusion*, 81, 84–90.
- Taylor, P. D., Crewe, T. L., Mackenzie, S. A., Lepage, D., Aubry, Y., Crysler, Z., Finney, G., Francis, C. M., Guglielmo, C. G., Hamilton, D. J., Holberton, R. L., Loring, P. H., Mitchell, G. W., Norris, D. R., Paquet, J., Ronconi, R. A., Smetzer, J. R., Smith, P. A., Welch, L. J., & Woodworth, B. K. (2017). The motus wildlife tracking system: A collaborative research network to enhance the understanding of wildlife movement. *Avian Conservation and Ecology/Ecologie et Conservation Des Oiseaux*, 12(1), 1–11. <https://doi.org/10.5751/ace-00953-120108>
- Torney, C. J., Morales, J. M., & Husmeier, D. (2021). A hierarchical machine learning framework for the analysis of large scale animal movement data. *Movement Ecology*, 9(1), 6.
- van Rij, J., Wieling, M., Baayen, H. R., & van Rijn, H. (2020). *itsadug: Interpreting time series and autocorrelated data using GAMMs* (Version 7.1.0) [Computer software]. <https://cran.r-project.org/web/packages/itsadug/>
- Walton, Z., Samelius, G., Odden, M., & Willebrand, T. (2018). Long-distance dispersal in red foxes *Vulpes vulpes* revealed by GPS tracking. *European Journal of Wildlife Research*, 64(6), 64.
- Wikelski, M., Kays, R. W., Kasdin, N. J., Thorup, K., Smith, J. A., & Swenson, G. W., Jr. (2007). Going wild: What a global small-animal tracking system could do for experimental biologists. *The Journal of Experimental Biology*, 210(Pt 2), 181–186.
- Wood, S. N. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear



models. *Journal of the Royal Statistical Society. Series B, Statistical Methodology*, 73(1), 3–36.

Wyckoff, T. B., Sawyer, H., Albeke, S. E., Garman, S. L., & Kauffman, M. J. (2018). Evaluating the influence of energy and residential development on the migratory behavior of mule deer. *Ecosphere*, 9(2), e02113.

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Gottwald, J., Royauté, R., Becker, M., Geitz, T., Höchst, J., Lampe, P., Leister, L., Lindner, K., Maier, J., Rösner, S., Schabo, D. G., Freisleben, B., Brandl, R., Müller, T., Farwig, N., & Naus, T. (2022). Classifying the activity states of small vertebrates using automated VHF telemetry. *Methods in Ecology and Evolution*, 00, 1–13. <https://doi.org/10.1111/2041-210X.14037>