BATScan: A radar classification tool reveals large-scale bat migration patterns

Yuval Werber¹ | Hadar Sextin² | Yossi Yovel²,³ | Nir Sapir¹

Abstract

1. Bat movement and behaviour are still mostly understudied over large scales. High-altitude, nocturnal activity makes visual identification of bats from the ground virtually impossible, dramatically hindering our ability to study their movement ecology. Despite the wide use of radar in aeroecology, its application to study specific taxa is limited due to incomplete target classification abilities. BATScan is a bat classifier for vertical-looking radar data, which enables identifying bats and characterizing their unique aeroecology.

2. We constructed the classifier using data from 10 radar deployments, covering a wide range of habitats on a central bird migration flyway over a 7-year period, comprising ~18 million observations. We analysed animal migration above the Hula Valley, home to over 30 species of bats spanning a range of 5–150 g in size and exhibiting a variety of ecological characteristics. We distinguished bat-labelled radar echoes for training according to phenology, morphology and movement ecology of bats, birds and insects. Several non-bat datasets were constructed and joined to train classifiers under increasing levels of difficulty. Class imbalance in the resulting training data was handled using a generative adversarial network for up-sampling the much smaller bat dataset. The resulting classification tool reached a high level of accuracy and precision, and was further scrutinized with an extensive set of ecological validations.

3. Bats perform seasonal migrations over long distances, but little is known about the spatial and temporal characteristics of this movement, and the ability to study it at a large scale has so far been limited. We present the Israeli BATScan dataset, containing over 60,000 bat observations spanning the entire country and representing multiple habitats. Using this data, we produce an unprecedented large-scale, highly detailed documentation of the yearly movements of bats on a major migration flyway, and distinguish this pattern from bird migration over space and time.

4. So far, radar aeroecology dealt primarily with birds, increasingly with insects, and only rarely with bats. We present BATScan, a classification tool that can incorporate bats into the framework of radar aeroecology to finally enable a comprehensive description of animal aeroecology.
1 | INTRODUCTION

The aerial habitat holds both opportunities and challenges for the science of ecology. On one hand, the airspace is practically limitless and free of obstacles, a perfect arena for in-depth, large-scale ecological analysis (Kunz et al., 2008). It serves as habitat for countless organisms, which take part in diverse ecological processes (Diefhl et al., 2017). Many vertebrates and invertebrates migrate, forage, compete, commute, mate and even sleep while airborne (Brunton, 2018; Frick et al., 2017; Rattenborg, 2017).

On the other hand, this habitat is out of our reach, and we must rely on technological means to thoroughly study it (Robinson et al., 2010). To grasp the largely unexplored ecological processes in this vast airspace, we depend on machinery and software to detect, visualize, document and analyse the organisms inhabiting the aerial habitat. Aero-ecologists study the products of these instruments, which can either be obtained by tracking single individuals over large areas and long periods using animal-borne devices, or by constantly monitoring a fixed volume of air, thus tracking innumerable individuals over short periods in a predetermined frame of reference (Phillips et al., 2019). Using radars for animal detection started about 80 years ago (Lack & Varley, 1945; Plank, 1956; Tolbert et al., 1958) and is widely implemented for ‘Eulerian sampling’ aeroecological research (Schmaljohann, 2020).

Vertical-looking radar (hereafter VLR) can provide a fixed, habitat point of view of the airspace, enabling long-term studies of aeroecological processes (Jeffries et al., 2013). The radar emits electromagnetic pulses and records the returning echoes which can then be analysed to produce multiple parameters regarding a passing animal’s shape, size and movement (Figure 1).

The detailed documentation of all animals that pass in a radar’s detection range describes the distribution of flying organisms in time and altitude.

The main limitation of wildlife radar is the difficulty of accurately classifying detected objects (Nohara et al., 2007; Rosa et al., 2016). Given the amount of modern ecological VLR data and the insights they are intended to provide, manual classification of individual echoes is generally impractical. In recent years, developments in computational software and hardware have facilitated automatic classification using machine learning algorithms of large VLR datasets (Rosa et al., 2016; Yeşil et al., 2019; Zaugg et al., 2008).

Separating bats from other biological VLR targets has been impossible so far due to their morphological and behavioural similarities to birds (Bruderer & Popa-Lisseau, 2005; Kunz et al., 2007; Rosa et al., 2016), and bats are often lumped with birds in radar data classifiers, including in those using machine learning algorithms (e.g. Zaugg et al., 2008). This is because supervised machine learning algorithms require large, independently verified datasets for training the algorithms, to learn to classify unfamiliar data. Training datasets must be compiled using accurate methods like manual human visual identification of the animals themselves (Niemü & Tanttu, 2018; Rosa et al., 2016) or expert classification of the data based on prior knowledge and experience (Zaugg et al., 2008). Visual identification is almost impossible for high-flying nocturnal bats, and thus bat training data must be compiled based on expert assessment, which may include animal behaviour, phenology and morphology.

A bat classifier for VLR data will enable research of large-scale and long-term bat ecology. Bats are highly mobile organisms, which play a role as pollinators, seed dispersers, insect pest regulators and nutrient cyclers over multiple geographical and temporal scales, making them ‘mobile links’, connecting distant components of the biosphere (Castillo-Figueroa, 2020; Gnaspini & Trajano, 2000; Lacher et al., 2019; Lundberg & Moberg, 2003; Muscarella & Fleming, 2007; Tremlett et al., 2019; Voigt et al., 2015).

Despite their importance, very little is known about different properties of long-range bat movement. Bats perform seasonal migrations spanning thousands of kilometres which, so far, were only studied using Lagrangian sampling methods (single individuals tracked over large areas and long time periods, Bach et al., 2022; Caprio et al., 2020; Cryan et al., 2014; Hutterer et al., 2005; Lagerveld et al., 2014; O’mara et al., 2014, 2019; Petit & Mayer, 2000; Russell et al., 2005; Sullivan et al., 2012; Villa & Cockrum, 1962; Wilkinson & Fleming, 1996). These methods rely on sparse, sporadic observations of usually small number of individuals, which provide a very limited view of bat migration, preventing comprehensive, large-scale analyses (Fleming, 2019). Implementing radar aeroecological approaches in bat research is expected to substantially advance our understanding of bat movement, behaviour and ecology.

Here we use BATScan, a bat classification algorithm we developed, to separate bats from birds, insect and inanimate reflectors in the airspace and study their ecology to produce the first Eulerian (short-term monitoring of all objects in a limited area) description of bat migration. We specifically describe and analyse bat seasonal phenology and altitudinal distribution and migratory movements, which are deduced from their movement directionality. Our work focuses on the aerial movement of bats over the Hula Valley, a major avian flyway where seasonal bat movements have previously been documented (Levin et al., 2013), throughout the full annual cycle. It represents the first systematic and comprehensive description of the abundance and distribution of high-flying bats anywhere in the world.

We present a bat classifier for VLR based on data obtained from multiple deployments of BirdScan MR1 radars. Our data processing scheme included novel use of deductive ecological knowledge...
and implementation of multiple AI tools. We overcame the challenge of producing a clean ‘Bat’ labelled dataset based on decades of ornithological monitoring in the Hula valley, and expert knowledge of behavioural and ecological patterns of avian fauna in the region. We isolated a period of 2 weeks when no bird migration takes place, hence avian nocturnal activity is negligible. We then compiled a list of all bird species present in Israel at this time, and assessed the avian contamination risk posed by each species to our bat dataset. Finally, we applied a rigorous filtering process to ensure a minimal amount of non-bat objects in the dataset.

We tested several types of classifying algorithms and reached over 90% accuracy while constantly maintaining minimal bird and insect cross-classification contamination. Standard diagnostic evaluations were complemented with a series of validations based on realistic ecological assumptions and the well-established BirdScan MR1 classifier. These were used to (1) validate bat-insect separation (2) assess false-positive rate based on the classification of diurnal bats detections (which are misidentified non-bat targets as no diurnal bat activity takes place in this area) and (3) additionally assess whether movement-related parameters of the bat class correspond to actual bat movement properties. The result is a robust machine learning-based classifier for separation of bats from other objects detected by VLR. We demonstrated the application and utility of this developed classifier by producing the Israeli BATScan dataset, describing large-scale phenology and directionality (>60,000 observations) of bats over 10 locations representing a wide spectrum of habitats. A wider implementation of the classifier is expected to revolutionize our understanding of bat aeroecology at a global scale.

2 | MATERIALS AND METHODS

All data processing, training and testing were done in R version 4.2.1 (R core team, 2022). A detailed description of methodology is given in Supporting Information. Data were collected using the BirdScan MR1 VLR, a 25kW, X-band, 9.4GHz vertical-looking pulse radar system, which can detect bat-sized objects up to ~800m above ground. Detected targets are automatically characterized in terms of flight track (heading, altitude, speed), wing-flapping characteristics and size. Each target is also classified to a few primary bird classes,
insects and non-biological targets (Zaugg et al., 2008; Supporting Information 1). The entire study is radar based, did not involve any contact or handling of animals and did not require any permit or ethical approval.

2.1 Hypothesis and rationale

Classifying bats using radar requires a research hypothesis and rationale due to our general inability to actually gather large-scale measurements of positively identified bats in flight. Wingbeat kinematics of bats were found to differ from these of birds (Riskin et al., 2008; Tian et al., 2006). Furthermore, bat wings mainly comprised living tissue and have a high water content, the main electromagnetic reflective substance in biological targets, which may also affect wing-related echo parameters. We accordingly hypothesized that bat radar echoes should differ from those of other flyers in wing reflection and kinematic properties, and proceeded to base our classifying approach on wing-flapping-related parameters that would be effected by such differences.

Wings have discernible effects on radar reflections from animals (Addison et al., 2022), and wing-flapping has noticeable signature in VLR, which is used for target classification (Rosa et al., 2016; Zaugg et al., 2008). We used WFF, average pulse width, average pulse length (hereafter together PP), and the pause rate (Figure 1) to describe wing-flapping characteristics, and flight altitude, speed and radar cross section (RCS, related to target size, Chilson et al., 2017; O’marca et al., 2014; Mirkovic et al., 2016), which interact differently with wing-flapping characteristics in birds and bats (Ellington, 1991; Grodzinski et al., 2009; Norberg & Rayner, 1987; Pennycuick, 2008).

We produced four training datasets, corresponding to four different classifiers (1. Full, 2. No speed, 3. No WFF, 4. No PP) using different combinations of radar parameters (Supporting Information 3.1). The radar’s ability to calculate parameters depends on echo strength and duration. Some parameters were only calculated for some of the targets; hence, we tested several parameter combinations to maximize data usage.

2.2 Training data construction

Machine learning techniques rely on externally validated datasets to characterize different classes and separate them in multidimensional space. Here we use ecological heuristics to produce ‘Bat’ and ‘Non-bat’-labelled data from the Israeli BirdScan radar dataset. The dataset is comprised of 18 million detections from 10 deployments of varying duration over 7 years throughout Israel, representing a diversity of habitats from extreme desert to agricultural wetlands and Mediterranean ecosystems (Table S1; Figure S1). Most of our data were collected at a site located in the Hula Research Centre, Hula Valley, Israel (35°43’ E, 33°03’ N), a major stopover site for migrants on the Eurasian-African flyway (Collins-Kreiner et al., 2013) and a regional biodiversity hotspot.

2.2.1 Bat-labelled data

It is practically impossible to visually or acoustically identify individual bats aloft and couple these identifications with radar signals. Here, we filtered the entire Israeli BirdScan dataset according to a specific set of conditions and radar parameter ranges which, to the best of our understanding, should contain bats almost exclusively. To assess the robustness of our approach against avian contamination of the bat dataset, we compiled a list of the 260 bird species found in Israel during the summer based on information from the Israeli bird ringing centre (https://www.birds.org.il/he/species-checklist).

We assigned each species with its WFF either from published literature (Bruderer et al., 2010), or calculated according to Rayner (1995) based on local body mass measurements of each species. The potential for contamination was then assessed based on WFF, activity pattern (during non-migratory season), occurrence and abundance in the Hula Valley region, as well as using information about the behaviour of four specific nocturnal species that might pose a contamination risk based on all the above-mentioned properties. These include ground-dwelling species which are highly unlikely to be found in relevant altitudes (two species of small owls, Athene noctua, Otus scops, which hunt near or on the ground, and the Eurasian stone-curlew, which forages and closely guards territories on the ground), and the little swift (Apus affinis) which is closely monitored by our group in a different project, suggesting that its night activity is rare and does not take place in the vicinity of the radar. We can thus conclude with high certainty that the risk of bird contamination of our bat dataset is very low.

Bat-labelled data (Figure 2a; Supporting Information 4.1) were taken from the 2018 to 2021 Hula Valley radar dataset and consisted of nocturnal targets only, during non-migratory season (which were filtered out based on directional analysis of each night’s vertebrate detections using the rayleigh.test function from R package ‘circular’, Agostinelli & Lund, 2022) when birds are largely inactive during the night. We removed signals classified as ‘insect’, ‘aeroplane’ and ‘flock’ by the BirdScan classifier, as these are very unlikely to be generated by single vertebrates (Zaugg et al., 2008). We kept targets with WFF range of 6–18 Hz which should represent all bat species in our dataset according to (1) an inverse relation to the size of the largest (Rousettus aegyptiacus) and the smallest (Pipistrellus pipistrellus) bat species in this area, respectively, using the formula given in Norberg and Norberg (2012) and (2) published WFF data of Rousettus aegyptiacus and a bat species of similar dimensions to P. pipistrellus (Bullen & McKenzie, 2002; Carpenter, 1986). We increased the upper WFF limit by ~25% to accommodate for increased WFF in foraging flight (Aldridge, 1987). Notably, this is still well outside the expected flapping frequency range of nocturnal insects in the area based on WFF estimations of different taxa, and size ranges of those in the Hula Valley region (Schaefer, 1976). We restricted RCS to a range which corresponds to animal sizes of up to 0.25 m², corresponding to Rousettus aegyptiacus, the largest species of bat in the area. RCS is a radar measure related to target size (amount of reflected
FIGURE 2 Training data construction: (a) ‘Bat’ training data compilation: We used several filters on the Hula Valley dataset to obtain a minimally contaminated bat dataset. Night data of animal targets from non-migratory periods during the summer of 2018–2021 were filtered according to WFF (wing-flapping frequency), RCS (Radar Cross Section) and altitude. The number of figurines reflects the conceptual effect of each filtering stage. (b) ‘Non-bat’ training data compilation: We used a gradual filtering approach which resulted in 5 nested datasets of diminishing size and increasing similarity to bat data in terms of location, season and radar parameters. These datasets were combined so that bat-similar data were replicated in the final non-bat dataset, increasing the percentage of bat-like non-bat data (Parameter restriction, Hula: 5 times, Parameter restriction, Israel: 4 times, etc.). (c) Training data assembly workflow: Illustration of the training data construction process from the full Israel BirdScan dataset, through parameter selection, bat and non-bat filtering, generative adversarial network up-sampling and final assembly of four training datasets corresponding to four classifiers based on different parameter combinations.
radiation) that ranges from actual object size to ~0 depending on various technical factors. (Supporting Information 3.1) but never exceeds actual size, so we only used an upper threshold of RCS values. Finally, we restricted altitude on a technical basis to above 50 m for avoiding ground clutter, and below 800 m, which is the upper detection threshold of the MR1 for the smallest insectivorous bat in the region *P. pipistrellus* (Shi et al., 2021).

### 2.2.2 Non-bat-labelled data

The 'non-bat' data (Figure 2b; Table 1) were extracted from daytime data of all classes and locations. A large proportion of day data is easily distinguishable for bats (small insects, large raptors, etc.). We aimed to create a data set that accentuates cases with parameter distributions similar to those of bats, as these would be the challenging instances for the classifier. Non-bat data similar to bat data (in terms of chosen parameters, mainly passerine-sized signals from birds and possibly also large Lepidopterans) is the most important for training a classifier to classify bats. However, it is naturally rare in a dataset containing the full spectrum of animal sizes and morphologies, including insects, large birds, etc. Accordingly, we joined several sub-datasets extracted from daytime radar data of all classes and locations (Supporting Information 4.2). Each sub-set was compiled using different filtering restrictions, to create a layered 'non-bat' dataset (Table 1) where observations which are more similar to bat samples are represented at a disproportionality higher frequency, by up to five times their actual incidence (Figure 2b; Table 2). This approach enabled training a robust classifier for separating bats from similar, non-bat, radar targets.

Finally, 'bat' and 'non-bat' training datasets were joined to form the overall training data, and these were divided to four separate datasets according to parameter availability in the data, to fit the parameter combinations of the selected classifiers (Figure 2c).

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Sub-dataset</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bat</td>
<td>—</td>
<td>12,583</td>
</tr>
<tr>
<td>Non-bat</td>
<td>Parameter restricted, Bat similar Hula</td>
<td>9,114</td>
</tr>
<tr>
<td></td>
<td>Parameter restricted, Bat similar Israel</td>
<td>15,720</td>
</tr>
<tr>
<td></td>
<td>General population, Hula, summer</td>
<td>22,3802</td>
</tr>
<tr>
<td></td>
<td>General population, Hula</td>
<td>18,53,463</td>
</tr>
<tr>
<td></td>
<td>General population, Israel</td>
<td>66,67,378</td>
</tr>
<tr>
<td>Non-bat</td>
<td>Total</td>
<td>87,69,477</td>
</tr>
</tbody>
</table>

### 2.2.3 Handling class imbalance

Classification algorithms are sensitive to the relative amount of labelled samples of each class introduced during the training stage. Here, the resulting 'Bat' and 'Non-bat' datasets were severely imbalanced, with 'non-bat' observations vastly outnumbering 'bat' observations, and thus heavily biasing the classification (Tables 2 and 3; Figure 2c). We handled this imbalance by up-sampling the bat dataset using a generative adversarial network (GAN) to produce balanced datasets with an equal number of observations in both classes (Supporting Information 4.3). A GAN generates artificial instances based on an original set by confronting two neural networks in an iterative process which gradually increases artificial sample accuracy (Zhai et al., 2022; Supporting Information 4.3). GAN data generation was done for each of the four training sets using the 'gan-GenerativeData' package in R over 10,000 iterations with discrimination probabilities of 0.95 (Goodfellow et al., 2014). We verified the adequacy and necessity of implementing GAN, which is complex and computer intensive, by training parallel sets of classifiers with data that were not up-sampled, or up-sampled using SMOTE (synthetic minority oversampling technique), which is considered a standard tool for up-sampling in imbalanced classification problems (Fernández et al., 2018).

### 2.3 Classifier construction

We used the bagged trees (BT) algorithm to produce four classifiers, one for each parameter combination (Table 5; Supporting Information 5). We trained the classifiers on 90% of the training data, keeping 10% for testing, using the 'train' function in the 'caret' R package (Kuhn, 2008) with the 'treebag' method, the 'Accuracy' metric and 10 times cross-validation. The GAN-generated bat samples were removed from the test datasets before testing to assess our algorithm's ability to handle the bias caused by the natural imbalance in the data. In addition to standard classifier diagnostics, we performed a series of ecological validation analyses (e.g. insect contamination assessment, diurnal bat classification, radar parameter evaluation of the bat class) to further assess whether our classifiers were indeed correctly classifying bats and accurately separating them from other objects (see Section 3.2).

### 2.4 Migration analysis

Bat classification allowed us to examine bat presence and directionality as a proxy for migration through the Hula Valley and to compare it to analogous passerine patterns (based on the BirdScan classifier). We calculated nightly mean flight azimuth and its standard error for each taxonomic class every night between August 2018 and November 2021 using the rayleigh.test function in the 'circular' R package (version 0.4-93; Supporting Information 6).
TABLE 2 Training data proportions: Original bat and non-bat target counts, count and proportion of artificial generative adversarial network (GAN) generated samples for each of the classifiers chosen for testing.

<table>
<thead>
<tr>
<th>Classifier</th>
<th>Bat samples</th>
<th>% generated</th>
<th>Non-bat</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full classifier</td>
<td>34,936</td>
<td>98</td>
<td>35,538</td>
</tr>
<tr>
<td>No speed</td>
<td>2,33,304</td>
<td>99</td>
<td>2,35,969</td>
</tr>
<tr>
<td>No WFF</td>
<td>4,48,814</td>
<td>99.8</td>
<td>4,49,416</td>
</tr>
<tr>
<td>No PP</td>
<td>92,160</td>
<td>98</td>
<td>93,985</td>
</tr>
</tbody>
</table>

Abbreviations: PP, pulse/pause data; WFF, wing-flapping frequency.

TABLE 3 Parameter composition of the classifiers selected for testing. Classifiers 2–4 were able to handle data lacking specific parameters allowing to considerably increase the amount of data we can classify because many radar targets lack some parameters and were consequently unsuitable for classifier 1. Nevertheless, all selected classifiers perform satisfactorily.

<table>
<thead>
<tr>
<th>Classifier</th>
<th>Altitude (m)</th>
<th>Speed (m/s)</th>
<th>Radar cross section (m²)</th>
<th>WFF (Hz)</th>
<th>Average pulse length (s)</th>
<th>Average pause length (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full classifier</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>No speed</td>
<td>2</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>No WFF</td>
<td>3</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>No PP</td>
<td>4</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Abbreviations: PP, pulse/pause data; WFF, wing-flapping frequency.

Distinguishing (-) from (+) and marking the fact that the parameter (column) was not used in the classifier (row). Grey shade indicates parameters omitted from corresponding classifier.

3 | RESULTS

3.1 | Classification

All four selected parameter spaces (Table 3), chosen according to our hypothesis and validated in an initial exploratory analysis, enabled successful bat/non-bat separation, with the full parameter space performing best (balanced accuracy, arithmetic mean of accuracy and specificity which accounts for data imbalance in the classified data—94.5%) and the parameter space without pulse-pause information being least successful (balanced accuracy—85.8%; Table 4).

Classifiers trained with the GAN up-sampled data were consistently superior to classifiers trained without up-sampled data or to classifiers trained with up-sampled data generated by the SMOTE procedure (Table 5).

3.2 | Validation

We performed several validation analyses including:

3.2.1 | Daytime classification

Given the nocturnal nature of bats in the study area, any ‘Bat’ classified during the day likely represents a false-positive classification. We performed this analysis on daytime data excluding the last hour prior to sunset (Calculated using the ‘sunset’ function in bioRad R package, Dokter et al., 2019) from two locations: Hula Valley and Sde-Boker, 262 km south of the Hula Valley, in the Negev Highlands, within the desert region of Israel. Notably these data were not included in previous stages of construction, training or testing the classifiers.

Except for Classifier 4 (No PP), all classifiers identified no more than 6% of day observations as bats, consistently less than half the proportion of bats identified at night (Table 5). In Classifier 4 (No PP), the lack of PP information probably prevented the separation of bats from passerines and produced similar proportions of bats during day and night. We decided to forego the use of Classifier 4 (No PP) based on these results. This validation supports our hypothesis that proper bat classification likely depends on complex properties of wing-flapping kinematics that differ between bats and birds.

3.2.2 | Insects

The purpose of this validation step was to assess the classifier’s robustness to insect contamination in the ‘Bat’ class. The BirdScan easily distinguishes insects from other radar targets, and the class is considered very reliable (Zaugg et al., 2008). We analysed the composition of ‘Bat’ and ‘Non-bat’ BATScan classes of the different classifiers in terms of BirdScan’s classification results and revealed a notable absence of insects in bat-classified data from all classifiers except Classifier 3 (No WFF). This classifier lacks the WFF parameter (Figure 3), and is clearly susceptible to classifying insects as bats.

This points to the importance of wing-flapping information for reliable class separation, in this case when separating bats from insects, and we consequently decided to forego the use of Classifier 3 (No WFF) at this stage.
3.2.3 | Distribution of unrestricted parameters

We compared airspeed, phenology and flight movement-related radar parameters to study the separation between bat and non-bat classified observations. We deliberately chose parameters that were not restricted during filtering in the preparation of the training data to look at the spontaneous performance of the classifier and validate that the created classes were separated independently of the filtering process. We compared bat and non-bat distributions of radar-generated parameters (average pulse length, average pause length, pulse length/pause length ratio and airspeed) and our calculation of pause rate (number of pauses/duration of signal).

Figure 4 displays the distribution of selected parameters which differ between classes and validate the existence of a distinct ‘Bat’ class. Pulses and pauses seem to be longer for bats compared to birds and shorter compared to insects (these parameters do not necessarily represent actual flapping patterns in insects, but they do represent distinct reflective properties of the class, which are nonetheless comparable, with pulse longer than pause, and with a slower pause rate. The analysis also suggests a narrower airspeed range for bats, which makes sense in light of the smaller taxonomic diversity and narrower size range of this class compared to non-bats. Bats are faster compared with insects, in line with their superior ability for active propulsion in the aerial medium.

### TABLE 4 Classifier diagnostics: Standard diagnostic of chosen classifier accompanied by further ecological validations.

Although all classifiers perform satisfactorily in terms of standard diagnostics, we disqualified Classifier 3 (No WFF) due to insect contamination and Classifier 4 (No PP) was disqualified due to inability to distinguish between bats and passerines. For full class compositions refer to Table 3.

<table>
<thead>
<tr>
<th>Classifier</th>
<th>#</th>
<th>Accuracy</th>
<th>95% CI</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Insect contamination, %</th>
<th>Day bat proportion, %</th>
<th>Day passerine proportion, %</th>
<th>Night bat proportion, %</th>
<th>Night passerine proportion, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full classifier</td>
<td>1</td>
<td>0.99</td>
<td>0.986–0.992</td>
<td>0.9</td>
<td>0.99</td>
<td>0.07</td>
<td>6</td>
<td>54</td>
<td>14</td>
<td>72</td>
</tr>
<tr>
<td>No speed</td>
<td>2</td>
<td>0.98</td>
<td>0.976–0.982</td>
<td>0.79</td>
<td>0.98</td>
<td>0.05</td>
<td>3</td>
<td>61</td>
<td>9</td>
<td>71</td>
</tr>
<tr>
<td>No WFF</td>
<td>3</td>
<td>0.98</td>
<td>0.986–0.987</td>
<td>0.85</td>
<td>0.98</td>
<td>8.83</td>
<td>6</td>
<td>52</td>
<td>16</td>
<td>72</td>
</tr>
<tr>
<td>No PP</td>
<td>4</td>
<td>0.97</td>
<td>0.972–0.976</td>
<td>0.74</td>
<td>0.97</td>
<td>0.16</td>
<td>9</td>
<td>12</td>
<td>23</td>
<td>32</td>
</tr>
</tbody>
</table>

Abbreviations: CI, confidence intervals; PP, pulse/pause data; WFF, wing-flapping frequency. Grey shade indicates poor performance in validation, which led to classifier disqualification.

3.3 | Bat aeroecology

Bats have distinct phenological characteristics compared to birds and insects (Korz et al., 2007) and we consequently explored time-series patterns of classified data from the Hula Valley radar to produce the first large-scale description of bat phenology throughout the yearly cycle. Activity plots of the bat class in the Hula Valley radar depict two phenological features which are more pronounced than in other taxa (Figure 4a). First, the absence of activity during winter corresponds to the strong reduction of activity, as well as emigration and hibernation, in many bat species inhabiting this area (Levin et al., 2015). This overall reduced appearance is evident also in birds and insects but it is much more pronounced in bats compared to the other two taxa. Second, the noticeable sharp increase at mid-June corresponds to the emergence of young resident bats that were born in the spring and begin their independent flights in this time of the year (Levin et al., 2013). The mean nightly bat flight directionality conforms with expected seasonal migration patterns (Figure 5). We found that expected seasonal migration patterns (Figure 5). We found that the mean nightly bat flight directionality conforms with expected seasonal migration patterns (Figure 5). We found that the mean nightly bat flight directionality conforms with expected seasonal migration patterns (Figure 5). We found that the mean nightly bat flight directionality conforms with expected seasonal migration patterns (Figure 5).
TABLE 5 Effect of up-sampling on the chosen classifiers: Each classifier was trained using the raw data and data up-sampled using SMOTE or generative adversarial network (GAN) and the results are displayed. The main advantage of GAN is manifested by a smaller proportion of non-bats targets that were classified as bats (false positive), as seen in the confusion matrices. Thus, classifiers trained with GAN up-sampled data produced considerably less contaminated bat classes.

<table>
<thead>
<tr>
<th>Classifier</th>
<th># Up-sampling</th>
<th>Reference</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>Accuracy (95% CI)</th>
<th>Balanced accuracy</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full classifier</td>
<td>None</td>
<td>Bat</td>
<td>33</td>
<td>4</td>
<td>0.99 (0.987-0.994)</td>
<td>0.77</td>
<td>0.55</td>
<td>0.99</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Non-bat</td>
<td>27</td>
<td>3,549</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full classifier</td>
<td>SMOTE</td>
<td>Bat</td>
<td>47</td>
<td>95</td>
<td>0.97 (0.964-0.975)</td>
<td>0.87</td>
<td>0.78</td>
<td>0.97</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Non-bat</td>
<td>13</td>
<td>3,458</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full classifier</td>
<td>GAN</td>
<td>Bat</td>
<td>54</td>
<td>31</td>
<td>0.99 (0.985-0.992)</td>
<td>0.945</td>
<td>0.9</td>
<td>0.99</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Non-bat</td>
<td>6</td>
<td>3,522</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No speed</td>
<td>None</td>
<td>Bat</td>
<td>88</td>
<td>6</td>
<td>0.99 (0.987-0.991)</td>
<td>0.74</td>
<td>0.48</td>
<td>0.99</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Non-bat</td>
<td>94</td>
<td>9,392</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No speed</td>
<td>SMOTE</td>
<td>Bat</td>
<td>142</td>
<td>134</td>
<td>0.98 (0.979-0.984)</td>
<td>0.88</td>
<td>0.78</td>
<td>0.985</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Non-bat</td>
<td>40</td>
<td>9,264</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No speed</td>
<td>GAN</td>
<td>Bat</td>
<td>144</td>
<td>159</td>
<td>0.98 (0.976-0.982)</td>
<td>0.887</td>
<td>0.79</td>
<td>0.983</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Non-bat</td>
<td>38</td>
<td>9,239</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No WFF</td>
<td>None</td>
<td>Bat</td>
<td>13</td>
<td>5</td>
<td>0.998 (0.9985-0.9991)</td>
<td>0.6</td>
<td>0.21</td>
<td>0.99</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Non-bat</td>
<td>47</td>
<td>44,936</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No WFF</td>
<td>SMOTE</td>
<td>Bat</td>
<td>28</td>
<td>195</td>
<td>0.995 (0.994-0.996)</td>
<td>0.73</td>
<td>0.46</td>
<td>0.995</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Non-bat</td>
<td>32</td>
<td>44,746</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No WFF</td>
<td>GAN</td>
<td>Bat</td>
<td>51</td>
<td>611</td>
<td>0.986 (0.985-0.987)</td>
<td>0.91</td>
<td>0.85</td>
<td>0.986</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Non-bat</td>
<td>9</td>
<td>44,330</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No PP</td>
<td>None</td>
<td>Bat</td>
<td>71</td>
<td>23</td>
<td>0.99 (0.989-0.992)</td>
<td>0.63</td>
<td>0.26</td>
<td>0.99</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Non-bat</td>
<td>195</td>
<td>23,573</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No PP</td>
<td>SMOTE</td>
<td>Bat</td>
<td>208</td>
<td>1,408</td>
<td>0.938 (0.934-0.941)</td>
<td>0.86</td>
<td>0.78</td>
<td>0.94</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Non-bat</td>
<td>58</td>
<td>2,218</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No PP</td>
<td>GAN</td>
<td>Bat</td>
<td>197</td>
<td>539</td>
<td>0.974 (0.972-0.976)</td>
<td>0.85</td>
<td>0.74</td>
<td>0.97</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Non-bat</td>
<td>69</td>
<td>23,057</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: CI, confidence intervals; PP, pulse/pause data; WFF, wing-flapping frequency.

resulting in an overall shorter migration season. In comparison to bird migration, bat activity was overall less directional (also during migration periods, Figure 5), possibly due to non-migratory activity (e.g. foraging) that takes place during the night, which is absent in nocturnal migrating passerines.

The seasonal directionality of bats is mostly evident at elevations between 200 and 600 m, a narrower range than that observed in passerines (Figure 5, bottom). Furthermore, the altitudinal separation reveals that bat activity markedly diminishes with elevation during the winter months, depicted as a reduction of point sizes and in many cases also complete absence of bat activity at the edges of the plots in the higher altitude bins.

3.4 | Israel BATScan data

Based on these tests and analyses, we formulated a general bat classification protocol for BirdScan MR1 data. We decided not to include Classifiers 3 and 4 (no WFF and no PP, respectively), based on their performance that was assessed in the validation stage. We filtered night-time observations from our full dataset according to parameter availability and used the appropriate classifier (Table 4) for each observation based on the availability of the speed parameter (which is missing in Classifier 2). The resulting dataset from applying the protocol on all Israeli MR1 data consists of ~430,000 observations (about 2% of the data collected by all BirdScan radars in Israel), roughly 15% (63,630) of which classified as bats (Table 6).

4 | DISCUSSION

The objective of radar aeroecology, and what makes it unique, is the ability to continuously describe the elusive aerial habitat and its inhabitants over a large airspace and in fine detail. Without it, we can merely scratch the bottom of this habitat, home to billions of organisms, and supplier of crucial ecosystem functions and services to humanity (Bauer et al., 2017; Voigt et al., 2018). An ability to clearly and automatically differentiate between birds and bats with radar data will greatly promote research and practice in aeroecology, conservation biology, ecological management and applied aspects of bird and bat ecology.

We critically supervised training data assembly, AI algorithm implementation, diagnostics and evaluations, adapting each stage to our specific problem and hypothesis regarding the effects of wing
reflectivity, wingbeat kinematics and size related parameters on organism reflectivity and classification. Phenology and morphology were central in our data assembly and parameter selection. Thus, processing non-bat targets which are morphologically and phenologically similar to bats was central to our approach. Non-bat radar data comprise a huge spectrum of targets, of which bat-similar targets comprise a small fraction. This proved problematic for training a robust classifier, and we believe that artificially increasing the proportion of bat-similar targets in the non-bat data helped refine the separation and allowed the classifier to learn how to distinguish them properly.

We believe that it is safe to conclude that the final product of our classification scheme is indeed a well-separated dataset with an acceptable amount of error, well within typical ranges (Rosa et al., 2016; Zaugg et al., 2008). High accuracies and low error rates typically serve as a sufficient indication of classifier performance, but due to the lack of ground truth data in our process, we further validated the separation based on known aspects of bat behaviour and movement, and the well-established BirdScan classifier (Zaugg et al., 2008).

Our choices throughout the classification process aimed at producing a bat class as clean and uncontaminated as possible, such that some bat echoes might have been removed in the process. We followed a conservative approach whenever a choice had to be made during the construction of the training datasets (migration filtering, radar parameter ranges), data usage and classifier selection. We negated insect contamination in bat classified data based on the BirdScan class composition (Figure 3), such that Classifier 3, contaminated with ~8% insects, was disqualified. The negligible number of classified bats in daytime data indicates a low false-positive rate (Table 5) and proves that our classifiers capture the separation between the groups and implement it properly. The phenomenology of bats that we described using classified data fits very well with known phenomenology of bats in the area, including hibernation of some of resident species (Levin et al., 2015) and absence of migratory species during the winter, as well as the start of independent flights of young bats in the early summer (Figure 4a; Levin et al., 2013). These patterns point to the reliability of BATScan.

The differential performance of each classifier serves as an indication of the importance of each parameter to the classification process (Table 5). The superiority of the full classifier indicates the importance of the information provided by each of the radar parameters. The fact that the absence of secondary wing-flapping information (pulse and pause length) caused the strongest reduction in classification performance, and that this classifier was eventually disqualified based on ecological validation, confirms our hypothesis regarding the importance of wing-related properties, likely associated with morphology and wingbeat kinematics, in separating bats from other taxa.

**Figure 3** BirdScan class composition of BATScan classified data according to the applied classifier: All classifiers mainly classify bats as ‘passerines’ BirdScan class, with smaller amounts of other classes. When data about target WFF (wing-flapping frequency) is absent (Classifier 3, No WFF), insect contamination is introduced into the bat data, which led to the decision not to include it in the final classification protocol.

**Figure 4** (a) Phenology of the three main aerial taxa in the Hula Valley Israel, bat, birds and insects. Movement Traffic Rate (MTR) is a measure of animal flow through the airspace, which corrects for biases due to changes in the scanned volume of the BirdScan at different altitudes. We log transformed Nightly MTR to negotiate multiple orders of magnitude in the data. Bats were classified with BATScan while non-bat targets were further classified to birds and insect using the BirdScan classifier. The sharper decrease in bat activity during the winter corresponds with hibernation of local bats and the emigration of migratory bats from the area during this time of the year. The noticeable peak in MTR during mid-June corresponds with an increase in the bat population due to juvenile bat pups that begin their independent flights at this time of the year and with migration influx from the south, specifically of *Rhinopoma microphyllum*. (b) Unrestricted distributions of wing movement parameters and airspeeds of bats, birds and insects that were recorded by BirdScan radars. Bats are more similar to birds than to insects in all aspects. Between bats and birds, bats have a longer pause, slightly longer pulse, larger pulse/pause ratio and a slower pause rate. Bats seem to be characterized by overall somewhat lower airspeeds compared with birds, while the airspeeds of insects were considerably lower than those of the two vertebrate groups.
In VLR, each observation usually describes an individual animal passing through the detection range. The current BATScan dataset is comprised of ~430,000 observations, 63,630 of which are of bats. The data were collected in various habitats, including deserts, wetlands, agricultural areas, plains, hills and mountains during a 7-year period. This may possibly be the largest dataset of individual flying bats ever collected, and it will constantly expand as BirdScan radars continue to operate in Israel and elsewhere. This dataset can
FIGURE 5  Migration analysis: Top: yearly phenology: Average night orientation on the north-south axis. Point colour corresponds to directional standard error of directionality, point size indicates sample size (number of bats) normalized to maximal nightly sample size for each taxonomic class separately. Middle: Directionality variability for each group throughout the year, demonstrating the time-shifted migration in bats. Bottom: Average directionality in 100 m altitude bins. Point colour corresponds to directionality standard error of each point, point size indicates sample size relative to maximal nightly sample size for each class and altitude bin separately.
greatly assist in understanding multiple aspects of bat aeroecology such as aerial habitat selection, dispersal, foraging, migration, etc. BATScan is constructed to handle BirdScan MR1 data, but the process can easily be adapted to other radar platforms by replicating the multi-step training and validation workflow described here. This successful endeavour of bat classification serves as a stepping-stone towards development of similar classifiers for other radar platforms. This can be achieved using similar construction processes, or based on conclusions drawn from BATScan classified data, or with BirdScan MR1 radars serving as ground truth for other systems, for example weather radars (Nilsson et al., 2018).

A reliable bat classifier will give a new perspective into bat research, literally enabling us to see in the dark and study aerial ecology, movement and behaviour of bats at multiple spatial and temporal scales. Implementing BATScan on the currently deployed network of BirdScans will shed insight into large processes like bat migration, foraging and the response of bats to different meteorological conditions in Europe, the Middle East and North America.

An additional beneficial outcome is the reduction of bat contamination in data related to other taxa. A quick calculation suggests that in the Israel BirdScan dataset, signals of nocturnal passerines may contain up to 10% bat contamination during busy migration periods and much higher proportions of up to almost 100% outside of migration seasons when passerine nocturnal activity is scarce. Reducing bat contamination will improve passerine-related trends and conclusions obtained from the data, to better understand the specific aeroecological properties of each animal group.

BATScan classification revealed novel insights regarding bat migration which have never been described in this area. Spring migration continues well into the summer, with evident northward directionality in June corresponding to known insectivore bat species movement into the area (Levin et al., 2013). The onset of bat autumn migration is accordingly shifted forward in time, becoming pronounced in September. Our analysis suggests bat migration takes place at a narrower altitude range compared to passerine migration (Figure 5), but this pattern may be affected by the relatively high abundance of non-migratory bat movement in lower altitudes (which makes the overall movement less directional). The potential of this method for studying large-scale movement patterns of bats is clear. VLR’s main strength lies in its ability to produce fine scale, long-term datasets, which are currently unobtainable for bats, preventing comprehensive understanding of bat phenology, abundance and distribution anywhere in the world. By distributing VLRs across different regions and within ecologically important sites, bat migration patterns can be described in high temporal resolution at multiple spatial scales. BATScan enables this type of analysis for the first time.

BATScan also has far-reaching conservational significance, as local radars are central in the ability to monitor and survey the consequences of human developments into the airspace for aerial wildlife. Currently, radars are largely unable to provide practical insights related to bats in this regard.

## 5 | CONCLUSIONS

Radar aeroecology is key for habitat scale study of the aerial wildlife. The field currently lacks the ability to identify bats, leaving their behaviour and whereabouts in the dark and negatively impacting the integrity of non-bat information due to bats being classified as other taxa. Accelerated human development into the airspace for transportation, accommodation, communication and energy production calls for the implementation of appropriate tools to reliably document all the inhabitants of the aerial habitat. BATScan, the first operational bat classifier, will greatly improve conservation efforts, promote sustainable development and facilitate exciting avenues of research.

## AUTHOR CONTRIBUTIONS

Yossi Yovel and Nir Sapir conceived the idea, designed the general methodology and oversaw the project and manuscript writing. Hadar Sextin conducted the initial exploration and produced initial results. Yuval Werber wrote all code except for initial exploration, acquired and constructed the datasets and led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

## ACKNOWLEDGEMENTS

We wish to acknowledge the Israel Science Foundation (grants no. 2333/17 and 1653/22) and the University of Haifa for purchasing the Hula BirdScan radar and supporting the study, and the Hula Research Centre for housing the radar. We thank David Troupin, Birgen Haest and Baptiste Schmid, and Dominik Kleger, Fabian Hertner and Felix Liechti (Swiss Birdradar Solution AG) for help and advice throughout the project. We thank Itai Bloch for his input regarding nocturnal activity of swifts in the Hula Valley. We further thank Mor Taub and Aya Goldstein from Tel Aviv University and Emmanuel Lourie from the Hebrew University for help and advice in preparing the manuscript. We thank EDF Renewables Israel and Blue Sky Energy, Asaf Mayrose and Zev Labinger (Houbara Environmental Consultancy), David Meninger (Geoteva Environmental Consultancy), Eyal Shochat
(Geoteva Environmental Consultancy and Ben-Gurion University), Itamar Giladi and Ofir Altshtien (Ben-Gurion University) and Oded Keinan (Arava Environmental Science Education Center) for help in data collection and sharing.

CONFLICT OF INTEREST STATEMENT
All authors declare no conflict of interests.

PEER REVIEW
The peer review history for this article is available at https://www.webofscience.com/api/gateway/woes/peer-review/10.1111/2041-210X.14125.

DATA AVAILABILITY STATEMENT
Israel BATScan data sample, classifiers, codes and other relevant datasets are available at https://figshare.com/projects/BATScan_data_MEE/150732.

ORCID
Yuval Werber https://orcid.org/0000-0002-9858-1872
Yossi Yovel https://orcid.org/0000-0001-5429-9245
Nir Sapir https://orcid.org/0000-0002-4777-0515

REFERENCES
Brunton, D. F. (2018). Swimming as a potentially important emergency


SUPPORTING INFORMATION

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

Appendix S1. Supplementary methods.

---