proceedings

Proceedings

Detecting birds and insects in the atmosphere using machine learning on NEXRAD radar echoes. †

Precious Jatau [1,*], Valery Melnikov [2] and Tian-You Yu [3]

[1] Advanced Radar Research Center, Cooperative Institute for Mesoscale Meteorological Studies, and School of Electrical and Computer Engineering, University of Oklahoma, Norman, OK 73019, USA; preciousjatau@ou.edu
[2] Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma and National Severe Storms Laboratory, Norman, OK 73072, USA; valery.melnikov@noaa.gov
[3] Advanced Radar Research Center, School of Electrical and Computer Engineering, and School of Meteorology, University of Oklahoma, Norman, OK 73019, USA; tyu@ou.edu

* Correspondence: preciousjatau@ou.edu; Tel.: 405-371-2750
† Presented at the 4th International Electronic Conference on Atmospheric Sciences, online, and 16-31 July 2021.

Abstract: NEXRAD radars detect biological scatterers in the atmosphere, i.e., birds and insects, without distinguishing between them. A method is proposed to discriminate bird and insect echoes. Multiple scans are collected for mass migration of birds (insects) and coherently averaged along their different aspects to improve the data quality. Additional features are also computed to capture the dependence of bird (insect) echoes on their aspect, range, and spatial locality. Next, ridge classifier and decision tree machine learning algorithms are trained on the collected data. For each method, classifiers are trained, first with the averaged dual pol inputs and then different combinations of the remaining features are added. The performance of all models for both methods, are analyzed using metrics computed on a held-out test data set. Further case studies on roosting birds, bird migration and insect migration cases, are conducted to investigate the performance of the classifiers when applied to new scenarios. Overall, the ridge classifier using only dual polarization variables was found to perform consistently well on both the test data and in the case studies. This classifier is recommended for operational use on the US Next-Generation Radars (NEXRAD) in conjunction with the existing Hydrometeor Classification Algorithm (HCA). The HCA would be used first to separate biological from non-biological echoes, then the ridge classifier could be applied to categorize biological echoes into birds and insects. To the best of our knowledge, this study is the first to train a machine learning classifier that can detect diverse patterns of bird and insect echoes, based on dual polarization variables at each range gate.

Keywords: machine learning; weather radar; aeroecology; algorithm; birds; insects

Citation: Lastname, F.; Lastname, F.; Lastname, F. Title. *Proceedings* **2021**, 65, x. https://doi.org/10.3390/xxxxx

Received: date
Accepted: date
Published: date

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. Introduction

The Next-Generation Radar (NEXRAD) network consists of 160+ S-band polarimetric Doppler weather radars (WSR-88D), deployed across the continental US, Alaska, Hawaii and Puerto Rico. Each WSR-88D measures six variables comprising of three single polarization variables and three dual polarization variables. The single polarization variables are the radar reflectivity factor (Z) which is proportional to the power of the received signal, Doppler velocity (V_r) which is determined from the power-weighted mean Doppler frequency shift of targets within the radar sampling volume and spectrum width (σ_V) which measures the variability of Doppler velocities within the sampling volume [1]–[3]. The dual polarization variables include differential reflectivity (Z_{DR}), the logarithmic ratio of the reflectivity factors from the Horizontal (H) and Vertical (V) polarizations, differential propagation phase shift (Φ_{DP}), the difference in phase shift between H and V polarizations and cross correlation coefficient (ρ_{HV}) which measures the diversity in type, shape and/or orientation of scatterers in the sampling volume [1].

In addition to weather echoes, the WSR-88D can detect biological scatterers such as birds, bats and insects [4], opening potential applications for broad scale studies of their behaviour. For example, large-scale radar monitoring can improve our understanding of the spread of avian diseases by allowing a detailed mapping of migratory flyways [5], [6]. Additionally, bird strikes are a serious aviation hazard for low-level flights [6]. NEXRAD can be used as a bird surveillance system, thereby improving aviation safety [6], [7]. Another application is the identification of wind tracers. Insects have been found to be mostly driven by the wind during flight while birds are active fliers and can contaminate the wind derived from radar. Previous research has been dedicated to retrieve velocities contaminated by birds, using the features of reflectivity and Doppler velocity fields [8]–[10]. Therefore, correctly separating insect echoes from bird echoes can improve the quality of radar wind products.

Many advances have been made in characterizing hydrometeor types [11]–[13]. However, the classification of biological echoes is still an active research field[7], [14]–[16]. A major obstacle to classifying biological echoes is that the shapes of birds and insects are strongly non-spherical [4]. Moreover, polarimetric measurements have a strong dependence on their size, shape and orientation [4], [17]. Thus, even in single-species ensembles, polarimetric quantities could have high variance depending on the azimuthal orientation [18]. This sometimes leads to similar measurements for bird and insect echoes, making it difficult to differentiate them. For example, the Z_{DR} of Purple Martin colonies have been found to range from -4 to 6 dB [19] while insects have been found to have Z_{DR} between 2 and 9 dB [20].

Various methods have been explored to detect biological echoes with radars[14]–[16], [21]–[25], though much less work has been done in distinguishing bird and insect radar echoes. Nonpolarimetric radar was used in [26] to discriminate these echoes by measuring their radar cross sections within close ranges from radar. However, only two cases were examined with this approach. A fuzzy logic algorithm was also developed for separating birds and insects echoes in [7]. But the use of Z as an input complicates the resolution of densely aggregated insects and sparse groups of large birds.

Machine learning models have been trained for detecting roosting birds in weather radar echoes focused on identifying their distinct toroidal shape. Convolutional neural networks were used in [27] to detect whether an individual radar image contained at least one Purple Martin or Tree Swallow roost, with correct predictions made about 90% of the time. Another machine learning system was developed in [28] that locates roosts within images and tracks them across frames. Although these methods are useful, they are designed to detect one orientation of birds while using the entire radar image as an input. They cannot be applied to a single range gate and cannot be used in situations where birds are not roosting.

We propose a machine learning model that can classify diverse orientations of bird and insect echoes, from a single radar range gate. Two supervised machine learning methods are investigated: ridge classifier and decision tree. Dual polarization radar scans containing separate large scale bird and insect migration were collected (Section 2). Next, the migrating bird (insect) echoes are segmented using blob coloring and then their textures were computed (Section 3). Velocity azimuth display (VAD) is applied to change the measurement coordinates from being relative to the radar to be relative to the target and multiple bird (insect) scans are averaged to reduce contamination by other echoes in section 3. The averaged scans are used to derive training inputs for the classifiers. The next sections summarize machine learning methods used (Section 4) and the metrics for evaluation (Section 5). Both machine learning methods are trained, first on only dual polarization variables and then on different combinations of the remaining features (Section 6). Their performances are evaluated using metrics computed on test data (Sections 7). Further case studies (section 8) are conducted to analyze performance on new scenarios from different WSR-88D radars. Finally, our conclusions and recommendations are presented in Section 9.

2. Data collection

2.1. Selection of bird and insect scans

The radar resolution volume is much larger than biological targets. As a result, it is common for a single volume to contain a mix of birds, insects and weather. Ideally we would want a homogenous composition of echoes. Furthermore, biota echoes can cover a large area (hundreds of kilometers). So it is impossible to inspect and label each volume. These creates difficulties in obtaining ground truth. Our approach was to collect multiple scans of mass bird (insect) migration in clear air (this term is used for radar observations free from precipitation), where we expect to obtain the highest possible number of range gates containing birds (insects). We leveraged some known features of biological echoes to accomplish this step. A substantial part of nocturnal echoes in spring and fall have been found to be migrating birds [20]. Such migration is characterized by a large area of echoes centered around the radar site with velocities having the same direction (highly aligned). Further analysis of birds (Purple martins) has shown that they have modes around 0 dB for Z_{DR} and 110° for Φ_{DP} [29]. Insects on the otherhand are commonly observed in clear air during warm seasons in Oklahoma, reaching peak intensity in the late afternoon [20]. They have also been found to have higher Z_{DR} than birds, with most echoes saturating at the 8 dB limit of the WSR-88D and lower Φ_{DP} [29]. These properties were used to select 90 clear air scans in which 45 PPIs are dominated by migrating birds and the other 45 scans by insects. All scans were collected from KTLX (located near Oklahoma City, Oklahoma) at the 0.5° elevation. Each scan was also chosen such that the majority of echoes were due to biological migration activity. This is important for the subsequent extraction of migration echoes by blob coloring. Finally, all gates with range less than 10 km were excluded to reduce contamination by ground clutter.

2.2. Selection of radar variables for machine learning algorithms

None of the single polarization variables are used in training our models. In general, birds are larger than insects. Since Z depends on the size of targets within the radar resolution volume, it could be a good differentiator. However, Z also depends on their population. In other words, a large Z value could be due to a sparse group of larger birds or a dense aggregation of insects. Because of this ambiguity in interpretating Z, it was excluded from the inputs to the model. Similarly, birds typically fly faster than insects. However, biological targets leverage the underlying wind field to aid their flight. As a result, passively flying insects on a windy day could migrate with larger velocities than actively flying birds in a mild wind field. Furthermore, radial projection and aliasing complicates the interpretation of V_r. Thus, V_r is excluded as an input to the model though it is used for VAD analysis to recover measurements from the target's aspect. Signal-to-noise ratio from biological scatterers are frequently low for the reliable measurement of spectrum width (σ_V). Due to this high noise contamination, σ_V is also excluded.

Dual polarization variables have been used for the identification of biological echoes [7], [14]–[16], [20], [29]. In this work, they are used in training our model to distinguish between birds and insects. Using only dual polarization variables also has the advantage of ensuring temporal coherence. Biological echoes are predominant at the lowest antenna elevation scan of 0.5° in clear air. At this elevation, the WSR-88D completes two sweeps, about 30 seconds apart. The first (surveillance) sweep measures the dual polarization variables and Z. The second (Doppler) sweep measures the legacy single polarization variables. Combining variables from both sweeps could introduce errors. Assuming a target flies along a radial at a speed of 10 ms^{-1}, it would have migrated 300 m (more than the length of one range gate) between both sweeps.

3. Feature processing to prepare inputs

In this section, further processing is performed on the collected bird (insect) scans to prepare inputs for training and evaluating the classifiers. All scans in the data set are for highly aligned migration cases. First, the texture of each dual-pol variable is computed for each scan. Next, blob coloring and minor region removal are used to extract only range gates containing migrating birds (or insects), followed by VAD analysis to find their heading. The next step is critical. Ideally, we would desire a bird migration scan to be purely comprised of bird echoes. However, they usually also contain few insect echoes. Similar scenarios often occur for insect migration scans as well. We propose a way of coherently averaging multiple scans, along range and the target's aspect to improve reduce this contamination in our training data.

3.1. Texture

Many image operations are performed on a local section defined by a window. Such windows are usually described with respect to a reference pixel, where the result of any computation is output. In our case the reference pixel is the middle one. Textures are the result of one of such operations, that characterizes the spatial variation of radar variables in the two-dimensional fields i.e. azimuth and range directions [13], [29]. We calculated texture using an 8-connected window, which a 3x3 grid of pixels with the reference at the center. Mathematically, at a given radar gate with range r and azimuth angle ϕ, the mean absolute deviation of a variable x from its neighbor gates within the window is calculated as

$$\Delta x_{r,\phi} = \frac{1}{N-1}\sum_{i=-1}^{1}\sum_{j=-1}^{1}\left|x_{r,\phi} - x_{r+i,\phi+j}\right| , \quad (1)$$

where i is the range gate offset, j is the azimuth offset, and N is the window size. Calculations were performed only when all the surrounding gates contained echoes.

3.2. Blob coloring and minor region removal to extract migration echoes

Blob coloring is an image processing method used to identify connected groups of pixels with the same value [30], [31]. It is applied to create a mask that indicates which range gates contain migrating echoes. Let us define some relevant terms before describing the algorithm. All definitions are with respect to a binary image where a pixel either contains a target (pixel value 1) or background (pixel value 0). A region (or blob) is a group of contiguous pixels with the same value. Two types of windows were applied in this study. The first is the previously described 8-connected window. The second window type is the 4-connected which refers to a reference pixel, and the neighbours directly above, below, left and right. Another operation performed is dilation, which involves iterating a window over an image and setting the pixels of interest at each step as the OR of all pixels within the window. The result is an expanded region of target pixels.

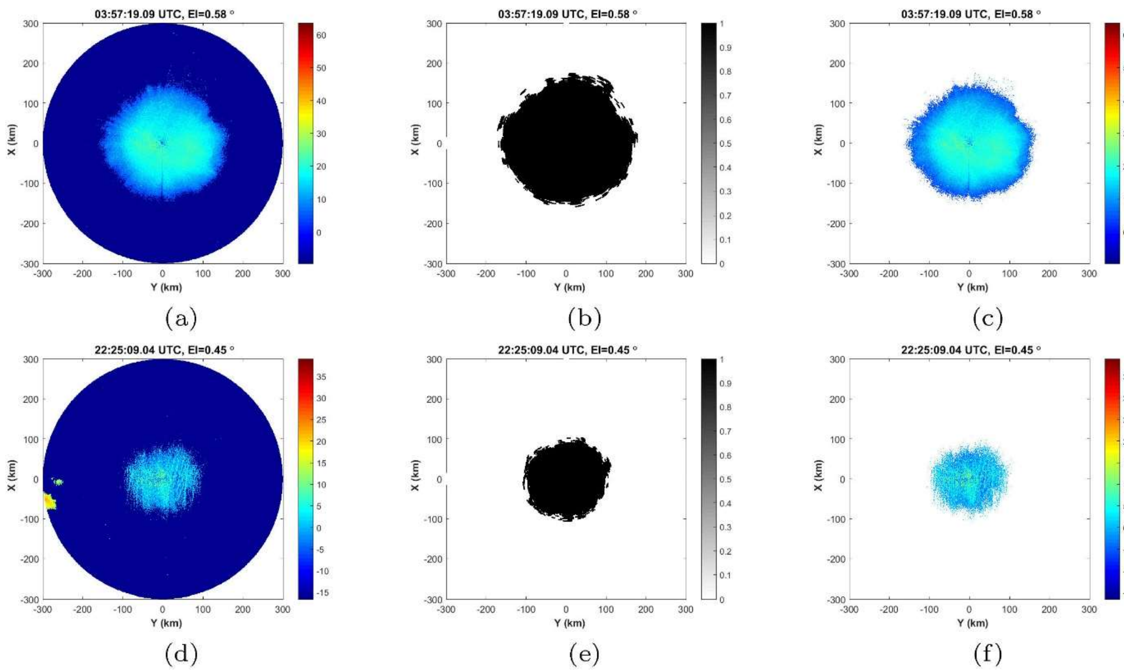


Figure 1. Blob coloring with minor region removal to extract large scale migration echoes. (a) Reflectivity (in dBZ) of bird migration echoes. (b) mask of bird migration echoes. (c) gates containing birds extracted using mask. (d) Reflectivity (in dBZ) of insect echoes (e) mask of insect echoes (f) gates containing insects extracted using mask

Data for bird and insect migration were collected for clear air days, which are characterized by a large area of biological echoes centered around the radar. An example for bird migration Z in clear air is shown in Figure 1a within a maximum range of approximately 150 km. Z is only chosen for demonstration, it is not used at any other point of this study. The data matrix for this scan can be considered as an image I where rows correspond to ranges and columns correspond to azimuth angles. The blob coloring with minor region removal algorithm is implemented as follows. First, the radar image I is binarized by setting all gates containing echoes to 1 while the remaining gates are set to 0. The second step involves dilating the binarized image twice to connect echoes at the fringe of the migration blob that might be isolated. The dilated image J is given as

$$J = (I \oplus B) \oplus B, \tag{2}$$

where $\oplus$ is the dilation operator, and B is the 8-connected window. In the third step, a region labelling algorithm [30] is applied to identify the different target regions in J. Next, minor region removal [30] is applied, where the largest target region is retained and the remaining target regions are set to background pixels. Often, this major region contains few isolated holes of background pixels. These holes are plugged, by complementing the image, repeating the blob coloring with minor region removal algorithm and re-complementing the image [30]. The resulting mask M is a binary image with one solid target blob (shown in Figure 1b) indicating the region containing migration echoes. The final image K (Figure 2c) is extracted by the element wise multiplication of the map M and the original image I. This is expressed as

$$K = I \odot M, \tag{3}$$

where $\odot$ represents the multiplication operation. This image would contain the migrating birds. The same procedure is repeated for insect cases. Figure 1d shows Z for insects with a minor precipitation region west of the radar. The generated map excludes this minor region (Figure 1e). The final extracted echoes would contain insects (Figure 1f).

3.3. Reference with respect to target azimuth

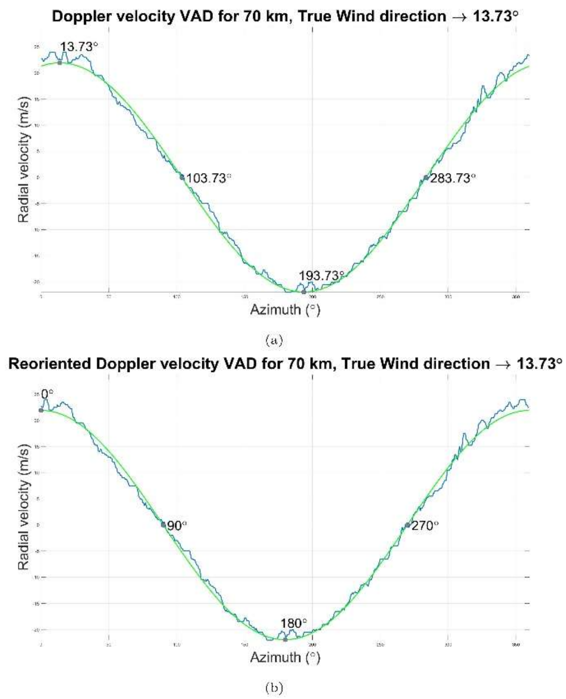


Figure 2. VAD analysis to reorient velocity to be relative to the target aspect for 70 km range gates. The blue line shows filtered V_r while the green line is the sine fit. (a) Initial VAD finds targets to be migrating toward 13.73° (b) Reoriented fit shifts measurements so migration is toward 0°.

Because of the non-spherical shape of biological targets, their radar returns would depend on the angle from which they are observed, hereafter referred to as their aspect. As such, methods for identifying biological echoes will have to account for this dependence. Cases of wide spread alignment can leverage traditional VAD [2], [32], [33] or azimuthal patterns in the correlation coefficient [18] to recover aspect information. We used VAD to rotate the variables, so they become a function of their aspect azimuth (ϕ_{aspect}). First a sinusoid model is fit to V_r at every range,

$$\widehat{V_r}(\phi) = |V|cos(2\pi f\phi + \delta), \tag{4}$$

where $\widehat{V_r}$ is the fitted radial velocity, ϕ is the radar's azimuth, $|V|$ is the magnitude of velocity along the migration direction, f is frequency and δ is a phase offset (in degrees). It is assumed that the wind field is uniform at every height so $f \approx \frac{1}{360}$ cycles/degree. The migration direction is defined as where the targets are heading. It is obtained as the radar azimuth that maximizes $\widehat{V_r}$,

$$\phi_{migration} = argmax_{\phi}\widehat{V_r}(\phi), \tag{5}$$

This direction captures measurements from the tail aspect. Scattering from other azimuthal aspects can be deduced by the lag from $\phi_{migration}$ as

$$\phi_{aspect} = \phi - \phi_{migration}, \tag{6}$$

such that a ϕ_{aspect} of 0° represents the tail region of biota, 90° represents the left wing region and 180° represents the head region.

An example for this procedure is shown in Figure 2. Figure 2a shows the VAD at range 70 km for one of the scans in the training set. The blue line is the filtered velocity obtained by applying a 10th order one dimensional median filter on V_r. The green line is the fitted $\widehat{V}_r$. Migration was found to be toward 13.73°. The radial velocity w.r.t to the target $V_r(\phi_{aspect})$, shown in Figure 2b, is obtained by shifting $V_r(\phi)$ to the left by 13.73°. Migration would be toward ϕ_{aspect} of 0°. This process is applied at every range ring to find the migration direction and rotate all dual polarization measurements and their textures. All measurements are now relative to the aspects of the targets

3.4. Averaging bird and insect cases

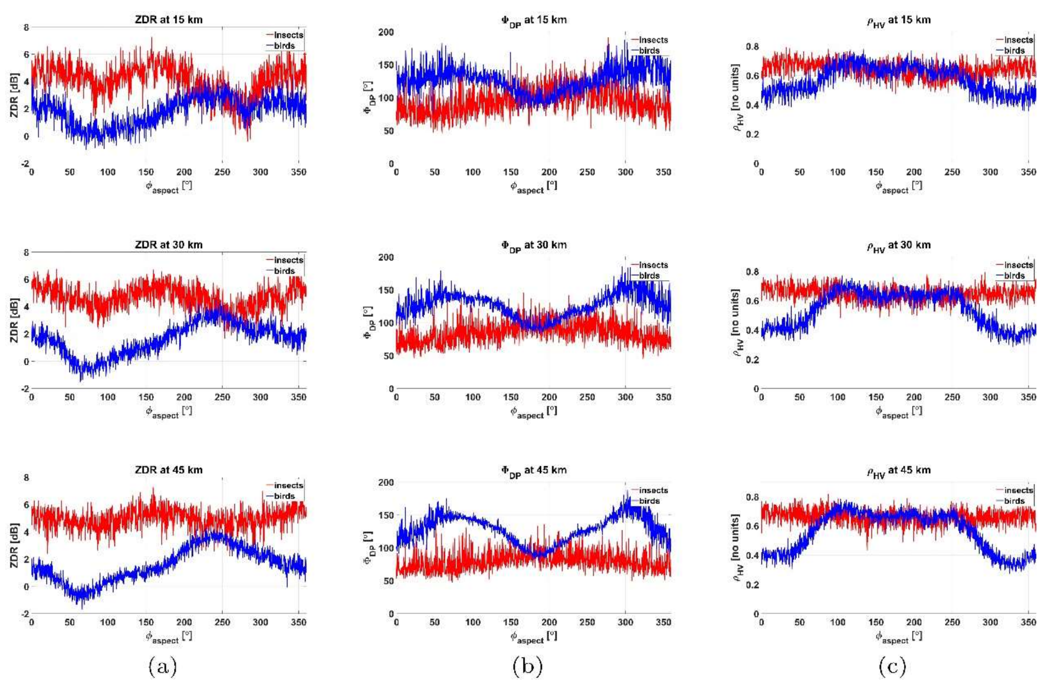


Figure 3. Averaged dual polarization variables as a function of ϕ_{aspect} for the training set. (a) Z_{DR} vs ϕ_{aspect} (b) Φ_{DP} vs ϕ_{aspect} (c) ρ_{HV} vs ϕ_{aspect}. Birds are in blue and insects red. Rows represent ranges of 15, 30 and 45 km from the radar.

To reduce the contamination of bird migration cases by insect echoes and vice versa, multiple scans are averaged. Following blob coloring and rotation of the collected scans and their textures, they are grouped into 3 batches. Let us call them batches A, B and C. Each batch contains 15 randomly selected scans per class (a total of 30 scans per batch). We will focus on A though all steps discussed equally apply to B. Each scan will have different azimuths, so we created a new range and aspect azimuth grid both starting at 0 and with resolutions of 250 m and 0.5° respectively. All scans were interpolated to this common grid. The new 15 scans for birds (insects) are then averaged. In the last step, all range gates in the resulting averaged scans from A and B are combined to form the training set, containing 1,711,624 samples: 57% bird and 43% insect cases. Batch C is used as the test set. It is not averaged so that it represents the kind of measurements we expect from the WSR-88Ds. The test set contained 9,402,821 range gates with 60% bird and 40% insect cases.

Some visualizations of the averaged training cases are shown in Figure 3. The blue curve is for birds and the red for insects. Each plot is for a dual polarization variable against the target aspect at specific ranges. From top to bottom, rows correspond to measurements 15, 30, and 45 km from the radar. From left to right, columns correspond to Z_{DR}, Φ_{DP} and ρ_{HV} respectively. The averaging procedure shows that dual-pol variables have a strong dependence on ϕ_{aspect} and exposes clear delineations between birds and insects. The results are also consistent with previous literature. Analysis in [19] found that echoes attributed to birds (Purple Martins) had Z_{DR} between -4 and 6 dB. In our case, the averaged Z_{DR} (shown in Fig. 4a) for birds is generally low, between -2 and 4 dB. The highest value is around 230° (between the head and right wing) and the lowest around 75° (between the tail and left wing). Insects were found to have high Z_{DR} (up to 10 dB) in [20]. Our averaged insect Z_{DR} is also generally higher with most gates between 3 and 6 dB. Interestingly, the values dip below the bird Z_{DR} values between ϕ_{aspect} of 230° and 300°. Φ_{DP} (shown in Figure 3b) for birds are generally higher than insects, with peaks around 50° and 300°. This is consistent with the observed symmetry of Φ_{DP} about the direction of migration [20]. Insects have lower Φ_{DP} values. ρ_{HV} (Figure 3c) for bird migration have been observed to have low values corresponding to tail-on viewing angles and high values for head-on angles [18], [19]. This can be seen in the sinusoid-like pattern in Figure 3c, with high values (around 0.7) between 60° and 250° and low values (around 0.4) otherwise. Insects generally have higher ρ_{HV} than birds with a mean value around 0.7.

After the averaging procedure, both the training and test data sets are normalized. The mean and standard deviation for each variable was computed from the 60 scans in batches A and B. They are used to normalize each variable by mean centering and scaling by their standard deviation. This ensures that all variables are on the same scale. The same procedure was applied to normalize their textures.

3.4. Input features for classifiers

The normalized dual polarization variables and their textures are used as input features for the classifiers. Additionally, inspection of the data revealed that variables varied gradually with range and ϕ_{az}. Thus, two new discrete features were created to capture this variation. The first is range interval, which refers to 10 km bins. The second is sector, which refers to 20° sectors computed from ϕ_{az}. All echoes collected in this study was from 10 to 230 km, so $range\ interval$ would contain 22 elements. The first element is ignored since we did not consider ranges below 10 km. $sector$ would contain 18 elements.

4. Machine learning methods

Our goal was to train an algorithm for distinguishing bird from insect echoes, that could be implemented operationally on NEXRAD. Traditionally, fuzzy logic has been used for classifying weather radar echoes. However, we opted for a supervised machine learning (ML) approach mainly because they predict probabilities for each range gate in addition to predicting output classes. They can also be easily updated as new data is available since they learn a model that minimizes prediction errors on the training data.

More complex neural networks have been successfully applied to detect [27] and track roosting birds [28], however they were not designed to make classification on a single radar range gate, and rather use a rendered image of a full radar scan as input. They are also trained to specifically detect birds emerging from their roosting sites. As such, these networks cannot be generalized to other patterns of bird activity or types of biological echoes. In this study, we investigate a supervised ML approach for distinguishing birds and insects, that can use inputs from a single range gate, are able to provide a probability that a range gate contains birds (or insects) and are easy to retrain as more data is collected. We explored using both the ridge classifier and decision tree.

The ridge classifier learns a linear combination of input variables that achieves the best separation between classes in the output. We used the SGDClassifier [34] in scikit-learn. For a single range gate, the function is given as

$$f(\boldsymbol{x}) = \boldsymbol{w}^T x + b, \tag{7}$$

where **w** is the weight vector , **x** is the input vector and b is a bias term. The goal is to find parameters that minimizes the log loss error given by

$$L\big(y_i, f(x_i)\big) = \log\Big(1 + \exp\big(-y_i f(x_i)\big)\Big), \tag{8}$$

where y_i is the label for each training example. A scaled L2-norm of the weights is also added to the above loss to stabilize learning by penalizing any explosion of the weights. The final loss function is given by

$$C(w, b, \alpha) = \frac{1}{n}\sum_{i=1}^{n} L\big(y_i, f(x_i)\big) + \alpha|w|_2^2, \tag{9}$$

where n is the number of training examples and α controls the effect of the weight penalty. The learning process uses stochastic gradient descent [35] on w and b, and a search on α to find values that minimize $C(w, b, \alpha)$.

Our second technique, decision trees, learn rules to recursively partition data so that samples with the same labels are grouped together. We used the DecisionTreeClassifier [34] from scikit-learn. Within the context of decision trees, an attribute A is a question asked about the data e.g is $Z_{DR} > threshold$? Answers to this question, like True or False, are called values V_k and are used to partition the data set. There are also two classes c_j containing p positive examples (birds) and n negative examples (insects). The entropy of an attribute measures its homogeneity. It is defined as

$$E\left(p(c_j), \ldots, p(c_m)\right) = \sum_{j=1}^{m} -p(c_j)\log_2 p(c_j), \tag{10}$$

where $p(c_j)$ is proportion of the jth class. High entropy indicates a uniform distribution over classes while low entropy indicates the dominance of some classes. Information gain measures the reduction in entropy for a given split. It is defined as

$$G(A) = E\left(p(c_j)..p(c_m)\right) - \sum_{k=1}^{l} \frac{n_k + p_k}{n + p} E\big(p(c_j)..p(c_m)\big|V_k\big), \tag{11}$$

where n_k and p_k are the number of positive and negative examples respectively in the kth split. In order words, $G(A)$ is the difference between the entropy before a split and the mean entropy after the split. Decision trees learn by finding attributes that maximizes $G(A)$.

5. Metrics

We used four metrics to assess our classifiers. They are accuracy (ACC), true positive rate (TPR), true negative rate (TNR) and area under curve (AUC). Table 1 below shows the confusion matrix for our classification problem. Birds are used as the positive class, so true positives (TP) are birds that are correctly classified as birds, false positives (FP) are birds classified as insects, false negatives (FN) are insects classified as birds and true negatives (TN) are insects correctly classified as insects. Each instance corresponds to a range gate.

Table 1. Confusion matrix for bird detection

	True labels	
Classifier output	Birds	Insects
Birds	TP	FP
Insects	FN	TN

Accuracy is the proportion of the whole data set that is correctly predicted. TPR is the proportion of correct predictions only on bird cases. Similarly, TNR is the proportion of correct predictions for the insect cases. They are calculated as shown in the following equations

$$ACC = \frac{TP+TN}{TP+FN+FP+FN}, \tag{12}$$

$$TPR = \frac{TP}{TP+FN}, \tag{13}$$

$$TNR = \frac{TN}{FP+T}, \tag{14}$$

Binary classifiers usually predict a probability (or score) for the positive class and then a threshold is applied to obtain the final class. The Receiver operating characteristics (ROC) curve plots TPR against the false positive rate (FPR) for varying probability thresholds [36]. FPR is $1 - TPR$. The goal of the ROC curve is to find an intermediate threshold that maximizes TPR and minimizes TNR. The Area under curve (AUC) metric summarizes the area under the ROC curve [36]. Good classifiers should have an AUC close to 100%.

6. Model training and validation

Classifiers are sensitive to the class distribution of the training set. Thus, we applied class weights [34] to samples to balance their effect on the loss function. For each machine learning method, 8 models are trained using different combinations of inputs. First, a base model is trained on only dual polarization variables and then different combinations of the remaining features are added to investigate their effect on performance. It should also be noted that not all inputs features can always be obtained from the radar scan. For example, sector is calculated using a sinusoid fit to the velocity of migration echoes. These echoes are mostly composed of a single species moving in a particular direction. In cases with diverse species moving in different directions, the sinusoid fit will not be possible and sector cannot be recovered. Velocity aliasing could also prevent the recovery of sector. Similarly, the calculation of texture for a given gate depends on the presence of targets in all neighboring gates. In an alternative scenario, this calculation will not be possible. In such cases however, the base model can always be used.

K fold cross validation [37] was used to tune the model hyperparameters. In this method, the data set is divided into K folds, model training is performed on K-2 folds, validation on 1 fold and testing on 1 fold. Since we already held out a test set (batch C), training was performed on K-1 folds and validation on 1 fold. The whole process is repeated K times where each fold is used as training and validation once. A total of 5 folds were used. After cross validation, the hyperparameters that have the best performance are chosen for each model. Final training is performed using the selected hyperparameters and the full training set. An ROC curve is then generated and a critical threshold found, such that it maximizes TPR and TNR. This threshold would be used to convert predictions into classes. The training process is stochastic so each run produces slightly different results. To have a robust assessment of performance, 30 independent training runs are repeated for each model. All the trained models are then evaluated on the test data. Confidence intervals for each metric is calculated using the bootstrapping percentile method where each metric is computed from an iteratively chosen random sample of the test data [27], [38]. We computed each metric for 100 iterations based on 1000 randomly chosen samples. The 100 metrics for 30 repeated runs form a total of 3000 estimates. The confidence interval is found as the 2.5% and the 97.5% points of the distribution [27], [38].

Table 2. 95% confidence interval for the ridge classifiers' and decision trees' metrics on migration data. The possible inputs are Dual-Pol(DP), their textures(ΔDP), sector($sect$) and range interval(RI).

	ACC	TPR	TNR	AUC
Ridge classifier				
Dual-Pol	0.814 - 0.858	0.832 - 0.892	0.778 - 0.844	0.868 - 0.911
Dual-Pol + texture	0.849 - 0.891	0.874 - 0.924	0.808 - 0.874	0.909 - 0.943
Dual-Pol + sector	0.812 - 0.858	0.822 - 0.886	0.784 - 0.85	0.869 - 0.911
Dual-Pol + range interval	0.815 - 0.86	0.836 - 0.894	0.774 - 0.844	0.869 - 0.912
Dual-Pol + texture + sector	0.849 - 0.891	0.87 - 0.924	0.81 - 0.874	0.91 - 0.943
Dual-Pol + sector + range interval	0.813 - 0.858	0.826 - 0.884	0.78 - 0.848	0.869 - 0.912
Dual-Pol + texture + range interval	0.849 - 0.891	0.872 - 0.926	0.81 - 0.872	0.91 - 0.944
Dual-Pol + texture + sector + range interval	0.85 - 0.891	0.87 - 0.922	0.812 - 0.876	0.91 - 0.944
Decision trees				
Dual-Pol	0.8 - 0.856	0.786 - 0.892	0.762 - 0.852	0.872 - 0.92
Dual-Pol + texture	0.778 - 0.852	0.786 - 0.902	0.718 - 0.846	0.866 - 0.925
Dual-Pol + sector	0.751 - 0.815	0.658 - 0.8	0.79 - 0.88	0.831 - 0.884
Dual-Pol + range interval	0.75 - 0.82	0.668 - 0.794	0.79 - 0.882	0.792 - 0.868
Dual-Pol + texture + sector	0.713 - 0.804	0.628 - 0.79	0.768 - 0.866	0.793 - 0.873
Dual-Pol + sector + range interval	0.684 - 0.762	0.518 - 0.68	0.814 - 0.894	0.711 - 0.796
Dual-Pol + texture + range interval	0.714 - 0.813	0.7 - 0.828	0.676 - 0.844	0.751 - 0.871
Dual-Pol + texture + sector + range interval	0.669 - 0.796	0.572 - 0.732	0.71 - 0.87	0.702 - 0.832

7. Performance

The 95% confidence interval for the model metrics are shown in Table 2. All the ridge classifiers are predictive with $ACC > 0.81$, $TPR > 0.82$, $TNR > 0.77$ and $AUC > 0.86$. Sector is expected to greatly improve results however it's addition to ridge classifiers cause marginal changes to performance. It slightly improves TNR, slightly reduces TPR and does not seem to have a noticeable effect on AUC. This could be because the training data was already averaged along the aspect angles creating a clearer delineation between both classes, so that classification can be effectively performed without sector. Recall that sectors are 20° ϕ_{az} bins. Addition of range interval marginally changes performance, improving ACC, TPR and AUC, and reducing TNR. The addition of texture generally improves the model metrics.

The decision tree models are also predictive for TPR, TNR and AUC but perform significantly worse on TPR with some models having values around 0.5. This seems to coincide with models using range interval and/or sector. A possible cause could be that it's binary decision making tends to prefer classifying whole range intervals (or sectors) as one class in contrast to ridge classifiers that only learns a probability adjustment. Using smaller range intervals and sectors might mitigate this problem. Like the ridge classifiers, incorporation of texture generally improves the model metrics. However, this might not generalize to non-migration cases. Recall that labels were provided based on the dominant migrating taxa. Textures have the effect of averaging measurements derived over a 3x3 neighborhood, so would emphasize the dominant class leading to better metrics for migration cases. However, for non-migration cases with a heterogeneous mix of scatterers, texture could lead to mis-classifications.

Both models were compared using an independent two sample t-test with a significance level of 0.05. The null hypothesis was that both metric distributions have the same mean. The ridge classifiers proved to be the better performing method with higher means on at least 3 metrics. Based on these results the ridge classifier was selected as a better method. All further discussions would be focused on this classifier. The best ridge classifier uses dual polarization variables, texture, sector and range interval as inputs with $AUC > 0.91$. It is possible though, that the improvement caused by the added features could be the classifiers over-fitting to migration cases. Thus, additional studies on a diverse variety of cases (presented in the next section) are required to understand the effect of these features on performance.

8. Case studies and discussions

In this section, the performance of the ridge classifiers is further tested on six cases. The aim of these analysis is to verify that the classifier's detections are consistent with the available ground truth and to observe the effect of different features on their performance. For operational use, we recommend that the ridge classifiers be used as a sub-classifier for the HCA. This configuration is applied to cases in this section. The first case analyzes an event of mass bird migration from the test data set (collected from KTLX) to explore the effect of the different inputs. The second case contain groups of bird roosts, insects and weather echoes collected from KHTX (located in Huntsville, AL). Ground truth was available from previous literature [29] so this would be a good test of the classifiers' accuracy. It also tests the classifiers on a different WSR-88D radar. In the third case, the classifier was tested on an event of birds observed fleeing their nests shortly after an earthquake in Oklahoma. The next case investigates another bird roost from KMOB (Mobile, AL) where ground truth was obtained from previous literature [27], [39]. The final case studies swarms of insects observed from six NEXRAD radars across the southern US, evaluating the potential of the classifier to be applied for broad scale surveillance of biological echoes.

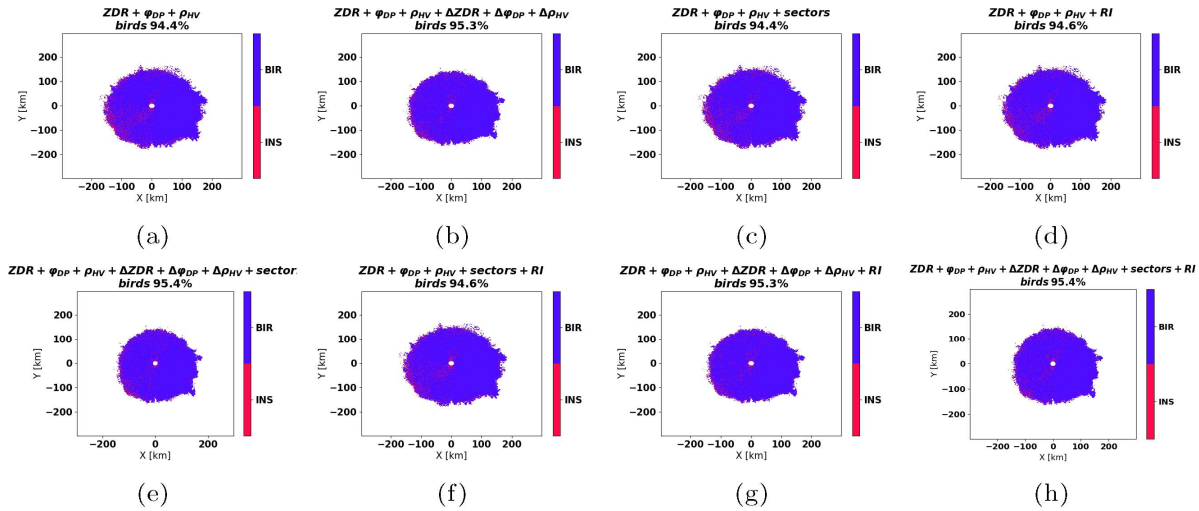


Figure 4. Ridge classification results for bird migration observed with KTLX radar at 04:13 UTC on 2 May 2015. BIR represent birds and INS represent insects.

8.1. Mass migration of biota, KTLX

The first case was collected from KTLX at 04:13 UTC on 2 May 2015 containing a swarm of migrating birds. It is one of the PPIs in the test set, so all the input variables are available. Sector could be recovered because migration was highly aligned. The classifier predicts a probability of each range gate containing bird echoes. The critical threshold (~0.5) is applied to binarize these probabilities to 0's (insects) and 1's (birds) and obtain the output class. The ridge classifier outputs are shown in Figure 4. Birds are colored blue while insects are colored red. Across the 8 ridge classifiers, birds are detected in $94.4\% - 95.4\%$ of classified gates, consistent with our hypothesis of birds as the main source of echoes. Overall, the performance of the base model barely changes as the remaining features (texture, sector and range interval) are added, with at most a 1% difference in proportion of birds detected. A similar study (not shown here) was performed on an insect swarm case in our test data. This case was collected from KTLX at 17:08 UTC on 11 July 2019. The classifiers detected an insect majority of $92.3\% - 92.7\%$ in classified gates. Again, there is a minute difference at most 0.4% for changing input features. This suggests that some features might be playing a redundant role.

8.2. Bird roosts from KHTX and KTLX

The second case study was conducted on data collected from the KHTX radar (located at Huntsville, AL) at 11:15 UTC on 11 August 2015. This case assesses the ridge classifiers on bird roosts from a different WSR-88D radar. The Z scan is shown in Figure 5b with three groups of echoes, labelled based on analysis conducted in [29]. The first group contains two colonies of purple martins engaging in their morning roosts, verified by ground observers from the Purple Martin Conservation Society [29]. The roosts are located north-west and west of the radar. Insects, surrounding the radar location were also identified by their comparatively low mean airspeed of $1.8 ms^{-1}$ and concentration at low heights [29]. The air speed was calculated by subtracting of wind speed vectors, obtained from balloon sounding from ground speeds obtained from radial velocity [18]. The final group were weather echoes identified by Z_{DR} near 0 dB, ρ_{HV} around 1 and Φ_{DP} near the calibration offset of 60° [29].

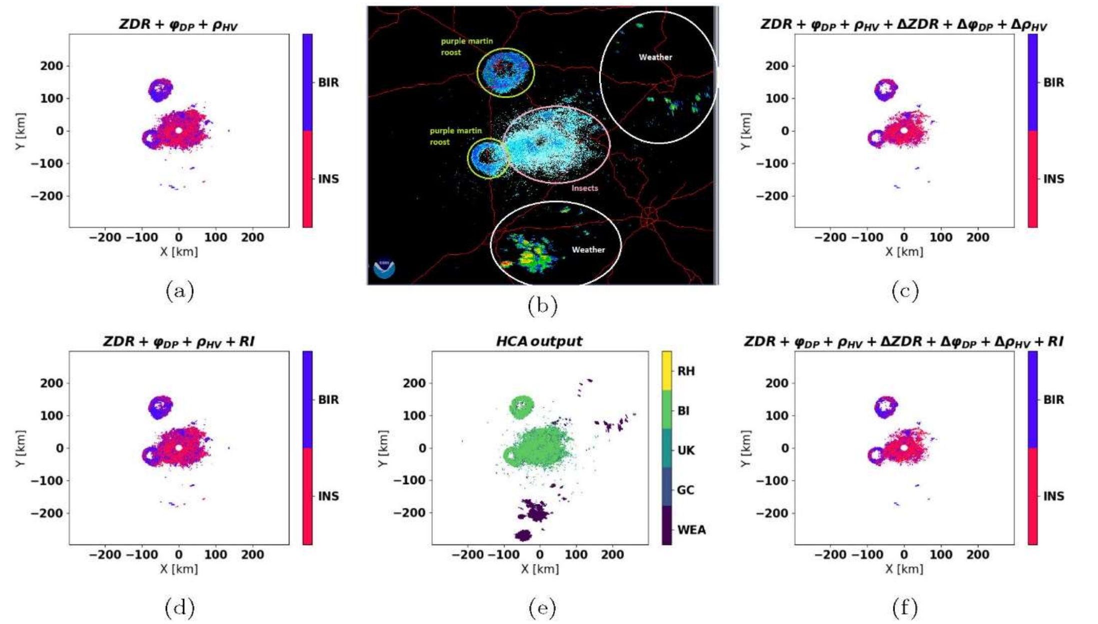


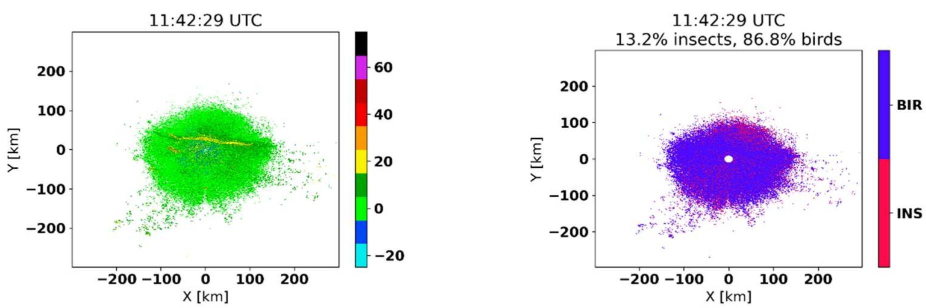
Figure 5. Bird roosts observed with KHTX at 11:15 UTC on 11 August 2015. (a) RC using dual polarization variables. (b) 0.5° Z scan showing bird roosts (west and north-west), insects (around the radar) and weather echoes (north-east and south). (c) RC using dual polarization variables and their textures. (d) RC using dual polarization variables and range interval. (e) HCA. (f) RC using dual polarization variables, their textures and range interval.

The HCA is applied to classify range gates into the five groups: weather WEA, biological BI, unknown UK, ground clutter GC and range folded echoes RH. The results (shown in Figure 5e) can be seen to corroborate the labels provided in [29]. Next, the ridge classifiers are applied on the biological class. Sector could not be recovered here because of the presence of diverse scatterers with different velocities. The base classifier (Figure 5a) and the classifier with range interval (Figure 5d) identify the roosts as bird dominated and the insect region as insect dominated. These results are consistent with the available ground truth. The models with texture (Figure 5c and 5f) identifies the insect region but mis-classify large parts of the roosting birds as insects, most noticeable where the western roost intersect with insect echoes.

Overall, the base classifier's detections matches the labels provided demonstrating the efficacy of the classifier on new cases of biological activity and a different NEXRAD radar. The addition of range interval does not have a noticeable effect on performance while texture seems to degrade performance on fine and hollow features like bird roosts. A similar case containing bird roosts (figure not included here) collected from KTLX at 11:47 UTC on 8 August 2017, was processed with the ridge classifier. Again, the base classifier detected the roosts as bird dominated and the addition of other features did not improve the result. This is further evidence that the remaining features might not be necessary. For the sake of algorithmic simplicity, we decided to choose the base classifier. The analysis in the remaining case studies would be focused on this classifier.

8.3. Birds escaping their nests during an Earthquake, KTLX

Broad scale movement of biological echoes is commonly in response to natural phenomena. The next case study is for one of such occurrences. An earthquake occurred in Oklahoma on 29 October 2015 at 11: 39 UTC, resulting in a splash of birds leaving their nests observed on the KTLX radar. The reflectivity of echoes recorded two minutes after the quake is shown in Figure 6a. Notice a line of high reflectivity values tracing the movement of the birds away from their nests. This line progresses southward in the next few scans. The base ridge classifier (shown) in Figure 6b detects a bird majority, with 86.8% of echoes classified as birds further corroborating ground observations.

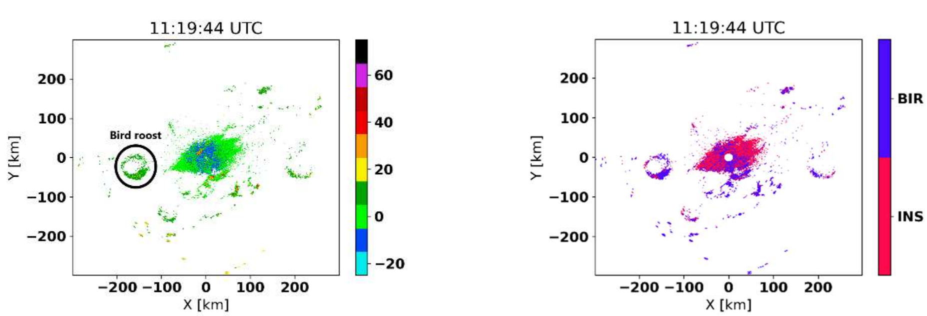


(**a**) (**b**)

Figure 6. Birds leaving their nests in response to an earthquake in Oklahoma on 29 October 2015 at 11: 39 UTC, observed by KTLX (**a**) Reflectivity showing a splash of birds leaving their nests (**b**) RC detect a bird majority with 86.8% of echoes classified as birds.

8.4. Bird roosts from KMOB

The fourth case study involves a bird roost observed by KMOB (located in Mobile, AL) on 4 July 2015 at 11:19 UTC. This case is one of many labelled manually by Kelly and Pletschet by searching radar imagery from one hour prior to local sunrise till 30 minutes after local sunrise, an effort that required examining 70,000 – 140,000 images per year[27], [39]. Figure 7a shows the reflectivity for this case with the observed bird roost enclosed in the black circle. The base ridge classifier (Figure 7b) detects birds as the main cause of this roost.

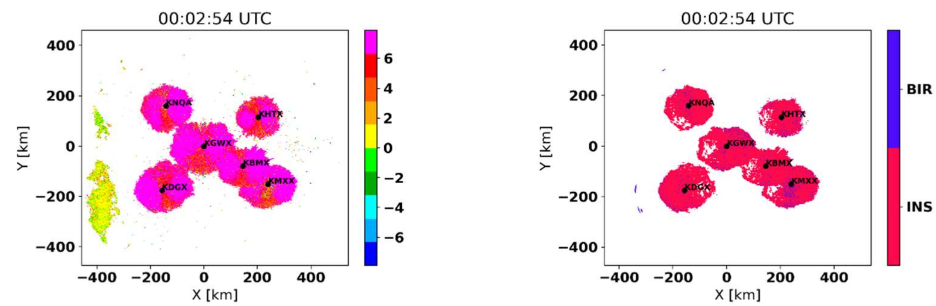


(**a**) (**b**)

Figure 7. Bird roost observed by KMOB (located in Mobile, AL) on 4 July 2015, 11:19 UTC (a) Reflectivity image showing bird roost to the west of KMOB (**b**) RC detects roosts to be mainly comprised of birds.

8.5. Insects observed over southern United States

The final case study is performed on a snapshot of the Southern United States on 19 April 2016 at 00:00 UTC (approximately 22 mins before local sunset). The snapshot includes returns from six NEXRAD radars. They are KNQA (located in Memphis, TN), KHTX (Huntsville, AL), KGWX (Columbus AFB, MS), KBMX (Birmingham, AL), KDGX (Jackson Brandon, MS) and KMXX (Maxwell AFB, AL). The ZDR of this snapshot is shown in Figure 8a below. Insects were identified around the radars for this case by their well-known dumb-bell patterns in ZDR and Z [29] . Furthermore, airspeed analysis using the 00 UTC Birmingham, AL, sounding on KBMX yielded airspeeds of 2.39 m/s in the lowest kilometer of airspace, indicating insect presence [29]. The base classifier (shown in Figure 8b) detects these echoes around the radar as being predominantly insects, matching observations in previous research. The holes in the detected insect swarms is an artifact of the prefiltering step using the HCA.



(**a**) (**b**)

Figure 8. Snapshots of insects over southern United States on 19 April 2016 at 00:00 UTC (approximately 22 mins before local sunset) (**a**) ZDR shows the characteristic dumb-bell pattern with high values horizontally across and lower values vertically across the insect swarms (**b**) RC detects insect swarm as predominantly insects.

9. Conclusions

NEXRAD's detection of birds and insects offers much promise for a variety of applications. In this work, we developed a classifier for distinguishing bird and insect radar echoes based on dual polarization variables. Unique challenges were faced during data collection due to complex scattering off their non-spherical bodies. This was addressed by leveraging cases of large-scale migration with an aligned heading to change measurement coordinates from being relative to the radar to be relative to the body aspect of biota. The mean flight direction, which would measure scattering off the tail, was found by VAD analysis and then measurements from other aspects are deduced by the lag off this mean heading. Another issue is the difficulty in labelling training data sets because of the frequent collocation of birds, insects and other non-biological echoes in the radar sampling volume. We addressed this by averaging 15 alignment calibrated bird (insect) migration scans to reduce the effect of the less dominant class.

The data preparation pipeline is summarized in the following steps. First, 45 scans containing mass migration in clear air were collected for each class. Blob coloring with minor region removal was applied to extract the blob of migration echoes and their textures computed. Extracted migration echoes are then rotated to become relative to the target's aspect. The rotated scans are grouped into 3 batches, each containing 15 scans per class. All 15 scans in two of the batches are averaged to reduced contamination. This is done for both classes. Gates from the 4 resulting averaged scans are used as training samples. The last unaveraged batch is used as the test set. All samples in both sets are assigned a unique range and aspect identifier. The range identifier, $range\ interval$, is the range index when the data is grouped into 10 km range intervals while the aspect identifier, sector, is the index from grouping into 20° target azimuth bins. These identifiers along with the dual pol variables and their textures make up the candidate input feature set.

Two machine learning methods were explored: ridge classifier and decision tree. Eight models were trained for each method, starting with a base model using only dual polarization variables and then adding other input features. Four metrics were used for evaluating the classifiers on the test data set. They are accuracy (ACC), true positive rate (TPR), true negative rate (TNR) and area under curve (AUC). A comparison of the metrics from both methods showed that the ridge classifiers performed better than decision trees in at least 3 out of 4 metrics. Based on this, the ridge classifiers were selected for classifying bird and insect radar echoes. All the ridge classifiers are predictive with $ACC > 0.81$, $TPR > 0.82$, $TNR > 0.77$ and $AUC > 0.86$. The addition of other features show an improvement of about 4% on these metrics, however later evidence suggests that this is probably due to over-fitting to cases of large scale migration.

Further qualitative case studies was conducted to assess the effect of the different inputs to the classifier on a bird and insect migration scan from the test set. The ridge classifiers detected birds in $94.4\% - 95.4\%$ of range gates for the bird scan and insects in $92.3\% - 92.7\%$ of gates for the insect scan, consistent with our hypothesis of the source of these echoes. The addition of the remaining features to the base model has a bare effect on performance, with an increase of at most 1% and 0.4% in the proportion of birds and insects detected respectively. This suggests that the additional input features might be superfluous. The classifiers were also evaluated on diverse cases of biological activity across NEXRAD. The training data was collected from KTLX, so the ability to detect biological patterns from other WSR-88Ds would provide strong evidence that it can be applied on the network. Furthermore, the training data did not contain bird roosts. Thus, the ability to detect roosting birds would be evidence in favour of the generality of the classifier. The next case explored bird roosts collected by the KHTX (located in Huntsville, AL). Previous studies [29] provided ground truth for bird roosts, insects and weather echoes for this scan. Sector could be recovered here because of the heterogenous mix of scatterers. The detections of the base classifier and the one with range interval match the provided ground truth. The addition of texture seems to degrade performance on the roosts, probably because it is less suited for capturing finer features. The classifier was also evaluated on a similar case from KLTX containing four bird roosts identified by observing the expanding ring over time. Again, the base classifier and the one with range interval detect all the roosts as birds, while the addition of texture degraded performance. Overall, the tests conducted show no evidence of improvements from adding features to the base classifier. For the sake of algorithm simplicity, the base ridge classifier is selected as the best model for our classification task.

Sometimes biological activity is a cue to underlying seismic events. In the next case, the base ridge classifier is tested on a splash of birds (observed by KTLX) fleeing their nests in response to an earthquake in Oklahoma. The classifier detects 86.8% of echoes to be from birds. This demonstrates a potential of use this classifier as a tool to study natural events of common interest to humans and birds/insects. For the fourth case study, the classifier was tested on a bird roost from KMOB where ground truth labels are known from previous research [27], [39]. The base classifier detects the roost as birds. The final case study demonstrates the use of the classifier for large scale surveillance on NEXRAD. Here, swarms of insects were observed across the southern United States just before local sunset using 6 NEXRAD radars. The insects were identified in previous literature by their characteristic dumb-bell pattern in Z and ZDR, and low mean airspeeds in the lowest 1 km of airspace. The classifier detects these swarms as inspects.

In our test cases, the base ridge classifier has been demonstrated to correctly classify different orientations of biological echoes across NEXRAD. As such, we recommend this classifier could be implemented on the network, as a sub-classifier on the HCA's biological class. The biggest challenge to developing biological classifiers is obtaining the ground truth. For future research, we hope to collect more ground truth data using other sensors to validate our classifier. We also hope this research encourages other data collection and verification efforts for biological radar echoes.

Author Contributions: Conceptualization, P.J. and V.M.; methodology, P.J.; software, P.J.; validation, V.M. and T.Y.; formal analysis, P.J.; investigation, P.J.; resources, V.M. and T.Y.; data curation, P.J.; writing—original draft preparation, P.J.; writing—review and editing, V.M. and T.Y.; visualization, P.J.; supervision, V.M. and T.Y; project administration, V.M.; funding acquisition, V.M. All authors have read and agreed to the published version of the manuscript

Funding: Funding for this study was provided in part by the NOAA/Office of Oceanic and Atmospheric Research under NOAA-University of Oklahoma Cooperative Agreement NA110OAR4320072

Institutional Review Board Statement: Not applicable

Data Availability Statement: The data set for this project was downloaded from the National Center for Environmental information (NCEI).

Acknowledgments: The authors would like to thank K. Horton, J. Kelly and P. Stepanian for helpful contributions in the data collection for this project. Additionally, we thank P. Zhang for constructive comments in the preparation of this paper, and the NEXRAD Radar Operations Center for funding this research.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

[1] M. R. Kumjian, "Principles and Applications of Dual-Polarization Weather Radar. Part I: Description of the Polarimetric Radar Variables," 2013.

[2] R. J. Doviak and D. S. Zrnić, *Doppler Radar and Weather Observations*. Academic Press, 1993. [Online]. Available: https://books.google.com/books?id=sWljQgAACAAJ

[3] R. E. Rinehart, *Radar for Meteorologists*. Rinehart Publications, 2004.

[4] J. Gauthreaux Sidney A., J. W. Livingston, and C. G. Belser, "Detection and discrimination of fauna in the aerosphere using Doppler weather surveillance radar," *Integrative and Comparative Biology*, vol. 48, no. 1, pp. 12–23, Jun. 2008, doi: 10.1093/icb/icn021.

[5] A. T. Peterson and R. A. J. Williams, "Risk mapping of highly pathogenic Avian Influenza distribution and spread," *Ecology and Society 13*, 2008, [Online]. Available: http://www.ecologyandsociety.org/vol13/iss2/art15/

[6] A. M. Dokter, F. Liechti, H. Stark, L. Delobbe, P. Tabary, and I. Holleman, "Bird migration flight altitudes studied by a network of operational weather radars," *Journal of The Royal Society Interface*, vol. 8, no. 54, pp. 30–43, 2011, doi: 10.1098/rsif.2010.0116.

[7] P. K. Jatau and V. Melnikov, "Classifying bird and insect radar echoes at S-band," Jan. 2019. [Online]. Available: https://ams.confex.com/ams/2019Annual/mediafile/Manuscript/Paper351588/Precious_Valery_AMS_paper_final.pdf

[8] S. Bachmann and D. Zrnić, "Spectral Density of Polarimetric Variables Separating Biological Scatterers in the VAD Display," *Journal of Atmospheric and Oceanic Technology*, vol. 24, no. 7, pp. 1186–1198, 2007, doi: 10.1175/JTECH2043.1.

[9] P. Zhang, S. Liu, and Q. Xu, "Identifying Doppler Velocity Contamination Caused by Migrating Birds. Part I: Feature Extraction and Quantification," *Journal of Atmospheric and Oceanic Technology*, vol. 22, no. 8, pp. 1105–1113, Jun. 2005, doi: 10.1175/JTECH1757.1.

[10] S. Liu, Q. Xu, and P. Zhang, "Identifying Doppler Velocity Contamination Caused by Migrating Birds. Part II: Bayes Identification and Probability Tests," *Journal of Atmospheric and Oceanic Technology*, vol. 22, no. 8, pp. 1114–1121, Jun. 2005, doi: 10.1175/JTECH1758.1.

[11] D. S. Zrnic and A. v Ryzhkov, "Polarimetry for Weather Surveillance Radars," *Bulletin of the American Meteorological Society*, vol. 80, no. 3, pp. 389–406, 1999, doi: 10.1175/1520-0477(1999)080<0389:PFWSR>2.0.CO;2.

[12] H. S. Park, A. v Ryzhkov, D. S. Zrnić, and K.-E. Kim, "The Hydrometeor Classification Algorithm for the Polarimetric WSR-88D: Description and Application to an MCS," *Weather and Forecasting*, vol. 24, no. 3, pp. 730–748, 2009, doi: 10.1175/2008WAF2222205.1.

[13] V. Chandrasekar, R. Keränen, S. Lim, and D. Moisseev, "Recent advances in classification of observations from dual polarization weather radars," *Atmospheric Research*, vol. 119, pp. 97–111, 2013, doi: https://doi.org/10.1016/j.atmosres.2011.08.014.

[14] S. Gauthreaux and R. H. Diehl, "Discrimination of Biological Scatterers in Polarimetric Weather Radar Data: Opportunities and Challenges," *Remote. Sens.*, vol. 12, p. 545, 2020.

[15] B. Radhakrishna, F. Fabry, and A. Kilambi, "Fuzzy Logic Algorithms to Identify Birds, Precipitation, and Ground Clutter in S-Band Radar Data Using Polarimetric and Nonpolarimetric Variables," *Journal of Atmospheric and Oceanic Technology*, vol. 36, no. 12, pp. 2401–2414, 2019, doi: 10.1175/JTECH-D-19-0088.1.

[16] A. Kilambi, F. Fabry, and V. Meunier, "A Simple and Effective Method for Separating Meteorological from Nonmeteorological Targets Using Dual-Polarization Data," *Journal of Atmospheric and Oceanic Technology*, vol. 35, no. 7, pp. 1415–1424, 2018, doi: 10.1175/JTECH-D-17-0175.1.

[17] E. S. Bridge *et al.*, "Technology on the Move: Recent and Forthcoming Innovations for Tracking Migratory Birds," *BioScience*, vol. 61, no. 9, pp. 689–698, Jun. 2011, doi: 10.1525/bio.2011.61.9.7.

[18] P. M. Stepanian and K. G. Horton, "Extracting Migrant Flight Orientation Profiles Using Polarimetric Radar," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 12, pp. 6518–6528, 2015, doi: 10.1109/TGRS.2015.2443131.

[19] M. S. van den Broeke, "Polarimetric Radar Observations of Biological Scatterers in Hurricanes Irene (2011) and Sandy (2012)," *Journal of Atmospheric and Oceanic Technology*, vol. 30, no. 12, pp. 2754–2767, 2013, doi: 10.1175/JTECH-D-13-00056.1.

[20] D. S. Zrnic and A. v Ryzhkov, "Observations of insects and birds with a polarimetric radar," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 36, no. 2, pp. 661–668, 1998, doi: 10.1109/36.662746.

[21] V. A. Drake and D. Reynolds, *Radar Entomology: Observing Insect Flight and Migration*. CABI, 2012.

[22] K. R. Hardy and I. Katz, "Probing the clear atmosphere with high power, high resolution radars," *Proceedings of the IEEE*, vol. 57, no. 4, pp. 468–480, 1969.

[23] T. J. Lang, S. A. Rutledge, and J. L. Stith, "Observations of Quasi-Symmetric Echo Patterns in Clear Air with the CSU–CHILL Polarimetric Radar," *Journal of Atmospheric and Oceanic Technology*, vol. 21, no. 8, pp. 1182–1189, 2004, doi: 10.1175/1520-0426(2004)021<1182:OOQEPI>2.0.CO;2.

[24] R. F. Contreras and S. J. Frasier, "High-Resolution Observations of Insects in the Atmospheric Boundary Layer," *Journal of Atmospheric and Oceanic Technology*, vol. 25, no. 12, pp. 2176–2187, 2008, doi: 10.1175/2008JTECHA1059.1.

[25] V. M. Melnikov, R. R. Lee, and N. J. Langlieb, "Resonance Effects Within S-Band in Echoes From Birds," *IEEE Geoscience and Remote Sensing Letters*, vol. 9, no. 3, pp. 413–416, 2012.

[26] W. J. Martin and A. Shapiro, "Discrimination of Bird and Insect Radar Echoes in Clear Air Using High-Resolution Radars," *Journal of Atmospheric and Oceanic Technology*, vol. 24, no. 7, pp. 1215–1230, 2007, doi: 10.1175/JTECH2038.1.

[27] C. Chilson, K. Avery, A. McGovern, E. Bridge, D. Sheldon, and J. Kelly, "Automated detection of bird roosts using NEXRAD radar data and Convolutional Neural Networks," *Remote Sensing in Ecology and Conservation*, vol. 5, no. 1, pp. 20–32, 2019, doi: 10.1002/rse2.92.

[28] Z. Cheng *et al.*, "Detecting and Tracking Communal Bird Roosts in Weather Radar Data," 2020.

[29] P. M. Stepanian, K. G. Horton, V. M. Melnikov, D. S. Zrnić, and S. A. Gauthreaux Jr., "Dual-polarization radar products for biological applications," *Ecosphere*, vol. 7, no. 11, p. e01539, 2016, doi: 10.1002/ecs2.1539.

[30] A. C. Bovik, "Chapter 4 - Basic Binary Image Processing," in *The Essential Guide to Image Processing*, A. Bovik, Ed. Boston: Academic Press, 2009, pp. 69–96. doi: https://doi.org/10.1016/B978-0-12-374457-9.00004-4.

[31] P. M. Stepanian, P. B. Chilson, and J. F. Kelly, "An introduction to radar image processing in ecology," *Methods in Ecology and Evolution*, vol. 5, no. 8, pp. 730–738, 2014, doi: https://doi.org/10.1111/2041-210X.12214.

[32] K. A. Browning and R. Wexler, "The Determination of Kinematic Properties of a Wind Field Using Doppler Radar," *Journal of Applied Meteorology*, vol. 7, no. 1, pp. 105–113, 1968, doi: 10.1175/1520-0450(1968)007<0105:TDOKPO>2.0.CO;2.

[33] S. A. Gauthreaux and C. G. Belser, "Displays of Bird Movements on the WSR-88D: Patterns and Quantification," *Weather and Forecasting*, vol. 13, no. 2, pp. 453–464, 1998, doi: 10.1175/1520-0434(1998)013<0453:DOBMOT>2.0.CO;2.

[34] F. Pedregosa *et al.*, "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[35] L. Bottou, "Large-scale machine learning with stochastic gradient descent," 2010.

[36] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861–874, 2006, doi: https://doi.org/10.1016/j.patrec.2005.10.010.

[37] R. Kohavi, "A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection," 1995, pp. 1137–1143.

[38] B. Efron and R. Tibshirani, "Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy," *Statistical Science*, vol. 1, no. 1, pp. 54 – 75, 1986, doi: 10.1214/ss/1177013815.

[39] E. S. Bridge *et al.*, "Persistence and habitat associations of Purple Martin roosts quantified via weather surveillance radar," *Landscape Ecology*, vol. 31, no. 1, pp. 43–53, 2016, doi: 10.1007/s10980-015-0279-0.