RESEARCH ARTICLE

Ecological Solutions and Evidence | BRITISH ECOLOGICAL SOCIETY

Ensemble predictions are essential for accurate bird migration forecasts for conservation and flight safety

Bart Kranstauber[1] | Willem Bouten[1] | Hans van Gasteren[1,2] | Judy Shamoun-Baranes[1]

[1]Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Amsterdam, The Netherlands

[2]Royal Netherlands Air Force, Breda, The Netherlands

Correspondence: Bart Kranstauber, Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, P.O. Box 94248, 1090 GE Amsterdam, The Netherlands. Email: b.kranstauber@uva.nl

Abstract

1. Accurate predictions of the abundance of migrating birds are important to avoid aerial conflicts of birds, for example, with aviation or wind power installations.

2. Here we develop a predictive model, using bird migration intensity extracted from operational weather data. We compare baseline phenological models to models incorporating both local and remote weather conditions using an ensemble approach. Single models are compared to ensemble models (average prediction of top 10 models). The models were evaluated by omitting single years from our 10-year dataset.

3. In general, we find that wind conditions, in addition to seasonal and diurnal dynamics, are key for accurate predictions. The spring and fall migratory seasons differ, both with respect to the selected environmental variables and the contribution of the environmental model compared to the phenological model. In fall, the accumulation of migrants due to strong headwinds is an important predictor of migration.

4. Because of the lower daily variation in migration intensity in spring, the phenological model performs better compared to fall. In fall, weather conditions contribute more to accurate predictions of migration intensity than in spring.

5. Overall, the ensemble approach produces more accurate predictions outperforming specific environmental models. We therefore recommend that ensemble models be used in operational settings such as flight planning to reduce bird aircraft collisions during intense bird migration.

KEYWORDS
bird migration, bird strikes, ensemble model, flight safety, radar monitoring, weather radar, wildlife management

1 | INTRODUCTION

Each spring and fall, billions of birds migrate between breeding and non-breeding ranges all over the globe (Dokter et al., 2018; Nussbaumer et al., 2020). During their migratory movements, birds encounter diverse anthropogenic activities that have direct consequences for migrants and humans. Artificial lights can attract birds and thus influence migration routes (Van Doren et al., 2017). Migrants may collide with tall structures such as higher buildings, power lines and wind turbines, resulting in mortality (Loss et al., 2014). Or birds may avoid large wind parks resulting in increased flight costs. Collisions between birds and aircraft can also have direct consequences for human safety, especially during take-off or low-level flights (Van Gasteren et al., 2019). To reduce the impact of these human–wildlife conflicts, there is a strong interest in predicting temporal patterns of migration (Shamoun-Baranes et al., 2017). Predictions can be used to inform policy related to light pollution, temporary shutdown of wind turbines or temporary changes in aviation planning (e.g. halting take-off or landing, changing altitude or holding exercises). Accurate predictions are thus of acute interest.

To accurately predict migration, insights into the major drivers of migration are needed. The most frequently used drivers for modelling migration intensity fall into two major categories: phenological trends and weather (Erni et al., 2002; Van Belle et al., 2007). Seasonal timing of the annual cycle driven by, for example, day length, is considered one of the key factors influencing the general timing of migration (Gwinner & Helm, 2003). Ultimately birds aim to synchronize migration with peaks in resource availability. Seasonal timing may therefore differ among species and regions. In addition, diurnal patterns are also known to influence migration activity. Migratory species are often categorized as diurnal or nocturnal. For example, passerines, one of the most abundant migrant groups, predominantly migrate during the night. Thus circadian migratory trends relate to sunrise and sunset.

Besides phenological trends, the weather has been shown to influence migratory behaviour and thus the number of birds passing an area (Alerstam, 1978; Richardson, 1990). Wind can influence ground speeds, airspeeds, flight routes and flight altitudes of birds during migration, supporting or impeding flight (Liechti, 2006). Advantageous wind conditions make it possible to increase the distance covered without increasing the flight speed (Safi et al., 2013). Therefore, birds are more likely to migrate under these conditions (Åkesson & Hedenström, 2000). Precipitation has been found to correlate negatively with migration intensity and departure decisions (Erni et al., 2002; Richardson, 1990). Other variables can relate to weather systems or a set of combined weather conditions (air pressure and change in air pressure) and temperature (Manola et al., 2020). To summarize, a range of weather variables have been related to migratory intensity and used for predictions of migration (Alerstam, 1978; Van Belle et al., 2007; Van Doren & Horton, 2018).

Predictive models should also account for differences in migratory dynamics and processes between seasons. There are several reasons migration may differ between spring and fall. The intrinsic motivation influencing migratory behaviour may differ among seasons. In spring, birds are thought to have a competitive advantage by arriving earlier in the breeding area. Several studies found a higher migration speed in spring compared to fall (Horton et al., 2016; Nilsson et al., 2013), although seasonal weather patterns can partially explain these differences (Kemp et al., 2010). Additionally, the demographic composition of individuals differs between fall and spring and thus potentially the response to external factors. In fall, there is a higher proportion of inexperienced birds which might respond differently to weather. Furthermore, weather patterns differ between seasons (Kemp et al., 2010). This means birds might have different selectivity for favourable conditions in different seasons.

To develop successful predictive models, one needs observations of migrating birds. Observations of migratory intensity should cover multiple years to separate phenological trends in bird migration from weather effects. This is desirable as weather conditions differ among years, and predictions should be independent of the specific conditions in 1 year (Van Belle et al., 2007). A weather radar is a powerful tool to remotely, continuously and systematically quantify migration densities above given airspace (Dokter et al., 2011), and for several countries data archives are available to enable the analysis of several years of data. It has been shown that measurements from the weather radar are a strong correlate of bird strikes in civil aviation (Nilsson et al., 2021). Another advantage is that weather radars cover a larger area compared to visual observations and dedicated bird radars. They thus provide insight into the trends of migration across a larger region.

To gain insight and confidence into the predictions of migration, it is important to understand the biological basis of the prediction. For various machine learning techniques, these insights might be harder to provide (James et al., 2013), although interpretable and explainable machine learning is a quickly developing and active field of research (Roscher et al., 2020). In most statistical modelling techniques, the relations between environmental conditions and the final predictions are easier to explore and understand.

The aim of this study is to develop predictive models for bird migration in spring and fall which will be used to structure flight planning and reduce the impact of bird strikes on military aviation. These models will be used to issue BirdTAMs (Bird-notice-to AirMen) for aviation (Van Gasteren et al., 2019). BirdTAMs range from one to eight and correspond to a vertically integrated per density of birds per km^2 or density per altitudinal bin (per km^3). From level five warnings are issued, and from level seven onward flight restrictions are issued up to heights where levels drop below five. We want to emphasize predicting moments of peak migration as these have the highest impact on flight safety. We use long-term migration intensity data extracted from weather radar. Given the range of possible environmental factors influencing migration intensity, one combined environmental model could result in overfitting. Overfitting can be solved by creating an ensemble of models or model average. By combining results from multiple smaller models into one aggregated prediction, the prediction error is reduced (Dormann et al., 2018). For such an ensemble model, the recommendation is to use omitted data for the evaluation of the performance.

We apply a two-step modelling approach; we first develop a phenological model to represent the seasonal and daily long-term migration dynamics and then test the added value of including weather variables. To evaluate model performance, we iteratively omit four single years for which we calculate predictions. We evaluate these predictions recognizing that phenological trends are part of the expectation, and we compare phenological models to models including environmental variables. For estimating the benefits of ensemble modelling, we compare single models with predictions made with an ensemble model. We investigate changes in both the prediction accuracy and the consistency of the predictions, while doing so we investigate differences between spring and fall. Finally, we discuss the benefits and limitations of using an ensemble prediction for aviation safety and other potential applications.

2 | METHODS AND MATERIALS

Several environmental conditions have been suggested to correlate with the number of migrants. Some of the most frequently used variables are wind conditions, occurrence and abundance of rain, surface pressure and air temperature (Alerstam, 1978; Van Belle et al., 2007; Van Doren & Horton, 2018). Therefore, selecting the optimal combination of environmental independent variables is a complex and important task. Identifying an optimal set of environmental variables simplifies implementing an operational model. To evaluate the performance of models based on various environmental conditions, we use model selection using a separate testing dataset. The results of this are used to identify the best performing model (for schematic model illustration see Supporting Information Section 1). Simultaneously, we average the top models to create an ensemble of models.

2.1 | Bird densities

To model the migration intensity, we rely on vertically integrated density (birds/km^2) estimates derived from weather radar using a radar cross section of 11 cm^2 per bird (Dokter et al., 2011). The data were extracted from 2008 to 2019, the years 2012 and 2013 were omitted because of poor performance in separating rain from birds in these years. At least, an additional period with the combined length of 43.3 days would need to be filtered out for those 2 years compared to the 21.6 days in total filtered now. Spring season was defined from 15 February until 30 April and fall from 1 August until 30 November. We combine data from two C-band radars in the Netherlands successively deployed in close proximity (before October 2016, De Bilt, ODIM: NLDBL, longitude/latitude: 5.178°, 52.102°; after October 2016, Herwijnen, ODIM: NLHRW, longitude/latitude: 5.138°, 51.837°). The distance between these radars is 29.6 km, and they cover the same central (flat) area of the Netherlands (Supporting Information Section 2). Both radars observed at a 5-min interval. We calculate the vertically integrated bird density using the software package *vol2bird* within a range between 5 and 25 km from the radar (Dokter et al.,

2011). For quality control, vertical profile time series of peak nights were visually inspected, omitting periods of rain and other non-bird reflections that were not filtered out by *vol2bird*. Combined 149.1 h (distributed across 32 distinct time periods) of data were omitted in spring and 368.8 h (across 43 periods) in fall. For vertical integration, the lowest altitudinal bin, between 0 and 200 m, was omitted as birds were identified inconsistently at this height and for our purpose of flight safety this range is less important. In total, the dataset contains 193,731 records in spring of vertically integrated densities for 5 min intervals and 323,166 records in fall (with a range of, respectively, 12,121–21,465 and 28,369–34,647 per year).

2.2 | Environmental conditions

We annotate every 5-min radar measurement of bird density with environmental conditions, in three different categories: local weather, remote weather and an index for the accumulation of birds due to poor weather conditions (Table 1). Because we develop a model with the explicit purpose of predictive modelling, we opt to not include measurements of the past migration densities in our approach for two reasons. First, they are not always available, for example, when there is maintenance on the weather radars. Second, as a prediction horizon of 2–3 days (to allow for planning of mitigation) is required there will be no information available for bird densities in the intermediate period. This means if recent bird densities were included in the model as predictors they would need to be predicted themselves. This would reduce the value of these measurements for predictive purposes, and therefore they are not included.

Weather conditions were linearly interpolated in time from the ECMWF ERA5 dataset (Hersbach et al., 2020). We extracted data from the location of the radar site and the conditions at locations where birds could depart from. The local conditions were averaged in a circle with a radius of 25 km around the radar at the surface level and linearly interpolated from pressure level data to the average in an altitude range from 400 to 1000 m as most migration in both seasons is concentrated below 1000 m. In spring, we selected two departure locations reflecting two streams of migration through the Netherlands: the south-east of the United Kingdom and the north-west coast of France (Supporting Information Section 2; Bradarić et al., 2020). For these locations, we averaged weather conditions in a circle with a radius of 75 km. In fall, we selected four departure locations in Norway, Sweden, Denmark and north-east Germany. For all remote locations, we averaged surface pressure and 10 m wind conditions interpolated in time for the 90-min following sunrise or sunset to reflect the departure period throughout the day and night. We do not use environmental data from times later in the night for those remote locations. We always include both east-west and north-south wind conditions at the same time to allow the model to describe any wind condition.

Periods of unfavourable migratory conditions can cause an accumulation of migrants that would have departed weather conditions permitting. This leads to an especially high number of migrants during

TABLE 1 The environmental variables included in the ensemble model.

Description	Season	Location	Spring weights	Fall weights
Wind at 10 m (NS and EW)	Both	Local	10 (5–13%)	8 (7–9%)
Wind at 100 m (NS and EW)	Both	Local	8 (4–15%)	9 (5–10%)
Wind at height (NS and EW, 400–1000 m)	Both	Local	17 (13–21%)	9 (5–18%)
Cloud cover	Both	Local	2 (2–5%)	1 (1–2%)
Precipitation rate	Both	Local	6 (6–10%)	7 (6–8%)
Surface temperature	Both	Local	9 (5–18%)	1 (1–6%)
Temperature at height (400–1000 m)	Both	Local	34 (16–35%)	1 (1–7%)
Pressure	Both	Local	1 (0–5%)	3 (3–4%)
24-h change in surface pressure	Both	Local	0 (1–3%)	8 (6–16%)
Sun angle	Both	Local	16 (13–16%)	9 (4–9%)
Sun angle change	Both	Local	13 (10–13%)	4 (5–11%)
Accumulation to cloud cover	Both	Local	5 (3–5%)	2 (2–4%)
Accumulation to rain	Both	Local	2 (0–4%)	3 (1–8%)
Accumulation to unfavourable wind	Both	Local	4 (0–7%)	1 (0–1%)
Accumulation to unfavourable wind and rain	Both	Local	4 (1–3%)	3 (1–7%)
Change in accumulation due to cloud cover	Both	Local	3 (0–4%)	1 (1–2%)
Change in accumulation due to rain	Both	Local	2 (0–2%)	3 (1–5%)
Change in accumulation due to unfavourable wind	Both	Local	9 (10–16%)	41 (28–44%)
Change in accumulation due to unfavourable wind and rain	Both	Local	4 (3–7%)	28 (15–31%)
Pressure Norway	Fall	Remote		7 (3–13%)
Pressure Germany	Fall	Remote		3 (0–2%)
Pressure Denmark	Fall	Remote		8 (3–12%)
Pressure Sweden	Fall	Remote		4 (3–7%)
Pressure France	Spring	Remote	1 (0–5%)	
Pressure United Kingdom	Spring	Remote	0 (0–2%)	
Surface wind Norway (NS and EW)	Fall	Remote		0 (0–1%)
Surface wind Germany (NS and EW)	Fall	Remote		3 (1–6%)
Surface wind Denmark (NS and EW)	Fall	Remote		1 (0–3%)
Surface wind Sweden (NS and EW)	Fall	Remote		0 (0–1%)
Surface wind France (NS and EW)	Spring	Remote	2 (1–3%)	
Surface wind United Kingdom (NS and EW)	Spring	Remote	14 (5–10%)	

Note: The rightmost two columns represented the summed weights for spring and fall of the ensemble models, where these variables occur in. Between brackets, we report the range of weight for the four datasets where 1 year has been omitted. Values between 5–10% have been highlighted in grey, 10–20% in light red and the highest values in red. For wind, we always included both the north-south and east-west components together.

the first days after environmental conditions improve. To capture this effect, we calculate an index for the hypothetical number of birds ready to depart, according to Erni et al. (2002), for day and night separately. This work requires a binary definition of whether the average conditions per day or night were favourable or not for migration. We selected three different environmental variables for unfavourable conditions to calculate three different accumulation variables. We suppose migration would cease when: (1) headwinds along the average migration direction (spring: 57°, fall: 229.3°, for full dataset) would significantly hinder migration (over 7 m/s; Kemp, 2012), (2) when rain exceeds an average of 1 mm/d or (3) when the average total cloud cover exceeds 0.9 1/h. These thresholds have been identified by comparing the dis-

tribution of environmental conditions for the top 10% of migratory nights and the lower 50% of nights. We add a fourth accumulation variable, representing the combined conditions when either rain or wind prevented migration. We explored accumulation indexes based on five different fractions in initial models. A fraction of $\theta = {}^{3}/_{4}$ performed best in these exploratory models; therefore, this was used for the fraction of birds that carry over from the previous nights while accumulating. As the work of Erni et al. (2002) calculates an index for the number of birds ready to depart, we furthermore include the time derivative representing the birds that are departing. The change in accumulation is calculated as the difference in accumulation between the current with the previous day or night.

2.3 | Modelling

To describe the migration density, we use generalized additive models with a quasi-Poisson distribution family (Wood et al., 2015). These models do not assume a specific relationship between the independent variables and the dependent variable. By using the quasi-Poisson error family with a log link function, we account for the distribution of the density data, with only positive values and a higher variance at higher values.

We first fit a model that captures the general phenological trends in migration. This model comprises a tensor spline comprising three variables: day of the year, allowing to capture seasonal trends, solar elevation to capture circadian effects relative to sunset and sunrise and time derivative of solar elevation to distinguish between sunset and sunrise (calculated using a numerical derivation over the surrounding minute (rad/min)). We implement seasonal trends this way as it is easier for the model to capture effects relative to sunset and sunrise compared to using time of day as an independent variable.

To emphasize accurate prediction of peak migration, we introduce weighting of the residuals for deviance calculations while fitting and evaluating the models:

$$weight = 1 + bird\ density^{0.75}. \tag{1}$$

Weighing in this way is equivalent to an observation being made as many times as the weighing factor. The weighting scheme was decided on by inspecting the phenological models and balancing the ability of the model to capture peak periods while not emphasizing migration peaks that occur in single years (e.g. 3–4 October 2016; Supporting Information Sections 3 and 6.2.3). We use this weighting throughout our analysis both while fitting the models and while evaluating them.

Using the phenological model as a basis, we fit models that include environmental variables. Each model consists of a tensor spline with the predictions of the phenological model, representing the expected seasonal migration, together with one environmental variable or a pairwise combination of two variables (Table 1). Variables related to the solar position were included to capture interactions between the time of day and environmental conditions. We exclude interactions between conditions at different remote sites as these seem biologically irrelevant and interactions between different wind conditions as these would result in high-dimensional models. This procedure results in 261 different environmental models in spring and 335 in fall. These numbers vary per season as the number of remote locations included in the models differs.

To identify the best model from these different environmental models, we use cross-validations. The dataset was split 10 times for cross-validation datasets with a 70:30 division. We excluded regular spaced continuous sections from the time-ordered data for cross-validation to avoid the influence of temporal correlations. For each environmental model, we calculate the deviance by the sum of squares to the excluded 30% of data. Using these deviances, we have two strategies to select models for final predictions that we compare, either select the best performing model or select the top 10 models

for each cross-validation. To calculate final predictions, we compare both strategies.

We investigated the importance of the various environmental predictors by exploring how frequently they get selected in the cross-validations. By investigating the percentage of models where variables occur in, we summarize which environmental conditions contribute most. The range indicates the variability in this measure for the four datasets where a year has been omitted.

2.4 | Prediction evaluation

To evaluate the overall performance of our approach, we omit single years from our dataset. We selected 2009, 2014, 2016, and 2019, as these are distributed over the study period. Predictions of the ensemble model and models that have been performing best in at least one cross-validation are made for these single omitted years, and deviances of predictions are calculated. This allows for to evaluate the model independently of the data used for model development (Dormann et al., 2018). To put these deviances in context, we calculate the ratio between these deviances and the deviance calculated from predictions based on the phenological model only. This evaluation allows to investigate the added value of including weather and accumulation input compared to the base expectation of the phenological model. We furthermore explore receiver operating characteristic (ROC) curves to visualize the predictive performance for BirdTAM density thresholds (10, 20, 40 birds/km^2) of the phenological and ensemble model (Fawcett, 2006).

3 | RESULTS

The phenological model, when applied to the full dataset captures the seasonal and diurnal trends of migration (Figure 1). Both seasons show a high density of birds just after civil dusk as well as civil dawn in fall. The seasonal timing of these peaks occurs in the second half of March in spring and at the beginning of October during fall. The peak of migration is higher in fall compared to spring. Another noticeable feature is the high density of migrants just before midnight in spring (Figure 1a). When fitting a gamma distribution to the nightly mean density for the peak periods (second half of March, first half of October), we find that the means for fall nights are more spread (spring: $\alpha = 1.57\ \beta = 0.17$; fall: $\alpha = 1.12\ \beta = 0.11$). We used a gamma distribution as it fits the skewed nature of the data.

The environmental conditions make it possible to describe times with higher and lower migration intensities than expected based on the phenological model alone (Figure 2). By exploring weights in the ensemble, we investigate the importance of environmental conditions (Table 1). Wind conditions are represented in 51% (35–62%) of the total ensemble weight in spring and 30% (25–38 %) of the weight in fall. In spring wind conditions in the United Kingdom turn out to be important, contributing 14% (5–10%) of all weights. Other important factors are the accumulation variables, which are included in 33% (23–39%) of
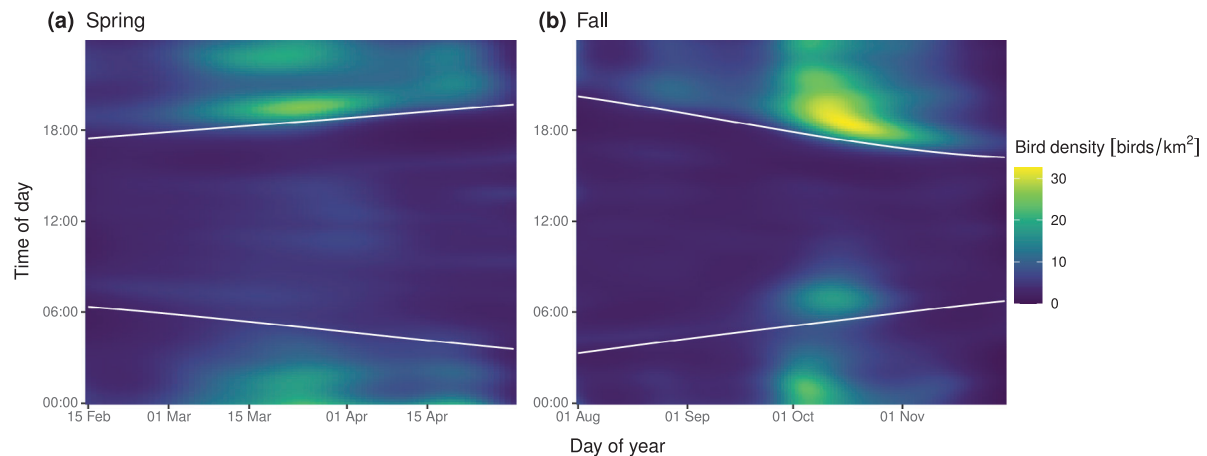
FIGURE 1 The seasonal migration trends as fitted over all the data. The phenological trends are visualized over time of year and time of day. The model is fitted as a function of time of year, solar elevation and the time derivative of solar elevation. The peaks of migration after civil dusk and dawn (white line) are visible during the peak migratory periods at the end of March and the beginning of October

the spring models and 68% (59–71%) of the fall models. Most of these are the derivative of the accumulation, respectively, 18% (17–25%) and 66% (56–65%) of the total weight. All combined remote conditions get weighted less in spring (17% (6–19%)) compared to fall (26% (16–37%)). Surface pressure and the 24-h change in surface pressure either remote or local occur hardly in the spring models (1% (1–8%)), while they are more common in fall (29% (22–43%)).

The average performance of the ensemble model, measured as the improvement over the phenological model, is better than the single environmental model (Figure 3). Furthermore, the variation in performance between the models that were best in one of the cross-validations compared to the ensemble is much higher. On several occasions, single models perform even worse than the phenological model, which is a sign of over-fitting. In contrast, the ensemble models always perform better than the phenological model. The average deviance ratio for the ensemble model in fall is lower compared to spring, indicating that environmental variables contribute more to accurate predictions in fall. The more accurate predictions of the ensemble model in fall compared to spring are also seen in the ROC curves for predicting thresholds of migratory intensity (Figure 3; Supporting Information Section 4). In spring, the difference between the ensemble models and phenological models is lower and the area under the curve is lower (Figure 3b,d).

Performance characteristics can also be used to identify how frequently warnings (BirdTAM > 5) for aviation safety would be raised, to do this we use ROC curves. Using the criterion that 75% of the time (i.e. of the 5-min observations) where the bird density exceeds 10 birds/km^2 should be correctly identified (true positive fraction), we calculate the percentage of time the warning should be raised using the ensemble model. In spring that would be 23.9% of the time and in fall 14.5% of the time. The spring warning period is considerably longer compared to fall; this relates to the difference in model performance (Figure 3; Supporting Information Section 4). Furthermore, the number of false positives also relates to the rarity that the threshold is

exceeded in the observed data (11.1% of the time in spring and 7.4% of the time in fall).

4 | DISCUSSION

Based on our findings, we recommend an ensemble model over single models when considering predictive applications such as providing warnings for aviation flight planning. The average performance is better, and the performance is less variable. Wind conditions are the most important contributor to environmental models. This contribution is through the accumulation of birds and direct effects.

In contrast to other models (Van Belle et al., 2007; Van Doren & Horton, 2018), we model both seasons separately to be able to compare them. There are large differences between spring and fall both in the prediction quality and the selected environmental variables. This warrants modelling the seasons as separate processes. These differences are already noticeable in the phenological model and to some extent can be explained by regional migration dynamics. The midnight peak in spring, for example, likely relates to migrants arriving from the United Kingdom, several hours after departure (Figure 1; Bradarić et al., 2020). To investigate this further, we take a detailed look at the average migration directions between the 10th and 25th of March for times with a high density of migrants (20 birds/km^2). This is more northerly (20:00–21:00 UTC, 57.1°) at the beginning of the night compared to later in the night (23:00–00:00 UTC, 72.7°) when migrants with a more easterly direction from the United Kingdom arrive. Response to weather also differs among seasons. In spring, air temperature at elevation is an important variable (Table 1). This corresponds to earlier findings in the United States for spring migration (Van Doren & Horton, 2018). In fall, there is high importance of the accumulation of variables. This can be explained by the poorer migratory conditions in fall where fewer days are suitable for migration (Kemp et al., 2010) which results in higher variability of the density of migrants between days. This means
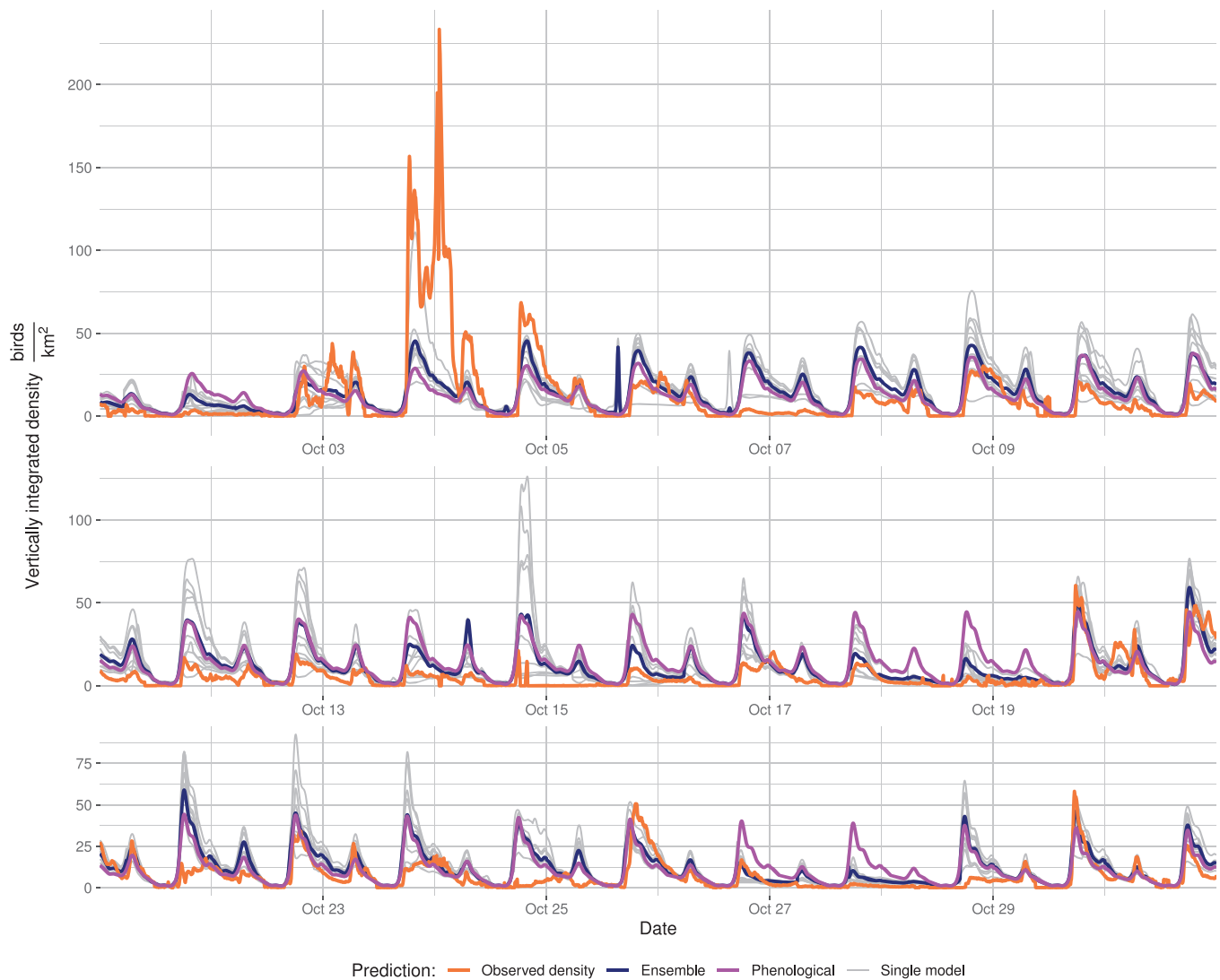
FIGURE 2 An example of the predictive model applied to the omitted October data of the year 2016. The orange line is the observed density. The purple line is the prediction by the phenological model, while the dark blue line corresponds to the ensemble model. Grey lines correspond to single environmental models

that environmental models have a larger impact in the fall compared to spring (Figure 3). It is important to realize that the phenological model also implicitly captured regional climatic trends which influence phenology. Small differences between the performance of the phenological and environmental models can have consequences for application too as the phenological model does not change between years. The phenological model can be used for planning of downtime and maintenance weeks or even years ahead as it provides a long-term expectation of periods of high migration densities.

By exploring the deviation between the predictions and observations for omitted years, it is possible to explore possible improvements. On the night of 3–4 October 2016, there is a migration peak that is not predicted accurately (Figure 2). However, when looking at the source data (Supporting Information Section 5), this night has exceptional migration densities across a large proportion of western Europe (Nilsson et al., 2019). On no other night of the 10-year dataset, a migra-

tory peak of this magnitude occurred (Supporting Information Sections 5 and 6); and only if similar peaks are part of the training data, we can expect that such peaks are predicted well. Another exceptional event is the high density of migrants in the second half of April 2019 (Supporting Information Sections 5 and 6). We think this relates to the ongoing easterly wind in this period displacing migrants that would normally migrate through Germany (see Lensink (2002)). To capture these effects, additional variables such as the accumulation of easterly winds can be helpful. These once in 10-year events show the importance of long training datasets and the necessity for suitable data archives (Shamoun-Baranes et al., 2021).

Earlier predictive models are a powerful tool for planning military aviation exercises (Van Gasteren et al., 2019). These models were based on two air defence radars, while a network of weather radars in north-west Europe is currently used to issue BirdTAMs. By improving on existing models through expanding the training dataset, and the
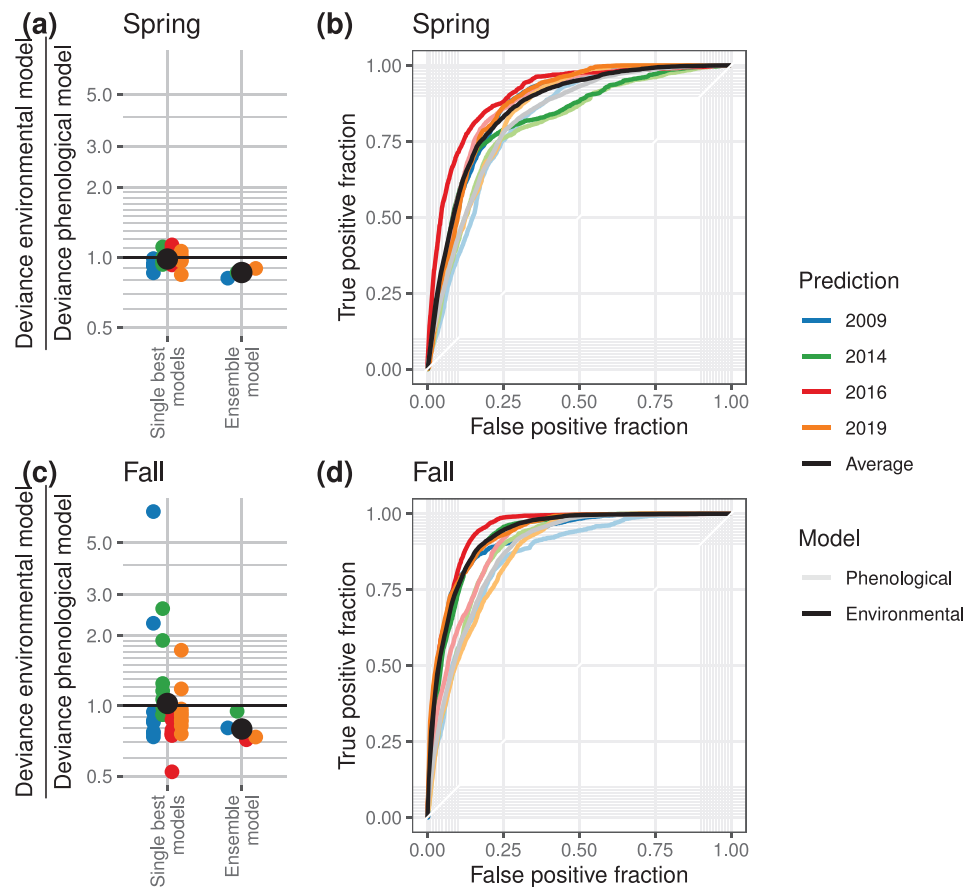
FIGURE 3 The performance of the different modelling strategies compared to the phenological model. Colours correspond to the different years for which data have been omitted. (a, c) Deviance ratios below one show that the environmental model performed better than the phenological model. The black point shows the geometric mean. The ensemble models on the right side have a lower deviance ratio and thus perform better. For the single models, there are multiple points per year, as different models performed best in each of the 10 cross validations. (b, d) The ROC curves for the ensemble models predicting a threshold of 10 birds/km^2. Curves are plotted both for the phenological (pale colours) and environmental model (bold colour). The black line is the average between the different ROC curves. Using these plots, we can identify the effectiveness of the model for identifying high-density migration events (e.g. in fall with the right threshold in the environmental model 75% of the occurrences of more than 10 birds/km^2 can be identified while only in 10% of fall cases with a low density of migrants the warning is raised, that is false negatives/positives)

environmental data used we expect these models to be of a higher value because they are based on the same data source (weather radars) and a 10-year long dataset. From studying the model, we gain ecological insight into the differences between spring and fall. Understanding the difference helps us improve on the models and improve future predictions. Similar models could also be valuable for civil aviation to support decisions (Nilsson et al., 2021) and to support shutdown or curtailment decisions for wind turbines (Marques et al., 2014).

The experiences from these predictive models are not unique to predicting migration for military aviation. Similar structural differences between seasons and selection of environmental variables are likely for models specific to predicting migration with the aim to mitigate the risk of birds colliding with wind turbines. Understanding the structure of the predictability, for example the quality of a phenological model, can be important for planning of replacement capacity when wind turbines are shut down.

CONFLICT OF INTEREST

The authors do not have a conflict of interest to declare.

AUTHOR CONTRIBUTIONS

DATA AVAILABILITY STATEMENT

The dataset has been derived from weather radars of the KNMI (NL/DBL: https://dataplatform.knmi.nl/dataset/radar-tar-volume-debilt-1-0 and NL/HRW: https://dataplatform.knmi.nl/dataset/radar-tar-vol-full-herwijnen-1-0) and annotated using weather data from the ECMWF ERA 5 dataset (https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5 (Hersbach et al., 2020)). The dataset used is available from Kranstauber et al. (2022).

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1002/2688-8319.12158

ORCID

Bart Kranstauber https://orcid.org/0000-0001-8303-780X
Willem Bouten https://orcid.org/0000-0002-5250-8872
Hans van Gasteren https://orcid.org/0000-0002-4810-3899
Judy Shamoun-Baranes https://orcid.org/0000-0002-1652-7646

REFERENCES

Åkesson, S., & Hedenström, A. (2000). Wind selectivity of migratory flight departures in birds. *Behavioral Ecology and Sociobiology*, 47(3), 140–144. https://doi.org/10.1007/s002650050004

Alerstam, T. (1978). Analysis and a theory of visible bird migration. *Oikos*, 30(2), 273–349. https://doi.org/10.2307/3543483

Bradarić, M., Bouten, W., Fijn, R. C., Krijgsveld, K. L., & Shamoun-Baranes, J. (2020). Winds at departure shape seasonal patterns of nocturnal bird migration over the North Sea. *Journal of Avian Biology*, 51(10), e02562. https://doi.org/10.1111/jav.02562

Dokter, A. M., Farnsworth, A., Fink, D., Ruiz-Gutierrez, V., Hochachka, W. M., Sorte, F. A. L., Robinson, O. J., Rosenberg, K. V., & Kelling, S. (2018). Seasonal abundance and survival of North America's migratory avifauna determined by weather radar. *Nature Ecology & Evolution*, 2(10), 1603–1609. https://doi.org/10.1038/s41559-018-0666-4

Dokter, A. M., Liechti, F., Stark, H., Delobbe, L., Tabary, P., & Holleman, I. (2011). Bird migration flight altitudes studied by a network of operational weather radars. *Journal of the Royal Society Interface*, 8(54), 30–43. https://doi.org/10.1098/rsif.2010.0116

Dormann, C. F., Calabrese, J. M., Guillera-Arroita, G., Matechou, E., Bahn, V., Bartoń, K., Beale, C. M., Ciuti, S., Elith, J., Gerstner, K., Guelat, J., Keil, P., Lahoz-Monfort, J. J., Pollock, L. J., Reineking, B., Roberts, D. R., Schröder, B., Thuiller, W., Warton, D. I., … Hartig, F. (2018). Model averaging in ecology: A review of Bayesian, information-theoretic, and tactical approaches for predictive inference. *Ecological Monographs*, 88(4), 485–504. https://doi.org/10.1002/ecm.1309

Erni, B., Liechti, F., Underhill, L., & Bruderer, B. (2002). Wind and rain govern the intensity of nocturnal bird migration in central Europe - A log-linear regression analysis. *Ardea*, 90, 155–166.

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8), 861–874. https://doi.org/10.1016/j.patrec.2005.10.010

Gwinner, E., & Helm, B. (2003). Circannual and Circadian Contributions to the Timing of Avian Migration. In P. Berthold, E., Gwinner, & E. Sonnenschein (Eds.), *Avian migration* (pp. 81–95). Springer. https://doi.org/10.1007/978-3-662-05957-9_5

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., … Thépaut, J.-N. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730), 1999–2049. https://doi.org/10.1002/qj.3803

Horton, K. G., Van Doren, B. M., Stepanian, P. M., Farnsworth, A., & Kelly, J. F. (2016). Seasonal differences in landbird migration strategies. *The Auk*, 133(4), 761–769. https://doi.org/10.1642/AUK-16-105.1

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning*. Springer Texts in Statistics, Vol. 103. Springer New York. https://doi.org/10.1007/978-1-4614-7138-7

Kemp, M. U. (2012). *How birds weather the weather: Avian migration in the mid-latitudes* [PhD Thesis]. University of Amsterdam. https://hdl.handle.net/11245/1.379371

Kemp, M. U., Shamoun-Baranes, J., Van Gasteren, H., Bouten, W., & Van Loon, E. E. (2010). Can wind help explain seasonal differences in avian migration speed? *Journal of Avian Biology*, 41(6), 672–677. https://doi.org/10.1111/j.1600-048X.2010.05053.x

Kranstauber, B., van Gasteren, H., Bouten, W., & Shamoun-Baranes, J. Z., Judy. (2022). *Dataset for predictive bird migration model*. figshare, http://doi.org/10.21942/uva.19416215.v1

Lensink, R. (2002). Patronen in de trek over Nederland. In *Vogeltrek over Nederland (1973-1993)* (pp. 331–369). Schuyt & co.

Liechti, F. (2006). Birds: Blowin' by the wind? *Journal of Ornithology*, 147(2), 202–211. https://doi.org/10.1007/s10336-006-0061-9

Loss, S. R., Will, T., Loss, S. S., & Marra, P. P. (2014). Bird–building collisions in the United States: Estimates of annual mortality and species vulnerability. *The Condor*, 116(1), 8–23. https://doi.org/10.1650/CONDOR-13-090.1

Manola, I., Bradarić, M., Groenland, R., Fijn, R., Bouten, W., & Shamoun-Baranes, J. (2020). Associations of synoptic weather conditions with nocturnal bird migration over the North Sea. *Frontiers in Ecology and Evolution*, 8, 542438. https://doi.org/10.3389/fevo.2020.542438

Marques, A. T., Batalha, H., Rodrigues, S., Costa, H., Pereira, M. J. R., Fonseca, C., Mascarenhas, M., & Bernardino, J. (2014). Understanding bird collisions at wind farms: An updated review on the causes and possible mitigation strategies. *Biological Conservation*, 179, 40–52. https://doi.org/10.1016/j.biocon.2014.08.017

Nilsson, C., Dokter, A. M., Verlinden, L., Shamoun-Baranes, J., Schmid, B., Desmet, P., Bauer, S., Chapman, J., Alves, J. A., Stepanian, P. M., Sapir, N., Wainwright, C., Boos, M., Górska, A., Menz, M. H. M., Rodrigues, P., Leijnse, H., Zehtindjiev, P., Brabant, R., … Liechti, F. (2019). Revealing patterns of nocturnal migration using the European weather radar network. *Ecography*, 42(5), 876–886. https://doi.org/10.1111/ecog.04003

Nilsson, C., Klaassen, Raymond H. G., & Alerstam, T. (2013). Differences in speed and duration of bird migration between spring and autumn. *The American Naturalist*, 181(6), 837–845. https://doi.org/10.1086/670335

Nilsson, C., Sorte, F. A. L., Dokter, A., Horton, K., Van Doren, B. M., Kolodzinski, J. J., Shamoun-Baranes, J., & Farnsworth, A. (2021). Bird strikes at commercial airports explained by citizen science and weather radar data. *Journal of Applied Ecology*, 58, 2029–2039. https://doi.org/10.1111/1365-2664.13971

Nussbaumer, R., Benoit, L., Mariethoz, G., Liechti, F., Bauer, S., & Schmid, B. (2020). *Modelling the flow of nocturnal bird migration with year-round European weather radar network*. bioRxiv, 2020.10.13.321844. https://doi.org/10.1101/2020.10.13.321844

Richardson, W. J. (1990). Timing of bird migration in relation to weather: updated review. In E. Gwinner (Ed.), *Bird migration* (pp. 78–101).

Springer Berlin Heidelberg. http://link.springer.com/chapter/10.1007/978-3-642-74542-3_6

Roscher, R., Bohn, B., Duarte, M. F., & Garcke, J. (2020). Explainable Machine Learning for Scientific Insights and Discoveries. *IEEE Access*, 8, 42200–42216. https://doi.org/10.1109/ACCESS.2020.2976199

Safi, K., Kranstauber, B., Weinzierl, R., Griffin, L., Rees, E. C., Cabot, D., Cruz, S., Proaño, C., Takekawa, J. Y., Newman, S. H., Waldenström, J., Bengtsson, D., Kays, R., Wikelski, M., & Bohrer, G. (2013). Flying with the wind: Scale dependency of speed and direction measurements in modelling wind support in avian flight. *Movement Ecology*, 1(1), 4. https://doi.org/10.1186/2051-3933-1-4

Shamoun-Baranes, J., Bauer, S., Chapman, J. W., Desmet, P., Dokter, A. M., Farnsworth, A., Haest, B., Koistinen, J., Kranstauber, B., Liechti, F., Mason, T. H. E., Nilsson, C., Nussbaumer, R., Schmid, B., Weisshaupt, N., & Leijnse, H. (2021). Weather radars' role in biodiversity monitoring. *Science*, 372(6539), 248–248. https://doi.org/10.1126/science.abi4680

Shamoun-Baranes, J., van Gasteren, H., & Ross-Smith, V. (2017). Sharing the Aerosphere: Conflicts and Potential Solutions. In P. B. Chilson, W. F. Frick, J. F. Kelly, & F. Liechti (Eds.), *Aeroecology* (pp. 465–497). Springer International Publishing. https://doi.org/10.1007/978-3-319-68576-2_18

Van Belle, J., Shamoun-Baranes, J., Van Loon, E., & Bouten, W. (2007). An operational model predicting autumn bird migration intensities for flight safety. *Journal of Applied Ecology*, 44(4), 864–874. https://doi.org/10.1111/j.1365-2664.2007.01322.x

Van Doren, B. M., & Horton, K. G. (2018). A continental system for forecasting bird migration. *Science*, 361(6407), 1115–1118. https://doi.org/10.1126/science.aat7526

Van Doren, B. M., Horton, K. G., Dokter, A. M., Klinck, H., Elbin, S. B., & Farnsworth, A. (2017). High-intensity urban light installation dramatically alters nocturnal bird migration. *Proceedings of the National Academy of Sciences*, 114(42), 11175–11180. https://doi.org/10.1073/pnas.1708574114

Van Gasteren, H., Krijgsveld, K. L., Klauke, N., Leshem, Y., Metz, I. C., Skakuj, M., Sorbi, S., Schekler, I., & Shamoun-Baranes, J. (2019). Aeroecology meets aviation safety: Early warning systems in Europe and the Middle East prevent collisions between birds and aircraft. *Ecography*, 42(5), 899–911. https://doi.org/10.1111/ecog.04125

Wood, S. N., Goude, Y., & Shaw, S. (2015). Generalized additive models for large data sets. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 64(1), 139–155. https://doi.org/10.1111/rssc.12068

SUPPORTING INFORMATION

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