An operational model predicting autumn bird migration intensities for flight safety

van Belle, J.; Shamoun-Baranes, J.; van Loon, E.E.; Bouten, W.

Published in:
Journal of Applied Ecology

DOI:
10.1111/j.1365-2664.2007.01322.x

Journal of Applied Ecology 2007
44, 864–874

An operational model predicting autumn bird migration intensities for flight safety

J. VAN BELLE,*† J. SHAMOUN-BARANES,* E. VAN LOON* and W. BOUTEN*

**Computational Biogeography and Physical Geography, Institute for Biodiversity and Ecosystem Dynamics, Universiteit van Amsterdam, Nieuwe Achtergracht 166, 1018 WV Amsterdam, the Netherlands; and †Nature Bureau, Support Operations Branch, Directorate of Operations, Royal Netherlands Air Force, PO Box 8762, 4820 BB Breda, the Netherlands*

Summary

1. Forecasting migration intensity can improve flight safety and reduce the operational costs of collisions between aircraft and migrating birds. This is particularly true for military training flights, which can be rescheduled if necessary and often take place at low altitudes and during the night. Migration intensity depends strongly on weather conditions but reported effects of weather differ among studies. It is therefore unclear to what extent existing predictive models can be extrapolated to new situations.
2. We used radar measurements of bird densities in the Netherlands to analyse the relationship between weather and nocturnal migration. Using our data, we tested the performance of three regression models that have been developed for other locations in Europe. We developed and validated new models for different combinations of years to test whether regression models can be used to predict migration intensity in independent years. Model performance was assessed by comparing model predictions against benchmark predictions based on measured migration intensity of the previous night and predictions based on a 6-year average trend. We also investigated the effect of the size of the calibration data set on model robustness.
3. All models performed better than the benchmarks, but the mismatch between measurements and predictions was large for existing models. Model performance was best for newly developed regression models. The performance of all models was best at intermediate migration intensities. The performance of our models clearly increased with sample size, up to about 90 nocturnal migration measurements. Significant input variables included seasonal migration trend, wind profit, 24-h trend in barometric pressure and rain.
4. *Synthesis and applications*. Migration intensities can be forecast with a regression model based on meteorological data. This and other existing models are only valid locally and cannot be extrapolated to new locations. Model development for new locations requires data sets with representative inter- and intraseasonal variability so that cross-validation can be applied effectively. The Royal Netherlands Air Force currently uses the regression model developed in this study to predict migration intensities 3 days ahead. This improves the reliability of migration intensity warnings and allows rescheduling of training flights if needed.

Key-words: bird migration, bird strikes, cross-validation, prediction, radar ornithology, robustness, weather influence

Journal of Applied Ecology (2007) **44**, 864–874
doi: 10.1111/j.1365-2664.2007.01322.x

Correspondence: Jelmer van Belle, Computational Biogeography and Physical Geography, Institute for Biodiversity and Ecosystem Dynamics, Universiteit van Amsterdam, Nieuwe Achtergracht 166, 1018 WV Amsterdam, the Netherlands (fax + 31 76 5447354; e-mail jelmer.van.belle@hetnet.nl).

Introduction

Collisions with birds pose serious safety risks for aviation and by 1999 had resulted in the loss of at least 52 civil and 283 military aircraft (Allan 2002). Although such accidents are rare and 65% of the bird strikes do not result in damage at all, the global cost of bird strikes, including the effects of cancelled flights and delays, are estimated at 1–1·5 billion USD year^{-1} (Allan 2002). Most bird strikes take place during take-off and landing, but chances of bird strikes *en route* increase dramatically during peak periods of bird migration (Blokpoel 1976; Sodhi 2002; Dekker, van Gasteren & Shamoun-Baranes 2003). For military training flights, which often occur at low altitudes and during nocturnal hours, these increased risks are particularly important; in contrast with civil aviation, rescheduling of military training flights is often possible. Therefore, in order to improve flight safety, several air forces issue bird migration intensity warnings, which result in flight restrictions and occasionally flight cancellations (for example see https://www.notams.jcs.mil/common/birdtam.html, accessed 24 April 2007). Migration intensity warnings for the Royal Netherlands Air Force (RNLAF) are based on radar measurements of bird densities (bird echoes km^2) from the north of the Netherlands (Buurma 1995). Spatial extrapolation over the rest of the Netherlands and adjacent countries is based on expert judgement of observed bird flight directions and knowledge of seasonal patterns. Multiday forecasts of migration intensities were not available before this study. The great impact of migration warnings on flight operations, their spatial uncertainties and abrupt temporal changes have prompted the RNLAF to improve the reliability of the warnings and investigate the potential of migration forecasts.

Several migratory flyways pass the Netherlands. In autumn, the most important flyway consists of passerines migrating from Scandinavia towards southern Europe and Africa, crossing the east of the Netherlands in a south-westerly direction (Berthold 2001). Another very important flyway consists of birds from north-east Europe wintering in the British Isles. These birds cross the north of the Netherlands in a west-south-westerly direction and usually follow the western coast of the Netherlands to cross the North Sea further south (Lack 1963). A few times each autumn, massive numbers of migrants cross the Netherlands in a southerly direction. According to their arrival times in the second half of the night, these birds apparently depart from Norway and cross the North Sea in a direct flight, landing in the northern part of the Netherlands (Buurma 1987).

Since the late 1950s, radars have been used to quantify bird migration (Sutter 1957; Buurma 1995; Bruderer 1997; Gauthreaux & Belser 1998) and to model the relationships between bird migration and weather with multivariate statistics (Nisbet & Drury 1968; Alerstam & Ulfstrand 1972; Able 1973). In several countries regression models have been developed to prevent collisions of aircraft with birds (Blokpoel 1969; Geil *et al*. 1974) and to assess the impact of wind turbines on migratory bird populations (Gauthreaux & Belser 2003; Barrios & Rodrígues 2004). Wind, rain, sea-level pressure, temperature, cloud cover and derivatives of all of these are the common predictors of migration (Richardson 1990; Zehnder *et al*. 2001; Erni *et al*. 2002; Schaub, Liechti & Jenni 2004). However, the relative contributions of these meteorological variables to predictions of migration densities differ strongly between studies. This raises the question: can weather-based models be extrapolated beyond the calibration domain to predict bird migration intensities in new situations?

In order to improve migration warnings for the RNLAF, our first objective was to determine how well existing regression models from other locations on the south-westerly migratory flyway in Europe predict migration intensities over the north of the Netherlands. We tested this by applying three existing models that describe the relationships between weather and migration intensities to Dutch weather data and measured migration densities. Each model consists of input variables and their coefficients. Our second objective was to develop a regression model that would provide more robust and precise nocturnal migratory intensity predictions for the Netherlands than the existing models. In ecological studies focusing on the influence of meteorological conditions on the spatial and temporal distribution of organisms, the variability in the data and the size of the data set are known to have a strong influence on model robustness (Beck 2002). Therefore our third objective was to determine the minimum calibration sample size for creating robust models.

Materials and methods

MIGRATION INTENSITY DATA

The S-band (10-cm wavelength) 20-MW medium-power radar (Thomson ARES, Thomson ICF, Bagneux, France) used in this study is situated in the north-east of the Netherlands and detects birds for up to 150-km distances. The measurement window is located 50–60 km from the radar, 20–60 km from the northern coast of the Netherlands, with its centre at 52°54′N and 6°11′E, and has a total surface of 850 km^2. The sampled altitudes in the measurement window range from 100 to 1100 m.

Nocturnal autumn migration intensities were measured from 1 August until 15 November of 2001, 2002 and 2003. During 10 min of each hour, bird echoes were extracted from the radar image with the automated tracking algorithm of ROBIN (radar observations of bird intensities and numbers, developed by TNO Defense, Security and Safety, The Hague, the Netherlands). Migration intensity was defined as the number of bird echoes per km^2, automatically corrected for the area with unwanted echoes from land or rain ('clutter') within the image (Buurma 1995). The altitude range of the radar beam is generally greater than the upper limit of

Journal of Applied Ecology, **44**, 864–874

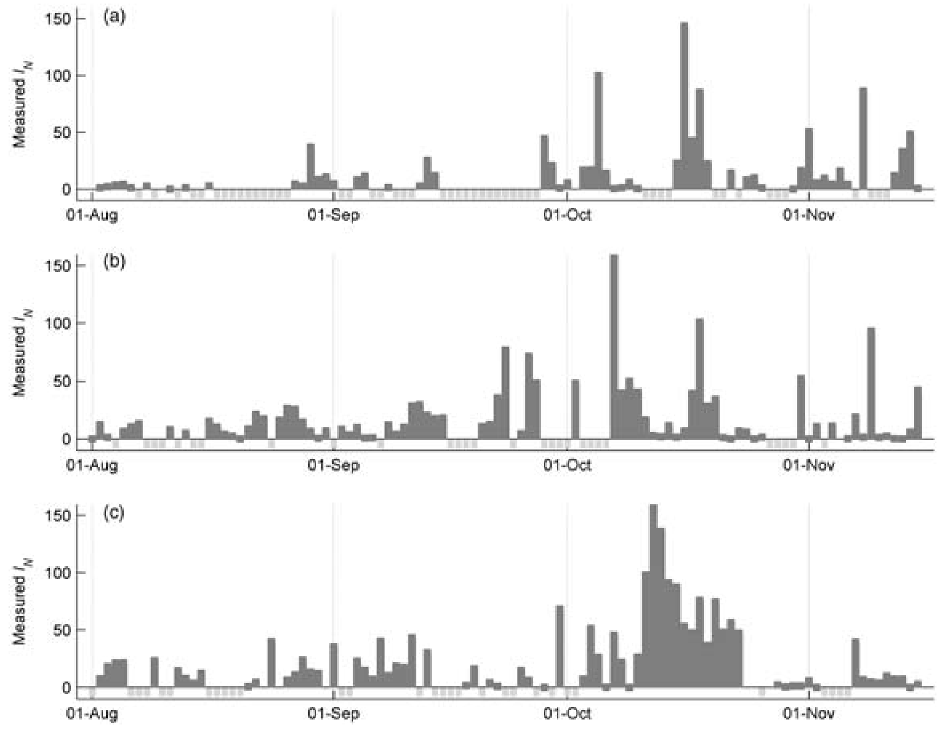

Fig. 1. Measured total migration intensities (I_N) for 2001 (a), 2002 (b) and 2003 (c), not log-transformed. For 7 October 2002 and 12 October 2003, peaks are not shown (167 and 352 bird echoes km^{-2}, respectively). Light grey bars represent missing data.

migration, but within the measurement window the altitude distribution cannot be determined. Therefore we expressed migration intensity as echo density per unit area (km^2) in the 100–1100 altitude band, rather than per unit volume, as is common in radar ornithology.

Gaps in the data set were generally because of radar downtime, particularly in September 2001, and intense rain clutter. During intense rain, resulting in 100% rain clutter detection, migration intensity was set to 0. In several of the hourly measurements (< 0·01% of all measurements), bird echoes were wrongly classified as rain because of the high aerial bird densities (exceeding 50 echoes km^{-2}). Migration intensities were then corrected manually from the clutter numbers. Another phenomenon that affected data quality was 'anomalous propagation', which occurred when the radar beam was trapped below a low-level temperature inversion layer. As a result, the air was sampled at lower altitudes than usual, leading to larger land clutter areas and wrongly sampled bird densities. Measurements with anomalous propagation were excluded from the analysis.

Before analysis, the radar data were pre-processed. Nights with more than 50% missing data were omitted from further analysis. Nights with complete hourly measurements were used to establish a standard hourly nocturnal trend. Missing values in the data set were then interpolated using the measured hourly intensities of that night and the standard nocturnal trend, adjusted for the length of the night. Finally, the total nocturnal migration intensity used in all subsequent analyses was the sum of hourly nocturnal migration intensities between sunset and sunrise. This resulted in 214 measurements of total nocturnal migration intensity I_N (echoes km^{-2} night^{-1}; Fig. 1; yearly summaries Table 1). To meet normality requirements for the regression analyses, migration intensities were transformed to their natural logarithm (Zar 1984).

Table 1. Summary of bird measurement data: mean number of nights, means and standard deviations of migration intensities

Year	*n*	Mean	σ
2001	58	17·71	27·48
2002	79	19·52	27·08
2003	77	28·35	46·15

Mean nocturnal migration intensities, measured during 1989–95 with the same radar, were used to establish a seasonal migration intensity trend. This trend served as a baseline for our models and was used as an independent input variable (I_b) in addition to meteorological variables that were expected to explain deviations from the seasonal trend.

METEOROLOGICAL DATA

Meteorological surface measurements of barometric pressure (P), temperature (T), wind speed (S_w), wind directions (α_w) and precipitation (R) were taken from the meteorological station Eelde (53°10′N and 6°35′E), 25 km north-east of the centre of the migration measurement window (source Netherlands Meteorological Institute KNMI, De Bilt, the Netherlands). In addition, wind speeds and directions at 1000, 925 and 850 hPa pressure levels (approximately 0, 750 and 1500 m altitude, respectively) were taken from sounding data measured at Emden (53°21′N and 7°13′E), 65 km north-east of the centre of the migration measurement window (source NOAAs National Oceanic & Atmospheric Administration, Washington, US) radiosonde database; http://raob.fsl.noaa.gov/, accessed 24 April 2007).

All sounding and surface data were collected at 23:00 UTC (Universal Time Code). Additionally, hourly precipitation amounts from Eelde were integrated over the period between sunset and sunrise. Twenty-eight weather variables were tested. These included several derivatives of surface and altitudinal wind speeds and directions, barometric pressure, temperature and precipitation. A description of all input variables that were tested and how they were derived is included in Table 2.

TESTS AND BENCHMARKS

To quantify the prediction performance of models, we used the root mean squared error (RMSE) and compared the 95 percentile range of predictions with the 95 percentile range of measurements. We compared the 95 percentile range because of the technical limitations influencing measurements at extremely low and high migration intensities (see migration intensity data for more details). The RMSE of each model was also compared with two benchmarks: (i) prediction of bird migration by the measured intensity of the previous night (RMSE = 1·41) and (ii) prediction of bird migration intensity by the baseline seasonal trend (I_b, RMSE = 1·23).

EXISTING MODELS

To test whether existing models can be extrapolated to new locations, we applied three log-linear regression models from southern Sweden (Zehnder *et al*. 2001), Denmark (Geil *et al*. 1974) and southern Germany (Erni *et al*. 2002) to the Dutch bird migration and meteorology data from 2001 to 2003. These studies explained 0·44–0·70 of the variance in local migration intensity or traffic rate.

The model for southern Sweden, M_S, is based on nightly means of migration traffic rates (birds km^{-1} h^{-1}), obtained with a passive infra red camera in fixed upright position, recorded from 7 August to 28 October 1998 (Zehnder *et al*. 2001). This model relates the natural logarithm of nocturnal migration traffic rates, I_S, to barometric pressure, P (hPa), 24-h change in barometric pressure, ΔP (hPa), wind speed, S_w (m s^{-1}) and tail wind speed, W_{t225} (m s^{-1}):

$$I_S = 26{\cdot}98 - 0{\cdot}11S_w + 0{\cdot}06W_{t225} - 0{\cdot}02P + 0{\cdot}03\Delta P \qquad \text{eqn 1}$$

The model for Denmark, M_D, is based on 4 years (1968–71) of radar data from an airfield surveillance radar with a wavelength of 20 cm (Geil *et al*. 1974). We applied their September model to the Dutch data from September. M_D relates the natural logarithm of migration intensity, I_D (bird echoes km^{-2}), in the first half of the night to P (hPa), ΔP (hPa), temperature T (°C) and lateral wind drift W_{d200} (m s^{-1}), defined as the wind speed component perpendicular to a mean flight direction of 200°:

$$I_D = -50{\cdot}12 + 0{\cdot}054P + 0{\cdot}105T + 0{\cdot}0291\Delta P + 0{\cdot}469W_{d200} \qquad \text{eqn 2}$$

The model for Germany, M_G, is based on conical scans with a pencil beam of a small-scale 3-cm wavelength tracking radar during one autumn season, 1 August to 29 October 1987 (Erni *et al*. 2002). For this model, four migration density measurements (bird echoes km^{-2}) per night were summed to calculate nocturnal migration intensity. The natural logarithm of migration intensity (I_G) is related to unfavourable wind (W_{u7}), nightly proportion of hours with rain (R_{pG}), the effect of accumulated nights with rain (Acc_r) and two short-term trend variables that account for the increasing (I_{Gu}) and decreasing (I_{Gd}) migration intensities at the start and the end of the season:

$$I_G = 5{\cdot}41 - 0{\cdot}04I_{Gu} - 0{\cdot}15I_{Gd} + 0{\cdot}12W_{u7} - 0{\cdot}60R_{pG} + 0{\cdot}13Acc_r \qquad \text{eqn 3}$$

Quantifying bird migration intensity with different measurement techniques and for different periods may lead to different absolute numbers. Therefore we corrected each model by subtracting the mean differences between predicted and observed values. Theoretically, these correction factors are multiplicative for non-logarithmic measurements and therefore additive for the logarithmic models. Consequently, the correction does not affect the coefficients that reflect the relative importance of the explanatory variables in the models.

INPUT VARIABLES FOR NEW MODELS

We developed new regression models for the Dutch bird density and weather data to acquire more robust and precise predictions of migration intensity over the Netherlands than existing models yielded. The baseline migration intensity (I_b, echoes km^{-2} night^{-1}) was included as an input variable to represent the seasonal migration intensity trend (see the Migration intensity data). We analysed 28 weather variables, including those from the existing models. Here we present only the variables included in our final models. Explanatory variables included the total nocturnal rainfall (mm) (R_t), the proportion of hours with rain intensity > 0·01 mm h^{-1} between sunset and sunrise (R_p), P, T and their 24-h increments (ΔP and ΔT, respectively). We quantified the influence of wind direction and speed on the bird's flight direction and ground speed by calculating the wind profit, which is the distance the wind would carry

Table 2. Description of input variables tested in predictive models

Variable	Name	Units	Description
T	Temperature	°C	Surface temperature
ΔT	24-h temperature increment	°C day^{-1}	24-h surface temperature increment
$\Delta T2$	24-h temperature increment for the second half of the season	°C day^{-1}	24-h surface temperature increment, from 15 September until 15 November. Until 15 September, value was set to 0
Td	Deviation from average temperature	°C day^{-1}	Deviation from average surface temperature for 1991–2000 at Eelde
P	Barometric pressure	hPa	Sea-level barometric pressure
ΔP	24-h pressure increment	hPa day^{-1}	24-h increment in sea-level barometric pressure
S_w	Surface wind speed	m s^{-1}	Surface wind speed
A_w	Surface wind direction	°	Surface wind direction
W_{t225}	Tail wind speed (225°)	m s^{-1}	Surface wind speed component towards a mean flight direction of 225° (Zehnder *et al.* 2001)
W_{d200}	Lateral wind drift (200°)	m s^{-1}	Surface wind speed component perpendicular to a mean flight direction of 200° (Geil *et al.* 1974)
W_{d180}	Lateral wind drift (180°)	m s^{-1}	Surface wind speed component perpendicular to a mean flight direction of 180° (Geil *et al.* 1974)
W_u	Unfavourable wind	–	Categorical variable (0,1) for unfavourable wind, based on wind profit. 1 for wind profits W_p < threshold; the threshold was optimized manually in steps of 0·1 m s^{-1}
W_{u7}	Unfavourable wind	–	Categorical variable for unfavourable wind, based on wind profit W_p (Erni *et al.* 2002). 0 for $W_p = -7$ m s^{-1}, 1 for $W_p < -7$ m s^{-1}
W_p	Mean wind profit	m s^{-1}	Mean of wind profits (equation 4) at 1000, 925 and 850 hPa pressure levels, calculated for a mean bird speed of 12 m s^{-1} and a mean flight direction of 223°
W_{p_m}	Maximum wind profit	m s^{-1}	Maximum of wind profits (equation 4) at 1000, 925 and 850 hPa pressure levels, calculated for a mean bird speed of 12 m s^{-1} and a mean flight direction of 223°
W_d	Mean lateral wind drift	m s^{-1}	Mean of wind speed components perpendicular to a mean flight direction of 223° and bird speed of 12 m s^{-1} measured at 1000, 925 and 850 hPa pressure levels
R_p	Proportion of hours with rain	–	Proportion of hours with rain intensity > 0·01 mm h^{-1} between sunset and sunrise
R_{pG}	Proportion of hours with rain for M_G	–	Proportion of hours with rain defined as 0 to 3 thirds (Erni *et al.* 2002)
R_t	Total nocturnal rainfall	mm	Sum of hourly nocturnal rainfall between sunset and sunrise
$R_{t\_\log}$	Natural logarithm of total nocturnal rainfall	mm	Natural logarithm of sum of hourly nocturnal rainfall between sunset and sunrise
R_{23}	Precipitation at 23:00 UTC	mm	Precipitation from 22:00 to 23:00 UTC
Acc_r	Accumulation effect of rain	–	Accumulation effect (0–1) as a result of rain (Erni *et al.* 2002 and equation 5)
Acc_w	Accumulation effect of wind profit	–	Accumulation effect (0–1) as a result of number of nights with unfavourable wind (equation 5)
I_b	Natural logarithm of baseline migration intensity	Bird echoes km^{-2} night^{-1}	Natural logarithm of seasonal migration intensity trend for the Netherlands based on measurements from 1989 to 1995
I_{Gu}	Increasing migration trend	–	Increasing migration trend in Germany (Erni *et al.* 2002)
I_{Gd}	Decreasing migration trend	–	Decreasing migration trend in Germany (Erni *et al.* 2002)
$\mathrm{D}^{\mathrm{doy}}$	Day of year	–	Day of year (1–365)
I_{b_nl}	Baseline migration intensity	Bird echoes km^{-2} night^{-1}	Seasonal migration intensity trend for the Netherlands based on measurements from 1989 to 1995

a bird towards its destination per unit time. Wind profits were calculated from the wind vector (consisting of wind speed, S_w, and wind direction, α_w) and the mean bird vector (bird speed, S_b, and bird direction, α_b): wind profit was defined as the length of the bird vector minus the length of the displacement vector (Erni *et al.* 2002):

$$W_p = S_b - \sqrt{S_b^2 + S_w^2 - 2 \cdot S_b \cdot S_w \cdot \cos(\alpha_w - \alpha_b)} \qquad \text{eqn 4}$$

The mean flight direction, α_b, and speed, S_b, were set to 223° and 12 m s^{-1}, respectively. The mean flight direction represents the mean of two predominant migratory directions over the Netherlands mainland based on expert knowledge and radar measurements. Using sounding measurements of wind, mean wind profit, W_p (m s^{-1}), was calculated as the mean of all wind profits up to pressure level 850 hPa. Note that W_p can attain negative values in headwind conditions, with strong lateral wind or with very strong tailwinds. The categorical variable for negative wind profits, W_u, was set to 1 for wind profits lower than a negative threshold; this threshold was optimized manually in steps of 0·1 m s^{-1}. Furthermore, a variable representing the potential accumulation of migrants because of previous nights with unfavourable weather was adopted from Erni *et al.* (2002), so that:

$$Acc_t = \frac{1}{3} \cdot Acc_{t-1} + \frac{2}{3} \cdot U_{t-1} \qquad \text{eqn 5}$$

where *Acc* is the accumulation effect, *U* is the variable for unfavourable weather, *t* relates to the current night and $t - 1$ relates to the previous night. U_{t-1} was set to 1 if unfavourable weather occurred during the previous night, thus increasing the potential accumulation of migrants (Acc_t). Independent accumulation variables were derived for the effect of rain (Acc_r) and the effect of unfavourable wind profit (Acc_w), based on $R_p > 0$ or $W_p < W_u$, respectively. All variables were normalized using the *z*-scores to enable comparison of the relative effects of the variables within each tested model.

MODEL CALIBRATION AND TESTING

The mean and variance of migration intensities (Table 1) did not differ significantly between years (ANOVA, $P < 0{\cdot}05$). Therefore we treated the data set as homogeneous. Multivariate log-linear regression models were identified and tested by applying cross-validation to different splits of the 3-year data set, following the guidelines for model building as put forward by Hastie, Tibshirani & Friedman (2001), Beck (2002) and Burnham & Anderson (2002). The steps in the cross-validation cycle are outlined as follows. First, the data set was split into a part for model calibration and a part for model validation. Next, all significant models up to a maximum of four explanatory variables were fitted using the calibration data. The four variables were combinations of the 28 available variables listed in Table 2 and 28 randomly permuted variables. Finally, the performance of each significant model was evaluated for the validation data, using the RMSE. The most adequate model was found by selecting those models with the lowest RMSE. In total, six data splits were made using the data in 2001, 2002, 2003, 2001 + 2002, 2001 + 2003 and 2002 + 2003 for calibration, while using the remaining data for validation. The models fitted on these different data splits were named M_1, M_2, M_{12} (etc.) when using the data for 2001, 2002, 2001 + 2002 (etc.) for calibration. By including a randomly permutated dummy variable for each explanatory variable, the above framework automatically checks for random significance. Permuted variables should never be selected as explanatory by this procedure. If any permuted variables are selected, the data set or model framework is invalidated (Good 2000). According to Stone (1977), Shao (1997) and Hastie, Tibshirani & Friedman (2001), leave-one-out cross-validation is equivalent to using an Akaike's Information Criterion (AIC) criterion in the calibration phase only, and leave-*n*-out cross-validation is equivalent to a Bayesian information criterion (BIC) criterion. By following the leave-*n*-out cross-validation procedure in our study as described above, model over-parameterization is prevented and the results are equivalent to the use of the BIC in model calibration.

To investigate the importance of calibration sample size on robustness of the input variables, we used a different approach for splitting data into calibration and validation sets. In this case the models were fitted on randomly selected calibration sets of different sizes: 30, 60, 90 and 120 data points. For each calibration set size, the data set was randomly split 2000 times into calibration and validation sets. Validation sets always contained 80 data points. Regression models were calculated and validated with each randomized set (8000 in total).

Results

MIGRATION INTENSITY PREDICTION BY EXISTING MODELS

Predictions of nocturnal migration intensity by existing models were better than the benchmarks (Table 3) after correction for the different methods of measuring migration intensity. The RMSE for each model equalled about 25% of the range of the measurements. The range of predicted values was much lower than that of the measurements; for example, only 32% of the measured range was predicted by M_G (Fig. 2).

MIGRATION INTENSITY PREDICTION BY NEW MODELS

The new regression models contained two to four significant input variables (Table 4): baseline migration

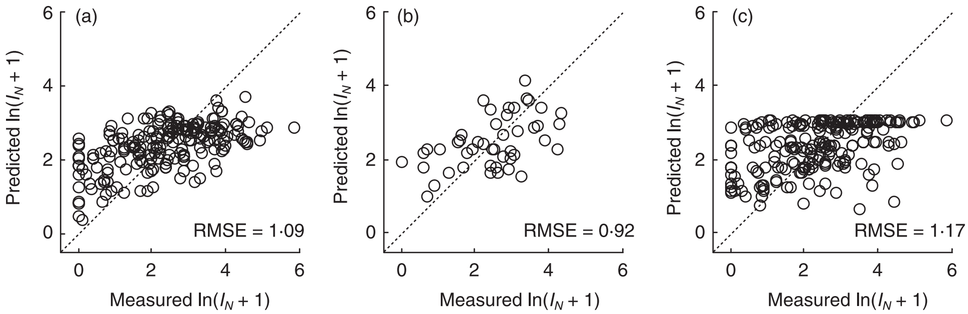

Fig. 2. Prediction performance of existing models: log-transformed predicted vs. measured migration intensities (I_N) for all years, by the models M_S from (a) Sweden, M_D from (b) Denmark, and M_G from (c) Germany, with original coefficients and added correction factors.

Table 3. Prediction performance of existing models (M_S, M_D and M_G) and benchmarks on Dutch data pooled for all years. RMSE, predicted (R_{pred}) and measured (R_{meas}) range of values (95% of logarithmic values), are given for model comparison. B_B and B_P represent the benchmarks based on the seasonal baseline migration intensity and migration intensity the previous night, respectively. The R_{meas} for M_D includes measurements from September only

	RMSE	R_{pred}	R_{meas}
M_S	1·09	0·98–3·31	0·00–4·62
M_D	0·92	1·19–3·83	0·42–4·30
M_G	1·17	1·05–3·07	0·00–4·62
B_B	1·23	1·00–3·02	0·00–4·62
B_P	1·41	0·00–4·62	0·00–4·62

Table 5. Prediction performance of new models: validation RMSE, mean difference between predictions and measurements (△), and predicted range of values (R_{pred})

	Validation period	RMSE	Δ	R_{pred}
M_1	2002	0·96	0·36	1·05–4·35
	2003	0·88	0·01	0·87–4·40
	2002 + 2003	0·92	0·19	0·98–4·40
M_2	2001	0·96	–0·36	–0·42–3·38
	2003	0·88	–0·42	–0·13–3·95
	2001 + 2003	0·91	–0·39	–0·17–3·92
M_3	2001	0·92	0·04	–0·09–3·96
	2002	0·92	0·44	0·20–4·75
	2001 + 2002	0·92	0·28	–0·09–4·58
M_{12}	2003	0·81	–0·24	0·30–4·16
M_{13}	2002	0·86	0·40	0·31–4·62
M_{23}	2001	0·89	–0·21	–0·38–3·67

intensity I_b, wind profit W_p, 24-h trend in barometric pressure ΔP and nocturnal proportion of hours with rain R_p. I_b and W_p were the most important input variables in all the new models; together, they explained at least 40% of the measured variance. ΔP was a significant variable for all subsets of years except for 2001. R_p was significant in all models based on 2 and 3 years and in M_2 but its contribution to the total explained variance was small compared with the contributions of other input variables. The coefficients of the input variables were of similar orders of magnitude between models. Comparison of the normalized coefficients showed that W_p had the largest effect on migration intensity: the contribution of W_p to the predicted values was on average 2·4 times as high as the contribution of ΔP, 3·5 times as high as R_p and 1·6 times as high as I_b.

Prediction performances of all new models were much better than the benchmarks and existing models (Table 5 and Fig. 3). Validation RMSEs were generally lower and predicted ranges were much closer to measured ranges (80–84% of the measured range; Fig. 4). Validation RMSEs of models calibrated with 2 years (RMSE 0·81–0·89) were lower, and differences between models were smaller, than for models based on single years (RMSE 0·88–0·96). Following validation of each model with independent years, it was clear that, on average, M_2 predictions, were too low for 2001 and 2003.

Table 4. New models: calibration periods, coefficients of normalized input variables and model fit to calibration data by adjusted R_2 and RMSE

		Input variables				Model fit	
	Calibration period	I_b	W_p	ΔP	R_p	R_2	RMSE
M_1	2001	0·593	0·741	–	–	0·45	0·90
M_2	2002	0·474	0·704	0·323	–0·274	0·64	0·75
M_3	2003	0·395	0·900	0·360	–	0·66	0·75
M_{12}	2001 + 2002	0·494	0·696	0·314	–0·218	0·57	0·82
M_{13}	2001 + 2003	0·485	0·823	0·307	–0·194	0·61	0·80
M_{23}	2002 + 2003	0·468	0·773	0·322	–0·233	0·64	0·78
M_{123}	2001 + 2002 + 2003	0·484	0·762	0·314	–0·216	0·62	0·81

EFFECT OF SAMPLE SIZE ON MODEL ROBUSTNESS

Model structure and validation performances were more robust for larger calibration sets (Table 6). Calibrations

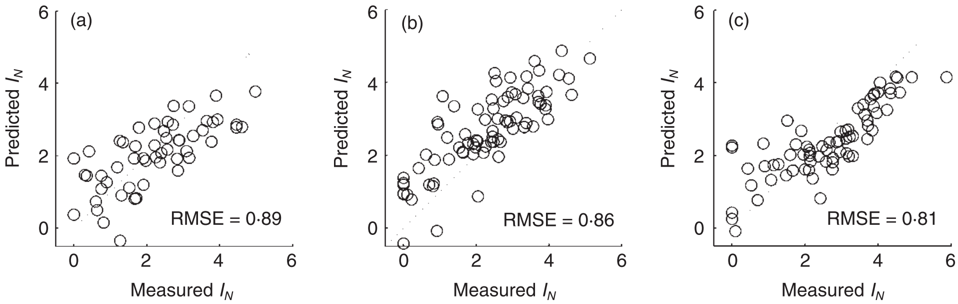

Fig. 3. Prediction performance of new models: log-transformed predicted vs. measured migration intensities (I_N). Predicted intensities for each year result from models calibrated with the two independent years: intensities for 2001 are predicted by M_{23} (a), 2002 by M_{13} (b) and 2003 by M_{12} (c).

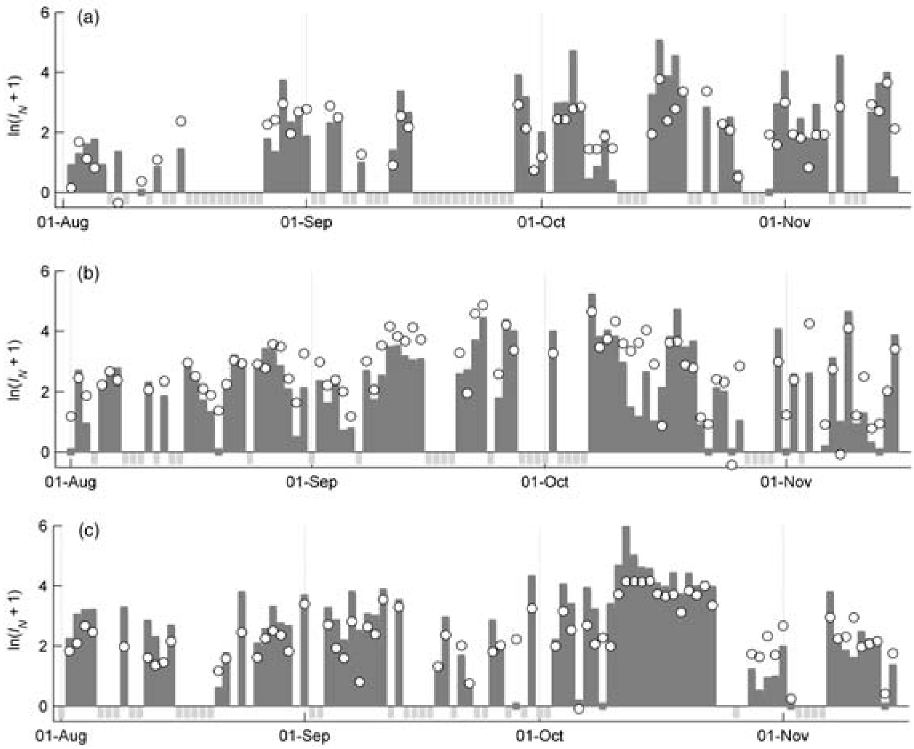

Fig. 4. Measured (bars) and predicted (circles) log-transformed migration intensities (I_N) over time. Predicted intensities for each year result from models calibrated with the two independent years: intensities for 2001 are predicted by M_{23} (a), for 2002 by M_{13} (b) and for 2003 by M_{12} (c). All migration intensities are given as natural logarithms + 1. Grey areas indicate dates with missing bird data.

with smaller data sets yielded fewer significant variables and a larger variation in variable significance. Subsamples of the data set, as small as 30 nights only, occasionally even yielded models without I_b or W_p, significant variables in all models with larger sample sizes. RMSEs of the validations were lower and more consistent with increasing calibration sample size: the mean RMSE decreased from 1·05 ± 0·19 to 0·89 ± 0·07.

OPERATIONAL MODEL

For operational use in the RNLAF, we calibrated our model with variables that were significant for all three models M_{12}, M_{13} and M_{23}, using the data of all 3 years (204 data points) in order to include as much information as possible. Normalized coefficients were roughly equal to the means of coefficients in M_{12}, M_{13} and M_{23}. As the aim of this model was to predict actual migration intensity, the model was converted to non-normalized input variables:

$$M_{123} = 1{\cdot}24 + 0{\cdot}84 I_b + 0.154 W_p + 0{\cdot}043 \Delta P - 0{\cdot}79 R_p \qquad \text{eqn 6}$$

This yielded a modelled range of 0·05–4·31, larger than that of any other model developed in this study, and a (calibration) RMSE of only 0·81.

Table 6. Effect of calibration sample size on model structure and performance. Fraction of occurrence of input variables, median number of significant input variables and the means and standard deviations of validation RMSE for sample sizes of 30, 60, 90 and 120 data points. For each sample size, regression models were calibrated with 2000 randomized samples. Validation data sets always consisted of 80 randomly selected data points. Only variables that occurred in more than 10% of the models are shown

	Calibration sample size			
	$n = 30$	$n = 60$	$n = 90$	$n = 120$
W_p	0·88	1	1	1
I_b	0·59	0·94	1	1
ΔP	0·18	0·44	0·63	0·80
R_p	0·11	0·19	0·30	0·47
R_t	0·13	0·25	0·25	0·21
P	0·14	0·19	0·24	0·19
ΔT	0·10	0·18	0·19	0·17
Acc_w	0·09	0·13	0·12	0·13
Number of variables	2	3	4	4
RMSE	1·05 ± 0·19	0·93 ± 0·083	0·90 ± 0·072	0·89 ± 0·069

Discussion

Existing regression models from other locations on the south-westerly migratory flyway predict migration intensities over the north of the Netherlands better than predictions based on migration intensity of the previous night or on the seasonal trend alone. However, new models are more robust and out-perform existing models. This study emphasizes the influence of the variation and size of calibration data sets for the predictive power and robustness of regression models.

Model performance was evaluated by comparing predictions to measured migration intensities. Therefore model results and data quality need to be addressed. In specific conditions, such as high rain clutter, anomalous propagation and very high bird densities, migration intensities were corrected, partly on the basis of expert judgement. Consequently, the quality of the migration intensity measurements is higher in the middle of the range than at the extremes. This is one of the possible causes for the larger discrepancies between observations and predictions at very high and very low migration intensities (Figs 3 and 4).

PREDICTION PERFORMANCE OF EXISTING MODELS

Although predictions by the existing models are better than the benchmarks, they are poor compared with the newly developed models. All existing models use different variables for the effect of wind. In the model for Germany (M_G), the effect of wind is represented by a categorical variable, based on a threshold for wind profit lower than -7 m s^{-1}. In M_G, wind can only reduce migration intensity, whereas in the other existing models wind can also have a positive effect on migration intensity. Moreover, the threshold of -7 m s^{-1} leads to only 12% of nights with unfavourable winds in our data set, whereas the German data set had unfavourable winds in 34% of all nights. The model for Denmark (M_D) only uses the lateral wind drift, whereas the model for Sweden (M_S) uses two variables, wind speed and the tail wind component. A second cause for the differences in performance is the use of variables that are not significant for the Dutch data set. For example, P in M_S and T in M_D are not significant after recalibration of the models with the Dutch data sets. Each existing model appears to be site and perhaps even time specific, therefore over-fitting of the models is probably a major cause of the differences in model performance on new data. This is particularly likely for M_S and M_G, as these two models are based on single years of measurement and are not validated with independent years.

PREDICTION PERFORMANCE OF NEW MODELS

As found in other multivariate ecological studies (Whittingham *et al.* 2006), it is clear that a single year of data is insufficient to identify the best and most robust model. The best newly developed models are based on subsets of 2 years. Model structures and prediction performance of newly developed models based on 2 years are very similar, but one specific year may show lower or higher migration intensities than other years. During validation of 2001, the independent model M_{23} performed even better than the calibrated model M_1, indicating that 2001 alone contained too little variation to develop a valuable regression model. On the other hand, M_3 fits the calibration data better than M_{23}, yet performance of M_3 in independent years is relatively poor. These examples nicely demonstrate the trade-off between model fit and general applicability of the models that is inherent to model selection.

The increased variation in model structure and performance with smaller data subsets shows that the calibration sample size has a pronounced effect on robustness of the models. The amount of variation within the calibration set also determines model robustness. For the number of variables and the variation contained in the data for this study, the model performance strongly improves with increasing sample size up to 90 data points and variation in model structure decreases with increasing sample size throughout the test.

INFLUENCE OF WEATHER

In this study, wind profit had the strongest influence on migration intensity. The large impact of wind on migration intensity is commonly found in the literature (Alerstam 1979; Gauthreaux & Belser 1998; Liechti & Bruderer 1998; Åkesson & Hedenström 2000; Erni *et al.* 2002). The effect of barometric pressure trends found in this study corresponds with the expected importance of synoptic weather for migration intensity (Richardson 1990). However, in contrast with most theories and studies on the relationships between rain and migration intensity (Richardson 1990; Erni *et al.*

2002; Schaub, Liechti & Jenni 2004), the effect of rain in our models is limited and its exact importance is hard to estimate. This may be related to the properties of the rain data and the bird data, the quality of which decreases in heavy rain. For example, the strength, amount and duration of rain can vary greatly over time and space, from short light showers to large-scale convective systems. These factors may influence migration differently, yet very few of these aspects are captured in hourly measurements of rain or their derivatives. Inter-annual variation in rainfall may also be influential. For example, in 2002 rain events were much more evenly spread over the season than in 2001 and 2003. This coincides with a relatively large effect of rain in M_2 compared with M_1 and M_3 (Table 4).

Several studies have shown that migration intensity is not only determined by the direct influence of weather but also by the effect of weather on the number of 'available' migrants (Rappole & Warner 1976; Fransson 1998; Erni *et al.* 2002; Pérez-Tris & Tellería 2002). Weather may have contrasting effects on the accumulation of migrants. For example, prolonged periods with unfavourable weather may cause birds to accumulate, having a positive effect on migration intensity. Conversely, extended periods of good weather may lead to decreasing numbers of migrants because of the lack of birds following peak migration. Similarly, rare but structural phenomena, like arrival from Norway over the North Sea (Buurma 1987), cannot be properly dealt with in linear regression models.

The number of weather variables tested and the similarities between new models indicates that further improvement of the models cannot be achieved with additional local weather variables alone. Model performance may be improved by including better variables accounting for the accumulation of birds, particularly for very high and very low migration intensities.

MODEL APPLICATION

As part of the Netherlands Bird Avoidance Model (http://ecogrid.sara.nl/bambas, accessed 24 April 2007), a new Internet-based decision support system to improve flight safety, the RNLAF has implemented the final predictive model for operational use (M_{123}; equation 6). During the migration season, the model is run 3 days ahead with forecasted meteorological variables for the north of the Netherlands (53°15′N, 6°59′E), available from NOAA's Air Resource Laboratory (http://www.arl.noaa.gov/ready/cmet.html, accessed 24 April 2007). The results are used to issue migration intensity forecasts 3 days in advance, which are updated daily, and may lead to rescheduling of flights if necessary.

For the RNLAF, the current mismatch between predictions and measurements of very low migration intensities is more problematic than for very high intensities. Although underestimated, the highest migration predictions are still well above any current standardized flight restriction levels, and are therefore avoided. The overestimation of low intensities requires further attention, as this may lead to false flight restrictions, which can be very costly.

In addition to flight safety, there are many other sectors that can benefit from migration intensity predictions. For example, migration predictions are of importance for the study and reduction of the impact of wind turbines on bird migration (Drewitt & Langston 2006; Hüppop *et al.* 2006), outbreaks and the spread of zoonotic pathogens such as West Nile virus and avian influenza (Reed *et al.* 2003; FAO 2005), and local efforts to reduce the obstruction to migration posed by tall buildings lit at night (e.g. Lights out NY, http://www.nycaudubon.org/NYCASBirdWatch/safeflightupdates, accessed 24 April 2007). The study of migration stopover ecology and particularly the arrival at and departure from stopover sites can greatly benefit from improved understanding of the relationship between weather and migratory activity (Schaub, Liechti & Jenni 2004; Bulyuk & Tsvey 2006). Furthermore, improved knowledge of environmental factors on bird migration is necessary to understand the evolutionary forces behind migration (Pulido, Berthold & van Noordwijk 1996) and the impact of climate change (Sparks 1999).

Extrapolation of regression models to new locations should be discouraged. Clearly, many existing models are only valid locally. Future research will be directed towards radar measurements from adjacent countries to get more insight into the spatial distribution of migrants, the relationships between migration intensities at different locations and the development of spatially explicit dynamic models of migration intensity. Cross-validation should be used to avoid overparameterization in developing models. To make effective use of cross-validation in model development, data sets should represent inter- and intraseasonal variability of the study site well.

Currently, the regression model developed in this study is in operational use in the Royal Netherlands Air Force. Every morning, new predictions of bird migration intensities for 3 days ahead are calculated from the weather conditions forecasted by NOAA. The model thus improves the reliability of the information needed for warnings for high bird migration intensities and, if necessary, rescheduling military flight training in the Netherlands.

Acknowledgements

We would like to thank L. Buurma of the Royal Netherlands Air Force for initiating this research, and H. Van Gasteren, also of RNLAF, for pre-processing all the radar data. Meteorological data were obtained from KNMI and NOAA. Furthermore we would like to thank B. Erni and the Vogelwarte Sempach for their ideas and the use of their original data for comparison, as well as P. Stephens, and two anonymous referees for constructive comments on a previous version of this manuscript.

This work was carried out in the context of the Virtual Laboratory for e-Science project (www.vl-e.nl). This project is supported by a BSIK grant from the Dutch Ministry of Education, Culture and Science (OC and W) and is part of the ICT innovation program of the Ministry of Economic Affairs (EZ).

References

Able, K.P. (1973) The role of weather variables and flight direction in determining the magnitude of nocturnal bird migration. *Ecology*, **54**, 1031–1041.

Åkesson, S. & Hedenström, A. (2000) Wind selectivity of migratory flight departures in birds. *Behavioural Ecology and Sociobiology*, **47**, 140–144.

Alerstam, T. (1979) Wind as a selective agent in bird migration. *Ornis Scandinavica*, **10**, 76–93.

Alerstam, T. & Ulfstrand, S. (1972) Radar and field observations of diurnal bird migration in south Sweden, autumn 1971. *Ornis Scandinavia*, **3**, 99–139.

Allan, J.R. (2002) The costs of bird strikes and bird strike prevention. *Proceedings of the 3rd National Wildlife*, *Special Symposium* (ed. L. Clark), pp. 147–153. Research Center, Fort Collins, Colorado, USA.

Barrios, L. & Rodrígues, A. (2004) Behavioural and environmental correlates of soaring-bird mortality at on-shore wind turbines. *Journal of Applied Ecology*, **41**, 72–81.

Beck, B. (2002) Model evaluation and performance. *Encyclopedia of Environmetrics* (eds A. El-Shaarawi & W. Piegorsch), pp. 1275–1279. Wiley & Sons, Chicester, UK.

Berthold, P. (2001) *Bird Migration: A General Survey*, 2nd edn. Oxford University Press, Oxford, UK.

Blokpoel, H. (1969) An attempt to forecast the intensity of nocturnal fall migration. *Proceedings of the World Conference on Bird Hazards to Aircraft* (ed. M.S. Kuhring), pp. 285–298. National Research Council of Canada, Kingston, Canada.

Blokpoel, H. (1976) *Bird Hazard to Aircraft*. Books Canada Inc., Buffalo, NY.

Bruderer, B. (1997) The study of bird migration by radar. I. The technical basis. *Naturwissenschaften*, **84**, 1–8.

Bulyuk, V.N. & Tsvey, A. (2006) Timing of nocturnal autumn migratory departures in juvenile European robins (*Erithacus rubecula*) and endogenous and external factors. *Journal of Ornithology*, **147**, 298–309.

Burnham, K.P. & Anderson, D.R. (2002) *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. Springer-Verlag, New York, NY.

Buurma (1995) Long-range surveillance radar as indicator of bird numbers aloft. *Israel Journal of Zoology*, **41**, 221–236.

Buurma, L.S. (1987) Patterns of high bird migration over the North Sea area in October. *Limosa*, **60**, 63–74.

Dekker, A., Van Gasteren, H. & Shamoun-Baranes, J. (2003) Eurbase. Progress report and first impressions on bird species. *Proceedings of the 26th International Bird Strike Committee* (ed. L.S. Buurma), pp. 225–238. International Bird Strike Committee (IBSC), Warsaw, Poland.

Drewitt, A.L. & Langston, R.H.W. (2006) Assessing the impacts of wind farms on birds. *Ibis*, **148**, 29–42.

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

FAO (2005) *Potential Risk of Highly Pathogenic Avian Influenza (HPAI) Spreading through Wild Water Bird Migration*. FAOAID Enews Special Issue. Avian Influenza Technical Task Force, FAO, Rome and Bangkok. http://www.fao.org/ag/AGAInfo/subjects/documents/ai/AVIbull033.pdf [last accessed 31 October 2006].

Fransson, T. (1998) Patterns of migratory fuelling in whitethroats *Sylvia communis* in relation to departure. *Journal of Avian Biology*, **29**, 569–573.

Gauthreaux, S.A. & Belser, C.G. (1998) Displays of bird movements on the WSR-88D: patterns and quantification. *Weather and Forecasting*, **13**, 453–464.

Gauthreaux, S.A. & Belser, C.G. (2003) Radar ornithology and biological conservation. *Auk*, **120**, 266–277.

Geil, S., Rabøl, J. & Noer, H. (1974) *Forecast Models for Bird Migration in Denmark*. Report from the Danish Defence Research Establishment, University of Copenhagen and Royal Danish Air Force. Danish Defence Research Establishment.

Good, P. (2000) *A Practical Guide to Resampling Methods for Testing Hypotheses*. Springer, New York, NY.

Hastie, T., Tibshirani, R. & Friedman, J. (2001) *The Elements of Statistical Learning; Data Mining, Interference and Prediction*. Springer, New York, NY.

Hüppop, O., Dierschke, J., Exo, K.M., Fredrich, E. & Hill, R. (2006) Bird migration studies and potential collision risk with offshore wind turbines. *Ibis*, **148**, 90–109.

Lack, D. (1963) Migration across the southern North Sea studied by radar. IV. Autumn. *Ibis*, **104**, 1–54.

Liechti, F. & Bruderer, B. (1998) The relevance of wind for optimal migration theory. *Journal of Avian Biology*, **29**, 561–568.

Nisbet, I.C.T. & Drury, W.H. Jr (1968) Short-term effects of weather on bird migration: a field study using multivariate statistics. *Animal Behaviour*, **16**, 496–530.

Pérez-Tris, J. & Tellería, J.L. (2002) Regional variation in seasonality affects migratory behaviour and life-history traits of two Mediterranean passerines. *Acta Oecologica*, **23**, 13–21.

Pulido, F., Berthold, P. & van Noordwijk, A. (1996) Frequency of migrants and migratory activity are genetically correlated in a bird population: evolutionary implications. *Proceedings of the National Academy of Sciences USA*, **93**, 14642–14647.

Rappole, J.H. & Warner, D.W. (1976) Relationships between behavior, physiology and weather in avian transients at a migration stopover site. *Oecologia*, **26**, 193–212.

Reed, K.D., Meece, J.K., Henkel, J.S. & Shukla, S.K. (2003) Birds, migration and emerging zoonoses: West Nile virus, Lyme disease, influenza A and enteropathogens. *Clinical Medicine and Research*, **1**, 5–12.

Richardson, W.J. (1990) Timing of migration in relation to weather: an updated review. *Bird Migration: Physiology and Ecophysiology* (ed. E. Gwinner), pp. 78–101. Springer Verlag, Berlin, Germany.

Schaub, M., Liechti, F. & Jenni, L. (2004) Departure of migrating European robins, *Erithacus rubecula*, from a stopover site in relation to wind and rain. *Animal Behaviour*, **67**, 229–237.

Shao, J. (1997) An asymptotic theory for linear model selection. *Statistica Sinica*, **7**, 221–264.

Sodhi, N.S. (2002) Competition in the air: birds versus aircraft. *Auk*, **119**, 587–595.

Sparks, T.H. (1999) Phenology and the changing pattern of bird migration in Britain. *International Journal of Biometeorology*, **42**, 134–138.

Stone, M. (1977) Asymptotics for and against cross-validation. *Biometrika*, **64**, 29–35.

Sutter, E. (1957) Radar als Hilfsmittel der Vogelzugforschung. *Ornitologische Beobachtungen*, **54**, 70–96.

Whittingham, M.J., Stephens, P.A., Bradbury, R.B. & Freckleton, R.P. (2006) Why do we still use stepwise modelling in ecology and behaviour? *Journal of Animal Ecology*, **75**, 1182–1189.

Zar, J.H. (1984) *Biostatistical Analysis*, 2nd edn. Prentice Hall, Princeton, NJ.

Zehnder, S., Åkesson, S., Liechti, F. & Bruderer, B. (2001) Nocturnal autumn bird migration at Falsterbo, South Sweden. *Journal of Avian Biology*, **32**, 239–248.

Received 19 July 2006; final copy received 7 March 2007
Editor: Philip Stephens