

MULTIVARIATE APPROACHES TO FORECASTING DAY-TO-DAY VARIATIONS

IN THE AMOUNT OF BIRD MIGRATION

W. JOHN RICHARDSON *

Richardson, W.J. 1974. Multivariate approaches to forecasting day-to-day variations in the amount of bird migration. p. 309-329 *In*: S.A. Gauthreaux Jr. (ed.), A conference on the biological aspects of the bird/aircraft collision problem, Clemson, Feb. 1974. Clemson Univ., Clemson, SC. 535 p.

ABSTRACT

Day-to-day variations in the amount of bird migration are closely related to weather conditions when considered from synoptic, univariate or multivariate viewpoints. Multivariate techniques are best suited for forecasting, but their limitations require evaluation. This paper describes applicable techniques and suggests approaches to overcoming the limitations.

Categorically- and continuously-scaled measurements of migration volume can be analyzed and predicted by multiple discriminant and multiple regression procedures respectively. Assumptions of linearity, normality and homoscedasticity in the regression situation can be checked by analysis of residuals. Violations are often found, but can usually be overcome by transformations. Violations of the additivity assumption also occur, but they rarely reduce forecasting accuracy significantly.

Procedures for dealing with non-stationary variables and predictor variables that are non-linearly related to migration volume are described. Models incorporating alternative but related sets of predictor variables are shown to be able to forecast migration volume with almost equal accuracy. Factor analysis can be very useful in reducing the number of predictors to a more manageable number, in avoiding overfitting, and in identifying causal relationships.

INTRODUCTION

The numbers of birds aloft are concentrated in space and time, and hence so is the bird hazard to aircraft (Richardson 1970). In some of these especially hazardous situations, it is both practical and desirable to modify aircraft operations on an hour-to-hour basis in order to avoid concentrations and to reduce the probability of collisions (Gunn and Solman 1968). If such a program is to be implemented, one must be able to recognize concentrations of birds in real time, or even better, to predict the locations and/or extent of concentration some hours in advance.

Both migrating and non-migrating birds may be concentrated in either space or time. This paper discusses numerical methods for forecasting the numbers of migrating birds aloft at various times. However the same techniques can be applied in developing procedures for forecasting other phenomena.

VARIABLES RELATED TO MIGRATION VOLUME

The amount of migration varies from day to day within a migration

*Present address: LGL Limited, environmental research associates
44 Eglinton Ave. W, Toronto, Ontario, Canada M4R 1A1

season over at least two and sometimes three orders of magnitude (e.g., Nisbet and Drury 1968; Haugh 1972; Able 1973). These variations are strongly correlated with variations in weather conditions. In general, birds tend to migrate when the winds are following relative to their 'preferred' directions of flight (Richardson 1971, 1972a). Thus in spring more birds are aloft when the winds are southerly than when they are northerly, while in autumn the converse is true. There are many interrelationships amongst weather variables, and hence it is not surprising that correlations with the volume of migration have been found for virtually all measurable weather variables (for reviews, see Lack 1960a; Richardson and Gunn 1971). Because of the strong relationships of weather conditions to the volume of bird migration, weather variables must be a major component in any forecasting system.

Various other variables are known or suspected to affect the volume of migration, at least on occasion. These include the date within the migration season, moon phase, number of days since the last major flight, and tidal patterns.

FORECASTING TECHNIQUES

Three types of approaches have been used to elucidate the relationships between migration volume and weather:

- (i) The *synoptic approach* involves relating the volume of migration to large-scale features of the weather, such as pressure systems, fronts and broad wind patterns.
- (ii) *Bivariate analysis* involves relating volume to local numerically measurable variables one at a time.
- (iii) *Multivariate analysis* examines the relationships of migration volume to many variables considered simultaneously.

Synoptic and univariate analyses are useful for describing the situations when birds tend to migrate and for understanding the adaptive significance of their migration timing systems. These techniques can be used to develop workable guidelines for predicting migration volume a few hours in advance (Blokpoel 1973, in prep.). However it is generally agreed that ultimately the best way of forecasting will be by applying multivariate models. These can be developed from multivariate analysis of daily observations of weather and the volume of migration. Multivariate procedures have three main advantages:

- (i) They permit simultaneous consideration of all the potentially useful predictor variables and their interrelationships.
- (ii) They provide more or less objective procedures for identifying the most reliable forecasting procedure.
- (iii) Multivariate forecasting models are readily applied because they require simple substitution of the values of predictor variables into one or more equations. Other forecasting procedures depend more upon the experience and judgement of the forecaster.

While there have been many synoptic and bivariate studies of migration volume, few multivariate studies have been done. Gruys-Casimir (1965) and Lack (1960b 1963a,b) applied simple forms of multiple regression to visual and radar data in Europe. Nisbet and Drury (1968) applied step-wise multiple regression to radar observations of spring migration over eastern Massachusetts, and Able (1973) used multiple correlation and multiple

discriminant analysis in a study of autumn migration over the southeastern United States. I have applied these and other techniques to radar data from Canada and Puerto Rico, and multiple regression is currently being applied to radar data from Europe.

Unfortunately, the numerous limitations of multivariate techniques are not widely understood or even recognized. Nisbet and Drury (1968) and Richardson and Haight (1970) identified several of these limitations. However their severity and effects on forecasting accuracy have not been evaluated previously.

The purposes of this paper are to describe multivariate techniques which I or others have found to be useful, to evaluate the limitations of these techniques, and to suggest approaches for overcoming those limitations. The data used to test and develop the procedures were derived from surveillance radar studies of migration in Nova Scotia and New Brunswick, Canada, and in Puerto Rico (Richardson 1971, 1972a, 1974, in prep.). The volume of migration was recorded each day and night on a 0 to 8 ordinal scale (Richardson 1972b).

MULTIVARIATE TECHNIQUES

Four standard multivariate techniques can profitably be applied in developing forecasting procedures for migration volume. *Multiple regression analysis* is the most widely used technique in this field. It is applicable when the dependent variable, here migration volume, is measured on a continuous scale. *Multiple discriminant analysis* is applicable when the dependent variable is categorical in nature (e.g., migration is either 'present' or 'absent', or 'below normal', 'normal' or 'above normal'). *Factor analysis*, which has not been used in any previous study of bird migration, can be used to identify the few basic environmental factors which are reflected in many measurable but inter-correlated variables. *Canonical correlation* is potentially useful as a means for considering the relationships of various predictor variables to several spatially or temporally adjacent measurements of migration volume; however it has not yet been applied to this problem, and is not discussed here. Multivariate techniques have a heuristic capability in addition to their roles in developing usable forecasting procedures and in testing hypotheses about relationships between predictors and migration volume. When many variables are measured for each of many cases, one frequently amasses so many data that it becomes impossible to recognize previously unsuspected relationships. This is especially true when the predictors are intercorrelated. Multivariate procedures aid in identifying these relationships, and thereby lead one to new hypotheses than can be tested with additional data.

The theory and implementation of multivariate procedures is discussed by Rao (1952), Harman (1967), Morrison (1967), Cooley and Lohnes (1971), and Sneath and Sokal (1973). Computer programs are available for each procedure (Nie et al. 1970; Cooley and Lohnes 1971; Rohlf et al. 1971; Dixon 1973).

MULTIPLE REGRESSION ANALYSIS

A multiple regression analysis fits a linear additive equation to the available data. This equation may then be used to forecast the

dependent variable, here migration volume, if the values of the various predictor variables are known or forecastable. The equation is of the form:

$$\# \text{ migrants} = C_0 + C_1 X_1 + C_2 X_2 + \cdots + C_k X_k$$

where there are k predictor variables, X_1, X_2, \dots, X_k . The combination of constants chosen by the analysis procedure is the one which maximizes forecasting accuracy for the cases of data upon which the analysis was based. The reliability of forecasts based on the equation is measured by the multiple correlation coefficient, R , or its square, the percentage of the day-to-day variance in volume accounted for by observed relationships between migration volume and the predictors. The relative contribution of each predictor variable in accounting for that variance is also found.

Multiple correlation, which has been used by Able (1973), is a form of analysis similar to multiple regression, with similar assumptions and limitations. It is not applicable as a forecasting technique.

Any multiple regression analysis assumes that the dependent variable is linearly and additively related to the predictors. Furthermore, the forecasting errors are assumed to be independent and normally distributed with constant variance across the range of values of each of the variables in the analysis. When these assumptions are inaccurate, the regression constants, multiple correlation coefficient, and forecasting capability of the equation are all unreliable. Methods for recognizing and overcoming violations of the assumptions are discussed below.

In most multivariate analyses there are strong intercorrelations amongst many of the predictor variables as well as correlations between predictors and the dependent variable. Different predictors are often measuring or being affected by the same underlying factor. For example, high volumes of northward migration are usually correlated with southerly winds, increasing temperature, humidity and cloudiness, and decreasing pressure, ceiling and visibility. These weather characteristics tend to occur together, all being associated with a high pressure area moving away to the east, a low approaching from the west, or both. Statistical procedures, multivariate or otherwise, cannot reliably distinguish the causal relationships from the 'spurious' (*sensu* Simon 1954) ones. However, the existence of intercorrelations amongst the predictors does not vitiate the usefulness of multivariate techniques, particularly for forecasting, provided that one recognizes that the models describe relationships rather than directly indicate causality (cf. Able 1973:1032).

Another consequence of the correlations amongst predictors is that inclusion of a few of them in the regression equation may give forecasts that are about as accurate as could be achieved by inclusion of all of them. 'Stepwise' regression techniques are used to find the smallest set of predictors that gives a forecasting ability similar to that achievable with all of the predictors. The procedure is to add variables to the equation one at a time until no other variable would, if included, produce a significantly improved forecasting capability (Draper and Smith 1966). At any one step, the variable added to the equation is the one which makes the greatest improvement in forecasting ability. In the example presented in Table 1 an equation including all 19 predictors accounted for 67.3% of the night-to-night variance densities, while an

TABLE 1

MULTIPLE REGRESSION ANALYSES OF THE VOLUME OF NOCTURNAL NORTHWARD
MIGRATION OVER EAST-CENTRAL NEW BRUNSWICK IN SPRING¹

	A. Model Including All Predictors	B. Stepwise Model ²	C. Alternative Stepwise ³ Model
Constant	+22.7539	- 8.1457	+ 0.5514
Magnetic Disturbance	+ 0.2073 ns	---	---
Ceiling	+ 1.4123 *	+ 1.2427 ***	Held Out
Visibility	- 0.1820 ns	---	---
Precipitation	- 1.4848 ns	---	- 1.7323 *
B. Press. -1000 mb	+ 0.0248 ns	---	---
B. Press. trend	- 0.2688 ns	---	---
Temperature trend	+ 0.1633 *	+ 0.2501 ***	Held Out
Temp. rel. to norm.	+ 0.0803 ns	---	+ 0.2935 ***
Rel. Humidity	+ 0.0661 (*)	---	---
Rel. Hum. trend	- 0.0445 (*)	---	- 0.0406 *
Opacity	+ 0.0200 ns	---	---
Hrs. since < 10/10	+ 0.0050 ns	---	---
SE comp. of wind	+ 0.1194 ns	+ 0.1749 *	Held Out
SW comp. of wind	+ 0.1929 ns	+ 0.1720 **	Held Out
SE comp. of wind	- 0.1546 ns	---	---
SE comp. trend	- 0.0314 ns	---	+ 0.1084 ns
SW comp. trend	+ 0.0068 ns	---	+ 0.1325 *
Day of year	- 0.3700 ns	+ 0.1688 ns	+ 0.1330 ns
(Day of year) ²	+ 0.0015 ns	- 0.0005 ns	- 0.0005 ns
Mult. Correl. Coeff.	0.820 ***	0.790 ***	0.713 ***
% Variance Explained	67.3	62.4	50.8
SE of estimated (#) ^{1.5}	3.42	3.36	3.86

¹ Based on radar observations of migrants moving NW, N, NE, and E (predominantly NE) one hour after sunset on 86 nights between 20 April and 5 June in 1970-71 (Richardson 1971). Volume on a 0-8 scale was raised to the 1.5th power prior to analysis to satisfy the normality and homoscedasticity assumptions. The constant and partial correlation to migration volume for each variable are listed. Scaling procedures for the predictors are described in the Appendix. Two-sided significance levels are given as follows: ns $P > .1$; (*) $.1 \geq P > .05$; * $.05 \geq P > .01$; ** $.01 \geq P > .001$; *** $P \leq .001$

² First date and date squared were considered as predictors. Thereafter variables were added to the model one at a time until none of the remaining excluded ones would, if included, significantly reduce the forecasting errors ($P \leq .1$). Once included, variables were retained in the model unless their partial correlation dropped below the $P = .2$ significance level. --- means not included.

³ Stepwise analysis identical to first, except that the four weather predictors included in the first analysis were not allowed to enter the model.

equation including only 6 predictors accounted for 62.4%. Indeed, inclusion of only the 4 weather variables and not the statistically insignificant date terms accounted for 62.2% of the variance. Thus stepwise analysis allows one to develop a simple, easily applied equation which gives forecasts little less accurate than those obtainable by considering many other variables.

Stepwise analysis has the additional advantage of reducing 'overfitting'. When there are more than about 15% or 20% as many predictors as cases, the regression model is likely to be unreliable. The extreme case of overfitting is that in which there are as many predictors as cases. Then the model will account for all of the variance in the data from which it was developed, but may be much less successful when applied to another set of data. In less severe cases, the multiple correlation coefficient and percentage of variance explained are overestimated, and the weights of the predictors are less reliable than their standard errors would suggest (Lane 1971). Stepwise analysis helps to keep the number of included predictors to a minimum, but does not totally eliminate overfitting. For example, I randomly divided the 86 nights considered in Table 1 into three groups of 29, 29 and 28 cases, and performed the stepwise analysis on each group, treating the date terms in the same way as the other 17 predictors. The percentages of variance 'explained' were 70.8, 72.3 and 83.4%, as opposed to 62.2% when all 86 cases were assessed together. The apparent increase in forecasting capability is artifactual, and would have been even more unrealistic if all 19 variables had been included in the model. Factor analysis, discussed below, provides another way of avoiding overfitting.

When interpreting the results of stepwise analyses, it is important to realize that correlated predictors can substitute for one another in the model (e.g., Table 1C). When one member of a correlated pair of variables enters the model, the partial correlation of the other to migration volume drops. The degree to which inclusion of one variable affects the partial correlation for another is proportional to the strength of correlation between the two predictors (i.e., the extent to which they are measuring the same underlying environmental factor). For example, the density of northward migration over New Brunswick in spring is correlated with ceiling, visibility, opacity and precipitation, which are strongly intercorrelated. Once one of these, ceiling, has entered the regression model, the residual importance of the others in accounting for variability in migration volume becomes insignificant (Table 2). When ceiling is not allowed to enter the model, precipitation and perhaps humidity trend substitute (Table 1C). The effects of these variables on volume cannot be separated, and one certainly cannot conclude that only ceiling is important to the birds. One might instead conclude on the basis of the information in Tables 1 and 2, that after factors such as wind have been considered, migration is denser in fair weather than in cloudy or stormy weather. Interpretational difficulties of this sort can be reduced by using factor analysis to identify the underlying environmental factors, and are in any event of little importance in arriving at an operationally useful forecasting ability.

TABLE 2

STEPWISE REGRESSION ANALYSIS OF THE VOLUME OF NOCTURNAL NORTHWARD
MIGRATION OVER EAST-CENTRAL NEW BRUNSWICK IN SPRING¹

Variable	Direction of Correl.	Step				
		1	2	3	4	5
VARIABLES INCLUDED						
Date and (date) ²	+, -	ns	ns	ns	ns	ns
Temperature trend	+		***	***	***	***
Ceiling	+			***	***	***
SW comp. of wind	+				**	**
SE comp. of wind	+					*
VARIABLES EXCLUDED						
Magnetic disturbance	0	ns	ns	ns	ns	ns
Ceiling	+	***	***	---	---	---
Visibility	+	***	**	ns	ns	ns
Precipitation	-	***	**	ns	ns	ns
B. Press. -1000 mb	+	(*)	(*)	ns	ns	ns
B. Press. trend	-	*	ns	(*)	ns	ns
Temperature trend	+	***	---	---	---	---
Temp. rel. to norm.	+	***	*	*	ns	ns
Rel. Humidity	-	***	(*)	ns	ns	ns
Rel. Hum. trend	-	***	ns	ns	ns	ns
Opacity	-	***	*	ns	ns	ns
Hrs. since < 10/10	-	**	**	ns	ns	ns
SE comp. of wind	+	**	(*)	*	*	---
SW comp. of wind	+	***	*	**	---	---
SE comp. of wind	0	ns	ns	ns	*	ns
SE comp. trend	+	(*)	ns	ns	ns	ns
SW comp. trend	+	***	ns	*	ns	ns

¹ Each column presents the partial correlations of the predictors to the 1.5th power of migration volume at one step of the analysis summarized in Table 1B. The upper section shows the variables included in the model; the lower section shows those not included. The significance levels, coded as in Table 1, refer to the partial correlations given that only those variables listed in the upper section are included in the model.

MULTIPLE DISCRIMINANT ANALYSIS

Multiple discriminant analysis finds constants for linear additive functions of the predictors which best discriminate between the various categories of the dependent variable, here migration volume. The functions are of the form:

$$\text{Discriminant Score} = C_1X_1 + C_2X_2 + \cdots + C_kX_k$$

The set of constants chosen for the first function is the one which gives scores having the largest possible ratio of between categories variance to within categories variance. When more than two categories are present, additional functions uncorrelated with the first can be computed; each may give additional information about the relationships of the predictors to migration volume. As in the case of regression analysis, a measure of the overall forecasting success can be computed (the canonical correlation coefficient), and the relative importance of the various predictors in leading to an accurate forecast is assessed.

One can use the discriminant function(s) to forecast the probability that a given combination of weather conditions will lead to a specific category of migration density, and to compute the most probable density category with a given set of environmental conditions. Formally, these forecasting operations are performed using the classification procedures described by Cooley and Lohnes (1971). In practice, it is usually unnecessary to use the fully numerical approach. One can compute the discriminant score(s) with a given set of predictors, and then forecast the density category as the one which has mean discriminant scores most similar to the score(s) computed for the occasion in question.

In order to assess the forecasting capability of a given discriminant model, one can use it to classify the cases from which the model was derived. One can then compare the 'predicted' category for each case with what was actually observed. This is routinely done using formal numerical classification techniques by the widely available BMD07M program (Dixon 1973), or by the Cooley and Lohnes (1971) CLASIF program. It is preferable to use the 'prior probabilities = n_i/n ' option. Considering the model for nocturnal SW movement over Nova Scotia in autumn (Table 3), the proportions of the 229 evenings classified accurately, inaccurately by one category (e.g., low density when really none), and inaccurately by two categories ('none' when 'high', or vice versa), were 67.2%, 28.8% and 3.9% respectively. The model for NE reverse movement had a lower canonical coefficient (0.575 vs. 0.714) and a correspondingly lower forecasting accuracy: 66.4%, 18.6%, and 15.0% respectively. The canonical coefficient for SE shorebird movement was intermediate (0.653), and so was the forecasting accuracy: 62.9%, 32.8%, and 4.3%. In view of the large sample sizes (at least for SW and NE movement) and the diversity of years and dates over which the data were collected, these success vs. failure rates are probably quite reliable as indicators of the error rates that would occur if the models were applied to new cases.

Discriminant analysis assumes linear additive relationships of the predictors to the discriminant scores, and it assumes that the variances and covariances amongst the predictors are similar for each category of density. These assumptions are discussed below.

TABLE 3

MULTIPLE DISCRIMINANT ANALYSES OF THE VOLUME OF NOCTURNAL MIGRATION
OVER NOVA SCOTIA IN AUTUMN, BASED ON
STEPWISE ANALYSES OF 18 PREDICTORS¹

	S, SW and W Movement	ESE - SSE Shorebird Movement	N, NE and E Reverse Movement
Dates Considered	31 July - 20 Nov.	31 July - 10 Oct.	31 July - 20 Nov.
No. Nights of Data	229	116	226
Canon. Correl. Coeff.	0.714***	0.653***	0.575***
Magnetic disturbance	---	---	---
Ceiling	---	---	---
Visibility	+ ns	---	---
Precipitation	---	- **	- *
B. Press. -1000 mb	+ ***	---	---
B. Press. trend	+ *	---	---
Temperature trend	---	---	---
Temp. rel. to norm.	- *	---	---
Rel. Humidity	---	---	---
Rel. Hum. trend	- (*)	---	---
Opacity	---	---	---
Hrs. since < 10/10	---	---	+ **
SE comp. of wind	---	- ***	+ ***
SW comp. of wind	- ***	+ ***	+ ***
Side comp. of wind ²	- (*)	- ***	---
SE comp. trend	---	---	---
SW comp. trend	---	- **	---
Day of year ³	- *	---	---

¹ Based on radar observations of migrants over central and western Nova Scotia one hour after sunset in the autumns of 1965 and 1969-71 (Richardson 1972a). Densities were originally measured on a 0-8 scale but were converted for these analyses to 'none', 'low relative to normal', or 'high relative to normal' because of year-to-year and week-to-week non-stationarities and because of bimodality. The second discriminant function in each of the three analyses gave an insignificantly ($P > .1$) improved forecasting capability, and is not presented here. The direction of the relationship between density and each predictor included in the models is given, as is the significance of each variable in achieving accurate forecasts (i.e., the significance of the correlation between the variable and the discriminant score). The directions of the relationships as listed here are opposite to the signs on the variable weights, not presented, since the mean discriminant score for above-normal density was in each of the analyses less than that for below-normal density. Significance levels are coded as in Table 1; scaling procedures for the predictors are given in the Appendix.

... continued

- ² Side component is SW for shorebirds and SE for the others.
- ³ Date squared not needed as a predictor since density category relative to normal on the date in question was the dependent variable.
-

Stepwise multiple discriminant analysis (Dixon 1973) has the same advantages over standard discriminant analysis as does stepwise regression over standard regression. However one must again bear in mind the fact that any one included variable may reflect the relationships of a variety of correlated predictors to migration density.

In interpreting the weightings of each predictor variable in a discriminant model, one must consider the order of the mean discriminant scores for each category of density (cf. Able 1973). When the mean score for high density is lower than that for low density, as it was in each of the analyses in Table 3, then a negative weighting on a predictor implies a positive relationship between density and that predictor, while a positive weighting implies a negative relationship. Conversely, when the mean score for high density is greater than that for low density, then the opposite is true.

Table 3 shows that each of three main types of movement over Nova Scotia in autumn tended to be most dense with fair weather, as represented in the models by some combination of good visibility, high pressure, and little precipitation. However the density of N, NE, and E (predominantly NE) reverse movement was positively related to cloud as well as negatively related to precipitation. This is understandable because reverse movement was densest in the warm sectors of low pressure areas. Each type of movement was very strongly related to wind direction, with the densest movements tending to occur with following winds. Table 3 and other evidence (Richardson 1972a, in prep.) shows that SW movement tended to be densest with N, NE & E winds in the central, eastern and southern portions of high pressure areas. SE shorebird movement tended to be dense with SW-NW winds immediately behind cold fronts, where the pressure is low or moderate, as well as in the northern and eastern sides of high pressure areas.

FACTOR ANALYSIS

Factor analysis has two characteristics which make it useful in elucidating density-weather relationships:

- (i) It reduces the number of predictors which are to be considered by regression or discriminant analysis. This eases interpretation and eliminates overfitting
- (ii) It permits one to identify underlying factors in the environment, each of which may be reflected in many of the original predictors. Standard factor analysis techniques result in uncorrelated factors.

Thus one can avoid the difficulties in interpreting relationships of density to correlated predictors by using the uncorrelated factors identified by factor analysis as the predictors in a regression or discriminant analysis. At the same time one avoids overfitting when sample sizes are marginal.

Factor analysis usually consists of two steps: extraction of 'principal factors' or 'principal components', followed by rotation of those factors within the dimensions of the original variables. Each principal factor extracted from the original variables is a linear additive function of those variables with different weights on each:

$$\text{Factor Score} = C_1X_1 + C_2X_2 + \cdots + C_kX_k$$

The constants are chosen in order to maximize the variance in factor scores over the cases being examined, within the constraint that the factors must be uncorrelated (orthogonal). When the original variables are strongly intercorrelated, a small number of factors can account for most of the variance amongst the original variables. Once the principal factors have been found, orthogonal rotation is used in order to concentrate the weighting on a few of the original variables, with little weighting on the others. The correlations of the original variables with the factor scores reveal the characteristic of the environment that is being measured by each factor.

Table 4 shows three main factors extracted from the 17 original predictors (excluding date) used in analyses of northward and north-eastward spring migration over New Brunswick (Table 1). The first factor is clearly a measure of fair vs. cloudy or stormy conditions, with strong weighting upon ceiling, visibility, precipitation, humidity, and opacity. Factor two discriminates between synoptic weather situations involving southerly or southwesterly winds, falling pressure, and high and rising temperature, and those involving northerly or north-easterly winds, rising pressure and low and falling temperature. It measures the polarity and strength of the east-west pressure gradient. Factor three discriminates cases with southeasterly side winds from those with northwesterly side winds. I evaluated each of these factors each evening and reran the regression analysis using these three factors as predictors (Table 5A). High volumes of migration were very strongly ($P < .001$) correlated with fair weather (Factor 1) and with synoptic conditions having southerly winds (Factor 2), but only marginally related to conditions with southeasterly winds (cf. Nisbet and Drury 1968, who found stronger relationships to SE than to SW winds for a similar population of birds). The factors are uncorrelated; hence causality can be inferred. The standard error of the estimate based on factors is little larger than that based on the original predictors (3.57 vs. 3.36; $F = 1.129$; $P > .1$). Thus the advantages of a factor approach are not offset by any significant loss of precision.

ASSUMPTIONS IN MULTIVARIATE ANALYSIS

Each of the techniques discussed above makes assumptions about the nature of the data and the form of the interrelationships amongst variables. Violations of these assumptions can lead to unreliable results. None of the previous multivariate studies has included any serious attempt to assess the accuracy of the assumptions.

ADDITIVITY

One assumes additive relationships between the predictors and the dependent variable (in regression) or the discriminant score. The effect

TABLE 4

STRUCTURE OF THREE FACTORS EXTRACTED FROM THE 17 ENVIRONMENTAL VARIABLES
USED IN ANALYSES OF MIGRATION OVER NEW BRUNSWICK ON 86 SPRING NIGHTS¹

	Correlation of original predictor with factor score		
	Factor 1	Factor 2	Factor 3
Magnetic Disturbance	0.053	0.175	- 0.065
Ceiling	0.874	- 0.071	0.001
Visibility	0.767	- 0.073	- 0.012
Precipitation	- 0.638	- 0.002	- 0.124
B. Press. -1000 mb	0.464	0.071	0.292
B. Press. trend	0.236	0.590	- 0.356
Temperature trend	0.368	- 0.616	0.058
Temp. rel. to norm.	0.274	- 0.642	- 0.128
Rel. Humidity	- 0.776	0.306	0.166
Rel. Hum. trend	- 0.616	0.190	0.216
Opacity	- 0.749	- 0.086	0.021
Hrs. since < 10/10	- 0.677	0.064	0.057
SE comp. of wind	- 0.005	- 0.310	0.837
SW comp. of wind	0.041	- 0.774	- 0.145
SE comp. of wind	0.224	- 0.098	- 0.358
SE comp. trend	0.190	- 0.097	0.695
SW comp. trend	0.104	- 0.599	0.127
Interpretation	Fair (+) vs. Cloudy or Stormy (-)	Conditions with Northerly (+) vs. Southerly (-) Winds	Conditions with SE (+) vs. NW (-) Winds

¹ Obtained using BMD08M program (Dixon 1973) by (i) principal factoring of the original variables, iterating from R^2 for communalities, and accepting factors with eigenvalues of 1.0 or more, followed by (ii) Varimax rotation using Kaiser normalization.

TABLE 5

STEPWISE MULTIPLE REGRESSION ANALYSES OF THE VOLUME OF NORTHWARD MIGRATION OVER EAST-CENTRAL NEW BRUNSWICK
IN SPRING USING THE FACTORS FROM TABLE 4 AS PREDICTORS

	A. Overall Model	B. Favourable Synoptic Weather	C. Intermediate Synoptic Weather	D. Unfavourable Synoptic Weather
Fair (+) vs. cloudy or stormy	+ ***	+ ***	+ ***	---
Synoptic Conditions with N(+) vs. S(-) winds	- ***	- *	- **	- ***
Synoptic conditions with SE(+) vs. NW(-) winds	+ (*)	---	+ (*)	---
No. of nights	86	39	26	21
Mult. Correl. Coeff.	0.748***	0.769***	0.800***	0.659***
% Variance Explained	56.0	59.1	64.0	43.4
SE of estimated (#)1.5	3.57	2.94	3.37	4.19

¹ Each column summarizes one analysis, giving the direction and significance of the partial correlations of the factors to the 1.5th power of migration volume. Significance coded as in Table 1; --- means variable did not enter the model.

Model A includes all 86 cases. Model B includes cases with more or less favourable weather (generally southerly winds with a downward pressure trend); model D includes cases with unfavourable weather (northerly winds and rising pressure). Model C includes 26 other cases. Note the differential weighting of the factors in different synoptic situations.

of each predictor upon the volume of migration is assumed to be independent of the values of other predictors. There are logical biological reasons why this assumption should not always be true. For example the relationship between wind speed and density might well be affected by wind direction. Strong winds are quite possibly more favourable than light winds when they are following, but they are certainly less favourable when the wind is opposing.

There is no simple and direct way of checking the accuracy of the additivity assumption. I have tried to assess the situation by comparing the models derived by separate analyses of data from days with synoptic weather that was favourable, unfavourable and indifferent. Table 5 presents one example of such a set of analyses. It would appear that different aspects of the weather are indeed most important to migrants in different weather conditions. For N and NE movement over New Brunswick on spring nights, given that the wind direction is more or less favourable, the strength of the synoptic pattern is apparently much less important in determining migration volume than is fair weather. Conversely, when the winds are opposing, the strength of the synoptic pattern, including wind speed, is of overriding importance. In intermediate conditions both factors seem to be important. Thus in the example presented in Table 5 and in other situations that I have examined, the response of migrants to various factors is not independent of the values of other variables, and hence the additivity assumption is violated.

The existence and nature of violations of the additivity assumption are of considerable interest in understanding the adaptive significance of birds' migration timing systems. However, these violations usually do not have a severe effect on forecasting accuracy. Of the ten categories of movement that I have examined using the partitioning technique of Table 5, only one partitioned model has given a standard error of estimate that approached being significantly less than the standard error from the associated overall model; that is model B in Table 5. Even there, a comparison of the standard errors for models A and B reveals only a very marginal difference ($P = 0.1$). Thus for forecasting purposes, the slight gain in precision obtainable by recognizing and correcting for non-additive relationships probably does not justify the extra difficulty in developing and applying the necessarily more complex procedures.

INDEPENDENCE

All techniques involving statistical inference assume that the units of observation give independent data. The accuracy of this assumption is always in doubt when the units are sequential parts of a time series. Virtually all studies of the effects of weather on migration volume have used 24-hour periods as the units of observation. While slight positive autocorrelations have been observed between densities on successive days (Nisbet and Drury 1968: 506; Richardson in prep.), these were the result of autocorrelation in the weather. After the effects of weather had been accounted for by multivariate models, the forecasting errors were not autocorrelated. Since it is the forecasting errors or residuals which are assumed to be independent, there is no evidence of significant violations of this assumption.

LINEARITY, NORMALITY AND HOMOSCEDASTICITY IN REGRESSION

The regression model assumes that the residuals (errors in forecasting) are normally distributed with constant mean and variance across the range of each predictor variable and the estimated dependent variable. Experience shows that in most situations in which multiple regression is appropriate, violations of one or more of these assumptions will occur. Serious violations are readily recognized by eye when the residuals from a regression analysis are plotted against each of the variables. These plots can be produced on the line printer by the same program which performed the regression analysis, as with the BMD02R program (Dixon 1973). No previous study of migration volume has used this simple but powerful technique. Any departure of the distribution from a band of constant height across the plot represents a violation. The nature of the departure reveals the type of violation and suggests how to overcome it (Anscombe and Tukey 1963; Draper and Smith 1966: Ch.3).

Linearity

When the mean of the residuals varies over the range of some predictor variable, a non-linear relationship exists between the dependent variable, in our case migration volume, and that predictor. The predictor must be rescaled such that it becomes linearly related to density. Alternatively, if the relationship is of some simple curvilinear form, it may be desirable to add higher-order terms involving that variable. For example, migration volume tends to be lower early and late in the migration season than at its mid-point. Provided that one does not extend the period of study too far, the relationship to date may be more or less parabolic. If so, it would be reasonable to include the date and the square of the date as predictor variables. When stepwise analysis is applied to data of this sort, it is usually necessary to make provisions for all terms based on a single parameter to enter the model together.

Wind direction and speed are particularly difficult to deal with because of the circular scale of directions and violations of the additivity assumption. Conversion of these two variables into components along and across the mean flight direction is a widely used and generally successful approach (see Appendix).

Homoscedasticity

When the dispersion in residuals varies over the range of the estimate of the dependent variable, one has a violation of the homoscedasticity assumption. A transformation of the dependent variable is necessary. When the distribution of the dependent variable is skewed to the right, there will be more variance in residuals for high estimates than for low estimates. This always occurs when migration volume is measured on an interval scale. The appropriate transformation will be one which compresses the upper end of the distribution more than it does the lower end (e.g., logarithm, square root, tenth root, etc.). Conversely, when the distribution of the dependent variable is skewed to the left, there will be more variance in residuals for low than for high estimates, and a scale-expanding transformation will be needed (e.g., square, antilogarithm, etc.). The best transformation to use

must be found by trial and error, redoing the analysis with alternative transformations until the residuals reveal no violations of assumptions. Analyses using inappropriate and appropriate scales for the dependent variable often suggest very different levels of significance for some of the predictors and produce very different degrees of forecasting success. However, provided that one selects the correct general type of transformation, its precise nature is rarely important. For example, 4th root, 10th root and logarithmic transformations usually indicate similar significance levels for each predictor and give a similar degree of forecasting accuracy (Nisbet and Drury 1968: 530; Richardson in prep.).

Nisbet and Drury (1968) selected as most appropriate the transformation which produced the maximal multiple correlation coefficient. This criterion usually but not always leads one to a model with homoscedastic residuals. It is preferable to look directly at the residuals, since they are the subjects of the assumption.

Heteroscedasticity of residuals when plotted against the predictor variables often disappears once an appropriate transformation of the dependent variable has been found. If not, a scale-compressing or scale-expanding transformation should be applied to the predictor concerned. This is most likely to be necessary when the values of the predictor variable are not normally distributed.

When transformations are used, statistical inferences are made about the transformed variables. One must be cautious in applying the conclusions to the untransformed variables. However, one should not hesitate to use simple scale-expanding and scale-compressing transformations when they are needed to meet the assumptions. There is rarely any reason for assuming that birds 'measure' a variable on the original rather than the transformed scale. Furthermore, the basic objective is to achieve an accurate forecasting procedure; experience shows that transformations are usually necessary for this to be achieved.

Normality

The residuals are assumed to be normally distributed. Violations of this assumption usually occur when the dependent variable is not normally distributed. These violations often disappear once the homoscedasticity assumption has been met by an appropriate transformation.

When the dependent variable has a bimodal distribution, no simple transformation is likely to satisfy the normality assumption. In this situation it is safest to categorize the densities and use discriminant rather than regression analysis. I have had to use this approach in most of the analyses of densities over Nova Scotia and New Brunswick (e.g., Table 3). Nisbet and Drury (1968) applied regression analysis to migration volumes which were bimodally distributed. While they recognized the problem, they did not assess the effect which it may have had upon the reliability of their results.

Ordinal scales

The volume of migration as detected by radar must often be recorded on an ordinal rather than an interval scale (see Siegel 1956: 22 for a

discussion of scaling procedures). Lack (1960b, 1963a, b) and I (this report; Richardson 1974, in prep.) have applied multiple regression to ordinal density data. Moon-watching and ceilometer studies suggest that ordinal scales are compressed at their upper end. Thus the relationships between values on the ordinal scales are presumably similar to those between values on interval scales after a square root, log or comparable transformation. Hence it is not surprising that ordinally-scaled densities require little or no transformation in order for the residuals to meet the normality and homoscedasticity assumptions (Table 1; Richardson 1974, in prep.). Providing that one checks the assumptions by analysis of residuals, multiple regression can be safely applied to ordinally scaled data, when the more desirable interval scaling cannot be attained.

ASSUMPTIONS IN DISCRIMINANT ANALYSIS

In addition to the additivity assumption, discussed above, discriminant models assume linear relationships between the predictors and the discriminant scores. In theory violations of this assumption can be recognized by examining the pattern of accurate and inaccurate forecasts across the range of each predictor. In practice this is rarely possible because it requires large sample sizes. As in the regression setting, violations of the linearity assumption can usually be corrected, once recognized, by rescaling or by addition of higher order terms.

Discriminant analysis of migration volume also assumes that the variances and covariances among the predictors are similar for each category of density. While serious violations of this assumption can be recognized by comparing corresponding elements of the variance-covariance matrices, appropriate solutions are less obvious. As in the regression setting, predictors which have non-normal distributions are likely to result in violations. Hence one useful guideline is to transform any non-normal predictor towards normality. Ceiling and visibility are variables which have seriously skewed distributions. The scales described in the Appendix were chosen in order to avoid this problem.

NON-STATIONARITY

When the observational units are distributed over a time series, one problem which may arise is the tendency for variables to vary systematically over the course of the observations (i.e., to be non-stationary; see Otnes and Enochson 1972: 395 for review). When analyzing migration volume, weather variables are often significantly correlated with date. Temperature is the most obvious case, but others such as pressure, wind speed, and cloudiness may also show a systematic trend over the course of the season. If one assumes that birds are adapted to respond to variations in weather about the normal values, then the appropriate form of any non-stationary predictor in the analyses is its departure from the normal value on the date in question. Only temperature has been recognized as non-stationary and treated in this fashion in previous multivariate studies. Normal values over the range of dates in a migration season may often be obtained by fitting a straight line to the observed daily values.

When the dependent variable is non-stationary, as is migration

volume over the course of the season, two approaches are possible. One can either convert to values relative to normal prior to the analysis, or one can use appropriate date terms as predictors. The former approach is used in Table 3; the latter in Tables 1 and 2.

EFFECTS OF EXCLUDED VARIABLES

It is sometimes not possible to include a potentially important predictor in a regression analysis because it is non-linearly related to the dependent variable, or because its value was recorded for only a fraction of the cases. One can perform the analysis of migration volume excluding the variable in question and then look for relationships between the residuals and that variable. Thereby one determines whether consideration of the extra predictor can account for any additional variance in the densities beyond what has already been accounted for by the other predictors. Nisbet and Drury (1968) applied this technique to the moon phase variable, and I have used it in order to assess the importance of synoptic weather and days of delay since the last intense flight, other factors being equal.

CONCLUSIONS

When carefully applied, multivariate techniques offer a powerful capability for producing operationally useful forecasts of migration volume. However the assumptions of these techniques should be considered more carefully than has been done in the past. The residuals should be examined in every regression analysis, and appropriate transformations should be made when the assumptions are violated. When they are appropriate, more use should be made of less familiar procedures such as discriminant and factor analysis. The heuristic capability of multivariate procedures should be exploited.

A sophisticated multivariate analysis usually requires much less time and effort than does the collection of the data. Thus an incomplete analysis of the data saves little effort and wastes information that has already been collected. An additional attraction of multivariate techniques is that once learned, they are applicable to a wide variety of problems.

ACKNOWLEDGMENTS

I thank the Telecommunications Maintenance section, Halifax International Airport, the personnel of Canadian Forces Stations Barrington, St. Margarets and Sydney, and the 140th AC & W Squadron, Puerto Rico Air National Guard, for their co-operation in obtaining the radar data. My wife Dorothy assisted in data compilation, and Drs. S. T. Emlen, W. W. H. Gunn and A. van Tienhoven offered valuable suggestions. This study was financed by the National Research Council of Canada through its Associate Committee on Bird Hazards to Aircraft.

LITERATURE CITED

- ABLE, K. P. 1973. The role of weather variables and flight direction in determining the magnitude of nocturnal bird migration. *Ecology* 54: 1031-1041.
- ANSCOMBE, F. J., and J. W. TUKEY. 1963. The examination and analysis of residuals. *Technometrics* 5: 141-160.
- BLOKPOEL, H. 1973. Bird migration forecasts for military air operations. *Canad. Wildl. Serv. Occas. Paper* 16: 1-18.
- COOLEY, W. W., and P. R. LOHNES. 1971. *Multivariate data analysis*. New York, J. Wiley and Sons.
- DIXON, W. J. 1973. *BMD: Biomedical computer programs*. Berkeley, Univ. of California Press.
- DRAPER, N. R. and H. SMITH. 1966. *Applied regression analysis*. New York, J. Wiley and Sons.
- GRUYS-CASIMIR, E. M. 1965. On the influence of environmental factors on the autumn migration of Chaffinch and Starling: A field study *Arch. Néerl. Zool.* 16: 175-279.
- GUNN, W. W. H., and V. E. F. SOLMAN. 1968. A bird warning system for aircraft in flight. Pp. 87-96 *in* The problems of birds as pests. (R. K. Murton and E. N. Wright, Eds.) *Institute of Biology Symposium No. 17*. New York, Academic Press.
- HARMAN, H. H. 1967. *Modern factor analysis*, 2nd ed. Chicago, Univ. of Chicago Press.
- HAUGH, J. R. 1972. A study of hawk migration in eastern North America. *Search-Agriculture* 2 (16): 1-60.
- LACK, D. 1960a. The influence of weather on passerine migration: a review. *Auk* 77: 171-209.
- LACK, D. 1960b, 1963a, b. Migration across the southern North Sea studied by radar. Part 2. The spring departure 1956-59. Part 4. Autumn. Part 5. Movements in August, winter and spring, and conclusion. *Ibis* 102: 26-57; 105: 1-54; 105: 461-492.
- LANE, N. E. 1971. The influence of selected factors on shrinkage and overfit in multiple correlation. Pensacola, Florida, Naval Aerospace Med. Res. Lab. Monograph 17.
- MORRISON, D. F. 1967. *Multivariate statistical methods*. New York, McGraw-Hill.
- NIE, N. H., D. H. BENT, and C. H. HULL. 1970. *SPSS. Statistical package for the Social Sciences*. New York, McGraw-Hill.
- NISBET, I. C. T., and W. H. DRURY, Jr. 1968. Short-term effects of weather on bird migration: A field study using multivariate statistics. *Anim. Behav.* 16: 496-530.
- OTNES, R. K. and L. ENOCHSON. 1972. *Digital time series analysis*. New York, J. Wiley and Sons.
- RAO, C. R. 1952. *Advanced statistical methods in biometric research*. New York, J. Wiley.
- RICHARDSON, W. J. 1970. Temporal variations in the volume of bird migration: a radar study in Canada. *Proc. World Conf. on Bird Hazards to Aircraft*, Kingston, Ontario: 323-334.
- RICHARDSON, W. J. 1971. Spring migration and weather in eastern Canada: a radar study. *Am. Birds* 25: 684-690.
- RICHARDSON, W. J. 1972a. Autumn migration and weather in eastern Canada: a radar study. *Am. Birds* 26: 10-17.

- RICHARDSON, W. J. 1972b. Temporal variations in the ability of individual radars in detecting birds. Nat. Res. Counc. Canada Assoc. Comm. on Bird Haz. to Aircraft Field Note 61: 1-68.
- RICHARDSON, W. J. 1974. Spring migration over Puerto Rico and the western Atlantic, a radar study. Ibis 116 (in press).
- RICHARDSON, W. J. and W. W. H. GUNN. 1971. Radar observations of bird movements in east-central Alberta. Canad. Wildl. Serv. Rep. Ser. 14: 35-68.
- RICHARDSON, W. J. and M. E. HAIGHT. 1970. Migration departures from Starling roosts. Canad. J. Zool. 48: 31-39.
- ROHLF, F. J., J. KISHPAUGH, and D. KIRK. 1971. NT-SYS. Numerical taxonomy system of multivariate statistical programs. Stony Brook, N. Y., Tech. Rep. State Univ. of New York at Stony Brook.
- SIEGEL, S. 1956. Nonparametric statistics for the behavioral sciences. New York, McGraw-Hill.
- SIMON, H. A. 1954. Spurious correlation: a causal interpretation. J. Amer. Statist. Ass. 49: 467-479.
- SNEATH, P. H. A. and R. R. SOKAL. 1973. Numerical taxonomy. San Francisco, W. H. Freeman.

APPENDIX

SCALING PROCEDURES FOR PREDICTOR VARIABLES

The scales for the predictors used in analyses of nocturnal migration are described below. These scales usually gave linear relationships to migration volume with no heteroscedasticity of forecasting errors across the ranges of the predictors. In analyses of diurnal migration, weather variables were measured at sunrise rather than sunset.

Magnetic disturbance.-- K index recorded at 20:00-23:00 AST, on a scale of 0 (0-5 gammas) to 9 (>500 gammas). One gamma = 0.00001 gauss.

Ceiling.-- Based on a 6/10 opacity criterion at sunset 1 = fog; 2 = < 1200 ft. but no fog; 3 = 1200-3900 ft; 4 = 4000-9900 ft.; 5 = 10,000-50,000 ft.; 6 = unlimited (i.e. < 6/10 total opacity).

Visibility.-- Square root of horizontal visibility in miles at sunset, to a maximum of 4.0 representing 16 or more miles.

Precipitation.-- 0 = none; 1 = drizzle or showers at sunset without widespread precipitation echo on radar display; 2 = same with precipitation echo on display, or continuous precipitation without widespread echo on display; 3 = continuous precipitation or thunderstorms with widespread echo on radar display.

Barometric pressure.-- In millibars minus 1000 at sunset. In autumn measured relative to normal because of non-stationarity over date.

Pressure trend.-- Pressure at sunset minus that six hours earlier, in mb.

Temperature trend.-- Sunset value minus sunset value previous evening in F°.

Temperature relative to normal.-- Sunset value in F° above (+) or below (-) normal sunset value on that date.

Relative humidity.-- Sunset value in per cent. In autumn measured relative to normal because of non-stationarity over date.

Relative humidity trend.-- Sunset value minus sunset value previous evening.

Opacity.-- Proportion of sky covered by opaque cloud at sunset, in tenths.

Hours since < 10/10 opaque.-- Number of hours prior to sunset during which it has been continuously overcast (0 = not overcast at sunset; 1 = overcast for one hour preceding sunset, etc.).

Following and side components of wind; absolute value of side component.-- Surface wind direction and speed at sunset in mph resolved trigonometrically along axes parallel and perpendicular to mean flight direction for the birds being studied. (Wind direction itself is measured on a circular scale and is not linearly related to migration volume). In autumn SW component measured relative to normal because of non-stationarity over date.

Wind component trend.-- Component at sunset minus that at sunset on previous evening.

Day of year.-- 1-365 scale. (In order to avoid round-off error, it is often helpful to use $(\text{date})^2/100$ in the analysis and then to adjust the constant accordingly).