A GENERALIZED DISCRIMINANT FOR SEXING FULMARINE PETRELS FROM EXTERNAL MEASUREMENTS

J. A. van Franeker[1,2] and C. J. F. ter Braak[1,3]

[1]*Institute for Forestry and Nature Research (IBN-DLO), P.O. Box 167,*
NL-1790 AD Den Burg (Texel), The Netherlands;
[2]*Institute for Taxonomic Zoology, University of Amsterdam, The Netherlands; and*
[3]*Agricultural Mathematics Group (GLW-DLO), Wageningen, The Netherlands*

Abstract.—Discriminant analysis can use morphometric differences between known male and female birds to predict the sex of unknown individuals in field studies. Geographic variation in size and shape often limits the predictive value of a discriminant function to the population from which it was derived. Specific discriminant functions for populations of five species of fulmarine petrels (Northern Fulmar, *Fulmarus glacialis*; Southern Fulmar, *F. glacialoides*; Antarctic Petrel, *Thalassoica antarctica*; Cape Petrel, *Daption capense*; and Snow Petrel, *Pagodroma nivea*) assigned 81 to 98% of birds in the samples to the correct sex, but the validity of each discriminant applied to alternative populations remained questionable. Our approach to overcome this limitation is to combine data from the different species into a single discriminant. Adequate performance of this generalized discriminant in samples of different species shows its validity for use in other populations of any of these species. The generalized function calculates the discriminant scores for individual fulmarine petrels as: $Y = HL + 2.38BD + 0.41TL - 0.21CL$, where HL is head length, BD is bill depth, TL is tarsus length and CL is bill length (measurements in millimeters). The cut point to split sexes is different in each sample and may be calculated directly from discriminant scores, without reference to sexed birds, by using a maximum-likelihood method. Depending on species, the generalized method results in 84 to 97% correct classifications and can be applied to other populations of fulmarine petrels without requiring samples of birds of known sex. *Received 19 November 1991, accepted 20 November 1992.*

Fulmarine petrels, like most seabirds, lack plumage characters by which sexes may be recognized. Small differences in mensural characters, however, may reveal sufficient dimorphism to distinguish sexes. From known correlations between sex and measurements in a sample of sexed birds, a discriminant analysis (Sokal and Rohlf 1981) can weight characters for their power to distinguish groups (sexes) of unknown individuals.

Many statistical packages supply computer programs for discriminant analysis. Computerized data processing has promoted the use of discriminant analysis. The method has been successfully applied to a wide variety of seabird species from different groups, including penguins (Scolaro et al. 1983, Gales 1988, Williams 1990, Williams and Croxall 1991), divers (Okill et al. 1989), tube-nosed birds (Dunnet and Anderson 1961, Brooke 1978, Copestake and Croxall 1985, Sagar 1986, Johnstone and Niven 1989), gulls (Shugart 1977, Ryder 1978, Fox et al. 1981, Nugent 1982, Coulson et al. 1983, Hanners and Patton 1985, Schnell et al. 1985), skuas (Hamer and Furness 1991) and also to freshwater and terrestrial birds (e.g. Anderson 1975, Green 1982). Various approaches are possible. The most usual for discrimination between sexes is to construct a single formula that calculates a discriminant score for each individual on the basis of its measurements. The cut point to partition scores into male and female groups is usually taken as the midpoint of the interval between the group means of sexed males and females.

Unfortunately, there are some drawbacks that prevent general usage of published sex discriminants. Many bird species show considerable variation in size over their geographical range. As a consequence, the cut point calculated for one population may not be appropriate for another. Also, geographic variation may involve shape, which could affect the weighting of characters in the discriminant formula. For example, Northern Fulmars (*Fulmarus glacialis*) from Spitsbergen are not only considerably smaller in overall size, but also have relatively shorter bills than Northern Fulmars from Britain: relative to head length, bill length is about 4% shorter in the Spitsbergen birds (van Franeker and Wattel 1982, unpubl. data). Wynne-Ed-

TABLE 1. Samples of birds used for discriminant function analysis. Total numbers, with numbers of sexed males and females in parentheses.

Species	Locality	Sex determined by	
		Dissection	Observation
Northern Fulmar	Netherlands	247 (117, 130)	—
Northern Fulmar	Jan Mayen	32 (12, 20)	—
Southern Fulmar	Ardery Island	27 (13, 14)	103 (51, 52)
Antarctic Petrel	Ardery Island	11 (6, 5)	66 (30, 36)
Cape Petrel	Ardery Island	30 (19, 11)	32 (16, 16)
Snow Petrel	Casey Station	—	32 (15, 17)

wards (1952) described further variation in bill shapes and in levels of sexual dimorphism in populations of the Northern Fulmar. Age-dependent variation in size or shape can be a further complicating factor (e.g. Coulson et al. 1983, Scolaro et al. 1983). The uncertainties induced by these types of variation urged authors to caution against the unchecked application of their sex discriminant to other populations (e.g. Fox et al. 1981, Nugent 1982, Witt et al. 1984, Gales 1988, Hamer and Furness 1991).

Since 1980 the first author has worked on several projects involving five species of fulmarine petrels: Northern Fulmar, Southern Fulmar (*F. glacialoides*), Antarctic Petrel (*Thalassoica antarctica*), Cape Petrel (*Daption capense*), and Snow Petrel (*Pagodroma nivea*). All studies required knowledge of the sex of birds handled in the field. In some field projects, birds sexed by dissection or by field observations allowed the construction of a discriminant function that could be applied to other birds within the population. However, in other study populations, it was not possible to sex an adequate sample of birds. In spite of a good general impression of sexual dimorphism in fulmarine petrels, we were unable to give a reliable prediction of sexes of birds in those populations where sexed individuals were missing. The same problem will also be encountered in future projects. Thus, we decided to reconsider all our data in an attempt to construct a reproducible method to discriminate sexes in populations of the fulmarine petrels without requiring birds of known sex. The method calculates (1) a single generalized discriminant from data on sexed birds of a number of different populations, and (2) population-specific cut points without reference to sexed birds. The predictive power of the method is demonstrated for populations of five species of fulmarine petrels.

METHODS

Samples of sexed fulmarine petrels are characterized in Table 1. Northern Fulmars, beached in the Netherlands between 1980 and 1988, were dissected for morphological and pollution-related studies (van Franeker 1983, 1985). Several other North Atlantic fulmar populations were studied, but an adequate sample of sexed birds was available only from Jan Mayen (van Franeker et al. 1986). Southern Fulmars, Antarctic Petrels, Cape Petrels and Snow Petrels were studied in 1984–1985 and 1986–1987 at the Australian base Casey (66°S, 110°E) and the nearby Ardery Island, Wilkes Land, Antarctica (van Franeker et al. 1990). For Antarctic species, two groups of sexed birds were available: birds sexed by dissection; and birds sexed by observing characteristic behaviors or external anatomy. Data on the latter group involved cloacal evidence (Serventy 1956, Boersma and Davies 1987), copulation position, egg laying, incubation shifts (Pinder 1966), known sex of partner and, for Snow Petrel, voice level (Bretagnolle 1990). Although generally reliable, a small risk for incidental misinterpretation of such observations may occur. Therefore, our calculations in this paper have often been made separately for dissected birds, and for all sexed birds (dissection + observation). Skins of dissected specimens from known breeding localities have been deposited in the care of the Zoological Museum of the Institute ot Taxonomic Zoology, University of Amsterdam.

Measurements.—Measurements of the following characters were taken from fresh corpses or live specimens (Fig. 1; Cramp and Simmons 1977): **bill length (CL)**, from edge of implantation of feathers to most distant part of curve of hook (accuracy ±0.1 mm); **bill depth (BD)**, from angle of gonys to dorsal surface of hook (±0.1 mm); **head length (HL)**, from supraoccipital to front edge of bill (±1 mm); **tarsus length (TL)**, from middle of midtarsal joint to distal end of tarsometatarsus (±0.5 mm); **wing length (WL)**, maximum flattened chord, from carpal joint to tip of longest primary (±1 mm); **tail length (TA)**, central tail feathers from point where emerging from skin to tip (±1 mm). Lengths of wing and tail were not always
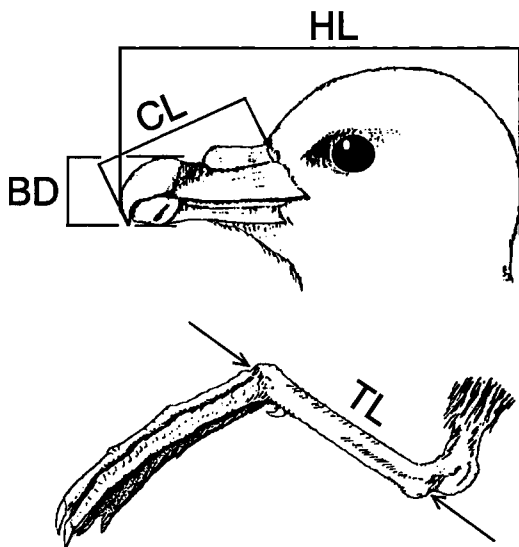
Fig. 1. Four measurements taken on fulmarine petrels: head length (HL); bill depth (BD); bill length (CL); and tarsus length (TL).

taken because of extreme wear or molt. Measurements of birds sexed by dissection are summarized in the Appendix. Body mass was not used in the analysis because of strong seasonal and short-term variability (cf. Johnstone and Niven 1989).

Discriminant analysis.—We developed a discriminant-function formula for calculating individual discriminant scores and a cutpoint to partition scores into male and female groups (Lachenbruch 1975). Formulas for separate populations were calculated by use of stepwise multiple-regression analysis (Genstat 5; Payne et al. 1987). Stepwise multiple regression ranks characters according to their discriminative power and supplies estimates for a constant and a regression coefficient b_i (i.e. the character's weight) for each of the characters ($i = 1, \ldots, n$) from which discriminant scores of birds can be calculated. To facilitate comparison between formulas for discriminant functions, we have omitted the constant and divided the coefficients by that of the first-ranked character. The discriminant score (*DS*) of birds is then calculated by:

$$DS = m_1 + w_2 m_2 + \ldots + w_n m_n, \qquad (1)$$

where m_i is the measurement of the bird for character i, and w_i is b_i/b_1 the adjusted coefficient. The adjusted coefficient for the first character is, thus, always recalculated to value 1. Scores of birds of known sex can be used to calculate the cut point to split the male and female components of the frequency distribution of discriminant scores. The cut point was calculated as the midpoint between the mean scores for males and females. As males are larger than females in all

fulmarine petrels (Appendix), the sign of the coefficients was chosen in such a way that a score of an unsexed bird above the cut point indicates a male, and values below indicate a female.

Discriminant analyses were carried out with both the original measurements and their logarithms. Logarithmic transformation did not improve results and is not discussed further. Multiple-regression analysis in Genstat supplies information on outliers, such as birds with malelike measurements but sexed as females. Outliers were checked for errors in data entry, but were not omitted from analyses.

Reliability of sex assignment by discriminant functions was tested by various methods. First, a self test was used in which performance of the discriminant function is tested on the material from which the function was calculated. As self tests ignore bias due to sample size, two other tests based on the principles of the V_1 validation procedure (Frank et al. 1965) were also applied. The cross test (Stone 1974) randomly attributes all birds in a sample to one of four groups. In each possible combination of these groups, the performance of a discriminant function calculated from three groups is tested on the remaining group. Thus, all birds are tested in a discriminant function derived from other birds. We also applied the jackknife test (Lachenbruch 1975, Dixon 1985), which is similar to the cross test, but calculates a discriminant function for all but one of the birds in the sample, evaluates that bird, and repeats the procedure for all birds in the sample. We consider the jackknife test the best indicator of performance because discriminants based on all individuals except from one will suffer least from small sample sizes. In small samples, self tests will tend to overestimate and cross tests will tend to underestimate the performance of a function. Nevertheless, we have also listed results of these tests, as they are used in literature on sex discriminants in birds.

Multisample discriminant analysis.—If samples of sexed birds in different populations are small, but different populations have a similar morphology (i.e. shape), it may be advantageous to estimate a generalized discriminant from the combined samples (cf. Rayens and Greene 1991). The generalized discriminant calculates a single set of weights of characters for use in all populations. The cut point of course has to be population-specific. A generalized discriminant can be calculated for a group of different populations within a species, or, as in our case, for a group of different species. Details of the method are as follows. In a population-specific discriminant for population p, the weights (equation 1) depend on the mean difference vector ($\mathbf{b}_p$) between the sexes and the pooled within-group covariance matrix ($\mathbf{S}_p$). In the formula, the vector of weights can be expressed as $\mathbf{S}_p^{-1}\mathbf{b}_p$ (Lachenbruch 1975). An analogous generalized discriminant can be obtained by $\mathbf{S}^{-1}\mathbf{b}$, where the within-group covariance matrix $\mathbf{S}$ and the mean difference

vector **b** are pooled across populations. This idea is not new; it is an option in the BMDP package (Dixon 1985:530–531 and 679–681). The method of pooling needs some further consideration. If the true within-groups covariance matrices are identical, an efficient method of pooling is to assign each individual equal weight in the calculation of **b** and **S**. However, when the true within-groups covariance matrices differ, such a way of pooling may be misleading: the populations that are overrepresented in the sample will unduly influence the generalized discriminant. To avoid this, each population can be given equal weight (i.e. **b** and **S** are averaged over the population-specific **b**$_p$ and **S**$_p$). The latter method was chosen for our data so as to not overemphasize sex differences in the large samples of Northern Fulmars.

Performance of multisample discriminant functions (MDF) was evaluated by calculating species-specific cut points for MDF scores and checking the performance for birds of known sex in each species. This procedure looks similar to the self test, but differs from it by the fact that the discriminant function was derived from a much larger (multispecies) sample of sexed birds than the (single species) sample that was tested. The cross test and jacknife test could not be applied because of species-specific cut points.

Cut-point calculation without reference to sexed birds.— Because of size variation, each population needs a population-specific cut point to partition scores into male and female groups. In the absence of sexed birds, the cut point has to be derived from the shape of the frequency distribution of discriminant scores. A discussion on different methods for decomposing mixtures of distributions into their components was given by Titterington et al. (1985). We have chosen to use the method of maximum likelihood based on the assumption that the distribution is a mixture of two univariate normals with possibly unequal variance. The calculations were performed with the expectation-maximization algorithm as detailed in example 4.3.2 of Titterington et al. (1985:84–87). This algorithm yields estimates of the means (μ_1, μ_2) and variances (σ_1, σ_2) of the normals. From these, the cut point is derived by solving for the point where the two normal densities intersect. In formula, the cut point is:

$$x_g = (\sigma_2^2 - \sigma_1^2)^{-1}\{\mu_1\sigma_2^2 - \mu_2\sigma_1^2 \\ + \sigma_1\sigma_2[(\mu_1 - \mu_2)^2 \\ + (\sigma_1^2 - \sigma_2^2)\ln \sigma_1^2/\sigma_2^2)]^{0.5}\}. \tag{2}$$

In the petrel species under consideration here, males are larger than females in overall size and in all individual characters (van Franeker and Wattel 1982; see Appendix). After calculation of the cut point, this knowledge suffices to assign sexes to birds in the sample. Information on VAX Genstat and personal computer programs for calculations of a generalized discriminant function from different samples and for the cut-point calculations are available from the first author.

RESULTS AND DISCUSSION

Population-specific sex discriminants.—Stepwise multiple regression of the measurements of Northern Fulmars from the Netherlands indicated the following sequence of characters in order of decreasing importance to a sex discriminant: head length, bill depth, tarsus length and bill length. Also, when repeating the analysis with wing length and/or tail length (smaller sample sizes because of missing data), head length proved to be the most discriminating character. Discriminant functions with from one to six characters, starting with head length and each time adding the next most important character, are shown in Table 2. Functions not using head length do not perform as well. A discriminant function using only bill length and bill depth ($CL + 2.519BD$; cut point = 81.44) resulted in 93.1% correct classifications. In appearance and performance, this function is very similar to the discriminants based on CL and BD that were developed by Dunnet and Anderson (1961) and MacDonald (1977) for Northern Fulmars breeding in Scotland. Such similarity was to be expected as the majority of Northern Fulmars beached in the Netherlands are of British origin (van Franeker 1979). Head length alone discriminates between sexes better than the combination of bill length and depth mentioned above. Adding characters to head length improves performance for Northern Fulmars beached in the Netherlands to about 98% according to all tests. More variables also enhanced bimodality in the frequency distributions of discriminant scores, which improves classification reliability for individual birds.

In discriminant function NL4 (Table 2) for Dutch beached Northern Fulmars, bill length is weighted negatively, whereas male bills average significantly larger than female bills. The correlation structure of the variables (as well as the sample size) indicates that the additional discriminative power of bill length in the formula lies not within its absolute value, but in its value relative to other characters, especially head length. Apparently, at a given head length, a male is likely to have a relatively smaller bill than a female having the same head length. Negative weights of other characters can be interpreted in a similar way. Okill et al. (1989)

TABLE 2. Discriminant functions for Northern Fulmars from The Netherlands. Function names given as codes indicating locality/species and numbers of characters used. For example, function NL3 for Dutch Fulmars based on 247 sexed individuals uses three characters of head length, bill depth and tarsus length to calculate discriminant scores as $HL + 1.057BD + 0.428TL$ and classifies birds with score over cut point 136.0 as males, and all lower values as females. Tests indicate that 97 to 98% of birds are assigned to correct sex by this function.

Function	n	Estimates for discriminant formula						Cut point	Percent correctly classified		
		HL	BD	TL	CL	WL	TA		Self test	Cross test	Jack-knife
NL1	247	1						95.2	95	95	95
NL2	247	1	0.919					110.6	97	97	97
NL3	247	1	1.057	0.428				136.0	97	98	97
NL4	247	1	0.935	0.365	−0.400			114.9	98	98	98
NL5	214	1	1.056	0.436	−0.457	−0.025		110.4	98	98	98
NL6	189	1	1.084	0.371	−0.510	0.015	−0.096	106.9	97	97	97

and Hamer and Furness (1991) also found negative signs of coefficients reflecting differences in shape rather than size.

For further analysis, we focus on the character combination of head length, bill depth, tarsus length and bill length (as in function NL4 in Table 2). Wing and tail lengths have been omitted because their contribution to the discriminant is small, and they contain missing values in all our samples. The other four characters were included because while head length is very important for Northern Fulmars from the Netherlands, this is not necessarily the case for other populations or species. Also, frequency distributions of discriminant scores tended to show stronger bimodality when more characters were used, which is important when determining cut points without reference to sexed birds.

Table 3 shows the results of discriminant analysis using the four selected characters for all samples/species. For some Antarctic species, two discriminants were calculated: one for birds sexed by dissection, and another for all sexed birds (sex determined by dissection or by observations). The functions based only on dissected birds have the disadvantage of being based on small sample size, but the advantage of certainty that the sex determinations are correct. The samples that combine all sexed birds are larger but may contain missexed individuals. Discriminant function JM4 for Northern Fulmars from Jan Mayen is not unlike the function for Dutch Northern Fulmars. Similarity in results for the two populations was expected, as both are assigned to the same subspecies (F. glacialis auduboni; van Franeker et al. 1986). Un-

known, however, is the reliability of these functions when applied to morphologically different subspecies from the High-Arctic Atlantic (F. g. glacialis) or from the Pacific (F. g. rodgersii; van Franeker and Wattel 1982, Wynne-Edwards 1952). Analyses for Southern Fulmar, Antarctic Petrel, Cape Petrel, and Snow Petrel ranked the characters HL, BD, TL and CL in the same sequence of importance as in Northern Fulmars, but estimates for regression coefficients fluctuated widely between species.

Strong fluctuations in regression coefficients were not limited to different species, but also occurred within a single species. Our discriminants derived from the large samples of birds sexed by dissection or observation were sometimes remarkably different from discriminants derived from dissected birds only. The correlation structure among variables, as well as the sample size, could account for a number of these differences. There were, however, no obvious differences in performance. We also tested the performances of formulas derived from the small sample of dissected birds to the large sample of all sexed birds and vice versa. In any particular sample, the two formulas produced almost equal results, in spite of the differences in weights of characters.

The observation that, within a species, two rather different discriminant formulas operated in a very similar manner was reason to do further tests to determine the value of our population-specific discriminant functions. Therefore, we tested the performance of the formula derived from our largest sample (NL4 of Dutch Northern Fulmars) to all other samples. Cut points were calculated as the midpoint between

TABLE 3. Population-specific discriminant functions using head length, bill depth, tarsus length and bill length. Functions derived from samples of birds sexed by dissection are named (e.g. TH4) and, when sexed by dissection or observation, OBS is added to the designation (e.g. TH4 + OBS).

Function	n	Estimates for discriminant formula				Cut point	Percent correctly classified		
		HL	BD	TL	CL		Self test	Cross test	Jack-knife
Northern Fulmar (Netherlands)									
NL4	247	1	0.935	0.365	−0.400	114.9	98	98	98
Northern Fulmar (Jan Mayen)									
JM4	32	1	1.763	0.503	−0.585	130.5	97	100	97
Southern Fulmar									
FU4	27	1	4.876	1.748	0.776	306.4	96	96	96
FU4 + OBS	130	1	24.31	1.755	3.174	724.5	90	86	89
Antarctic Petrel									
TH4	11	1	6.063	10.74	14.63	1,201.0	100	64	82
TH4 + OBS	77	1	4.633	0.712	−0.386	176.4	86	79	82
Cape Petrel									
DA4	30	1	2.355	0.333	−0.775	95.8	87	77	83
DA4 + OBS	62	1	1.164	0.208	0.819	126.7	86	76	81
Snow Petrel									
PA4 + OBS	32	1	3.096	−1.230	2.471	106.9	94	91	91

the group means of scores of known males and females. Test results for percentages of correct classifications were: Northern Fulmar Jan Mayen, 96.9%; Southern Fulmar, 92.6 and 86.9%; Antarctic Petrel, 90.9 and 83.1%; Cape Petrel, 86.7 and 83.9%; and Snow Petrel, 87.5% (when two figures given, first refers to dissected birds and second to enlarged sample of all birds sexed by dissection or observation). When comparing these results to jackknife test results in Table 3, formula NL4 for the Dutch Northern Fulmars classifies sexes in other species almost as well as population-specific discriminants. In Southern Fulmar and Snow Petrel the percentage of correct classifications by formula NL4 was somewhat lower than when using population-specific discriminants, but in Antarctic Petrel and Cape Petrel performance was improved.

The strong differences in character weightings within a population (Table 3) and the similar performances of widely different functions throw some doubt on the value of the character weights in our population-specific discriminants, at least for the Antarctic species. Apparently, our population samples are inadequate to give a consistent and optimal (better than general) description of sexual dimorphism within the populations studied. This is not re-ally a problem when working in these populations, but it does create considerable uncertainty when applying the functions to other populations that may be morphologically different. On the other hand, the results indicate that there may be sufficient similarities in morphology between different species of fulmarine petrels to construct a generalized discriminant for all these species. If we are able to construct a reliable general discriminant for different species, such a discriminant likely will perform adequately in unknown populations of each of those species because morphological variation between species is larger than variation within species. In our attempt to construct a multispecies discriminant, we are aware that it cannot improve on the performance of proper single-species discriminants based on an adequate variety of samples of different populations, or maybe even a single population. However, as in our case, adequate samples for such species-specific discriminants often are lacking. Our approach is to provide a tested general discriminant that can be used as long as tested species-specific discriminants are missing.

Generalized discriminant for fulmarine petrels.— All samples of birds sexed by dissection plus the sample of Snow Petrels sexed by observations (379 birds from six populations of five

TABLE 4. Generalized discriminant formulas for fulmarine petrels based on morphology of all species. Performance tested in samples of birds sexed by dissection (1) or in enlarged samples containing both dissected birds and birds sexed by observations (2) of: Northern Fulmar, Netherlands (n_1 = 247); Southern Fulmar (n_2 = 27 and n_2 = 130); Antarctic Petrel (n_1 = 11 and n_2 = 77); Cape Petrel (n_1 = 30 and n_2 = 62); and Snow Petrel (n_2 = 32).

					Percent correctly classified							
	Estimates for discriminant formula				Northern Fulmar	Southern Fulmar		Antarctic Petrel		Cape Petrel		Snow Petrel
Formula	HL	BD	TL	CL	1	1	2	1	2	1	2	2
GEN1	1				95	89	86	100	79	80	84	91
GEN2	1	2.29			96	93	89	100	83	84	84	91
GEN3	1	2.61	0.48		96	96	91	100	83	87	86	88
GEN4	1	2.38	0.41	−0.21	97	96	91	100	83	87	87	91

species; Table 1) were used in the calculation of a generalized multispecies discriminant for fulmarine petrels. Table 4 shows the general discriminant for the selected combination of four main characters (GEN4), as well as for combinations using only some of these characters. Performance of the generalized discriminants was tested on samples of all species. Cut points were calculated as the midpoint of the interval between group means of scores of known males and females in the tested sample. Values for these cut points are not shown in the table as they are specific for the sample and not for a species.

Sex discrimination by the GEN4 formula, with cut points calculated from sexed birds, equals the performance of population-specific formulas. Tests of performance of GEN4 (Table 4) show results almost equal to those of self tests, and often better than those of jackknife tests for population-specific discriminants (Table 3). A theoretical problem with test results from Table 4 is that generalized formulas were derived from the tested samples. As each sample is only partly responsible for the generalized discriminant, the effects are expected to be small. However, to validate our approach, independent tests were done on Northern Fulmars. Because two samples of Northern Fulmars were available, one of them could be omitted from the calculations of a generalized discriminant without loss of the morphological variability of all five species. A generalized discriminant for fulmarine petrels was calculated without using the sample of Dutch Northern Fulmars, so only using five samples and species with 132 sexed birds. The resulting formula was similar to GEN4 and in a performance test, 96% of Dutch Northern Ful-

mars were classified correctly as to sex. Similarly, a generalized formula calculated without the use of the Northern Fulmars of Jan Mayen provided 97% correct classifications when tested on the independent sample of Jan Mayen birds.

The increase in performance between just measuring head length (GEN1) and the more complicated use of GEN4 may seem small. Indeed, head length would be sufficient to give a general impression of relative numbers of males and females in many samples. However, many studies require accurate sex assignments for as many individuals as possible, with the least possible doubt in each individual. Formulas using more characters meet both demands by increasing performance and by enhancing bimodality in frequency distributions (Fig. 2). We recommend the use of formula GEN4.

Generalized formulas in Table 4 were constructed using the samples of birds sexed by dissection. When using the larger samples including birds sexed by observations, the formulas looked very similar and showed only minor differences in test results.

Cut points without reference to sexed birds.—It was our intention to provide a method for discriminating sexes in samples from unknown populations without reference to sexed birds. Therefore, we calculated cut points in GEN4 scores for all samples by means of the expectation-maximization-algorithm procedure. The resulting cut points differed only slightly from those calculated from scores of known males and females. The performance of the combination GEN4 formula with the cut point based on this algorithm is shown in Table 5 and compared to that of population-specific discrimi-

TABLE 5. Performance of GEN4 discriminant with population-specific cut points calculated without reference to sexed birds. Percentages of correct sex assignments compared with those for population-specific discriminants with cut points derived from sexed birds (Table 3). When two samples mentioned under same heading, the first refers to birds sexed by dissection and the second to all birds sexed by either dissection or observation.

Species (n)	Percent correctly classified by	
	Population-specific discriminant[a]	GEN4[b]
Northern Fulmar NL (247)	98	93
Northern Fulmar JM (32)	97	97
Southern Fulmar (27, 130)	96, 89	96, 91
Antarctic Petrel (11, 77)	82, 82	100, 84
Cape Petrel (30, 62)	83, 81	90, 87
Snow Petrel (32)	91	88

[a] Jackknife test (Table 3).
[b] Cut points calculated using expectation-maximization algorithm.

nants. Calculating cut points without reference to sexed birds sometimes caused a decrease in performance as compared to population specific functions, especially so in Northern Fulmars from the Netherlands. However, performance improved in some other samples and overall reliability is still between 84 and 97% depending on species. The fact that unequal sex ratios (e.g. in the Jan Mayen birds) did not have a detrimental effect on cut-point calculation provides support for the use of the expectation-maximization algorithm when it is needed.

Concluding remarks.—Population-specific discriminant functions for several species of fulmarine petrels showed fluctuations in the weights given to certain characters. Certainly, each population or each species will have its own morphological properties with a specific weighting of characters and a specific cut point that optimally splits males and females. However, depending on the morphological variability of a species, one might need samples of sexed birds from several populations to arrive at an optimal species-specific discriminant function. Our samples were shown to be insufficient to develop optimal species-specific functions for fulmarine petrels. A species-specific function that has been derived from, and tested in one population only, may produce uncertain results in other populations.

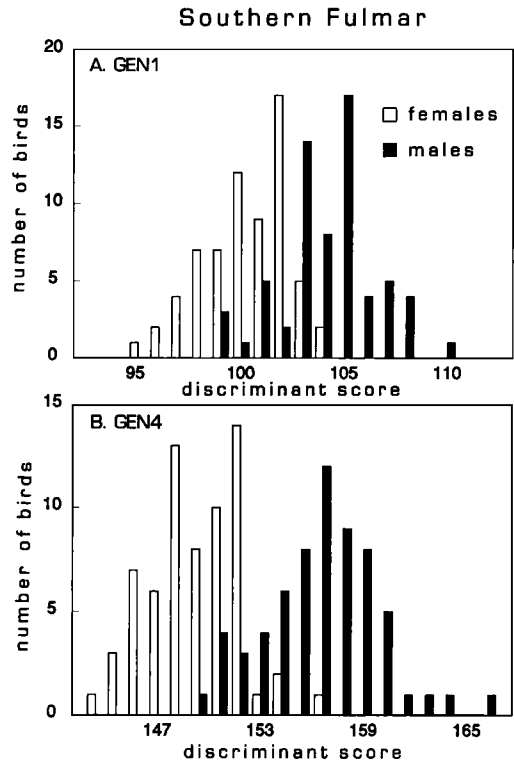By looking at similarities in morphology between different species, this problem could be

Fig. 2. Improved bimodality in frequency distributions of discriminant scores of Southern Fulmars when number of characters in function was increased: (A) GEN1 uses head length only; (B) GEN4 adds bill depth, tarsus length and bill length.

solved. Morphological variation between species is larger than variation within species. The fact that a generalized discriminant formula performed well in populations of five different species suggests that one would have similar predictive value for other populations of these species. By combining the generalized discriminant with the procedure involving the expectation-maximization algorithm to calculate cut points, samples of sexed birds from such populations are no longer essential.

It seems reasonable to assume a more or less constant level of sexual dimorphism within most species. In the two samples of Northern Fulmars from the Netherlands and Jan Mayen (both subspecies *F. glacialis auduboni*) similar levels of sexual dimorphism were found. According to Wynne-Edwards (1952), Northern Fulmars from Spitsbergen (*F. g. glacialis*) and the North Pacific Ocean (*F. g. rodgersii*) may be less dimorphic, but still considerably more so than the Southern

Fulmar. For Cape Petrels, Sagar (1986) constructed a sex discriminant for subspecies *D. capense australe* that assigned 82% of birds to the correct sex, which is similar to our findings for *D. c. capense* when using similar characters. Assuming that the sampled populations are more or less representative for levels of sexual dimorphism within each species, the GEN4 formula, combined with the cut-point calculation using the expectation-maximization algorithm, is likely to have a success rate of 85 to 90% for correct sex assignments in Antarctic Petrels, Cape Pigeons and Snow Petrels, as well as over 90% in Southern Fulmars and up to 97% in Northern Fulmars.

The remaining species of fulmarine petrels (*Macronectes* spp.) were not studied, but sexual dimorphism as described by Bourne and Warham (1966), Johnstone (1977) and Hunter (1984), suggests reliable results of generalized discriminants for Giant Petrels as well. Possibly, the generalized discriminant can be applied to other procellariid species when alternatives are lacking, but this needs further testing.

Evidently, our statement that the generalized discriminant can be applied to any population of fulmarine petrels with a predictable level of performance is only true when the sample used is from a homogeneous population. The sample from beached Dutch Northern Fulmars is largely of British (or southern North Atlantic) origin, but is heterogeneous in the sense that it contains an admixture of birds from High Arctic populations and of juveniles dying shortly after fledging. Both these groups differ in size and shape. As a result, the GEN4 procedure was less accurate than the population-specific discriminant (93 vs. 98% correct sex assignments). When dark-colored individuals, which definitely originate from High Arctic populations (van Franeker and Wattel 1982), and likely juveniles (as judged from anatomy and plumage development; van Franeker 1983) are removed from the Dutch sample, the generalized discriminant did better (96% correct sex assignments). In our analyses, we did not *a priori* remove specimens of this type from the Dutch sample because similar situations may occur when sampling unknown populations. In our opinion, the results indicate an acceptable level of error in such cases.

There may be exceptional situations of heterogeneous populations. We experienced this with Snow Petrels at Ardery Island (Windmill Islands, Wilkes Land, Antarctica), where two size morphs are breeding (Cowan 1981, van Franeker in prep.). The Snow Petrels around the nearby Casey Station (sample used in this paper) are all small-morph birds. In the case of Snow Petrels of Ardery Island, field observations of pair bonds, as well as the shape of the frequency distribution of discriminant scores from the GEN4 formula, clearly indicated that the sample was not homogeneous. As far as possible, such circumstantial evidence should be checked before applying any sex-discriminant analysis. A situation like the one on Ardery Island may be considered highly exceptional and does not invalidate the application of the generalized discriminant formula and cut-point calculation to unknown populations.

The approach of a generalized discriminant might be valuable for other bird groups as well. For example, there are several publications on population-specific discriminant functions for different species of gulls. Most can only be applied to the study populations from which they were derived; those authors that explicitly discuss the possible use of their discriminant in other populations warn that usefulness of their function has to be tested in each other population with a sample of birds of known sex (Fox et al. 1981, Nugent 1982). When comparing the various functions for gulls, it is remarkable that head length and bill depth are weighted in a ratio of: about 1:2.6 in the Herring Gull (*Larus argentatus*; Fox et al. 1981), 1:2.3 in the Black-backed Gull (*L. dominicanus*; Nugent 1982), and 1:1.9 in the Laughing Gull (*L. atricilla*; Hanners and Patton 1985). Apart from being similar to the *HL:BD* ratio in fulmarine petrels (function GEN2), the concurrence in ratio for different *Larus* gulls suggests that a generalized formula for this gull genus would be possible. Integrated analysis of the data in such cases could mean that the predictive value of existing discriminant functions could be extended to populations or species for which specific functions are not yet available.

Our approach does not replace optimized functions for local populations or species, but is offered as a practical solution for the many cases where sexed birds cannot be obtained and where tested population- or species-specific discriminant functions are not available. Our method can be used to develop species-specific discriminants by combining samples from different populations. Lack of multiple samples for

individual species of fulmarine petrels forced us to start on a multispecies level. Because the generalized discriminant has been tested on different species, it has predictable performance in populations of these species. When more data become available for the separate species, we aim at developing species-specific discriminants for the fulmarine petrels.

LITERATURE CITED

ANDERSON, A. 1975. A method of sexing Moorhens. Wildfowl 26:77–82.

BOERSMA, P. D., AND E. M. DAVIES. 1987. Sexing monomorphic birds by vent measurements. Auk 104:779–783.

BOURNE, W. R. P., AND J. WARHAM. 1966. Geographical variation in the giant petrels of the genus *Macronectes*. Ardea 54:45–67.

BRETAGNOLLE, V. 1990. Behavioural affinities of the Blue Petrel *Halobaena caerulea*. Ibis 132:102–105.

BROOKE, M. DE L. 1978. The dispersal of female Manx Shearwaters *Puffinus puffinus*. Ibis 120:545–551.

COPESTAKE, P. G., AND J. P. CROXALL. 1985. Aspects of the breeding biology of Wilson's Storm Petrel *Oceanites oceanicus* at Bird Island, South Georgia. Br. Antarct. Surv. Bull. 66:7–17.

COULSON, J. C., C. S. THOMAS, J. E. L. BUTTERFIELD, N. DUNCAN, P. MONAGHAN, AND C. SHEDDEN. 1983. The use of head and bill length to sex live gulls Laridae. Ibis 125:549–557.

COWAN, A. N. 1981. Size variation in the Snow Petrel (*Pagodroma nivea*). Notornis 28:169–188.

CRAMP, S., AND K. E. L. SIMMONS (Eds.). 1977. The birds of the Western Palearctic, vol. 1. Oxford Univ. Press, Oxford.

DIXON, W. J. 1985. BMDP statistical software. Univ. California Press, Berkeley.

DUNNET, G. M., AND A. ANDERSON. 1961. A method for sexing living Fulmars in the hand. Bird Study 8:119–126.

FOX, G. A., C. R. COOPER, AND J. P. RYDER. 1981. Predicting the sex of Herring Gulls by using external measurements. J. Field Ornithol. 52:1–19.

FRANK, R. E., W. F. MASSY, AND D. G. MORRISON. 1965. Bias in multiple discriminant analysis. J. Marketing Res. 2:250–258.

GALES, R. 1988. Sexing adult Blue Penguins by external measurements. Notornis 35:71–75.

GREEN, P. T. 1982. Sexing Rooks *Corvus frugilegus* by discriminant analysis. Ibis 124:320–324.

HAMER, K. C., AND R. W. FURNESS. 1991. Sexing Great Skuas *Catharacta skua* by discriminant analysis using external measurements. Ringing & Migr. 12: 16–22.

HANNERS, L. A., AND S. R. PATTON. 1985. Sexing Laughing Gulls using external measurements and discriminant analysis. J. Field Ornithol. 56:158–164.

HUNTER, S. 1984. Breeding biology and population dynamics of giant petrels *Macronectes* at South Georgia (Aves: Procellariiformes). J. Zool. (Lond.) 203:441–460.

JOHNSTONE, G. W. 1977. Comparative feeding ecology of the giant petrels *Macronectes giganteus* (Gmelin) and *M. halli* (Mathews). Pages 647–668 *in* Adaptations within Antarctic ecosystems (G. A. Llano, Ed.). Proceedings of the Third SCAR Symposium on Antarctic Biology, Houston, Texas.

JOHNSTONE, R. M., AND B. E. NIVEN. 1989. Sexing Grey-faced Petrels by discriminant analysis of measurements. Notornis 36:261–265.

LACHENBRUCH, P. A. 1975. Discriminant analysis. Hafner Press, New York.

MACDONALD, M. A. 1977. The prelaying exodus of the Fulmar *Fulmarus glacialis* (L). Ornis Scand. 8:33–37.

NUGENT, G. 1982. Sexing Black-backed Gulls from external measurements. Notornis 29:37–40.

OKILL, J. D., D. D. FRENCH, AND S. WANLESS. 1989. Sexing Red-throated Divers in Shetland. Ringing & Migr. 10:26–30.

PAYNE, R. W., ET AL. 1987. Genstat 5 reference manual. Clarendon Press, Oxford.

PINDER, R. 1966. The Cape Pigeon, *Daption capensis* Linnaeus, at Signy Island, South Orkney Islands. Br. Antarct. Surv. B. 8:19–47.

RAYENS, W., AND T. GREENE. 1991. Covariance pooling and stabilization for classification. Comp. Statist. Data Anal. 11:17–42.

RYDER, J. P. 1978. Sexing Ring-billed Gulls externally. Bird-Banding 49:218–222.

SAGAR, P. M. 1986. The sexual dimorphism of Snares Cape Pigeons (*Daption capense australe*). Notornis 33:259–263.

SCHNELL G. D., G. L. WORTHEN, AND M. E. DOUGLAS. 1985. Morphometric assessment of sexual dimorphism in skeletal elements of California Gulls. Condor 87:484–493.

SCOLARO, J. A., M. A. HALL, AND I. M. XIMENEZ. 1983. The Magellanic Penguin (*Spheniscus magellanicus*): Sexing adults by discriminant analysis of morphometric characters. Auk 100:221–224.

SERVENTY, D. L. 1956. A method of sexing petrels in field observations. Emu 56:213–214.

SHUGART, G. W. 1977. A method for externally sexing gulls. Bird-Banding 48:118–121.

SOKAL, R. R., AND F. J. ROHLF. 1981. Biometry: The principles and practice of statistics in biological research, 2nd ed. W. H. Freeman and Co., New York.

STONE, M. 1974. Cross-validatory choice and assessment of statistical predictions. J. R. Statist. Soc. B. 36:111–147.

TITTERINGTON, D. M., A. F. M. SMITH, AND U. E. MAKOV. 1985. Statistical analysis of finite mixture distributions. John Wiley and Sons, New York.

VAN FRANEKER, J. A. 1979. Some notes on the Fulmar in the Netherlands. Ardea 67:49–53.

VAN FRANEKER, J. A. 1983. Inwendig onderzoek aan zeevogels [Dissection of seabirds]. Nieuwsbrief NSO 4:144–167.

VAN FRANEKER, J. A. 1985. Plastic ingestion in the North Atlantic Fulmar. Mar. Pollut. Bull. 16:367–369.

VAN FRANEKER, J. A., AND J. WATTEL. 1982. Geographical variation of the Fulmar *Fulmarus glacialis* in the North Atlantic. Ardea 70:31–44.

VAN FRANEKER, J. A., K. CAMPHUYSEN, AND F. MEHLUM. 1986. Status over Jan Mayens Fugler [The status of the birds of Jan Mayen]. Var Fuglefauna 9:145–158.

VAN FRANEKER, J. A., P. J. BELL, AND T. L. MONTAGUE. 1990. Birds of Ardery and Odbert islands, Windmill Islands, Antarctica. Emu 90:74–80.

WILLIAMS, T. D. 1990. Annual variation in breeding biology of Gentoo Penguins, *Pygoscelis papua*, at Bird Island, South Georgia. J. Zool. (Lond.) 222:247–258.

WILLIAMS, T. D., AND J. P. CROXALL. 1991. Annual variation in breeding biology of Macaroni Penguins, *Eudyptes chrysolophus*, at Bird Island, South Georgia. J. Zool. (Lond.) 223:189–202.

WITT, H.-H., E. DE JUANA, AND J. M. VARELA. 1984. Geslechtsspezifische Grössenunterschiede in einer westmediterranen Population des Gelbschnabelsturmtauchers (*Calonectris diomedea*). Vogelwarte 32:234–236.

WYNNE-EDWARDS, V. C. 1952. Geographical variation in the bill of the Fulmar (*Fulmarus glacialis*). Scott. Nat. 64:84–101.

APPENDIX. Measurements of fulmarine petrels sexed by dissection, except for Snow Petrels, which were sexed based on observations. Measurements in millimeters for: (CL) bill length; (BD) bill depth; (HL) head length; (TL) tarsus length; (WL) wing length; (TA) tail length. For each sex, data presented as $\bar{x} \pm$ SD (n), range, and with t-value indicated for comparison between sexes. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

Northern Fulmars from The Netherlands.—CL: ♂ 40.7 ± 1.67 (117), 36.9–45.1; ♀ 37.4 ± 1.36 (130), 33.5–39.9; t = 16.87***. BD: ♂ 17.6 ± 0.75 (117), 15.7–19.4; ♀ 16.0 ± 0.70 (130), 14.5–17.6; t = 18.07***. HL: ♂ 98.8 ± 2.33 (117), 91–103; ♀ 91.6 ± 2.07 (130), 86–97; t = 25.71***. TL: ♂ 55.8 ± 1.85 (117), 50.5–61.0; ♀ 51.7 ± 1.72 (130), 43.0–55.5; t = 18.25***. WL: ♂ 337.3 ± 8.46 (97), 316–360; ♀ 324.9 ± 8.71 (117), 301–347; t = 10.51***. TL: ♂ 122.1 ± 5.40 (90), 109–139; ♀ 118.9 ± 5.15 (109), 104–133; t = 4.27***.

Northern Fulmars from Jan Mayen.—CL: ♂ 39.6 ± 1.68 (12), 37.0–42.2; ♀ 37.3 ± 1.23 (20), 34.8–39.5; t = 4.57***. BD: ♂ 17.8 ± 0.57 (12), 16.4–18.3; ♀ 16.0 ± 0.50 (20), 14.9–16.6; t = 9.07***. HL: ♂ 99.3 ± 1.87 (12), 95–102; ♀ 92.4 ± 1.79 (20), 89–96; t = 10.34***. TL: ♂ 56.0 ± 2.02 (12), 53.5–59.0; ♀ 52.5 ± 1.20 (20), 50.5–55.5; t = 6.25***. WL: ♂ 337.7 ± 9.65 (6), 325–348; ♀ 329.2 ± 7.91 (15), 317–344; t = 2.09*. TL: ♂ 125.3 ± 2.55 (9), 122–129; ♀ 120.3 ± 4.53 (19), 111–128; t = 3.07**.

Southern Fulmars from Ardery Island, Wilkes Land, Antarctica.—CL: ♂ 45.6 ± 1.36 (13), 43.0–47.6; ♀ 42.8 ± 1.71 (14), 38.9–45.5; t = 4.72***. BD: ♂ 16.4 ± 0.59 (13), 15.1–17.1; ♀ 15.3 ± 0.48 (14), 14.4–16.1; t = 5.05***. HL: ♂ 104.0 ± 1.78 (13), 101–107; ♀ 100.2 ± 1.97 (14), 95–103; t = 5.25***. TL: ♂ 54.3 ± 1.55 (13), 52.0–56.5; ♀ 51.9 ± 1.06 (14), 50.0–53.5; t = 4.91***. WL: ♂ 354.1 ± 9.85 (13), 331–365; ♀ 346.5 ± 5.75 (14), 337–357; t = 2.47*. TL: ♂ 135.2 ± 4.88 (13), 125–143; ♀ 133.4 ± 3.27 (14), 127–140; t = 1.13.

Antarctic Petrels from Ardery Island, Wilkes Land, Antarctica.—CL: ♂ 37.4 ± 0.98 (6), 35.9–38.9; ♀ 35.4 ± 1.38 (5), 33.8–37.0; t = 2.79*. BD: ♂ 13.9 ± 0.47 (6), 13.5–14.8; ♀ 12.9 ± 0.53 (5), 12.5–13.8; t = 3.24*. HL: ♂ 96.7 ± 1.86 (6), 95–99; ♀ 92.8 ± 1.64 (5), 90–94; t = 3.61**. TL: ♂ 46.7 ± 1.51 (6), 44.5–48.5; ♀ 45.1 ± 1.56 (5), 54.5–47.0; t = 1.70. WL: ♂ 334.8 ± 7.83 (6), 327–348; ♀ 328.8 ± 5.26 (5), 322–335; t = 1.45 TL: ♂ 121.8 ± 4.79 (6), 116–129; ♀ 121.0 ± 6.48 (5), 114–127; t = 0.24.

Cape Petrels from Ardery Island, Wilkes Land, Antarctica.—CL: ♂ 31.7 ± 1.05 (19), 30.0–33.4; ♀ 30.2 ± 1.09 (11), 28.3–31.7; t = 3.59**. BD: ♂ 10.7 ± 0.42 (19), 9.7–11.5; ♀ 10.0 ± 0.42 (11), 9.4–10.8; t = 4.10***. HL: ♂ 81.6 ± 1.71 (19), 79–84; ♀ 78.4 ± 1.75 (11), 75–81; t = 4.93***. TL: ♂ 47.1 ± 1.50 (19), 44.5–50.5; ♀ 45.7 ± 1.45 (11), 43.0–48.5; t = 2.58*. WL: ♂ 283.6 ± 4.45 (19), 273–292; ♀ 278.3 ± 6.96 (11), 266–288; t = 2.55*. TL: ♂ 107.8 ± 2.79 (19), 102–112; ♀ 106.0 ± 3.10 (11), 103–111; t = 1.63.

Snow Petrels from Casey Station, Wilkes Land, Antarctica.—CL: ♂ 21.4 ± 0.49 (15), 20.4–22.1; ♀ 19.7 ± 0.83 (17), 18.2–21.2; t = 7.00***. BD: ♂ 9.3 ± 0.38 (15), 8.7–10.1; ♀ 8.6 ± 0.39 (17), 7.8–9.5; t = 5.05***. HL: ♂ 72.2 ± 1.32 (15), 70–76; ♀ 68.7 ± 1.93 (17), 64–72; t = 5.89***. TL: ♂ 34.4 ± 0.83 (15), 33.0–35.5; ♀ 33.9 ± 1.46 (17), 30.5–37.0; t = 1.15. WL: ♂ 273.9 ± 6.35 (15), 263–283; ♀ 264.4 ± 7.75 (17), 251–277; t = 3.76***. TL ♂ 116.5 ± 5.01 (15), 107–123; ♀ 115.0 ± 4.47 (16), 110–125; t = 0.88.