



A computer method for resolving mixed normal distributions

JOHN H. MORGAN

Chemin de Laval, F-11160 Cabrespine, France

Biometric data collected by bird ringers can often be a mixture of two normal (ie Gaussian) distributions, for example, due to sexual size dimorphism, population differences, or species differences. Here, I present a computer-aided solution for identifying the means, standard deviations and mixing proportion of two groups where the data are mixed. When distributions are unimodal and no other information exists, the program tests the two-component model generated *versus* a normal distribution. Further developments are possible and the program is made freely available for this purpose. When the method was applied to published wing length data used to sex birds, means were closely comparable with values deduced from sexing by dissection. For most wing length distributions examined, a two-component model is a better descriptor of the observed frequencies than a normal distribution. Statistical testing of the separate components by parametric methods will be efficient and more powerful than testing the unresolved distribution.

Biometrics, such as wing length and body weight, have proved useful for identifying groups within mixed populations. Examples include distinguishing the sexes in species where males and females have similar plumage characteristics (eg van Balen 1967, Follows 1969, Thorne 1975, Reynolds 1978, Larsson & Strandvik 1986, Bishop & Emmens 1991, Daunert 1993, Flury *et al* 1993, Dougall 1997, Scebba 2001); identifying similar species, where they occur together (eg Craik & Harvey 1984, Walinder *et al* 1988); and identifying distinct races which mix in the non-breeding season or on migration (eg Thake 1981, Summers *et al* 1988a, Hope Jones 1990, 1992, Shirihai & Colston 1992). Additionally, separating the component distributions of the sub groups reduces standard errors, making comparisons between data sets easier to interpret.

A simple method of resolving inter-mixed normal distributions is to plot cumulative frequencies of classed data, either as normal equivalent deviates (Sokal & Rohlf 1995) or on probability paper (Harding 1949, see also Bhattacharya 1967). The measure of a normally distributed trait will give a straight line when plotting the trait size against the cumulative total on probability paper; a mixture of two distributions, when mixed and plotted as one, will give a sigmoid curve plot (the probability curve, Fig 1). One or more points of inflexion in the curve indicate a transition from one mode to the next. Graphical methods suffer from inaccuracy when assessed by eye, so in this paper I describe an extension to the method, using a computer program developed to resolve cumulative frequencies into two

components. In the past, this method has been applied to bird wing length data (Griffiths 1968, Morgan 1972, Dougal 1997, Scebba 2001), which provides a convenient paradigm for discussing the method.

Starting with approximate parameters for the mixture distribution, the program finds a best fitting cumulative frequency plot by assuming a mixture of two components, each having a normal distribution. From the program, the mean, standard deviation and population proportion for each component can be estimated. To verify its operation, some data from the literature are re-examined (Reynolds 1978, Da Prato & Da Prato 1978, Dougall 1997).

METHODS

The computer program

For a two-component, normal mixture there are five parameters describing the probability curve: mean and standard deviation for each component, and the mixing proportion. The expected cumulative frequency distribution of the mixture is found by adding components together. The computer program generates parameter values for two subgroups, then calculates an expected distribution to test against that observed. The fit is improved iteratively, by incrementing parameter values and re-testing against the observed data. This continues to either a limit set in the number of iterations, or until an index of fit reaches a point programmed by the operator.

Email: jmorgan@ataraxie.fr

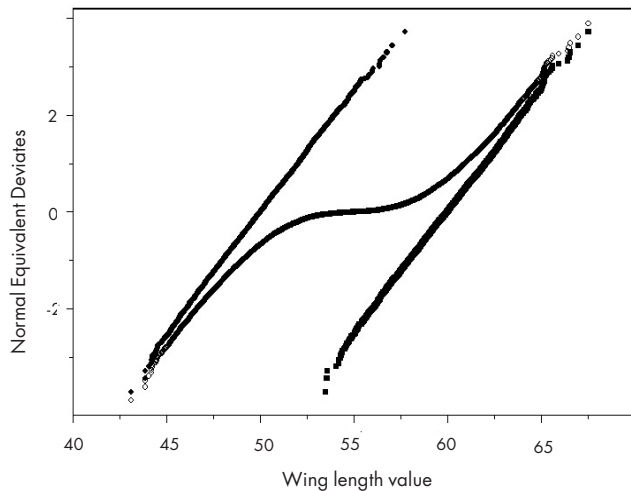


Figure 1. Normal quantile plot for randomly generated, normally distributed data. The two straight lines are data for wing lengths with means of 50 and 60 respectively (SD = 2 and N = 5,000, for both). The sigmoid curve is the plot for the combined 10,000 wing lengths when they are treated as one population. Normal equivalent deviates represent the number of standard deviations away from the mean value.

This empirical process may have some advantages over analytical methods (Everitt & Hand 1981, Titterton *et al* 1985), because of the ease with which certain criteria can be imposed by the operator *viz* the use of fixed values for individual parameters and inter-linking variances to make them covary. When used in the program, these will subsequently be reflected in the output. Reducing the number of parameters estimated from the data releases additional degrees of freedom for subsequent significance testing.

The computer program is written in Microsoft "QBasic". The copyrighted program can be obtained free from the author. An email address should be provided to which files can be sent. Some MS-DOS executable files needed to run it could also be provided. Detailed notes on program use accompany the listing. The author would be happy to see any developments of this software, providing that it remains the property of the public domain under a GNU General Public Licence.

Program operation

Before and during a program run, certain steps are carried out by the analyst: examination of the data for errors and inconsistencies; assessment for goodness-of-fit to a normal distribution; resolution using the graphical method (Fig 1) or otherwise obtaining initialising estimates; selection of a convergence criterion (CC) from one of four indices provided for this purpose; monitoring the closeness with which expected values approach observed values. These procedures are detailed in the notes supplied with the program.

If any parameters are to be constrained, an *a priori* decision based on available biological knowledge of the situation will have to be made and justified. For example, it is possible to set the mixing proportion at 50%, as usually anticipated for sex ratios. In addition, the fractional standard deviation FSD (coefficient of variation/100) for one component may be set as a ratio of the other. It seems reasonable to expect this ratio to be close to or equal to unity for male versus female biometric measurements. A user-variable ratio, rather than 1.0, is preferred because the FSD for wing length can be up to 4% higher in males than in females (Morgan & Shirihai 1997).

Incorporating such criteria into maximum likelihood methods for mixtures might seem to be a preferable alternative. It does not appear to have been developed, at least for morphometric data. (S N Freeman personal communication). Additionally, procedures for determining and correcting bias in analytical estimates, which often turn out to be intractable problems (Everitt & Hand 1981), will not be required when heuristic methods are used.

Artificial two-component distributions derived from standard normal tables took between 200 and 800 recursions to solve. Therefore, it is expected that the lowest index value encountered by the program during 500 - 1,000 cycles of program recursion will be close to a global minimum.

Once a set of minimising parameters has been established for any data set, the extent to which a sample fits the mixture model must be assessed. Output comprises parameter values, the expected frequency distribution, significance tests for two-parameter versus five-parameter models and an index derived from the Kolmogorov-Smirnov D-test for goodness-of-fit (D-fit, see Appendix).

An observed set not significantly different from normal may be compared to the two-component distribution by a likelihood ratio test. Any set showing probabilities greater than 0.1 may be disregarded, unless literature suggests that size dimorphism is present. Inspection of the expected distribution and its residuals from observation sometimes suggests further analysis and should always be carried out.

Comparison with published data sets

The program was used to resolve wing lengths from a series of datasets originally published to demonstrate the occurrence of sexual dimorphism. One set was used to identify a mix of subspecies. Some previously unpublished data are used to demonstrate how results from the program can be interpreted.

Autumn Chiffchaff *Phylloscopus collybita* in southern Britain
Reynolds (1978) divided the data into males and females by splitting the modes for wing length at the lowest point on a histogram. The means and standard deviations found, with proportion of the larger component fixed at 0.555, as implied by Reynolds (1978), are compared with published values in Table 1. The estimates given here should be more realistic than those in the original paper. The distribution fits very closely (D-fit = 0.20).

Lesser Redpoll *Carduelis flammea* in Scotland

In this study, Da Prato & Da Prato (1978) sexed birds by plumage characteristics. Unbiased estimates for the five parameters were therefore available to compare with program output (Table 1). Because parameter estimates given by the authors do not match the data in their histogram, the latter have been assumed correct here. D-fit only varies slightly for a wide range of parameter values due to the low number of individuals measured and the closeness of the means for each sex. The closest fit (D-fit = 0.15) was obtained with mixing proportion fixed at the implied published value (0.565). The recalculated means were similar to those presented in the original paper, but the standard deviations were larger.

Skylark *Alauda arvensis* in Scotland

The Skylark wing lengths measured by Dougall (1997) had been divided into males and females using a graphical plot. It is assumed that the author obtained means and standard deviations by fitting two population lines to the graph. Parameter values given or implied by Dougall generate a two component model that fits the data (D-fit = 0.50). Running the program with these as starting values gave a very close fit (D-fit = 0.25).

The revised parameters give different 95% confidence intervals to those in the original paper; that for males is wider, for females narrower. This generates a different overlap region for the sexes.

Spring Sand Martins *Riparia riparia* in Eilat, Israel

Data, collected during 1984-95 at the International Birdwatching Centre, Eilat (29° 5'N 0° 30'E), show the existence of a discrete subspecies *Riparia riparia eilata*. It is almost unknown other than on migration in late February to early March at Eilat (Shirihai & Colston 1992). There is very little size overlap with longer-winged subspecies that migrate through Eilat. The program gave close agreement with published values (Table 2).

Autumn Sedge Warblers *Acrocephalus schoenobaenus* in Dorset, UK

Data from the Radipole *Acrocephalus* Ringing Group (G Pepler personal communication) and Christchurch Harbour Ringing Station (Morgan 1972) were combined for first calendar year birds (n = 3,965), tested for normality (D-fit = 1.04, P < 0.01) and resolved. Observed data fitted the expected distribution very closely (Table 3, lines 1 and 2 : D-fit = 0.26, D-fit = 0.31 with equal FSDs). Coefficients of variation calculated for wing length means appear to be lower than any reported for passerines. This could not be checked exhaustively in the literature.

A poor fit (D-fit = 0.69) was obtained when the same data were resolved with male proportion fixed at 0.500 (Table 3, line 3). In view of this large change in degree of fit, an examination was made of class frequency residuals, and program runs were made with mixing proportions 0.550 to 0.650, FSDs set equal, male set to +/- 5% female FSDs, and switching between CCs.

Table 1. Comparison of published results with the results obtained from the computer based reanalysis of the data. Data are wing lengths (mm), presented as means with standard deviations (SD) and coefficients of variation (CV).

	Males			Proportion male	Females		
	Mean	SD	CV		Mean	SD	CV
Chiffchaff n = 273 (Reynolds 1978)							
Original analysis	62.1	1.39	2.2	0.555	57.2	1.22	2.1
Present method	62.18	1.28	2.1	0.555 (fixed)	57.1	1.08	1.9
Lesser Redpoll n = 214 (Da Prato 1978)							
Original analysis	71.33	1.27	1.8	0.565	68.59	1.52	2.2
Present method	71.39	1.17	1.6	0.565 (fixed)	68.56	1.32	1.9
Skylark n = 247 (Dougall 1996)							
Original analysis	114.3	2.26	2.0	0.490 (implied)	105.0	3.09	2.9
Present method	114.1	2.67	2.3	0.545	104.6	2.64	2.5

Table 2. Wing length data for Sand Martins ($n = 3,961$) trapped on migration through Eilat, in late February to early March (IBC, Eilat). Proportion is the proportion of the population which is of the subspecies *Riparia riparia eilata*. * denotes mid-range, the corresponding SD is estimated from the authors' data using "Rankits". CV is the coefficient of variation.

	<i>Riparia riparia eilata</i>			Proportion	Other subspecies		
	Mean	SD	CV		Mean	SD	CV
Present method	93.35	2.42	2.6	0.070	107.55	3.22	3.0
Original analysis	93.5*	2.2					

Table 3. Wing lengths of juvenile Sedge Warblers ($n = 3,965$) trapped in Dorset, UK, in 1971. * denotes 95% confidence interval for wing length. CV = 1.82 to 1.94. See text for details.

'Set = Male' is where the female FSD is kept equal to that of the males.

	Mean	Males		Proportion	Mean	Females	
		SD	CV			SD	CV
Original analysis	64.89	1.27	2.0	0.700	62.82	1.09	1.7
Original analysis	65.01	1.22	1.9*	0.630	62.99	1.18 (set = Male)	
Present analysis	65.15	1.31	2.0	0.500 (fixed)	63.41	1.28 (set = Male)	
Present analysis	64.94	1.29	1.9	0.650	62.98	1.16	1.9
Present analysis	65.22	1.26	2.0	0.500 (fixed)	63.30	1.22 (set = Male)	

In testing the effects of these constraints, residual comparisons were always largest in classes 64 mm, 65 mm and 67 mm. There were, in general, around 4% too few, 5% too many and 12% too few observations respectively. To gain insight, transferring 5% of readings from the 65 mm to the 64 mm class and transferring excess of observed over expected frequency from each of 66 mm and 68 mm to the 67 mm class formed an artificial data set. The reason for this set of choices is discussed below. A normal distribution with mean and standard deviation calculated from the adjusted data was a 'fit' (D-fit = 0.45), although rejected by a likelihood ratio test ($G = 22.36$, $df = 9$, $P < 0.01$).

Resolving these artificial data into two components gave very close fits, though parameters showed only small differences when compared to those from unadjusted data (Table 3, lines 4 and 5 compared with lines 1 and 2: D-fit = 0.15). A likelihood ratio test differentiated the two-component mixture model from a normal distribution ($G = 16.50$, $df = 4$, $P < 0.005$). Close fits were obtained by using the adjusted data either with proportion fixed at 0.5 (D-fit = 0.23) or with FSDs equal (D-fit = 0.36).

Wing lengths of Chiffchaff *Phylloscopus collybita* in Eilat, Israel

A bimodal data set of 7,817 measurements (13 springs, 12 autumns) was resolved into two components (Table 4). Close fits for the whole set (D-fit = 0.29) and for a spring subset (D-fit = 0.32) were found. Setting fractional standard deviations equal had no meaningful effect (D-fit = 0.34).

Wing lengths of Eastern Bonelli's Warbler *Phylloscopus orientalis* in Eilat, Israel

This species shows less size sexual dimorphism than Chiffchaff. The data set of 624 wing lengths measured on fully-grown birds in spring is unimodal but does not fit a normal distribution (D-fit = 1.57). The components for males and females (Table 4) gave a distribution that the data fitted very closely (D-fit = 0.21).

Wing lengths of Tree Pipit *Anthus trivialis* in Eilat, Israel

Tree Pipit has size sexual dimorphism (Cramp 1988), but a normal quantile plot of 1,276 spring birds does not diverge noticeably from a straight line, and a normal distribution derived from the sample parameters is a poor fit (D-fit = 0.69). The program gave a model that the data fitted very closely (D-fit = 0.29). The fit was equally close with proportion 0.5 and the FSDs set equal (Table 4). A likelihood ratio test failed to significantly differentiate either two-component model from normal ($G = 10.85$, $df = 8$, $P = 0.21$).

Wing lengths of Lesser Whitethroat *Sylvia curruca* in Eilat, Israel

Spring passage from East Africa is concentrated through the eastern Mediterranean. Subspecies identified at Eilat include eastern *S c blythi* and *S c halimodendri* as well as *Sylvia curruca curruca* and related forms from Europe (Shirihai 1996). The 5,802 wing lengths are a poor fit to a normal distribution (D-fit = 0.73). The computer program gave components that fitted the data very closely (D-fit = 0.20). The fit was less close (D-fit = 0.40) with mixing proportion held at 0.5 and the

Table 4. Wing length data for selected migrant passerines at Eilat, Israel, 1984-1995. Data from IBC, Eilat

	Mean	Males SD	CV	Proportion	Mean	Females SD	CV
Chiffchaff (n = 7,816)							
	63.68	2.03	3.2	0.455	58.33	2.01	3.4
Male and female FSDs held equal	63.52	2.12	3.2	0.485	58.18	1.94	
spring only (n = 4,885)	63.29	1.93	3.1	0.485	57.81	1.80	3.1
Eastern Bonelli's Warbler (n = 624)							
	68.94	1.40	2.0	0.560	64.63	1.27	1.96
Tree Pipit (n = 1,276)							
	89.80	2.16	2.4	0.520	86.14	2.03	2.4
Male and female FSDs held equal	89.86	2.14	2.4	0.500(fixed)	86.22	2.05	
Lesser Whitethroat (n = 5,802)							
	66.85	1.85	2.7	0.490	65.62	1.62	2.5
Male and female FSDs held equal	66.70	1.73	2.7	0.500(fixed)	65.52	1.69	

FSDs set equal (Table 4). Likelihood ratio indicates both two-component models are preferable to normal ($G = 19.60$, $df = 6$, $P < 0.005$ and $G = 20.1$, $df = 8$, $P < 0.01$).

Possibilities for further program development

The implied precision of point estimates as in Griffiths (1968), Morgan (1972), Dougall (1997) and Scebba (2001) cannot be justified if sample sizes are too small. Results suggest that sample sizes averaging below 25 per class may not give useful estimates when the modes are close (Table 1). The solution of artificial sets having standard amounts of displaced data might clarify this problem.

Using the estimates of population proportion and standard deviations, approximate standard errors for the means can be obtained. Data sets generated by random sampling (Monte Carlo method) could be resolved and these resulting parameters used to yield standard errors by "bootstrap" techniques. Extensions to separate mixtures of multiple components, and with multiple variables, seem feasible.

DISCUSSION

General

Wing length can be used for gauging an individual bird's size, allowing body mass to be calibrated so that it more accurately reflects condition (Gosler *et al* 1995). Also, when size dimorphism between the sexes is considerable, some, perhaps even all, individuals can be sexed by a combination of measurements, including wing length

eg Cetti's Warbler *Cettia cetti* and Corn Bunting *Miliaria calandra*. Separation of male and female components in bimodal wing length distributions has been used frequently to estimate parameters, most recently by Scebba (2001).

Division into three components has been used to estimate racial composition in mixed assemblages. Summers *et al* (1988b) use discriminant analysis and maximum likelihood methods to obtain estimates of the proportion of Icelandic and British Redshanks *Tringa totanus* wintering in Britain, with parameters for each sex of each subspecies being known beforehand. Providing such fixed values for the parameters would permit evaluation of proportions from a program such as that used here.

When the means of the components are less than two standard deviations apart, the mixture will most likely be unimodal (Behboodan 1970). However, the most efficient statistical procedures for testing differences between estimated parameters, or for comparing distributions, rely on data being normal. Standard errors can be computed efficiently and without bias. It is therefore of great utility to be able to extract the parameters of normal distributions from mixtures, providing there exists a sound basis for doing so. Resolving data sets into male and female components would allow smaller sample sizes to be used to compare, say, annual cohorts of juvenile passerines in late summer.

Creating models that locate additional components of mixtures due to size difference from causes other than sexual dimorphism is possible. However, the size of the class interval in the wing length distributions analysed

(1 mm) will restrict the method because additional parameters will be estimated from the data. With four components to resolve, 11 degrees of freedom are lost; equivalent to 11 classes in the frequency distribution. Only collections with marked differences in size between the components will generate sufficient frequency classes for univariate analysis. Developing multivariate methods to incorporate additional morphometrics, ones that are known to vary significantly with age, sex and population origin would be more efficient.

Individual species

Because Sand Martin subspecies *R r eilata* is morphometrically distinct from other subspecies, the analysis here effectively gives parameter estimates from a larger sample than in the original paper (Table 2). It also allows identification of individuals that had been overlooked at the time of capture. It is possible to identify birds belonging to *R r eilata* with any desired level of certainty, using criteria based on the parameters generated by the program.

Results from the analysis of Sedge Warbler juveniles in Dorset (Table 3) highlights both the usefulness and the shortcomings for this type of analysis. Ringing recoveries indicate that annual cohorts gathering in southern Britain in late summer originate from a limited geographical area (Mead & Clark 1993) and biometrics should therefore be free from any effect due to the west-east cline (Cramp 1992). Adult Sedge Warbler males are significantly longer-winged than females (Cramp 1992) and it might be expected that some difference exists in juveniles of each sex (ie birds carrying their first set of primary feathers from hatching, see Alatolo *et al* 1984). The means deduced (Table 2) are thought to be due to differences in wing length, despite the possible presence of systematic error by measurers. Note that the difference between means is comparable with a non-significant value that follows from data given in Cramp (1992).

Systematic error occurring in a substantial proportion of the measurements will affect results. If it leads to mis-recording for data classes lying between component modes, estimates of the population proportion will be affected. This situation will be most acute if one of the modes happens to coincide with a potential source of error. Results will need to be confirmed by other means.

If the result of analysing the Sedge Warbler data truly reflects the situation in a cohort of juveniles, then interest focuses on the proportion of males. Blackcaps *Sylvia atricapilla* captured in Belgium (Herremans 1989) and Oxfordshire (Holloway & Edwards 1989) had a proportion of males greater than 0.5. Explanations were found in the trapping conditions. It seems unlikely that

differentiation could occur so early in the season among juvenile Sedge Warblers. It is possible that in the UK population an excess of males may exist at fledging, although further data are needed.

For the Eilat data (Table 4), division into male and female wing length means is clearly possible for a number of species. Other useful results were obtained (Morgan & Shirihai 1997), despite the presence of both age-related and geographical variation in size.

Chiffchaffs on migration at Eilat have a mixed age structure (first year and adult) and are known to originate from a wide geographical area (Shirihai 1996, Morgan & Shirihai 1997). Both attributes result in significant differences in wing length between the subgroups (Cramp 1992). Feather abrasion during the winter months leads to differences in wing length between birds on autumn and spring passages (pers obs).

Eastern Bonelli's Warbler wing lengths resolved readily into two components with identical coefficients of variation for each of the sexes. However, as the small sample size implies, parameter estimates are subject to a wide margin of error, and the only significant inference is that males are longer winged than females. Comparing and contrasting with a similar data set for Western Bonelli's Warbler *P bonelli* would clearly be of interest.

The non-significant deviation from a normal distribution of Tree Pipit data may be due to clinal variation in wing length overriding the effect of the bimodality. Also, the data comprise both adult and first year birds which are known to differ in mean wing length (Cramp 1988). Two means obtained from the program are separated by 3.6 mm, and this is comparable with a difference between sexes that follows from data given in Cramp (1988). Thus, despite the non-significant result of a Likelihood Ratio test, the two-component model is preferable.

For the Lesser Whitethroat data, the best fit was markedly better than a normal distribution. There is only slight sexual size dimorphism in this species, and what little there is could be obscured in a mixed population such as that occurring in Eilat. However, the difference between the two component means derived was identical to a difference that follows from data given in Cramp (1992), and together with the positive result of the Likelihood Ratio test makes the two-component model preferable.

Examining these results suggests that after resolving annual cohorts into components for each sex, the coefficient of variation for the wing length might be used to assess the extent of geographical variation present in migrant populations. Also, detecting significant differences between wing length means in successive annual cohorts is more likely after resolution

into sexes, because standard errors for resolved components will be less than for unresolved sets. In some cases, however, it will be necessary to ensure data sets are free from systematic error (Morgan 2004).

Conclusions

It is possible to resolve mixtures of two normally distributed components more precisely than can be achieved by a graphical method alone. For acceptable levels of precision in the results, the size of data sets needs to be large (500+) whenever there is overlap between measurements so that modes corresponding to each mean do not resolve in a graphical plot. Resolution into components reduces standard error for means compared to that for the unresolved data. Monte Carlo methods might usefully be employed to estimate standard errors if parameter estimates are thought to be too imprecise.

Whenever the difference between component means is close to the level of measurement precision, there is little improvement in fit over the simpler, two-parameter normal. For this and other reasons, searching the data for evidence of additional modes will only be productive if measuring techniques can be reformed to achieve increased precision. Development of multivariate methods could be possible also if bird ringers can collect accurate additional biometrics.

Resolving wing length sets into component classes for species showing size polymorphism can yield useful results. However, interpretation may not be straightforward if systematic error occurs during the gathering of the data. The protocols used for taking biometrics in the field appear to need revision before the method can give justifiable results in some cases.

ACKNOWLEDGMENTS

I particularly thank Jeremy Greenwood and Steven Freeman of the British Trust for Ornithology for very helpful discussion and criticisms, Eric Rexstadt of the University of Alaska at Fairbanks, Andrew Gosler of the Edward Grey Institute at Oxford, and those referees who read and made helpful comments on earlier drafts. The help of the IBCE, CHRS and RARG as sources of data for this publication are acknowledged.

REFERENCES

- Alatolo, R.V., Gustafsson, L. & Lundberg, A.** (1984) Why do young passerine birds have shorter wings than older birds? *Ibis* **126**, 410-415.
- Behboodjan, J.** (1970) On a mixture of normal distributions. *Biometrika* **57**, 215-217.
- Bhattacharya, C.G.** (1967) A simple method of resolution of a distribution into its Gaussian components. *Biometrics* **23**, 115-129.
- Bishop, J. & Emmens, R.** (1991) Sexing of Sedge Warblers using wing length. *Report of Rye Meads Ringing Group* **12**, 43-47.
- Craik, J.C.A. & Harvey, S.M.** (1984) Biometrics and colour forms of Common Terns and Arctic Terns. *Ringing & Migration* **5**, 40-48.
- Cramp, S.** (ed) (1988) *The Birds of the Western Palearctic*. Vol V. Oxford University Press, Oxford.
- Cramp, S.** (ed) (1992) *The Birds of the Western Palearctic*. Vol VI. Oxford University Press, Oxford.
- Da Prato, E.S. & Da Prato, S.R.D.** (1978) Criteria for sexing Lesser Redpolls. *Ringing & Migration* **2**, 48-49.
- Daunert, J.** (1993) Zur flugellange oberlausitzer Nachtigallen (*Luscinia megarhynchos*). *Zoologische Abhandlungen* (Dresden) **47**, 205-212.
- Dougall, T. W.** (1997) Biometrics and sex ratios of Skylarks *Alauda arvensis* in south-east Scotland. *Ringing & Migration* **18**, 37-49.
- Everitt, B.S. & Hand, D.J.** (1981) *Finite Mixture Distributions*. Chapman & Hall, London.
- Follows, G.** (1969) Weights and measurements of Corn Buntings. *Ringers' Bulletin* **3**, 11-12.
- Flury, B.D., Airoidi, J.-P. & Biber, J.-P.** (1993) Gender identification in Water Pipit (*Anthus spinoletta*) using mixtures of distributions. *Journal of Theoretical Biology* **158**, 465-480.
- Gosler, A.G., Greenwood, J.J.D., Baker, J.K. & King, J.R.** (1995) A comparison of wing length and primary length as size measures for small passerines: A report to the British Ringing Committee. *Ringing & Migration* **16**, 65-78.
- Griffiths, J.** (1968) Multimodal frequency distributions in bird populations. *Bird Study* **15**, 29-32.
- Harding, J.P.** (1949) The use of probability paper for the graphical analysis of polymodal frequency distributions. *Journal of the Marine Biological Association UK* **28**, 141-153.
- Herremans, M.** (1989) Habitat and sampling related bias in sex-ratio of trapped Blackcaps *Sylvia atricapilla*. *Ringing & Migration* **10**, 31-34.
- Holloway, G.J. & Edwards, P.J.** (1989) Biometrics, capture date and sex ratio in relation to migration in Blackcaps. *Ringing & Migration* **10**, 108-112.
- Hope Jones, P.** (1990) The occurrence of large ('northern') Razorbills in British and Irish waters. *Ringing & Migration* **11**, 105-110.
- Hope Jones, P.** (1992) Wing lengths and weights of spring Wheatears *Oenanthe oenanthe* at Bardsey, Gwynedd. *Ringing & Migration* **13**, 162-166.
- Larsson, L. & Strandvik, P.** (1986) Wing length as a character for sexing Rock Pipits *Anthus spinoletta littoralis*. *Ringing & Migration* **7**, 71-74.
- Mead, C.J. & Clark, J.A.** (1993) Report on bird ringing in Britain and Ireland for 1991. *Ringing & Migration* **14**, 1-72.
- Morgan, J.H.** (1972) Acrocephalus '71. *The Birds of Christchurch Harbour* **13**, 43-44.
- Morgan, J.H.** (2004) Remarks on the taking and recording of biometric measurements in bird ringing. *The Ring* **26**, 61-68.
- Morgan, J.H. & Shirihai, H.** (1997) Passerines and Passerine Migration at Eilat. *International Birdwatching Centre Eilat Technical Publications* **6**.
- Reynolds, A.** (1978) Chiffchaffs at Rye Meads. *Ringing & Migration* **2**, 38 - 41.
- Scebba, S.** (2001) Biometrics and sex ratios of Skylarks *Alauda arvensis* during migration in southern Italy. *Ringing & Migration* **20**, 364-370.
- Shirihai, H.** (1996) *The Birds of Israel*. Academic Press, London.
- Shirihai, H. & Colston, P.R.** (1992) A new race of the Sand Martin *Riparia riparia* from Israel. *Bulletin of the British Ornithologists Club* **112**, 129-132.

- Sokal, R.R. & Rohlf, F.J.** (1995) *Biometry*. Freeman, New York.
- Summers, R.W., Corse, C.J., Smith, N.R. & Whitfield, D.P.** (1988a) The biometrics and wintering area of Icelandic Purple Sandpipers. *Ringing & Migration* **9**, 133-138.
- Summers, R.W., Nicholl, M., Underhill, L.G. & Petersen, A.** (1988b) Methods for estimating the proportion of Icelandic and British Redshanks (*Tringa totanus*) in mixed populations wintering on British coasts. *Bird Study* **35**, 169-180.
- Svensson, L.** (1992) *Identification Guide to European Passerines*. 4th edn. Svensson, Stockholm.
- Thake, M.A.** (1981) A simple method for estimating the racial composition of a sample of migrants away from the breeding grounds. *Die Vogelwarte* **31**, 12-13.
- Thorne, C.J.R.** (1975) Wing length of Reed Warblers. *Wicken Fen Group Report* **7**, 10-13.
- Titterton, D.M., Smith, A.F.M. & Makov, U.E.** (1985) *Statistical Analysis of Finite Mixture Distributions*. Wiley, NY.
- van Balen, J.H.** (1967) The significance of variation in body weight and wing length in the Great Tit *Parus major*. *Ardea* **55**, 1-59.
- Walinder, G., Karlsson, L. & Persson, K.** (1988) A new method for separating Marsh Warblers *Acrocephalus palustris* from Reed Warblers *A. scirpaceus*. *Ringing & Migration* **9**, 55-62.

APPENDIX

The statistic D-fit is calculated from the Kolmogorov-Smirnov statistic, 'D', uncorrected for continuity, by addition of $1/2n$, then multiplying by the square root of n , the sample size. D-fit's value increases monotonically with increasing divergence of the sample from the hypothesised distribution. In terms of conventional hypothesis testing, a D-fit value of 0.886 or greater corresponds to a probability of 0.05 or less. D-fit becomes 'highly significant' at 1.031 (equivalent to $P = 0.01$). In this paper, values of D-fit between 0.9 and 0.6 are described as a 'poor fit' to the data, and in the absence of other evidence the model may be considered inappropriate. For values 0.6 to 0.4 the model is said to be a regular 'fit', 0.4 to 0.3 a 'close fit' and below 0.3 a 'very close fit'. Values above 0.9 are taken to indicate that the expected frequency distribution does not fit the data well, and an alternative model is called for. Actual probabilities may be found, if required, by consulting statistical tables.

(MS received 11 April 2002; MS accepted 1 August 2004)