Errors in large-scale shorebird counts

Rappoldt, C.; Kersten, Marcel; Smit, C.

Published in:
Ardea

Publication date:
1985

ERRORS IN LARGE-SCALE SHOREBIRD COUNTS

CEES RAPPOLDT[1]), MARCEL KERSTEN[2]) & COR SMIT[1])

[1]) Research Institute for Nature Management, P.O. Box 59, 1790 AB Den Burg, The Netherlands
[2]) Zoölogisch Laboratorium, Kerklaan 30, 9751 NN Haren, The Netherlands

Received 20 December 1983, revised 25 August 1984

CONTENTS

1. INTRODUCTION

During the past 15 years bird counts have been carried out regularly in the Wadden Sea area. Results were obtained by counting the waders, ducks and geese on sandbanks, islands and coastal salt marshes during high tide. The first counts were essential for getting more insight in number and distribution of shorebirds in the area, of which particular parts were threatened by land reclamation (Rooth 1966, Spaans 1967, Boere & Zegers 1974, 1975, 1976). A sufficiently large number of counts also permits the study of trends and fluctuations in the numbers of birds present. Especially for this last purpose knowledge of the accuracy of the counts is necessary. Justified conclusions can be more easily drawn and inconclusive data more easily recognized. Knowledge of the errors also permits a better planning of future counting activities.

In an earlier paper we described a variety of counting experiments with an explorative character (Kersten *et al.* 1981). In this paper we shall improve some of our methods and use them to calculate errors in large-scale counts.

It should be noted that the methods described refer to "good counts", in the sense that no rough guesses of important numbers of birds occur. Throughout the paper, the data refer to shorebirds, and the Dutch Wadden Sea serves as an example for a large-scale area. The methods, however, are more widely applicable and are described step by step in the successive chapters. Mathematical details are brought together in an appendix.

2. ERRORS IN COUNTS OF A SINGLE FLOCK

2.1. INTRODUCTION

In this chapter a method is described to quantify the stochastic and systematic error in the counting result for a single flock of birds. The stochastic error describes the variability of the counting result and is measured as a relative standard deviation (RSD, the ratio of standard deviation over mean). The systematic error is the difference between the mean of many counts and the actual number of birds present. It is the predictable part of the error.

It is well known that the results of different observers, counting the same flock of birds, differ from each other. In our previous paper (Kersten *et al.* 1981) we analysed the results of such "counting experiments" for sitting flocks of birds. The RSD appeared to be independent of flock size. No differences between bird species were found. Measuring this RSD value for a certain flock, counting conditions as vegetation type and distance are the same for all observers. We therefore call this type of RSD the within-situation RSD, abbreviated as RSD_w. It represents the stochastic error within a certain counting situation.

The mean counting result of many observers (the observers-mean) may vary according to factors as vegetation type, distance, light conditions or spatial distribution of birds in the flock. The detailed analysis of these separate influences would be very impractical. Together they lead to "fluctuations" in the observers-mean. These fluctuations are treated in this paper as random ones. The between-situation RSD_b is then formally defined as the relative standard deviation of the observers-mean within the range of "normally" occurring counting conditions.

If one observer counts a single flock, the result will differ from the (hypothetical) mean of many observers and will be influenced by the specific counting conditions. So RSD_w and RSD_b need to be combined to get the total stochastic error in the counting result for a single flock RSD_{flock}:

$$RSD_{flock} = \sqrt{(RSD_w^2 + RSD_b^2)} \quad (1)$$

2.2. FIELD METHOD

To measure the errors field experiments were carried out as follows. Several observers counted a flock of birds in the same way as they are used to do during large-scale counts. This took a few seconds for a small flying flock up to about 10 minutes for a large sitting flock. Then the real flock size N, the number of birds actually present, was determined. For flying flocks this could only be done photographically. For sitting flocks of shorebirds we often made use of the tidal migration from the roost back to the mudflats. This migration may proceed very slowly. During several hours some tens of birds or single individuals fly out of the flock. Such small numbers can be accurately counted (see 2.4.3). A flock of Curlews, for instance, was counted in this way by three observers. Their results were 1768, 1790 and 1796, illustrating the precision of this method.

The above procedure has been repeated for a number of flocks with different sizes.

2.3. CALCULATION OF THE ERRORS

The systematic error is described by an expression of the form:

$$C = \alpha N^{\beta} \quad (2)$$

The formula gives the counting result C, averaged over observers and situations as a function of the real flock size N. The systematic error is the difference between N and C. The parameters α and β are estimated from the observed real flock sizes N_i and the means C_i of the counting results for each flock. The datapoints ($^e\log N_i$, $^e\log C_i$) are used in a regression analysis. The points will generally lay around the regression line. This variability is interpreted as being the result of differences in counting situation. The RSD_b is then estimated from the residuals in the regression analysis (see appendix). Further, for each flock, the results of the individual observers will differ from the mean C_i. From these differences the within-situation RSD_w is estimated.

A difference in flock size will generally imply a difference in relative systematic error (the systematic error expressed as a percentage). The difference between a flock of 100 and a flock of 1000 birds may well be significant with respect to the (relative) systematic error. The difference between 9100 and 10.000, however, is unlikely to be equally important. The use of logarithmic scales for C and N reflects this idea.

2.4. RESULTS

2.4.1. Sitting shorebirds

In Table 1 the field results are given of several observers, each counting a flock of sitting shorebirds. Further the real flock sizes N are given and a value of RSD_w, calculated for each flock separately. The RSD_w does not depend on flock size (correlation coefficient −0.01, n = 11), and an overall value of RSD_w may be calculated. The result is 25%, which roughly corresponds to the value of 20% reported in Kersten *et al.* (1981). In Fig. 1 the mean counting result C is plotted as a function of N. The regression line is given by:

$$C = 1.24\, N^{0.952} \quad (3)$$

An F-test has been used to test whether or not this differs significantly from C = N. No significant difference exists ($F_{2,9} = 0.996$; $P = 0.41$). This implies that expression (3) may not be used

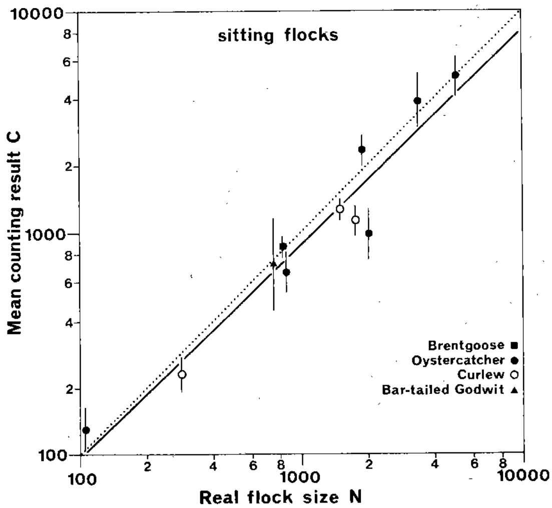

Fig. 1. The mean C of the counting results of several observers as a function of real flock size N for sitting flocks of shorebirds (data from Table 1). The regression line does not significantly differ from the dotted line C = N and may therefore not be used for the correction of counts. The lengths of the vertical lines around the points indicate the size of the within-situation RSD_w. The overall value for RSD_w is 25%. The between-situation RSD_b is 27% and the total stochastic error amounts to 37%.

Table 1. Real flock sizes N and counting results for flocks of sitting shorebirds. The real flock sizes have been determined making use of a slowly passing tidal migration back towards the foraging areas. Before the start of the migration the flock was counted on the high tide roost by a number of observers. The relative standard deviation calculated from the observers results for each flock is the within-situation RSD_w. The RSD_w does not depend on flock size. (Calculating the RSD_w values, not the numbers themselves but their natural logarithms were used (see appendix), which results in slightly different values)

Species	Real flock size N	Counting results of different observers						Within-situation RSD_w in %
Oystercatcher	105	110	115	125	182			23
Curlew	292	205	265					18
Bar-tailed Godwit	749	380	620	700	830	1400		47
Brentgoose	850	750	860	910	960			11
Oystercatcher	856	500	540	700	700	730	850	20
Curlew	1505	1170	1240	1430				10
Curlew	1784	945	1050	1185	1355			16
Oystercatcher	1891	2100	2110	2250	2300	3100		16
Oystercatcher	2009	600	940	960	1090	1200	1250	27
Oystercatcher	3400	3000	3400	4100	5700			28
Oystercatcher	5084	3700	4300	5150	5170	5550	6700	21

to estimate real flock sizes from counting results. We may conclude, however, that a possibly existing systematic error cannot be large. An error larger than several tens of percents would have been demonstrated by the data.

From the residuals in the regression analysis a between-situation RSD_b of 27% is found. The total stochastic error for sitting flocks then amounts to:

RSD_{flock} (sitting flocks) $= \sqrt{(25^2 + 27^2)} \approx 37\%$

If an observer counts, for instance, 1000 birds, this value for RSD_{flock} implies a stochastic error of 370.

Calculating the regression line it is assumed that also RSD_b is independent of N. A plot of the residuals against ${}^e\log N$, however, indicates a slight increase of RSD_b with number, leading to a value of RSD_{flock} of about 45% for N = 10,000. The difference with 37% is not significant and does not seem to be of any practical importance (see also chapter 5). Therefore the use of 37% for all flock sizes is recommended.

2.4.2. Flying shorebirds

In Table 2 and Fig. 2 the same analysis is made for flying flocks of waders and geese. As in the case of sitting flocks the regression line does not significantly differ from C = N ($F_{2,8} = 1.060$; $P = 0.39$) and the same discussion applies.

The within-situation RSD_w is 17%, which is of the same order of magnitude as the RSD_w for sitting flocks. The between-situation RSD_b appears to be very small in this case, only 3%. The

Table 2. Real flock sizes and counting results for flocks of flying shorebirds. The real flock sizes have been determined photographically. If more than one observer counted, a value for RSD_w had been calculated

Species	Real flock size N	Counting results of different observers						Within-situation RSD_w in %
Bar-tailed Godwit	60	55	70					17
Oystercatcher	63	64						—
Curlew	77	90						—
Dunlin	147	175						—
Oystercatcher	157	125	140					8
Oystercatcher	158	200						—
Brentgoose	411	450	480	500	500	550		7
Brentgoose	425	400	450	450	470	500	520	9
Bar-tailed Godwit	587	600						—
Brentgoose	850	500	580	850	900	950		29

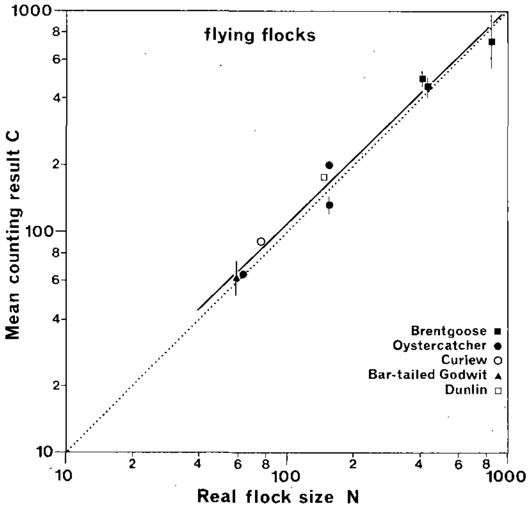

Fig. 2. The mean C of the counting results of several observers as a function of real flock size N for flying shorebirds (data from Table 2). Where no standard deviation is indicated, only one observer counted. The overall value of the RSD_w is 17%, the RSD_b is 3% and the total stochastic error RSD_{flock} amounts then to 17%. The regression line does not significantly differ from C = N.

total stochastic error amounts to:
RSD_{flock} (flying shorebirds) $= \sqrt{(17^2 + 3^2)} \approx$ 17%.

2.4.3. Starlings

As an illustration, we also applied the above described method to Starlings, counted during their daily flight to the sleeping roost. Han de Boer kindly put his unpublished data at our disposal.

Fig. 3 shows the mean result of three observers. The real flock sizes were determined photographically using a 27 DIN film. The counts are nearly exact for flock sizes less than about 50 birds. Above a flock size of 50, apparently another counting method is used, resulting in a strong underestimate. The difference with flying shorebirds might be due to the dusky conditions. The regression line is given by:

$$C = 1.079\ N^{0.902} \qquad (4)$$

This differs significantly from C = N ($F_{2,9}$ = 50.2; $P < 0.001$). The estimates for RSD_w and RSD_b are 18% and 10% respectively. These

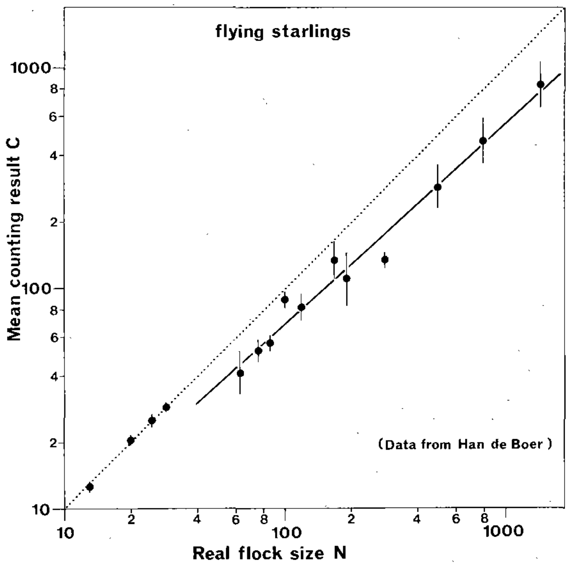

Fig. 3. The mean C of the counting results of several observers as a function of real flock size N for flocks of Starlings flying towards their sleeping roost. N was determined photographically. The regression line reflects a significant underestimate ($C = 1.079\ N^{0.902}$). The RSD values are RSD_w = 18%, RSD_b = 10% and RSD_{flock} = 21%. The counts of flocks smaller than about 50 birds appear to be very close to the real number.

Table 3. A numerical example illustrating the influence of different flock sizes on the stochastic error in a total counting result. The used value for RSD_{flock} is 37%. In the last column the relative errors are calculated using only the size of the large flock in the numerator of expression (8). The approximation is good enough, illustrating that relatively small flock sizes can be neglected in calculations of the stochastic error. In the last line of the table all flock sizes are equal. The RSD in the total count is very low in such cases

One large flock C_1	19 small flocks	Total count C_{tot}	Using (7) and (8)		Approximate $RSD(C_{tot})$
			ΔC_{tot}	$RSD(C_{tot})$	
10,000	—	10,000	3700	37	37
8000	105	10,000	3000	30	30
6000	211	10,000	2200	22	22
4000	316	10,000	1600	16	15
2000	421	10,000	1000	10	7
500	500	10,000	800	8	—

values are comparable with the results for flying shorebirds. The total stochastic error amounts to 21%.

2.4.4. Correction for the systematic error

The regression line for Starling counts may be used as a calibration line to correct counting results for systematic error. Expression (4) can be rewritten as:

$$N = 0.927\ C^{1.11} \qquad (5)$$

A counting result of 1000 birds (RDS = 21%, the above value for Starlings) leads to N = 1982. This figure, however, is merely an estimate. In the counting result and in the regression parameters stochastic errors are present. In the appendix is described how these errors can be combined to calculate the stochastic error for N. The result for the above example is 525, which is 26%.

3. ERRORS IN COUNTS OF MANY FLOCKS

3.1. ADDING FLOCK COUNTS

Suppose k flocks of birds are counted and the results are C_1, C_2, C_3, ... C_k. The stochastic errors $\triangle C_i$ (standard deviations) in these counts are then given by:

$$\triangle C_1 = RSD_1 \times C_1,\ \triangle C_2 = RSD_2 \times C_2, \ldots\ldots \triangle C_k = RSD_k \times C_k \qquad (6)$$

The total counting result is simply:

$$C_{tot} = C_1 + C_2 + C_3 + C_4 + \ldots\ldots\ C_k$$

For the stochastic error $\triangle C_{tot}$ in the total counting result holds:

$$(\triangle_{tot})^2 = (\triangle C_1)^2 + (\triangle C_2)^2 + (\triangle C_3)^2 + \ldots\ldots + (\triangle C_k)^2 \qquad (7)$$

Not the errors themselves, but their squares are added. Expression (7) is the basic rule for the stochastic error in the sum of independent counts. The RSD in the total counting result is defined as:

$$RSD(C_{tot}) = \triangle C_{tot}/C_{tot}$$

If the RSD_i in the flock counts C_i are all equal to RSD_{flock}, the full expression for the $RSD(C_{tot})$ becomes (making use of the expressions (6) and (7)):

$$RSD(C_{tot}) = \frac{\sqrt{(C_1^2 + C_2^2 + C_3^2 + \ldots\ldots + C_k^2)}}{C_{tot}} \times RSD_{flock} \qquad (8)$$

Since the fraction in the right part of this expression is always smaller than unity, the $RSD(C_{tot})$ is always smaller than the RSD_{flock}. The errors in the separate counts tend to neutralize each other. It can be shown that, for a certain number of flocks, the lowest value for $RSD(C_{tot})$ is reached when all flock sizes are equal. In that case $RSD(C_{tot}) = RSD_{flock}/\sqrt{k}$ (k equal flock sizes).

If the flocks strongly differ in size, an approximate RSD value may be calculated by using only the largest flock sizes in the numerator of expression (8). In Table 3 an illustrative numerical example is given. The sizes of the relatively small flocks hardly contribute to the error in the total number, which implies that they need not to be known. This property of stochastic errors is very useful for practical calculations. It also holds, if the smaller flocks are counted less accurately.

If a good calibration line is available, the systematic error in the total counting result can be calculated. For every flock C_i the real flock size N_i is estimated using the method described in section 2.4.4. Then the total number, corrected for the systematic error is simply:

$$N_{tot} = N_1 + N_2 + N_3 + \ldots\ldots\ldots + N_k.$$

Formally, speaking in terms of systematic error, this can be written as:

$$N_{tot} - C_{tot} = (N_1 - C_1) + (N_2 - C_2) + \ldots\ldots + (N_k - C_k).$$

Systematic errors are simply additive. It should be noted that the same correction formula is applied to all C_i. This implies that the N_i are not independent of each other. Therefore the stochastic errors $\triangle N_i$ (see 2.4.4) do not add up like the $\triangle C_i$ in expression (7). In the appendix a method is given to calculate the stochastic error in N_{tot}.

3.2. COUNTING BIRDS IN AN AREA

Counting the birds in an area is not necessarily the same as counting a number of well-separated single flocks. An observer has to find all birds, not only the flocks but also the dispersed individuals. Displacements have to be noted and discussed with nearby observers. These sources of error do not occur in the counts of single flocks. Errors associated with missing and displacements will be called area errors. The errors of a large-scale count may be calculated from the flock sizes only if the additional area errors

Table 4. The barrier islands Terschelling and Ameland have both been counted on two successive days. For species with mean counting result C in different ranges the table gives the mean of the observed RSD values. From reported flock sizes for the two islands (questionnaire, chapter 4) the RSD for the island counts can be calculated and for each of the species groups the mean result is given. For abundant species (C > 1000) the observed and calculated RSD values agree. For scarce species (10 < C < 100) the area errors dominate

	RSD Terschelling		RSD Ameland	
	observed	calculated from flock sizes	observed	calculated from flock sizes
10 < C < 100	59	—	78	—
100 < C < 1000	22	16	19	10
C > 1000	19	17	13	12

can be neglected. The validity of this assumption can be checked by counting the birds in an area two or several times, and by comparing then the measured RSD values with calculated ones.

For the Wadden Sea as a whole such a check is difficult to organize. The birds on individual barrier islands however can be easily counted on two successive days. The RSD values found are very inaccurate (only two counts can be compared). Therefore the results for different species are averaged. Table 4 gives such average RSD values for three different ranges of the mean counting result C. For relatively scarce species (10 < C < 100) the observed RSD values are large, even larger than 37%, the RSD_{flock} for sitting flocks.

For common species the RSD values are much lower and they are compared with expected values. The latter are calculated from the flock sizes for the two islands (see 4.2) and the results of chapter 2. For the appropriate sets of species (defined by the ranges of C) the average results are also given in Table 4. For C > 1000 the observed RSD values are only slightly larger. Thus, for abundant species the area errors hardly contribute to the island RSD. The island RSD is largely determined by the stochastic errors made during the counts of the flocks.

We interprete these results as follows. The area error which is most likely to occur is the missing of individuals or whole flocks. Such an error leads to a systematic underestimate of the real number and also to fluctuations in the counting result, since the degree of underestimate will not be constant. If this type of stochastic error dominates, the species is called scarce. The birds are easily missed. The stochastic and the systematic error are large and the counting result may just indicate an order of magnitude.

An abundant species will be counted partly as sparsely distributed individuals. This part of the count may be relatively inaccurate. For abundant species, however, the accuracy of the count is not affected by small, inaccurate numbers but it depends only on the count of the large flocks (see 3.1). This reasoning also suggests that, for abundant species (here: C > 1000), the systematic error is unlikely to be larger than the stochastic error for a single flock. Otherwise the fluctuations related to missing would cause an increase of the RSD value for an area (and the species would be scarce in the above mentioned sense).

4. ERRORS IN SHOREBIRD COUNTS OF THE DUTCH WADDEN SEA

4.1. INTRODUCTION

In this chapter stochastic errors are calculated for Wadden Sea counts of different species in different seasons. The systematic error could not be quantified accurately from our field data (see section 2.4.1 and 3.2). So corrections for Wadden Sea counts were not calculated (see also chapter 5).

The distribution of birds in the Wadden Sea during high tide is more or less constant from year to year, due to the constant pattern of tidal water movements and the stable distribution of foraging and resting areas. When errors are calculated for one count in a certain season, the results will be applicable to other counts as well. The flock sizes usually occurring in the Wadden Sea have been investigated by means of a questionnaire.

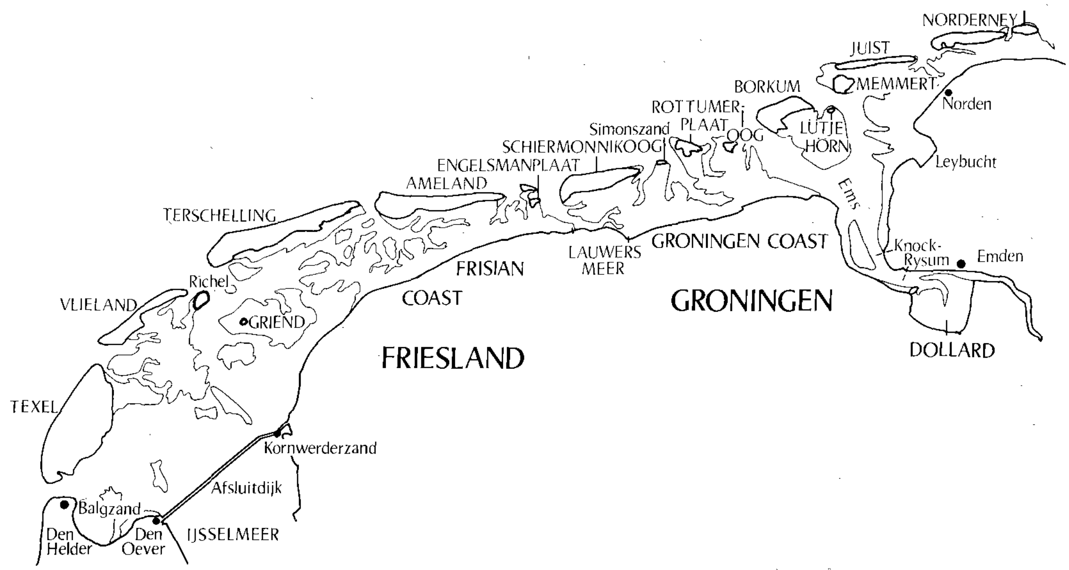

Fig. 4. The Dutch part of the Wadden Sea with the usual counting areas and some other names. About 100 observers are required for a shorebird count during high tide.

4.2. METHOD

The area is divided into its usual counting areas, viz. the 5 large islands, the 5 small ones and 4 parts of the mainland coast (Fig. 4). In each of these areas, high tide roosts are present with adjoining tidal flats. For the 14 counting areas a picture of the occurring flock sizes has been composed. To calculate errors for Wadden Sea counts the data are simply pooled.

Instead of using actually existing counts for the error calculation we constructed "standard counts" for three different seasons called late summer, winter and spring. Such a standard count gives, for all species considered, a characteristic distribution of numbers over the 14 counting areas. To construct the standard counts we used published Wadden Sea counts (Rooth 1966, Spaans 1967, Boere & Zegers 1974, 1975, 1977) and more recent unpublished ones (Zegers *in litt.*). The three seasons have been defined in a slightly different way for the different species in order to refer to interesting situations. The counts of most waders, for instance, show one or more peaks during August and September. For those species the late summer standard count has been defined as a regularly occurring peak value.

For each counting area, the three standard counts were presented to an experienced observer regularly counting in that particular area. We asked them to give, for each species, the sizes of the 5 largest flocks they usually have to deal with. We also asked them whether the flocks are normally counted in flight or sitting on the ground.

Of course the reported flock sizes cannot be accurate in any sense. No two counts of an area are identical. For a certain area, for instance, the standard count is 13,000 and the observer is asked which flock sizes occur. There may, for example, always be one flock of several thousands of birds (say 5000) and several flocks of many hundreds (up to 2000 say). In that case the count for the species considered may be characterized by the flock sizes 5000, 2000, 1000, 1000 and 500. The smaller flocks are not important for the error calculation.

Using the RSD_{flock} values for single sitting and flying flocks of shorebirds (see 2.4) and the results of chapter 3, the stochastic error in the standard counts can be calculated from the reported flock sizes.

4.3. RESULTS

In Table 5 calculated RSD values for shorebirds counts in the Dutch Wadden Sea are given together with the standard counts and the months to which these counts refer. The word "top" behind the month numbers does not mean a record number, but refers to the migration peak within the given period.

Pintail, Ringed Plover, Turnstone, Spotted Redshank, Greenshank and (in winter) Redshank occur in intermediate numbers (100—1000) in most of the 14 counting areas. This implies that area errors may be present causing the calculated RSD values to be too low (see section 3.2). In Table 5 this is indicated by the symbol >. Using, for instance, twice the value in the table, one still has a rough estimate of the stochastic error on Wadden Sea scale. When the word large is given instead of an RSD value, the standard counts in all counting areas are low and large unknown area errors will dominate.

The results for the abundant species are more interesting and will be discussed into further de-

Table 5. Calculated stochastic errors in counting results for the entire Dutch Wadden Sea, expressed as a percentage of the standard count C. The RSD values are calculated from reported flock sizes and the RSD values for single flocks of flying and sitting shorebirds. The symbol > means that additional area errors may be present. The word large indicates that area errors probably dominate. See text for remarks on the separate species

Species	Late summer			Winter			Spring		
	Months	Standard count C	RSD in %	Months	Standard count C	RSD in %	Months	Standard count C	RSD in %
Mallard	10/11	21,000	5	1/2	21,000	4	3/4	4500	6
Teal	10/11	19,000	12	1/2	5700	6	3/4	3000	15
Wigeon	10/11	117,000	6	1/2	55,000	7	3/4	21,000	10
Pintail	10/11	4800	>6	1/2	3500	>5	3/4	780	large
Shelduck	10/11	50,000	5	1/2	26,000	7	3/4	13,000	5
Brentgoose	11	29,000	6	1/2	12,000	11	4	52,000	9
Oystercatcher	9/10/11	228,000	5	1	167,000	5	4	52,000	6
Grey Plover	9/10 top	21,000	4	1	2600	large	5 top	24,000	5
Ringed Plover	9/10 top	4100	>5	1	low	large	5 top	2700	>9
Turnstone	7/8 top	5400	>7	1	2300	>8	4/5 top	2300	>9
Curlew	7/8 top	100,000	6	1	56,000	5	3/4 top	59,000	6
Bar-tailed Godwit	7/8 top	51,000	7	1	23,000	7	4/5 top	94,000	8
Redshank	7/8 top	46,000	8	1	7900	>5	3/4 top	14,000	5
Spotted Redshank	7/8 top	3100	>8	1	low	large	5 top	970	>6
Greenshank	8/9 top	6600	>6	1	low	large	5 top	690	large
Knot	8/9 top	52,000	9	1	41,000	15	4/5 top	35,000	14
Dunlin	8/9 top	291,000	5	1	95,000	4	4/5 top	184,000	5
Avocet	8/9 top	18,000	11	1	low	large	4	3900	4
Herring Gull	7/8/9	62,000	7	1	47,000	5	3/4	39,000	7

tail. Most Mallards are counted along the mainland coast in Groningen, in the Dollard area and on Texel. Flock sizes do generally not exceed 500 individuals (in spring 200), leading to RSD values around 5% for the total Wadden Sea number. In late summer and spring almost all Teals of the Dutch Wadden Sea are counted in the Dollard. And since half of the Dollard number is counted as a single flock, the RSD values are large. In winter, the Frisian Coast, Texel and Terschelling are also important (RSD = 6%). For Wigeon in late summer, 13 flocks have been reported between 3000 and 8000 in size (RSD = 6%). Most of these flocks occur along the Frisian Coast and on Terschelling. In winter this picture remains unchanged, although the numbers are smaller. In spring the error is determined by a few large flocks on Texel and in the Dollard area (RSD = 10%). Shelducks in late summer and winter also tend to concentrate in two counting areas. Flock sizes along the coast in Friesland and Groningen, however, do not exceed 2000. Many small flocks occur (or bands of floating birds along the coast, roughly described as a number of relatively small flocks). In spring, Shelducks are widely distributed along the Wadden Sea coast of mainland and islands (RSD = 5%). Brentgoose occurs in late summer in many of the counting areas (RSD = 6%). In winter, there are fewer important areas and the birds are counted in relatively large flocks (RSD = 11%). In spring a few very large flocks (5000—8000) along the Frisian Coast are responsable for the 9% RSD value. For Barnacle Goose no errors have been calculated. Its distribution is not stable. When necessary calculations can be carried out for the real counts.

Avocet occurs in late summer in large flocks in the Dollard area and along the Frisian Coast (RSD = 11%). In spring the flock sizes are much smaller (RSD = 4%). Knot is well known for its large flocks. In late summer Knots are present on most islands (RSD = 9%), but in winter and spring they only occur in the west part of the Wadden Sea. They are counted then as a few flocks that may exceed 10,000 in size (RSD = 15%).

For all other abundant waders (Oystercatcher, Grey Plover, Curlew, Bar-tailed Godwit, Redshank, Dunlin) and Herring Gull the RSD values are between 4 and 8 per cent. These species are all widely distributed in the Wadden Sea area and in the larger counting areas they mostly form more than a single large flock.

5. GENERAL DISCUSSION

The stochastic error of about 37% for a single sitting flock is probably larger than most observers realize. In spite of this large error a large-scale count of a common species may be accurate (in the sense of having a low RSD value). Compared to many other biological measurements an error of 5% to 10% is reasonable. Measurements of prey density, food intake or available foraging area are generally not more accurate. The level of abundance for which the single flock errors dominate will depend on the type of area and the counting procedure.

In the results for small study areas, the stochastic error can be of considerable importance. It may be worthwile in such cases to improve the accuracy by counting the birds in flight during tidal migration, by photographic methods or by counting the birds on their foraging areas.

The error calculations for Wadden Sea counts show that differences in RSD values can be ascribed to clearly recognizable properties of the flock size pattern. The calculated RSD values are only weakly dependent on the (correct or incorrect) details in standard counts and flock sizes. This implies that the results indeed may be applied to other, really existing counts.

Counts of a few individuals, spread out over an area, are generally extremely inaccurate. Also large systematic errors will be present in the results. Studying the ecology of abundant species, it may sometimes be worthwile to neglect the scarce species during the counts in order to preserve the accuracy of the interesting numbers.

The RSD values in this paper should not be interpreted as accurate ones, that may be used for critical statistical tests. Several reasons can be given: Differences between species with respect to RSD_{flock} have been neglected. The description of the counts to which the RSD values refer is subjective (normal counts, without rough guesses, carried out by normal observers). And the models used for error calculation contain simplifications. For instance, a small dependence of RSD_{flock} on flock size might exist (see also the appendix).

There is, however, a general justification for a rather rough treatment of stochastic errors. These errors are estimated only to permit a better interpretation of observations and they need not to be very accurate themselves. Accurate measurements of stochastic errors are generally a waste of time. For example, an RSD value of 37% gives an impression of the inaccuracy of counts of single sitting flocks. Whether this should be 25% or 50% is hardly interesting. Our picture of the counts would not be affected.

The above discussion does not apply to the systematic error. This error may be used for correction and it is generally worthwile to determine it accurately. Unfortunately we have not been able to quantify accurately the systematic error in counts of a single flock. That would require a much larger set of field data. If it exists, however, it is probably an underestimate of the real number with some tens of percents or less. The systematic error caused by missing of birds in an area forms a more serious problem. For the most abundant species of the Wadden Sea it was shown in section 3.2 that the systematic underestimate is unlikely to be larger than about 37% (the RSD value for single flocks). The underlying assumption is that missing of birds occurs irregularly and will therefore cause a large stochastic error as well. Further quantification of the missed fraction, however, is very difficult. It is probably easier to avoid the problem by improving the counting procedure. A subjective impression of the quality of the count remains important. A species as the Oystercatcher in the Wadden Sea is nearly completely counted as a number of large flocks on well known places. In such a case the systematic error in the single flock counts is probably the only one.

It should be noted that the use of corrected counts is limited to special cases, for instance, in the calculation of the total food intake by a population. Mostly, counts are used in a comparative way to study trends, fluctuations, differences etc. Then the stochastic error is the only relevant one.

The purpose of this paper has been to investigate the accuracy of shorebird counts. It seems practically impossible, however, to repeat a large-scale count of shorebirds. This is also true for a count of breeding birds in a large area or a longterm count of seabird migration along the coast. To get an impression of the accuracy, one has to develop a rough model to investigate how basic sources of error influence the final results. In our case of high tide counts, the errors for single flocks happened to be an important factor. This may also be true for other animals largely counted in groups. For non-colonial breeding birds a probability of missing and the use of binomial statistics might be useful. The use of a rough model of the counting procedure considerably limits the amount of field data that is required to get an impression of the accuracy of the observations.

6. ACKNOWLEDGEMENTS

The counting experiments in the field were carried out by Gerda Allersma, Han de Boer, Nellie van Brederode, Peter Brouwer, Arend van Dijk, Frans Kirkels, Theunis Piersma, Kees van Scharenburg, Jan Tinbergen, Joost Tinbergen, Carla Visser, Martin Wassink, Wim van de Zande and the authors. Financial support was given by Stichting Veth tot Steun aan het Waddenonderzoek. Hans van Biezen considerably improved the analysis of the field data on single flocks. The flock size questionnaire was filled out by Han de Boer, Nellie van Brederode, Sjoerd Dirksen, Meinte Engelmoer, Jan van 't Hof, Dick Jonkers, Gerwin Meyer, Meindert Otter, Piet Postma, Rob Schuckard, Jan Veen, Wim Weyman and Piet Zegers. Their patience with the often returning forms is highly appreciated. Piet Zegers also checked the standard counts. Han de Boer put his data on Starlings at our disposal and Dick Jonkers his double count of Terschelling. Mike Moser, Johan van Rhijn, Jelle Schilstra and Wim Wolff commented on the manuscript. We thank all these people for their help.

7. SUMMARY

Field experiments were carried out to examine the stochastic and systematic error in the counts of single flocks of birds. For flocks of sitting shorebirds the use of a relative stochastic error of 37% is recommended, independent of flock size. For flying flocks of shorebirds this is 17%. The systematic error in both cases could not be quantified accurately, but is smaller than some tens of per cents.

For abundant shorebird species the stochastic error in the result of a large-scale count can be calculated from the errors in the counts of the largest occurring flocks. The addition of counts of single flocks leads to a decrease of the (relative) stochastic error. Calculations for the Dutch Wadden Sea show a stochastic error of 5 to 10 per cent in the counts of abundant species. For these species, the systematic error caused by missing of birds is unlikely to be larger than the stochastic error for single flocks (about 37%).

For species which are common in one or a few parts of the Wadden Sea area only, the stochastic error may be somewhat larger than for the abundant species.

The counts of relatively scarce species, of which the individuals are widely spread over an area, will generally be very inaccurate. Large errors resulting from missing of bird during the count will dominate. The count may just indicate an order of magnitude.

8. REFERENCES

Boere, G. C. & P. M. Zegers. 1974. Wadvogeltelling in het Waddengebied in juli 1972. Limosa 47: 23—28.

Boere, G. C. & P. M. Zegers. 1975. Wadvogeltellingen in het Nederlandse Waddengebied in april en september 1973. Limosa 48: 74—81.

Boere, G. C. & P. M. Zegers. 1977. Watervogeltellingen in het Nederlandse Waddengebied in 1974 en 1975. Watervogels 2: 161—173.

Kersten, M., C. Rappoldt & C. Smit. 1981. Over de nauwkeurigheid van wadvogeltellingen. Limosa 54: 37—46.

Rooth, J. 1966. Vogeltelling in het hele Nederlandse Waddengebied in augustus 1963. Limosa 39: 175—181.

Spaans, A. L. 1967. Wadvogeltelling in het gehele Nederlandse Waddengebied in december 1966. Limosa 40: 206—215.

9. SAMENVATTING

Door middel van telexperimenten in het veld kunnen de toevallige en systematische fout in de aantalsopgave voor één groep vogels bepaald worden. Voor zittende groepen wadvogels bedraagt de toevallige fout 37% van de aantalsopgave. Deze fout is niet of slechts in geringe mate afhankelijk van de grootte van de groep vogels. Voor vliegende groepen wadvogels bedraagt de toevallige fout 17%. De systematische fout kon niet nauwkeurig gekwantificeerd worden, maar is niet groter dan enkele tientallen procenten.

Voor algemene wadvogelsoorten kan de toevallige fout in een grootschalige telling berekend worden uit de fouten in de tellingen van de grootste groepen in het gebied. Het optellen van aantalsopgaven voor afzonderlijke groepen leidt tot een vermindering van de (relatieve) toevallige fout. Voor het Nederlandse Waddengebied zijn berekeningen uitgevoerd, die betrekking hebben op karakteristieke aantallen en groepsgrootten in nazomer, winter respectievelijk voorjaar. Voor vrijwel alle soorten waarvan vele duizenden exemplaren in het gebied aanwezig zijn bedraagt de toevallige fout in het totaal getelde aantal 5 à 10%. De systematische ondertelling veroorzaakt door het missen van vogels ligt voor deze soorten vermoedelijk tussen de 0 en 37%. Voor Wintertaling, Kanoetstrandloper en Kluut is de toevallige fout groter (10 à 15%), doordat deze soorten slechts in enkele deelgebieden in groot aantal voorkomen.

Voor schaarse soorten, waarvan slechts enkele honderden exemplaren in het Waddengebied geteld worden, zijn de toevallige en systematische fout in het telresultaat veel groter. Dat wordt veroorzaakt door het missen van vogels. De tellingen zullen veelal slechts een indruk geven van de orde van grootte van het aantal aanwezige vogels.

10. APPENDIX

The used statistical methods are described below in a more formal way. Also a few assumptions and simplifications are explained, which were hard to present in the main text. Much of the background reasoning, however, can only be found in the above chapters, especially concerning the application to large-scale counts.

The model for single flock counting results

A regression model permits the estimation of the between-situation stochastic error, the overall stochastic error and the systematic error in counts of single flocks of birds. The model is given by:

$$Y_{ij} = \alpha' + \beta X_i + \epsilon_i + \varphi_{ij}. \qquad \text{(A1)}$$

The meaning of the symbols is as follows:

- i - The flock number, $i = 1 \ldots\ldots I$.
- j - Count number for flock i, $j = 1 \ldots\ldots n_i$, where n_i is the number of counts of flock i carried out by different observers.
- $Y_{ij} \equiv$ ${}^e\log C_{ij}$, where C_{ij} is the j-th counting result for flock i.
- $X_i \equiv$ ${}^e\log N_i$, where N_i is the number of birds actually present in flock i (real flock size).
- ϵ_i - The between-situation error terms, which are assumed to be independently and identically distributed as $N(0, \sigma_b^2)$. Differences in counting situation (light conditions, distance etc.) cause fluctuations in the counting result. Therefore σ_b^2 is called the between-situation variance.
- φ_{ij} - The within-situation error terms describing differences between observers counting the same flock. All φ are assumed to be independently and identically distributed as $N(0, \sigma_w^2)$.
- α', β - Regression parameters. By defining $\alpha \equiv \exp\alpha'$ the regression line can be expressed as:

$$C = \alpha N^\beta$$

We further define:

$$\bar{Y}_{i.} \equiv \Big(\sum_{j=1}^{n_i} Y_{ij}\Big)/n_i$$

$$\bar{Y} \equiv \Big(\sum_{i=1}^{I} \bar{Y}_{i.}\Big)/I$$

Note that $\bar{Y}$ is not necessarily equal to the mean of all Y_{ij}. For $\bar{Y}_{i.}$ and $\bar{Y}$ holds:

$$\text{var}\,\bar{Y}_{i.} = \sigma_b^2 + \sigma_w^2/n_i, \qquad \text{(A2)}$$

$$\text{var}\,\bar{Y} = \frac{\sigma_b^2}{I} + \frac{\sigma_w^2}{I^2}\sum_i \frac{1}{n_i} \qquad \text{(A3)}$$

Also if the same person provided counting results for different flocks, independence of the error terms is assumed, as if different persons were counting. This assumption in fact points towards the simplification of the error term in (A1). In principle separate between-situation, between-observer, situation/observer interaction and within-observer error terms could be included. Such a model, however, requires a very sophisticated set of field data. The simpler model meets our (pragmatic) requirements as well.

For counting results it is convenient to work with relative errors (RSD values). They are largely independent of flock size. Regression analysis, however, requires an absolute error being independent of flock size. The solution is the use of natural logarithms in the regression model. A relative stochastic error in an untransformed flock size approximately equals an absolute error in its natural logarithm. This means: $RSD_b \approx \sigma_b$ and
$RSD_w \approx \sigma_w$

The regression parameters α' and β are estimated by means of a standard straight line regression analysis using the points $(X_i, \bar{Y}_{i.})$. As long as $\sigma_b^2 > \sigma_w^2$ the use of weights is not necessary (see (A2) and remark below). So

$$\hat{\beta} = \frac{\sum_i (X_i - \bar{X})\,\bar{Y}_{i.}}{\sum_i (X_i - \bar{X})^2} \qquad \text{(A4)}$$

$$\hat{\alpha}' = \bar{Y} - \hat{\beta}\bar{X} \qquad \text{(A5)}$$

The within-situation variance is estimated by:

$$\hat{\sigma}_w^2 = \frac{\sum_i \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_{i.})^2}{\sum_i n_i - I} \qquad \text{(A6)}$$

Note that the X_i are not needed here. The between-situation variance σ_b^2 is estimated using the residual sum of squares RSS from the regression analysis:

$$E[RSS] = E\Big[\sum_{i=1}^{I}(\bar{Y}_{i.} - \hat{\alpha}' - \hat{\beta}X_i)^2\Big]$$

E[RSS] has to be written in terms of the variances σ_b^2 and σ_w^2. Straightforward calculation using (A2), (A4) and (A5) leads to:

$$E[RSS] = (I-2)\,\sigma_b^2 + \Big\{\sum_{i=1}^{I}\frac{1}{n_i}\Big(\frac{I-1}{I} - \frac{(X_i - \bar{X})^2}{\sum_j (X_j - \bar{X})^2}\Big)\Big\}\,\sigma_w^2 \qquad \text{(A7)}$$

In the special case of all n_i being equal to n this result simplifies to (compare with (A2)):

$$E[RSS] = (I-2)\,\{\sigma_b^2 + \frac{\sigma_w^2}{n}\} \qquad \text{(A8)}$$

Expression (E7) or (A8) is now used for the estimation of σ_b^2 by replacing E[RSS], σ_b^2 and σ_w^2 by RSS, $\hat{\sigma}_b^2$ and $\hat{\sigma}_w^2$ respectively. (RSS follows from the regression analysis $\hat{\sigma}_w^2$ and from (A6)). For var $(\hat{\beta})$ can be derived, using (A2) and (A-4):

$$\text{var}\,(\hat{\beta})\,.\,\sum_i (X_i - \bar{X})^2 = \sigma_b^2 + \frac{\sum_i \frac{1}{n_i}(X_i - \bar{X})^2}{\sum_i (X_i - \bar{X})^2}\,\sigma_w^2 \qquad \text{(A9)}$$

To estimate var$(\hat{\beta})$ and var$(\bar{Y})$ the estimated variances are used in (A3) and (A9).

In cases where both $\sigma_b^2 < \sigma_w^2$ and n_i values strongly vary, it may be useful to use weighted least squares to get a second and more accurate estimate of α' and β. Weights are then calculated using expression (A2). This method has not been applied here, since the regression lines for shorebirds cannot be used of correction.

Error calculations

One observer counts a flock of birds. The result is written as C^* (we write C^* in order to keep a clear distinction between this "new" count and the C_{ij} mentioned above; in the main text the "star" is omitted). The stochastic error in C^* is estimated using:

$$\text{var}\,C^* = C^{*2}RSD^2_{flock} =$$

$$= C^{*2}\{RSD_b^2 + RSD_w^2\} \approx C^{*2}\{\hat{\sigma}_b^2 + \hat{\sigma}_w^2\} \qquad \text{(A10)}$$

If the hypothesis ($\alpha' = 0$, $\beta = 1$) is rejected (F-test), the regression line may be used to correct C^* for the systematic error. The corrected count is:

$$N^* = \{C^*\exp(-\hat{\alpha}')\}^{1/\hat{\beta}}$$

Using (A5) this can be written as:
$^{e}\log N^* = \bar{X} + \hat{\beta}^{-1}\{^{e}\log C^* - \bar{Y}\}$
Since $\hat{\beta}$, C^* and $\bar{Y}$ are uncorrelated var N^* equals (using (A10)):

$$\text{var } N^* \approx \left(\frac{\delta N^*}{\delta C^*}\right)^2 \text{var } C^* + \left(\frac{\delta N^*}{\delta \hat{\beta}}\right)^2 \text{var } \hat{\beta} + \left(\frac{\delta N^*}{\delta C^*}\right)^2 \text{var } \bar{Y}$$

$$\approx \left(\frac{1}{\hat{\beta}}\right)^2 \{RSD_{flock} N^{*2} + N^{*2}(^{e}\log N^* - \bar{X})^2 \text{ var } \hat{\beta} + N^{*2} \text{var } \bar{Y}\}$$

If counting results for different flocks C_i^* are summed we have:
$C_{tot}^* \equiv \sum_i C_i^*$ and

$\text{var } (C_{tot}^*) = RSD_{flock}^2 \{\Sigma C_i^{*2}\}$.

A corrected total is simply the sum of all corrected counts N_i^*. And using the same method as for a single flock:

$$\text{var } N_{tot}^* \approx \left(\frac{1}{\hat{\beta}}\right)^2 [RSD_{flock} \sum_i N_i^{*2} + \{\sum_i N_i^* (^{e}\log N_i^* - \bar{X})\}^2 \text{var } \hat{\beta} + N_{tot}^{*2} \text{ var } \bar{Y}]$$

The stochastic error due to RSD_{flock} tends to decrease with increasing number of flocks. The terms with var $\hat{\beta}$ and var $\bar{Y}$ however, do not "average away", since the same regression line is used for the correction of all flocks.

Present address C. Rappoldt
Emmastraat 53, 1814 DM Alkmaar, The Netherlands.