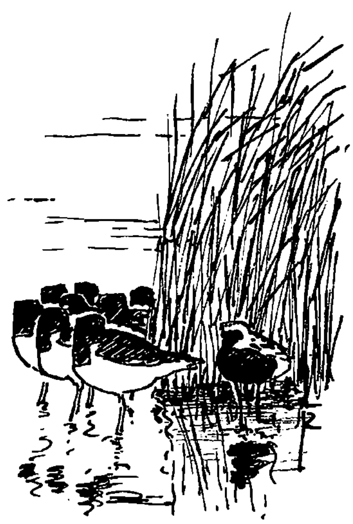
STATISTICAL ANALYSIS OF LONG-TERM MONTHLY OYSTERCATCHER *HAEMATOPUS OSTRALEGUS* COUNTS

JAAP VAN DER MEER[1], RICHARD N.M. DUIN[2] & PETER L. MEININGER[3]

Van der Meer J., R.N.M. Duin & P.L. Meininger 1996. Statistical analysis of long-term monthly Oystercatcher *Haematopus ostralegus* counts. Ardea 84A: 39-55.



Thirteen years of monthly counts of Oystercatchers from the Oosterschelde and Westerschelde, SW. Netherlands, contained approximately 13% missing values. Several log-linear Poisson models which assumed independence of the data were used to fill in (impute) these missing values. Yet these models either fitted poorly (when only main effects year, month, and site were considered) or were too complicated (when first order interaction effects were also taken into account) to provide a reliable description. Furthermore, the errors appeared to be dependent. Consequently, regression models which consider simple deterministic functions of time and temporal dependence between the errors, were used for a more parsimonious description of the underlying process. The model which assumed that after some point in time all monthly means are increased or decreased by some factor yielded the best fit. This factor differed among months. Using this model, a sudden change was estimated to have occurred in June/July 1983, three years before the completion of hydrotechnical works that resulted in a loss of about 17% of the intertidal area in the Oosterschelde. Winter and spring numbers have decreased, while summer and autumn numbers have increased.

Key words: Oystercatcher - *Haematopus ostralegus* - imputing - missing counts - Oosterschelde

[1]Netherlands Institute for Sea Research (NIOZ), P.O. Box 59, 1790 AB Den Burg, Texel, The Netherlands; [2]Rijksinstituut voor Kust en Zee, P.O. Box 20907, 2500 EX Den Haag, The Netherlands; [3]Rijksinstituut voor Kust en Zee, P.O. Box 8039, 4330 EA Middelburg, The Netherlands.

INTRODUCTION

Monthly counts of Oystercatchers *Haematopus ostralegus* have been made since the mid-1970s in the Oosterschelde and Westerschelde, two estuaries in the estuarine complex of the rivers Rhine, Meuse and Scheldt. The main objective of this monitoring programme was the description of possible changes over years in the use of the area by migrating and wintering birds. Changes are possibly related to the construction of a storm-surge barrier in the mouth of the Oosterschelde estuary and of two secondary dams on the eastern boundaries of the estuary in the years 1982-1986. These hydrotechnical works resulted in a loss of about 23 km^2 (17%) of the 137 km^2 intertidal area in the Oosterschelde (Schekkerman *et al.* 1994). The adjacent Krammer-Volkerak area became a fresh water lake and 33 km^2 tidal flats were lost, which might have had an impact on bird numbers in the Oosterschelde and Westerschelde. The intertidal area of the Westerschelde estuary (82 km^2) was not reduced.

In order to justify the choice of the most appropriate statistical method(s) for the analysis of this dataset, we thought that it might be helpful to start with a short review of the various methods that are currently being used. In spite of the apparent similarity in objectives (monitoring programmes of birds usually aim to assess the pattern of population changes over years within some defined region) and data (available data commonly consist of counts on a number of sites over a number of years, and sometimes data are available for several months from each year), quite distinct methods of statistical analysis have been used in a host of recent studies (Geissler & Noon 1981, Mountford 1982, Van Latesteijn & Lambeck 1986, Moses & Rabinowitz 1990, Beintema *et al.* 1993, Thomas 1993, Ter Braak *et al.* 1994, Underhill & Prŷs-Jones 1994). Apparently, which statistical analysis to apply to assess population changes is not a straightforward matter. In this short review, we will emphasize the objectives and assumptions of various methods. The objectives and assumptions were often not explicitly stated in the original papers, but are nevertheless of utmost importance in considering whether a method is appropriate for the analysis of a particular dataset. Bird monitoring projects and the accompanying methodological problems have received much attention recently (Sauer & Droege 1990), and various reviews have been published (Ter Braak *et al.* 1994, Underhill & Prŷs-Jones 1994). However, these reviews were restricted to the single problem of missing values, whereas our review addresses a wider range of problems. Finally, the methods most appropriate for the analysis of 13 years of Oystercatcher count data are applied.

SHORT REVIEW OF STATISTICAL METHODS

We start by noting that the choice of an appropriate method should be based upon (1) the statistical design of the monitoring programme, and (2) the type of inference that one aims to draw.

Concerning the statistical design, a major distinction is that between a census and a sample. In a census the entire region of interest is counted each period (e.g. year), where a region might consist of a number of geographically separate sites, e.g. all estuaries in Britain. In contrast, in a sample, only a selection of all sites is counted each year. The sites are usually selected at the start of the programme, and revisited in succeeding years.

Two types of inference

Two types of inference can be identified. Firstly, one might only be interested in the actual bird population across all sites. If a complete census was taken and all sites were counted (supposing for the time being that the counts were without measurement error) inferential statistics would not be needed. Confidence intervals and *p*-values are meaningless in this case. If, on the other hand, a sample of sites has been selected at random and counted without measurement error, the randomization introduced by the sampling design would provide the basis for inference. The error of the estimate for the size of the bird population would arise solely from the random sampling variation that is present when only a sample of sites is measured instead of all sites (Cochran 1977). This type of inference is therefore called randomization-based inference.

Basically, the route regression method applied in the North American Breeding Bird Survey follows this approach (Geissler & Noon 1981). The number of birds on randomly selected sites (routes) are counted each year. For each selected site separately, the time series is first summarized by the (weighted) slope of the linear regression over time. Calculation of the (weighted) slope for each selected site is simply an arithmetic exercise, without the need for any probability model. Next, the mean (weighted) slope of all sites is estimated. The variance of this estimate is solely determined by the sample size, the variance among sites, and the sampling fraction (i.e. the number of selected sites divided by the total number of sites). Other and less restrictive summaries of the time series can be used as well (see, for example, Roberts 1992).

The second type of inference regards the time series of the total number of birds across all sites as a single realisation of some underlying 'process'. Even when the total world population of a species has been counted, the question arises whether the observed temporal variation reflects 'real' changes in the underlying 'true' mean, or whether the time series can be sufficiently described by some random process with a constant 'true' mean. In the latter case, all variation might be due to a randomly varying environment, e.g. the alternation of mild and severe winters. For bird counts, it seems quite obvious that some sort of dependence will also be present between succeeding counts; if there are more birds than usual in one month, it seems likely that there will be more birds than usual in the next. Thus, one aims to decompose the time series into (1) a systematic trend, possibly with a seasonal component if monthly data are involved, (2) a random process exhibiting serial correlation, and (3) a purely random process, in which the data are independent and identically distributed. Measurement error, for example, may be described by a purely random process (but see Sauer 1994). There is no clear-cut receipt how to decompose the data into these three components, and far-reaching assumptions about the systematic part and the random part are necessary to enable estimation of the parameters (Judge *et al.* 1980). The second type of inference is therefore called model-based inference.

Van Latesteijn & Lambeck (1986) followed the model-based approach to analyse monthly Oystercatcher counts from the Roggenplaat, a tidal flat in the Oosterschelde, from 1964 to 1979. They first subtracted overall monthly means from the counts. Subsequently, the remaining difference was described in terms of two processes; an abrupt change in the systematic mean and a random process that accounted for serial correlation. The estimated time at which the abrupt change in numbers occurred, coincided exactly with the time of closure of the last gap in the dam being built in the adjacent Grevelingen estuary. In this case, serial correlation in the random part was taken into account. Usually, however, one relies on the assumption that all the observations are independent (e.g. Beintema *et al.* 1993). Such ignorance of the statistical dependence among the data may severely bias the results of hypothesis testing on trends in bird numbers. The same is true for the estimation of confidence intervals of the parameters of some model describing the systematic changes in the mean.

Two simple examples illustrate the difference between the two types of inference. First, since 1928 censuses of the number of breeding Grey Herons *Ardea cinerea* in Britain have been taken annually. The severe winters of 1947 and 1963 caused a marked fall in the numbers of breeding herons in the following spring, and the effects of these severe winters lasted for several years. In the first approach, an analysis of the heron data would be restricted to a description of the time series of the actual number of breeding herons. In contrast, the second approach would ensue a decomposition of the time series into a 'systematic' part and a 'random' part. As a result, all variation might be attributed to a random process with serial correlation, which implies that no lasting systematic change in the number of breeding herons occurred. In the second example, observations that were made of some variable y_{ij} at five sites ($i = 1, .., 5$) in each of ten years ($j = 1, .., 10$) are taken into account. Two such datasets are presented in Fig. 1. Only a linear trend over years is taken into account. Now consider two competing ways to test for a linear trend; first, a one sample t-test using the slopes b_i of each of the five sites and, second, a simple linear regression using for each year $\overline{Y}_{.j}$, the average over the five sites. The randomization-based approach aims to draw inference about the mean linear trend over all sites, and when the five sites are a sample selected at random from all sites, the most obvious test is the one sample t-test. In contrast, in model-based inference, one aims to examine whether the 'underlying' process exhibits a linear trend. Ignoring possible serial correlation (the assumption of independent observations may be unrealistic in practice, but this imperfection does not affect the core of the present argument), the appropriate test is the simple linear regression using for each year the average (or sum) over the five observed sites. The variation among sites is ignored, and, in fact, the test results only apply to the five study sites. For dataset A, with highly variable slopes among sites but with only minor deviations from linearity, the one sample t-tests reveals a non-significant result, i.e. the trend is not significant. Con-

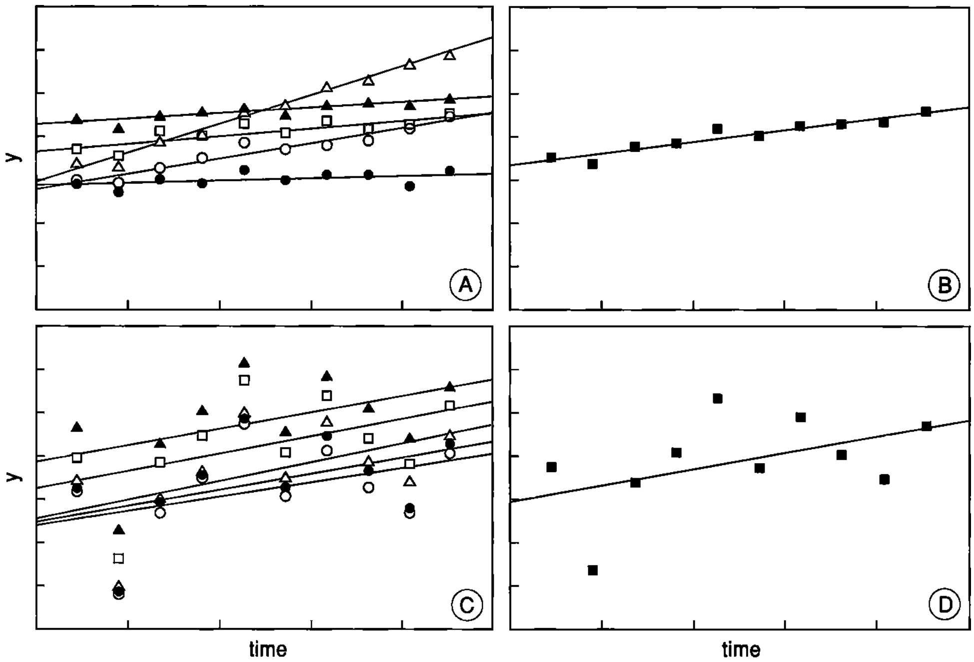


Fig. 1. Two artificial datasets of changes in time of some variable y observed at five different sites. Sites are indicated by symbols. Solid lines are simple linear regressions. The panels on the left show the data per site. Averages are shown in the right panels. In the first dataset (A, B) only minor deviations from linearity occur, but the variation of the regression slopes among sites is large. In the second dataset (C, D), sites show almost similar relationships with time, apart from a site-specific constant. Yet, large deviations from linearity occur. The statistical significance of the mean slope depends on which of the two accompanying error terms is used. The one sample t-test uses the variation of the regression slopes among sites. The simple linear regression uses the deviations from linearity. The choice of a test should be based on the statistical design and on the type of inference that one aims to draw.

versely, the slope of the linear regression through the means differs significantly from zero. For dataset B opposite results were obtained for both tests. Clearly, the statement by itself that a significant linear trend in time has been established, would not mean much.

Missing values, incomplete counts and measurement error

In practice, the statistical analysis of bird monitoring data is often hampered by the presence of many missing values (Ter Braak *et al.* 1994). For example, more than half of the 1200 sites × 23 years of wintering waterfowl counts in The Netherlands is missing (Beintema *et al.* 1993). Other examples of data sets with many missing values are the BTO heronries counts (Marchant *et al.* 1990) and the BTO birds of estuaries enquiry (Prater 1979, Kirby *et al.* 1991). In all these programmes, the aim is to conduct a complete census and to visit all sites annually. There is no true sampling of sites, and missing values at sites occur haphazardly. Using the terminology proposed by Little & Rubin (1987), one needs an imputation technique, in which missing data are replaced by model predictions.

The problem of missing data has recently attracted much interest and several methods have been proposed to deal with it. Opdam *et al.* (1987) used a log-linear Poisson model to impute missing Sparrowhawk *Accipiter nisus* counts. They assumed that the expected counts were a product of a year factor and a site factor. The observed counts were supposed to follow a Poisson-like distribution. Underhill & Prŷs-Jones (1994) used essentially the same method for imputing missing counts in the birds of estuaries enquiry. As counts were available for several months per year, they included a month factor in their model. Mountford (1982) and Thomas (1993) used a weighted average of ratio estimates across years. The route regression method (Geissler & Noon 1981) first summarizes the data of each site into a linear trend. Calculation of the slope is an arithmetic exercise. In fact, the presence of missing observations prevents the same calculations from being carried out for each site. Nevertheless, they still used the least-squares slope when data were missing. In terms of imputation techniques, they imputed missing observations by predictions of a least squares regression line.

The presence of incomplete counts and the assumption that each count is conducted without any error must also be considered. The latter assumption is made in randomization-based inference. Underhill & Prŷs-Jones (1994) developed a procedure for handling incomplete counts. The procedure lacks a formal justification and a clear interpretation, and is classified as an *ad-hoc* procedure by Ter Braak *et al.* (1994). The problem of the accuracy of a single census has been examined by Matthews (1960), Prater (1979), Kersten *et al.* (1981), and Rappoldt *et al.* (1985). Rappoldt *et al.* (1985) showed that the standard deviation of the random error is proportional to the mean. For large-scale shorebird counts this error was assessed to be approximately 5 to 10 percent. Clearly, the presence of missing, incomplete or erroneously measured data requires that assumptions are made beyond those of randomization, and all inference therefore is, at least to some extent, model-based.

From imputation to stochastic space-time process models

As noted above, the distinction between randomization-based inference, which is concerned with estimating the properties of the population of all sites (not to be confused with the population of birds) from a sample of sites, and model-based inference, that is concerned with estimating the properties of an underlying process, is of importance for interpreting statements on confidence intervals and hypothesis tests about trends in bird populations. Furthermore, the distinction between a census and a sample should be kept very clearly in mind.

If the actual bird population across all sites is the focus of interest, then the use of a model may be restricted to supplying imputed values and correcting incomplete counts. The next step depends upon the sampling design. When a census has been carried out, the completed data, and so including imputed values, can be summarized by some descriptive statistics. When only a sample of sites has been taken, the mean over all sites can be estimated from the completed data using randomization-based methods. A disadvantage of the imputation approach is that the imputed value is treated as known, and the uncertainty about the correct imputation model is not specified. However, the results from several alternative models may be compared to display the sensitivity of inference to the model choice (Little & Rubin 1987).

When emphasis is put on the underlying process, a model has to be selected and its parameters estimated. When a census has been carried out, the yearly counts (summed over all sites) can be modelled by, for example, some random process model that exhibits serial correlation. The occurrence of missing values may still cause problems in obtaining yearly counts. In that case, a two-step procedure may be followed. Firstly, the missing values are imputed. Next, parameters are estimated using the completed data. If data from a sample of sites are available, the time series preferably should be analysed by stochastic space-time process models, which accounts for both serial (temporal) and spatial dependence. Yet this is an almost unexplored field of research.

METHODS

Data

In the present paper we focus on monthly counts of Oystercatchers that were made between July 1978 and June 1991 in the estuarine complex of the rivers Rhine, Meuse and Scheldt. The intention in that programme was to make each month a complete census of all waterbirds present in the area, in order to detect possible changes in the use of the area by migrating and wintering birds. The data used in the present analysis are restricted to counts from the Oosterschelde and Westerschelde, the two most important estuarine areas in the region. In the Oosterschelde, four sub-areas were distinguished. The dataset (13 years × 12 month × 5 sites, i.e. 780 observations) contained 102 missing counts and a further 97 counts that were classified as incomplete. Most of the missing counts occurred in only two years. Between September 1983 and June 1985, counts were made only in January 1984, May 1984, August 1984, and January 1985. Over this period, counts were not available for 18 months, resulting in 90 missing values. In all the modelling which follows, 'bird years' instead of calendar years were used; these start in July and last until June of the next year.

Statistical analysis

There was no true sampling of sites, and therefore there was no need to make the inferential step from a sample of sites to the statistical population of all sites. Statistical models that were used, served two aims. Firstly, models were applied for the imputation of missing counts and the correction of incomplete counts. Secondly, models were used to describe the underlying process.

Following Opdam *et al.* (1987) and Underhill & Prŷs-Jones (1994), log-linear Poisson models were used. When only main effects are used, the assumptions are made that the expected number of birds at a site in a particular year and month can be described by the product of a site factor, a year factor, and a month factor. Hence, the differences among years, indicated by the year factor, are the same for all sites and months, and the seasonal trend, indicated by the month factors, is the same for all sites. These assumptions are probably unrealistic (Mitchell *et al.* 1988). Therefore, first-order interaction effects, which would test for the discrepancies with these assumptions, were considered too. In addition, a dummy variable indicating whether a count was incomplete was included in the analysis.

In a log-linear Poisson model, it is assumed that the data are independent and that

$$Var(y_i)=\sigma^2 E(y_i)$$

where $E(y_i)$ is the expected value of, in this case, the number of birds y_i. The discrepancy between the model and the data can be expressed by the deviance, which is itself closely related to the likelihood. The likelihood, roughly stated, is equivalent to the probability of obtaining the data if the model is true. Suppose that from two possible models, one model can be obtained by imposing constraints on the parameters of the other (e.g. all year factors are zero) and suppose further that this simplest model, which is said to be nested in the other, is true. Then, approximately, the difference in deviance (divided by the dispersion parameter σ^2) has a X^2_t distribution, where t is the difference in degrees of freedom between the two models. So if the difference in (scaled) deviances when a certain parameter is either included or excluded is too large, we would conclude that the parameter should be included in the model. This test is called the likelihood ratio chi-square test and the results might be used as a criterion for model selection. Because the expectation of X^2_t equals t, a difference in (scaled) deviance that greatly exceeds t, indicates a poor fit with the simpler model. The dispersion parameter σ^2 is estimated by the mean deviance of the most complicated model available, which is, in the present case, the one with all first order interaction effects included. The mean deviance is the deviance divided by its number of residual degrees of freedom. The assumption of independent observations was examined by analysis of the deviance residuals. See McCullagh & Nelder (1989) for a detailed discussion of log-linear models.

The imputed values of the log-linear Poisson model were compared to those produced by the obvious alternative model, which is a linear model that uses the logarithms of the counts (Moses & Rabinowitz 1990). The underlying assumption of a multiplicative lognormal error is consistent with a standard deviation of the random counting error that is proportional to the mean (Rappoldt *et al.* 1985). The predicted values $\hat{x}_i$ of the linear model, that uses natural logarithms of the counts, were back-transformed using the following transformation to approximately correct for the bias

$$y_i = exp\ (\hat{x}_i + \frac{s^2}{2})$$

where s^2 is the residual mean square.

Finally, using total counts per month, regression models were used in which a temporal dependence between the errors was taken into account, to provide a more parsimonious description of the underlying process. Both linear trends and step trends were considered. The total monthly counts were based on 678 observations and 102 imputed values. The total counts per month were log transformed. The errors were modelled by a variety of (seasonal) autoregressive-moving average processes (Chatfield 1989). The first order autoregressive process ($AR(1)$), for example, is given by $X_t = \alpha X_{t-1} + Z_t$, where X has zero mean and the Z_t's are independent and identically normally distributed. This implies that the deviation from the trend in the number of birds in month t depends on the deviation from the trend in month t-1. This makes the apparently reasonable assumption that if there are more birds than usual in a month, there will be more birds than usual in the next. For monthly observations we may also expect X_t to be correlated with X_{t-12}. This correlation, and the dependence on X_{t-1}, is accounted for by a first-order seasonal autoregressive process, denoted as $SAR(1) \times (1)_{12}$.

Both the log-linear Poisson model and the linear model (using log-transformed data) with independent errors were estimated with the regression procedure of the Genstat 5.1 programme (Payne *et al.* 1987). The models with correlated errors were fitted by maximizing the likelihood with the time series procedure of the Genstat 5.1 programme (Payne *et al.* 1987).

RESULTS

Log-linear Poisson models for imputing missing values

The log-linear Poisson model (assuming independent observations), in which all first-order interaction terms were taken into account (model i8) yielded the best fit (Table 1). This means that the year factors differed between sites and months, and that the trend across months differed among sites. Inclusion of a dummy variable that indicates incomplete counts did not improve the fit significantly. This means that counts that had been thought not to be complete did not actually have significantly lower numbers than complete counts. As the accompanying (back-transformed) estimated parameter value equalled 0.92, these incomplete counts were, on average, a non-significant 8% lower than the remaining counts. For 18 months in the years 1983 and 1984 no counts were available at all, and the accompanying year × month interaction parameters were therefore not able to be estimated. Hence the model which includes all first-order interaction effects (model i8, see Table 1) cannot be used to impute missing values for these 18 months. The model without the year × month interaction term (model i5, see Table 1) was used for these 18 months instead. This procedure is not ideal. In order to provide an impression of the reliability of the imputed values, actual counts and model predictions, obtained by combining the predictions of model i8 and model i5 (see Table 1), are plotted against time for each site separately in Fig. 2.

The relationships between the predictions of the log-linear Poisson models, the predictions of the linear model applied to the logarithms of the counts, and the observed values show that both models gave similar predictions (and thus imputed values) for each count (Fig. 3). Log-transformation of the counts resulted in only slightly higher predictions. This means that in the present case the choice for either a log-linear Poisson model or a linear model applied to the logarithms of the counts, did not make much difference.

Description of the corrected data

Plots of the corrected data (raw counts, but model predictions are used to impute missing values) against time for each site separately show the broad features of the data (Fig. 2). A major decrease in mid-winter numbers occurred in the eastern part of the Oosterschelde. With few exceptions, all mid-winter counts from the winter 1984/1985 onwards are under 20 000. Numbers were consistently higher in the preceding winters of 1978/1979 until 1983/1984, the difference between the two periods being about 10 000 birds. The change in bird numbers corresponds with the loss of one third (19 of 59 km^2) of the intertidal area in the eastern part due to the engineering works. The loss in the other three areas combined totalled only about 4 km^2. But, despite this minor loss of intertidal area, numbers in the central part of the Oosterschelde have increased over the years, both in summer and in winter (Fig. 2).

Data such as those presented in Fig. 2 are usually summarized further to assist interpretation. Useful summary statistics are the averages over months, i.e. the means of the corrected counts for each combination of year and site (Fig. 4A), averages over years (Fig. 4B), and averages over sites (Fig. 4C). These summary values clearly show that the central part of the Oosterschelde has be-

Table 1. Analysis of deviance of Oystercatcher counts. The factors year (y), month (m) and site (s) have 13, 12 and 5 levels, respectively. First order interactions are indicated by two letters (e.g. ym). For each model the residual degrees of freedom (*df*) and the residual deviance are given. Each model is tested against (*versus*) a more complicated model, with one or more extra factors included. The test statistic is the difference in scaled deviance. The dispersion parameter is set equal to 438186/444. So, for example, the test of model 1 *versus* model 4 reveals a difference in scaled deviance of (1079544-1055693)/(438186/444) = 24.12. The scaled deviance approximately has a X^2_t distribution, where t is the difference in degrees of freedom between the two models (e.g. 662 - 650 = 12). The difference in mean scaled deviance, that is the difference in scaled deviance divided by t, is given (e.g. 24.12/12 = 2.01). The results of the significance tests are shown in the last column, where N.S. means that $p > 0.05$, and *** means that $p < 0.001$.

model	factors	*df*	deviance	*versus*	difference in mean scaled deviance	*p*
i0	null	677	4 809 622	i4	140.88	***
i1	m+s	662	1 079 544	i4	2.01	N.S.
i2	y+s	661	3 881 606	i4	260.31	***
i3	y+m	654	1 930 345	i4	221.56	***
i4	y+m+s	650	1 055 693	i8	3.04	***
i5	y+m+s+ys+ms	558	711 649	i8	2.43	***
i6	y+m+s+ym+ms	492	670 394	i8	4.90	***
i7	y+m+s+ym+ys	488	534 250	i8	2.21	***
i8	y+m+s+ym+ys+ms	444	438 186	i9	1.99	N.S.
i9	y+m+s+ym+ys+ms+dummy	443	436 224			
i10	full	0	0			

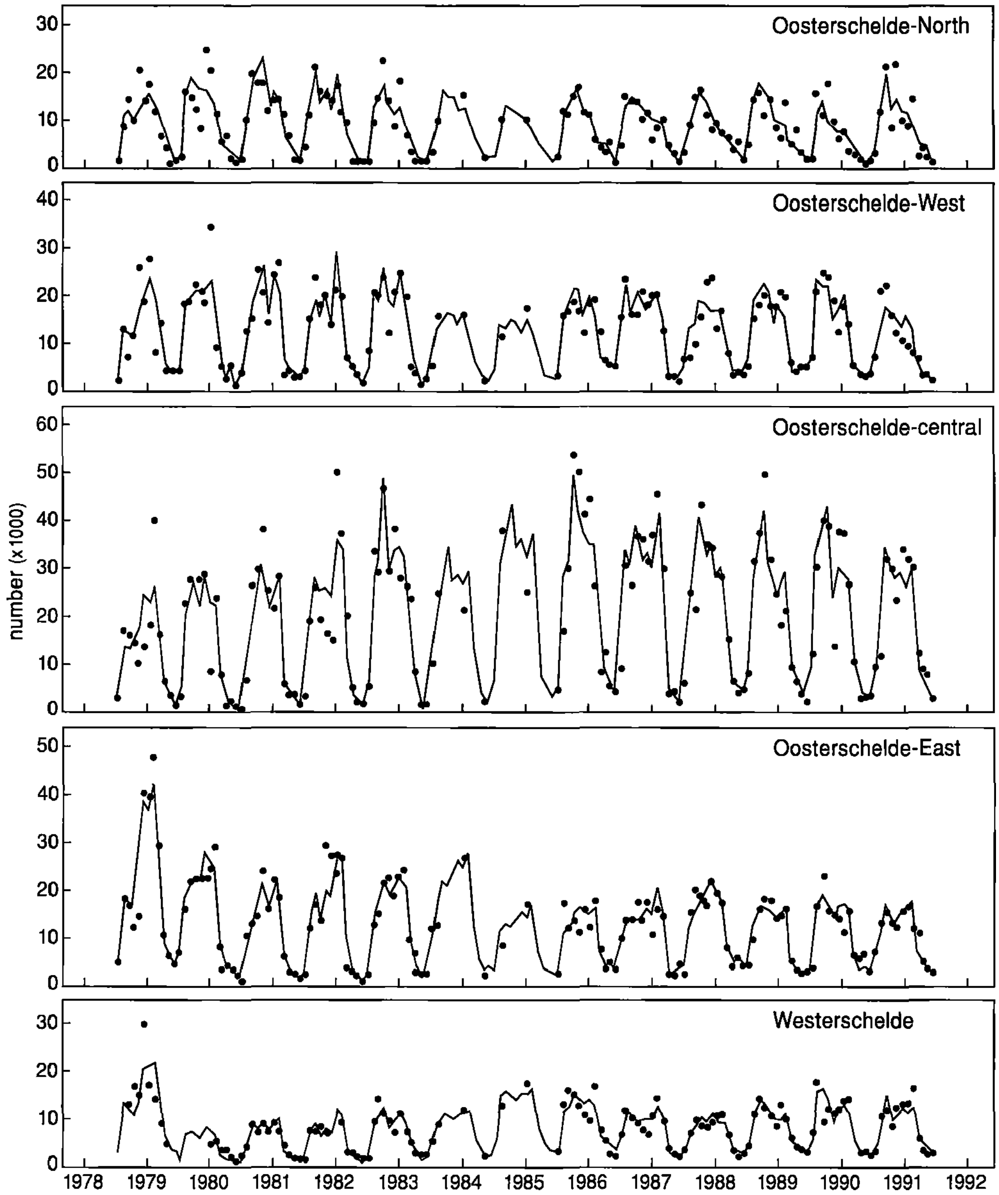


Fig. 2. Oystercatcher counts (dot) and model predictions (line) of the log-linear Poisson model i8. For 18 months in 1983/1984, and 1984/1985 predictions were obtained from model i5. See text and Table 1 for model descriptions.

come more important over the years, whereas the opposite is true for the eastern part (Fig. 4A). The seasonal pattern differed between all five sites. Numbers in autumn were relatively low for the Westerschelde and the eastern part of the Oosterschelde, whereas autumn numbers, particularly in October, were high for the central part of the Oosterschelde (Fig. 4B). In some months (for exam-

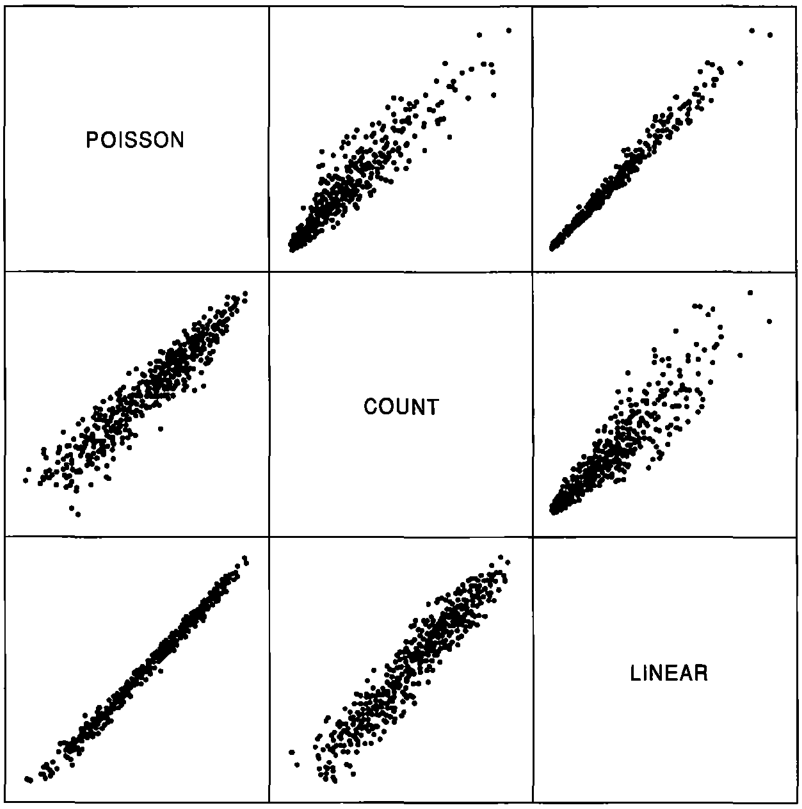


Fig. 3. Predictions of the log-linear Poisson model (model i8, see Table 1) in which all first order interactions are included (POISSON), predictions of the analogous linear model that uses log-transformed data (LINEAR), and the actual counts (COUNT) plotted against each other. The three panels in the upper-right triangle show untransformed data, and the three panels in the lower-left triangle show log-transformed data. To give an example, the lower left panel plots the log-transformed predictions of the linear model on the abscissa *versus* the log-transformed predictions of the log-linear Poisson model on the ordinate. The figure shows that the predictions of the two models are very similar, compared to the similarity between model predictions and the actual data.

ple July and October) numbers increased over the years, whereas in other months (for example January) numbers decreased (Fig. 4C).

The summary statistics of the corrected data (Fig. 2) shown in Figs. 4A, B & C, inevitably lead to a loss of information. If we ignore the problems associated with the few months where interaction parameters could not be estimated, Figs. 4A, B & C display the parameters of the first-order interaction log-linear Poisson model (model i8, see Table 1). The loss of information is quantified by the deviance (Table 1) and displayed by the differences between the data and the model predictions, as presented in Fig. 2. Otherwise stated, the model predictions in Fig. 2 and the summaries in Figs. 4A, B & C contain the same information. Yet, this is not to say that the Figs. 4A, B & C provide an easy interpretable summary. Some further summarizing might be needed. For example, to prevent Fig. 4C from becoming too crowded, averages over sites were only presented for four months. The trend over years for each month is further summarized by the linear component of the trend, as shown in Fig. 4D. The linear compo-

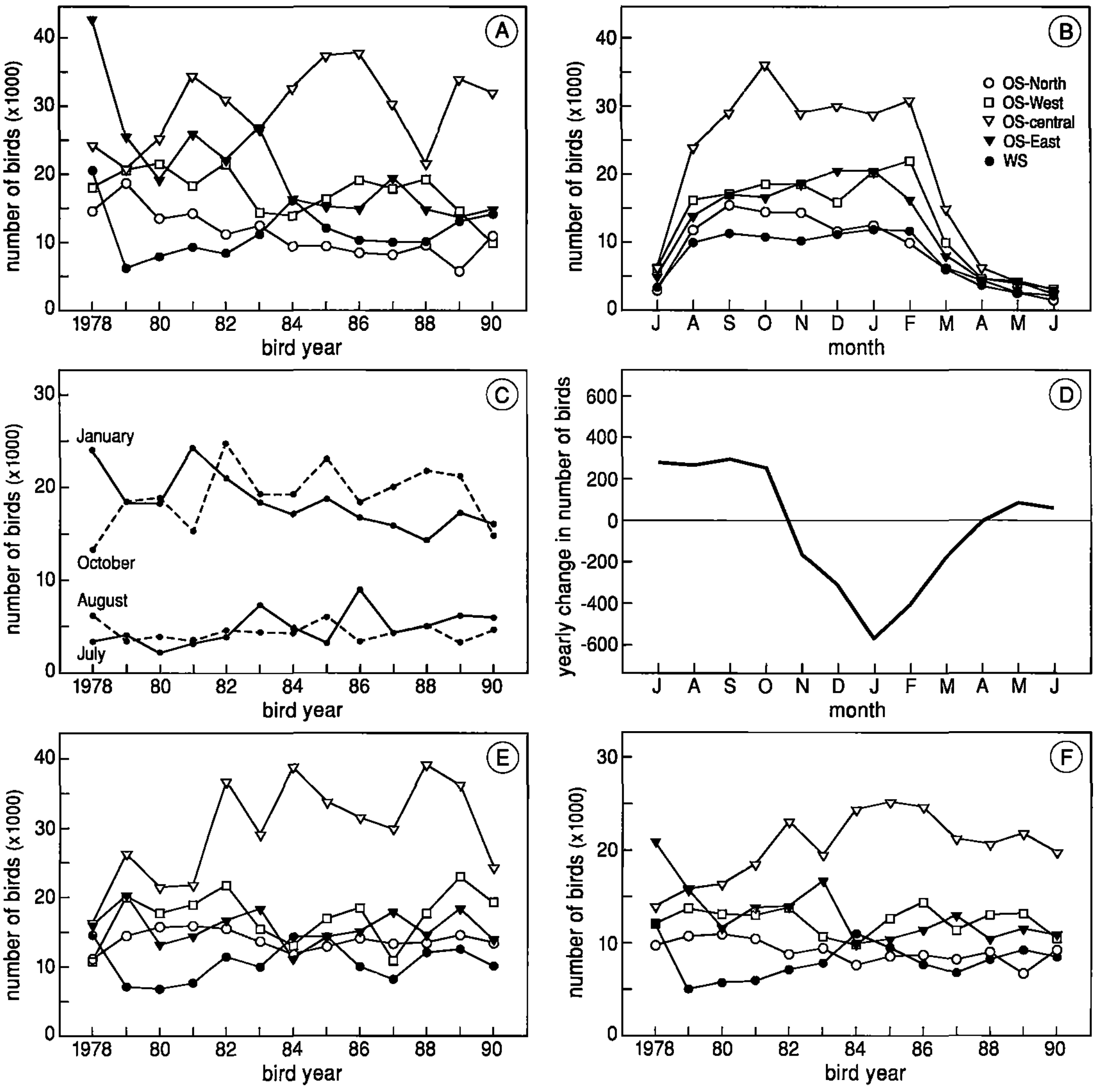


Fig. 4. Oystercatcher counts averaged over months, years, and sites, respectively. When missing values occurred, imputed values were used instead. (A) for each year × site combination the mean is taken over all 12 months; (B) for each month × site combination the mean is taken over all 13 years; (C) for each combination of four months and all years, the mean is taken over all five sites; (D) for each month the trend over years in the mean (taken over all five sites) is summarized by the slope of an ordinary least-squares regression of the mean versus year; (E) for each year × site combination the mean is taken over three late summer/early autumn months (August, September, October); (F) for each year × site combination the mean is taken over three winter months (December, January, February). Symbols refer to the five sub-areas (Figs. 2 & 4B), except for Fig. 4C where the lines refer to four different months.

nent shows an increase for the late summer/early autumn months, and a decrease for the winter months. The strongest decrease occurred in January. The increase in late summer/early autumn was solely due to an increase in numbers in the central part of the Oosterschelde. The other regions did not show any consistent trend (Fig. 4E). The decrease in winter could be attributed to decreases in the eastern, western and northern parts of the Oosterschelde (Fig. 4F). Winter numbers in the central part showed a slight increase over years. In the Westerschelde winter numbers did not change much.

Model diagnostics

No relationship between the deviance residuals and the estimated values could be detected for model i8 (see Table 1). But for two sites, a significant serial correlation was found between the residuals from succeeding counts. Values for the central part of the Oosterschelde and the Westerschelde were 0.22 and 0.26, respectively. Large negative cross-correlations among sites were also observed, the largest being between the central part of the Oosterschelde and the other four sites, with the western part -0.53, the northern part -0.44, the eastern part -0.39 and the Westerschelde -0.36. The cross-correlation between the northern part and the Westerschelde was the only positive cross correlation (0.12). All other cross correlations varied between -0.05 and -0.19. Negative cross-correlations suggest an exchange of birds between areas. Indeed, it seems probable that the most important exchanges would occur between the central part and the other regions, given its central geographical position. The appearance of small negative cross-correlations between sites must not be overinterpreted, because correlations in the order of -0.25 will arise, even when the observations are independent.

Table 2. *Time series models for monthly Oystercatcher counts. The relationship with time is described by month factors (m), and either by year factors (y), by a linear function with year, i.e. year is treated as a covariate (c), or by a step function (s). First order interactions are indicated by two letters, e.g. cm. For each model the model degrees of freedom (df), and the residual sum-of-squares (SS) are given. The error is modelled by a first-order autoregressive model ($AR(1)$) or by a seasonal first-order autoregressive model ($SAR(1) \times (1)_{12}$). Models were tested against (versus) a more complicated model, with one or more extra factors included. The results of the F-tests are given in the last column, where N.S. means that $p > 0.05$, and ** means that $p < 0.01$.*

model	factors	*df*	SS	*versus*	*F*	*p*
d1	m	144	7.256	d10	2.90	N.S.
d2	y+m	132	6.322			
d3	y+m+$AR(1)$	131	5.283			
d4	y+m+$SAR(1) \times (1)_{12}$	130	5.276			
d5	c+m	143	7.193			
d6	c+m+$AR(1)$	142	5.701			
d7	c+m+cm	132	5.736			
d8	c+m+cm+$AR(1)$	131	4.856			
d9	c+m+cm+$SAR(1) \times (1)_{12}$	130	4.832			
d10	s+m	142	6.971	d11	34,42	**
				d12	3.11	**
d11	s+m+$AR(1)$	141	5.603	d13	2.56	**
d12	s+m+sm	131	5.528	d13	26.06	**
d13	s+m+sm+$AR(1)$	130	4.605	d14	1.73	N.S.
d14	s+m+sm+$SAR(1) \times (1)_{12}$	129	4.544			

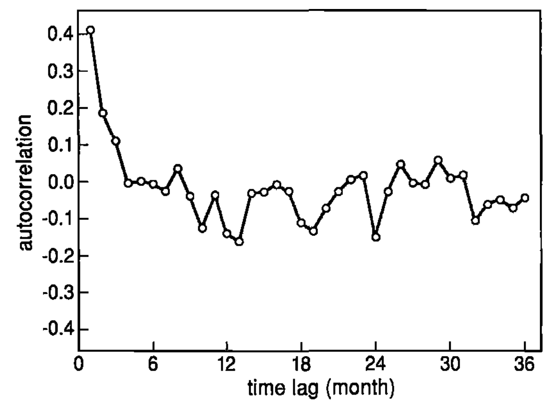


Fig. 5. Estimated serial correlation function (and twice the standard error) using the residuals from model d12 (see Table 2).

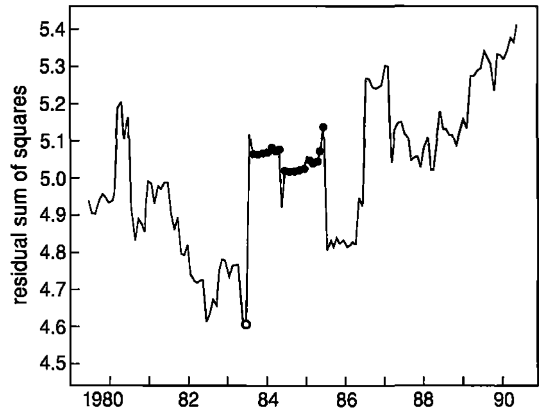


Fig. 6. Residual sum of squares versus the time of the abrupt change in mean numbers (model d13, see Table 2).

Models for dependent data

A variety of models for the corrected monthly counts (102 missing values were imputed, counts were summed over all sites) were fitted (Table 2). The relationship with time is described by (1) month factors, (2) month plus year factors; (3) month factors plus a linear relationship with year, where the slope of the linear relationship may differ among months, and (4) month factors plus a step function which indicates a sudden change at some point. The magnitude of the sudden change may also differ among months. The residuals of the last model showed large serial correlations (Fig. 5). These correlations suggested that a first-order autoregressive process ($AR(1)$) or a first-order seasonal autoregressive process ($SAR(1) \times (1)_{12}$) may be appropriate to model the errors. The step model (model d13, see Table 2), including a first-order autoregressive process for modelling the errors, appeared to give the best fit (Table 2). The time of the sudden change was estimated to have occurred between the counts made in June and July 1983 (Fig. 6). The estimated regression parameter of the autoregressive process equalled 0.41. The other parameters are displayed in Fig. 7. A simpler model in which the magnitude of the steps was constrained to be equal for each month resulted in a much poorer fit (model d11, see Table 2).

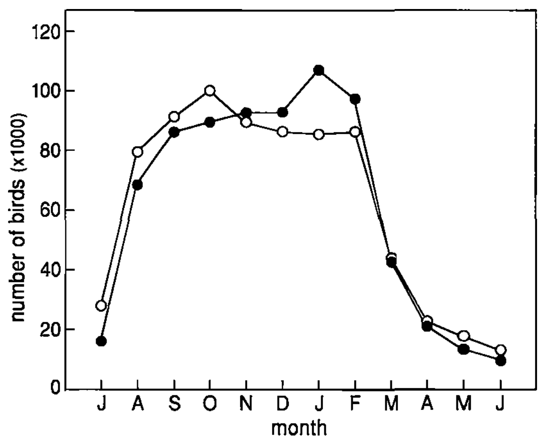


Fig. 7. Predicted number of Oystercatchers before and after the abrupt change in mean numbers (estimated in June/July 1983). Predictions were obtained from model d13 (see Table 2).

DISCUSSION

Statistical methodology

Statistical models for the Oosterschelde and Westerschelde monthly Oystercatcher censuses might serve two purposes: (1) to impute missing values and to correct for incomplete counts; and (2) to model trends in the underlying process.

The significant contribution (assuming independent errors) of all first-order interaction factors implies that a description of bird monitoring data simply in terms of year factors, month factors and site factors is of limited scope, even though this is the general practice (Moses & Rabinowitz 1990, Underhill & Prŷs-Jones 1994). Consequently, the log-linear Poisson model that includes (all) first-order interactions is probably more useful for imputing missing values. Yet, using the model with all first-order interactions to impute missing values was not entirely successful. The model could not provide predictions for some year × month combinations, because no counts were available at all. We used the model without the year × month interaction term (model i5, see Table 1) to impute values for these year × month combinations. Hence it is assumed that the year factors are equal for all months. The deviance of model i5 (see Table 1) was still significantly lower than the deviance of the main effects model (model i4, see Table 1), which makes the additional assumptions that the year and month factors are the same for all sites.

For descriptive purposes Poisson-like methods have the advantage over the assumption of lognormality in that no transformations of the data are required. The logarithmic transformation, for example, may be less appropriate because we are not interested in describing the geometric mean of bird counts. Underhill & Prŷs-Jones (1994) illustrate the inadequacy of the geometric mean by noting that counts of 700 and 3000 at two sites give the same geometric mean as counts of 150 and 14 000, while the total numbers of birds at the two sites differ considerably. The underlying process that generates the data may, however, be more appropriately described by a relationship between log(*count*) and the explanatory factors year, month, and site, plus added normal error. In fact, such a model corresponds better to the results obtained by Rappoldt *et al.* (1985), who found that the standard deviation was proportional to the mean. The assumption of Poisson-like distributed variables, on the other hand, requires that the variance is proportional to the mean. So it is not at all evident that the imputed values from a log-normal model are worse than those obtained from a Poisson model. The severe criticism by Underhill & Prŷs-Jones (1994) on the use of the geometric mean may be less relevant when the model has taken into account all systematic components and the remaining errors are independently and identically distributed. In such a case no serious misinterpretation of the model results need to occur. In practice, there may not be much difference between both imputation techniques, as can be seen from Fig. 3. The use of models with gamma-distributed observations might be a suitable alternative for data with constant coefficient of variation (McCullagh & Nelder 1989).

As has already been pointed out by Ter Braak *et al.* (1994), the use of Poisson regression has several advantages over the use of the EM-algorithm used by Underhill & Prŷs-Jones (1994) to estimate the maximum-likelihood parameters in their main effects model. The main advantage is that the model can be easily extended. In the present article, for example, we added interaction effects and a dummy variable coding for incomplete counts. Another option is to include observer effects (Sauer *et al.* 1994, Van der Meer & Camphuysen 1996). Furthermore, standard errors of estimates and statistical tests are provided and the numerical method of Poisson regression tends to reach more rapidly the maximum likelihood solution (Thisted 1988, Ter Braak *et al.* 1994). On the other hand, Poisson regression of larger datasets than ours may cause technical problems, as large amounts of computer memory are needed.

The assumption of independent errors was violated and significant serial correlation (indicating temporal dependence) and cross-correlations (indicating spatial dependence) were present among the residuals of the log-linear Poisson model that included all first-order interactions. Inclusion of both temporal and spatial dependence in log-linear Poisson regression seems worthwhile, but will generate a lot of statistical and numerical complications. The extension of log-linear Poisson models to account for time dependence is the subject of current statistical research, and several alternative approaches have been proposed (Liang & Zeger 1986, Engel & Keen 1994). We have chosen a simpler approach. All sites were grouped together, and the total monthly counts were log transformed. The log transformed data were assumed to be normally distributed. A variety of models, taking into account both a systematic change in the mean and temporal dependence, was fitted to the year × month data. A model similar to the one used by Van Latesteijn & Lambeck (1986) was also fitted. This model assumes a constant seasonal pattern. After some point in time, all monthly means are increased or decreased by some common factor. For the present dataset, this model proved to be unsatisfactory. A more complicated model, in which the amount of change may differ among months, gave a much better fit (model d13, see Table 2). It should be noted that the model also fitted better than the prevailing year effect and site effect model. The assumption that month factors are the same in all years could be relaxed because a rather simple year trend (a step function) was assumed. Some part of the remaining error was attributed to temporal dependence. Although the composition in trend and correlated error remains arbitrary, the model (model d13, see Table 2) has much appeal. The underlying process is modelled by a relatively simple trend in time, and the presence of temporal dependence, which means that when there are more birds than usual in a month, there will be more than usual in the next, seems quite likely for bird counts.

Changing bird numbers

The shift in the seasonal pattern of numbers was estimated to have occurred in June/July 1983. After this point of time the Oosterschelde and Westerschelde contained more birds than before in late summer and early autumn, but fewer birds in winter (Fig. 7). Lambeck (1991) found a similar change in the seasonal pattern when he compared data before 1984 with data after 1986. That the pattern of changes over time differed between regions (Figs. 4E & F) suggests possible biological explanations. The number of Oystercatchers in late summer/early autumn in the adjacent Krammer-Volkerak area during the pre-barrier period was about 5000-8000 birds (Meininger *et al.* 1984, Meininger *et al.* 1985). The late summer/early autumn increase in the central part of the Oosterschelde, which was approximately of the same magnitude, may have been caused by birds that had been displaced from their feeding area in the Krammer-Volkerak. Apparently, these birds were not able to settle in the nearer northern part of the Oosterschelde, which did not show much change in numbers for this time of year. Yet, a settlement of the displaced birds in the northern part resulting in the departure of birds already present, cannot be excluded. Another interesting point is that the late summer/early autumn Oystercatcher numbers in the eastern part of the Oosterschelde did not change, despite considerable (30%) losses in intertidal area due to the construction of the secondary dams. In winter, however, Oystercatcher numbers in the Oosterschelde did decrease. In the eastern part, numbers decreased by about 10 000 birds. This decrease coincided with the loss in intertidal area. The western and northern part of the Oosterschelde together showed a same decrease of about 10 000 Oystercatchers (Fig. 4F). In these regions, however, the amount of intertidal area hardly changed. The winter number of Oystercatchers in the central part increased by a few thousand, far less than the decrease in the other regions. Summing up, a decrease in the number of wintering Oystercatchers in the eastern part of the Oosterschelde corresponded to the decrease in intertidal area in that region. This decline (plus the decrease in the Krammer-Volkerak region) was not counterbalanced by an increase in the other parts of the Oosterschelde and Westerschelde. The numbers in the western and northern part did decrease as well, although the amount of intertidal area hardly changed within these regions.

Finally, it is interesting to note that the time of sudden change was estimated to have occurred in June/July 1983. This is more than three years before the completion of the Oosterschelde barrier and the secondary dams. Indeed, winter numbers in the eastern part of the Oosterschelde were already at a low level of about 15 000 birds in the pre-barrier winters of 1984/1985 and 1985/1986 (Fig. 4F). Winter numbers in the western part of the Oosterschelde showed a major decrease between the winter 1982/1983 and 1983/1984 (Fig. 4F). Winter numbers in the northern part of the Oosterschelde showed a decrease already in the early 1980s (Fig. 4F). Unfortunately, many missing values occurred between September 1983 and June 1985. Although the use of other models for imputation (i.e. linear models using log-transformed counts) yielded the same estimate of the time of the sudden change, the occurrence of so many missing values in the period close to this estimate sheds some doubt on the precision of the estimate. Nevertheless, our results suggest that changes in bird numbers that have occurred should not too easily be explained as a result of the engineering works that were performed.

ACKNOWLEDGEMENTS

The main basis for this article is the extensive series of bird counts organized by Rijkswaterstaat (RIKZ) and carried out by many professional and amateur bird watchers. The first authors' contribution was partly financed by Rijkswaterstaat, project DG 466. We thank Cajo J.F. ter Braak for valuable suggestions on a preliminary report. Cor Berrevoets and Rob Strucker assisted in data processing. Thanks to Ralph Clarke, Bruno Ens, John D. Goss-Custard, Rob H.D. Lambeck, Philip M. North, Richard Pettifor, Kees Rappoldt, and Leo Zwarts for constructive criticism on the manuscript.

REFERENCES

Beintema A., H. Buesink & L. van den Bergh 1993. Overwinterende watervogels in Nederland, 1967-89. Limosa 66: 17-24.

Chatfield C. 1989. The analysis of times series, 4th edition. Chapman, London.

Cochran W. 1977. Sampling techniques, 2nd edition. Wiley, New York.

Engel B. & A. Keen 1994. A simple approach for the analysis of generalized linear mixed models. Statist. Neerl. 48: 1-22.

Geissler P.H. & B.R. Noon 1981. Estimates of avian population trends from the North American breeding bird survey. Studies Avian Biol. 6: 42-51.

Judge G.G., W.E. Griffiths, R. Carter Hill & T-C. Lee 1980. The theory and practice of econometrics. Wiley, New York.

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

Kirby J.S., J.R. Ferns, R.J. Waters & R.P. Prŷs-Jones 1991. Wildfowl and wader counts 1990-91. Wildfowl Trust, Slimbridge.

Lambeck R.H.D. 1991. Changes in abundance, distribution and mortality of wintering Oystercatchers after habitat loss in the Delta area, SW Netherlands. Acta XX Congr. Int. Ornithol.: 2208-2218.

Liang K.Y. & S.L. Zeger 1986. Longitudinal data analysis using generalized linear models. Biometrika 73: 13-22.

Little R.J.A. & D.B. Rubin 1987. Statistical analysis with missing data. Wiley, New York.

Marchant J.H., R. Hudson, S.P. Carter & P.A. Whittington 1990. Population trends in British breeding birds. British Trust for Ornithology, Tring.

Matthews G.V.T. 1960. An examination of basic data from wildfowl counts. Proc. Int. Ornithol. Congr. XII: 483-491.

McCullagh P. & J.A. Nelder 1989. Generalized linear models, 2nd edition. Chapman & Hall, London.

Meininger P.L., H.J.M. Baptist & G.J. Slob 1984. Vogeltellingen in het Deltagebied in 1975/76-1979/80. Rijkswaterstaat report DDMI 84.23, Middelburg.

Meininger P.L., H.J.M. Baptist & G.J. Slob 1985. Vogeltellingen in het Zuidelijk Deltagebied in 1980/81-1983/84. Rijkswaterstaat report DGWM 85.001, Middelburg.

Mitchell J.R., M.E. Moser & J.S. Kirby 1988. Declines in midwinter counts of waders roosting on the Dee estuary. Bird Study 35: 191-198.

Moses L.E. & D. Rabinowitz 1990. Estimating (relative) species abundance from route counts of the breeding bird survey. In: J.R. Sauer & S. Droege (eds) Survey designs and statistical methods for the estimation of avian population trends: 71-79. Biol. Rep. 90(1) U.S. Fish Wildl. Serv., Washington.

Mountford M.D. 1982. Estimation of population fluctuations with application to the Common Bird Census. Appl. Statist. 31: 135-143.

Opdam P., J. Burgers & G. Muskens 1987. Population trend, reproduction, and pesticides in Dutch Sparrowhawks following the ban on DDT. Ardea 75: 205-212.

Payne R.W., P.W. Lane, A.E. Ainsley, K.E. Bicknell, P.G.N. Digby, S.A. Harding, P.K. Leech, H.R. Simpson, A.D. Todd, P.J. Verrier & R.P. White 1987. GENSTAT 5 reference manual. Clarendon Press, Oxford.

Prater A.J. 1979. Trends in accuracy of counting birds. Bird Study 26: 198-200.

Rappoldt C., M. Kersten & C. Smit 1985. Errors in large scale shorebird counts. Ardea 73: 13-24.

Roberts E.A. 1992. Sequential data in biological experiments. Chapman & Hall, London.

Sauer J.R. & S. Droege (eds) 1990. Survey designs and statistical methods for the estimation of avian population trends. Biol. Rep. 90(1) U.S. Fish Wildl. Serv., Washington.

Sauer J.R., B.G. Peterjohn & W.A. Link 1994. Observer differences in the North-American Breeding Bird Survey. Auk 111: 50-62.

Schekkerman H., P.L. Meininger & P.M. Meire 1994. Changes in the waterbird populations of the Oosterschelde (SW Netherlands) as a result of large-scale coastal engineering works. Hydrobiologia 282/283: 509-524.

Ter Braak C.J.F., A.J. van Strien, R. Meijer & T.J. Verstrael 1994. Analysis of monitoring data with many missing values: which method? In: E.J.M. Hagemeijer & T.J. Verstrael (eds) Bird numbers 1992. Distribution, monitoring and ecological aspects: 663-673. Statistics Netherlands, Voorburg.

Thisted R.A. 1988. Elements of statistical computing. Chapman & Hall, New York.

Thomas G.E. 1993. Estimating annual total heron counts. Appl. Statist. 42: 473-486.

Underhill L.G. & R.P. Prŷs-Jones 1994. Index numbers for waterbird populations. I. Review and methodology. J. appl. Ecol. 31: 463-480.

Van der Meer J. & C.J. Camphuysen 1996. Effect of observer differences on abundance estimates of seabirds from ship-based strip transect surveys. Ibis 138: 433-437.

Van Latesteijn H.C. & R.H.D. Lambeck 1986. The analysis of monitoring data with the aid of time-series analysis. Environ. Monitor. Assess. 7: 287-297.

SAMENVATTING

Vanaf het midden van de jaren zeventig zijn iedere maand langs de Oosterschelde en Westerschelde tellingen van het aantal Scholeksters uitgevoerd. Door de bouw van de in 1986 gereedgekomen stormvloedkering in de monding van de Oosterschelde en de aanleg van twee zogeheten compartimenteringsdammen aan de oostzijde van het estuarium is 23 km^2 van het in totaal 170 km^2 intergetijdegebied verdwenen. Bovendien is 33 km^2 intergetijdegebied in het nabijgelegen Krammer-Volkerak verloren gegaan. De tellingen bieden de mogelijkheid om mogelijke veranderingen door de jaren heen in het aantal in het gebied voorkomende Scholeksters te beschrijven. Dit verhaal behandelt tellingen die zijn uitgevoerd van juli 1978 tot juni 1991. De Oosterschelde werd opgedeeld in vier deelgebieden. De Westerschelde is als geheel in beschouwing genomen. Ongeveer 13 procent van de 780 tellingen (13 jaar, 12 maanden, 5 gebieden) is niet uitgevoerd. Statistische modellen, de zogeheten log-lineaire Poisson modellen, werden gebruikt om deze ontbrekende waarnemingen 'in te vullen'. Vervolgens konden de gecorrigeerde tellingen (dat wil zeggen de tellingen aangevuld met ingevulde waarden waar de telling ontbrak) worden samengevat in, bijvoorbeeld, het verloop door de jaren heen van het gemiddeld aantal aanwezige vogels in de wintermaanden per gebied afzonderlijk (Fig. 4F). Het best passende log-lineaire model voldeed als methode om ontbrekende waarnemingen 'in te vullen', maar schoot te kort als een geschikt model om een mogelijk onderliggend proces te beschrijven. Ten eerste waren in de tijd opeenvolgende fouten (een fout is het verschil tussen de werkelijke waarnemingen en de modelvoorspellingen) gecorreleerd. Ten tweede kende het model erg veel parameters en leek daarom bij voorbaat te ingewikkeld om een onderliggend proces te beschrijven. Daarom werden veel simpelere deterministische modellen (zoals, het model dat een lineaire toe- of afname in de tijd veronderstelde, of het model dat een eenmalige stapsgewijze verandering veronderstelde) die wel rekening hielden met de correlatie tussen opeenvolgende fouten, gebruikt om een eenvoudige beschrijving te geven van een mogelijk onderliggend proces. Het model waarin werd verondersteld dat na een bepaald tijdstip het aantal Scholeksters met een bepaalde factor toe- of afnam, bleek het best met de data overeen te komen. Deze factor verschilde per maand. De eenmalige stapsgewijze verandering zou hebben plaatsgevonden tussen juni en juli 1983. Dit was drie jaar voor de gereedkoming van de stormvloedkering en de compartimenteringsdammen. Na dit tijdstip zijn de aantallen Scholeksters in de winter en in het voorjaar afgenomen, maar in de zomer en in de herfst toegenomen. Helaas ontbraken juist in de periode na de zomer van 1983 veel tellingen, zodat aan de exactheid van de schatting van het tijdstip van verandering twijfels verbonden blijven.