BTO Research Report No. 287

WeBS Alerts III: Methodological and Presentational Considerations

Authors

Philip W Atkinson, Graham E Austin & Mark M Rehfisch

Report of work carried out by The British Trust for Ornithology
under contract to the WeBS Partnership

May 2002

The National Centre for Ornithology, The Nunnery, Thetford, Norfolk IP24 2PU
Registered Charity No. 216652

British Trust for Ornithology

WeBS Alerts III: Methodological and Presentational Considerations

BTO Research Report No 287

Philip W Atkinson, Graham E Austin & Mark M Rehfisch

Report of work carried out by The British Trust for Ornithology
under contract to the WeBS Partnership

Published in June 2012 by the
British Trust for Ornithology, The Nunnery, Thetford, Norfolk, IP24 2PU, U.K.

ISBN 978-1-908581-06-8

CONTENTS

Page No.

LIST OF FIGURES

Page No.

Page No.

1. INTRODUCTION

1.1 Scope of this Report

The WeBS Alerts System has undergone a lengthy gestation. When a WeBS Alerts system was first considered, it was felt appropriate that similar methodologies be applied to terrestrial species and waterbirds. With this in mind, the current Alerts methodology was devised and described in Gregory *et al*. 1999. This approach was applied to waterbirds at national, regional and Species Protection Area scales (Atkinson *et al*. 2000) and later at the SSSI scale (Atkinson *et al*. 2001).

This is the third report in the series and is the final report before the finalisation of the WeBS Alert methodology and its inclusion in the WeBS as a regularly used monitoring scheme. The previous two reports have identified several issues concerning Site Alerts which concern the statistical reliability of the Alert system. These have included the matching of WeBS site and designated site boundaries, the impact of using different numbers of months to produce the smoothed species indices and the effect of missing count data on the Alerts generated. The first of these is currently being addressed but the last two are issues that are discussed in the current report.

Waterbird assemblages were not addressed in the first two Alerts reports. This refers to the 20,000 threshold level above which an area may be designated as a Special Protection Area. There are various means of calculating assemblage Alerts and we explore the various options and also discuss the problems with data availability.

Alerts generate a great deal of information which needs to be presented in an easily digestible format. The Alert tables generated in the Alerts Reports were thought to be acceptable but it was difficult to obtain an overview of any geographic pattern in the Alerts generated, and a graphical way of showing the information is presented.

Therefore, this final report has three main aims:

1. Investigate the impacts of missing data, the differences in species monthly count variation and the number of months used to calculate the smoothed index on the overall reliability of the Alerts generated (Chapter 2).

2. Explore the various options for producing assemblage Alerts (Chapter 3).

3. Chapter 4 presents a way of graphically depicting Alerts.

2. ASSESSING THE IMPACT OF MISSING DATA, VARIABILITY IN COUNTS AND THE NUMBER OF MONTHS A SPECIES IS INDEXED OVER ON WeBS SITE ALERTS SYSTEM

2.1 Introduction

The Alerts system has been developed as a means of highlighting large changes in the size of terrestrial and waterbird populations by smoothing out short-term fluctuations in population trends in order to reveal the underlying long-term population trend. The Alerts methods have been tested (Gregory *et al*. 1999) and have been implemented for waterbirds at different spatial scales (Atkinson *et al*. 2000). However there has been little critical attention paid to the impact of missing data (which has to be imputed), the inherent variability in the count data and the number of months used to produce the smoothed index, on the reliability of the Alerts system.

At present, the same degree of smoothing has been applied to all species despite the fact that different species may inherently show different levels of seemingly stochastic variation in the time series, e.g. Wren *Troglodytes troglodytes* population indices are likely to be far more variable than, for example, Mute Swan *Cygnus olor* populations.

Coupled with the inherent variation shown by different species, the proportion of missing counts and the number of months over which the species is indexed will also affect the variation in the smoothed index values. For example, stochastic factors might be predicted to cause more variation in an index calculated over two months than one calculated over a longer period. As the sample size increases the relative variation in the index is therefore likely to decrease. In a similar manner, extremely high or low counts will have a greater impact on the index as the proportion of missing counts increases.

Given the non-parametric nature of the General Additive Models (GAMs) used to smooth the WeBS counts it is extremely difficult to understand analytically the relative impact that missing counts, inherent species variability and the number of months over which a species is indexed will have on the reliability of the WeBS Alerts system.

To understand the impact these factors and their interactions have on the smoothed index values, we have used a simulation approach to determine the relative impact of each of these measures on the overall reliability of the WeBS Alerts. Once these are understood it will then be possible to make recommendations regarding the acceptable level of missing counts in the dataset and the number of months over which Alerts should be calculated. Using a simulation approach rather than actual WeBS counts allows us to (a) remove site-specific variation, (b) impose predetermined declines in the population mean and (c) remove differences in count variance resulting from population sizes at different WeBS sites. In summary, simulation allows much more flexibility to test the impacts of the different variables on Alerts.

Ideally High Alerts (a -50% change, referred to as ‘Alert’ throughout the remainder of this report) will only be generated whenever a decline greater than 50% occurs and not by smaller decreases in abundance. However this is unrealistic as there is variation associated with counts and this is a greater issue with Site Alerts as site counts are inherently more variable than counts from a larger series of sites. For example, a population that is showing an overall long-term trend of -50% over 25 years will show annual variation around this mean and thus it is possible that Alerts may be generated when the decline in the population mean is less than the critical 50% threshold and vice versa. Within an Alert system it is desirable that the number of these 'False Alerts' are kept as low as possible. There are two main mechanisms through which these can be generated.

i. Due to the variance between monthly counts, Alerts may be triggered by chance even if the long-term trajectory shows no trend or a trend of less than -50% over 25 years. For example, if a species shows highly variable counts, it might be expected that in some years some very low counts would appear and in other years extremely high counts would appear. When calculating the population change over 25 years, an Alert may be triggered if a high count appears at the beginning of the series and a low one at the end. This type of False Alert would be characterised by a series of sporadic Alerts which may be triggered in some years and turned off soon after they are triggered. Highly variable counts may be obtained if counting is difficult (e.g. large Knot flocks can be easily missed or double counted on large estuaries), or could be a factor of the species biology. For example, some species may show nomadic tendencies within a winter or only occur irregularly on some sites in cold winters.

ii. The term False Alerts and the term False Alert is probably better suited to the situation of missing counts. In the Alerts methodology missing values are imputed using month, site and the smooth function of year factors. If, for example, a change of -49% is calculated using a full set of data, missing data in the first or last few years of the 25 year time series will be imputed. As these imputed values are estimates with corresponding confidence limits, it is quite conceivable that the imputed data will push the change to -50%, thus triggering an Alert. Alternatively, sets of data containing imputed values may not trigger an Alert even though the full data set would. When the population change with the full dataset is near the critical Alert threshold (-50% for High Alerts or -25% for Medium Alerts) the sensitivity of the models to generating False Alerts is higher. Intuitively, species which show more variation in monthly counts are likely to suffer more from this type of problem.

An understanding of the limitations and strengths of the methods used is essential to any system that highlights changes in waterbird populations. This chapter explores the effects of missing counts and species variability on Alerts and examines the number of months that should, from a statistical rather than biological point of view, be used to produce the smoothed index. For ease of interpretation, any following reference to the terms Alerts refers to a population change of more than -50% over 25 years. However the points made will also apply to other types of Medium and High Alert triggered over different time periods.

2.2 Methods

2.2.1 Power Laws and simulating WeBS data

The impacts of missing data on the WeBS Alerts may be investigated either by using complete time series data collected from WeBS sites or by simulating a data set which behaves in a similar fashion to real data. As the number of sites with complete data sets is small, simulating WeBS count data simulation was considered the better technique. Simulating data also allows the modeller to inflict precise declines on the data and explore the general principles behind the elasticity of the Alerts system while omitting any site specific factors that may influence existing WeBS counts.

Ecologists have long been modelling time series of populations and are often interested in the fluctuations that occur both within (temporal) and between sites (spatial heterogeneity). Power laws (Taylor 1961, Equation 1) often provide a remarkably good way of describing a vast range of natural populations by relating variance to mean population size. During the 1980s and 1990s this spawned a whole industry in looking at the properties of the power law relationship and its ecological implications for natural populations. For the majority of species, the power-law scaling parameter (α) lies between 1 and 2 and often lies at one extreme or other. Where $\alpha = 1$ the variance equals the mean and counts follow a random distribution around a mean as described by the Poisson distribution. As $\alpha \rightarrow 2$ the variance behaves more like the square of the mean indicating a clumped or contagious distribution. In the WeBS Alerts context, this means greater stochasticity in terms of population change between time periods for those species with higher counts.

$$V \propto M^{\alpha} \qquad \text{(Equation 1)}$$

Where V = sample variance, M = sample mean and α = scaling parameter.

The power-law theory can be applied to simulating WeBS data. At high population sizes WeBS data may be described reasonably using a negative binomial distribution which requires a known mean and variance. For example, Figure 2.1 shows the observed distribution of Knot counts on the Wash and Curlew on Morecambe Bay. For Knot there is a remarkably good fit between these two distributions and a reasonable one for Curlew. These are just two examples but the same pattern was found to hold across a wide range of species and sites.

To simulate a set of WeBS data, we needed to determine the power law relationship for some example species. For the purposes of this report, we have used wader species but the same principles will apply to other species of waterbird. We used the three standard WeBS months for waders (December, January and February) to calculate a series of means and variances for each winter on all SPAs that were designated for a particular species (e.g. Dunlin in Figure 2.2). Each point represents the mean and variance of the three monthly counts in one winter. A linear regression line was fitted through the data, the slope of which is equivalent to α in Equation 1.

To simulate a time series of WeBS data it was necessary to select a mean and calculate a variance using M and α in Equation 2. As there is error associated with the slope α we attempted to include this by randomly selecting residuals from the mean-variance regression based around the magnitude of the mean. These were categorised by the integer values of the mean. For example, if ln[Mean] equalled 8.3, then a residual was picked from at random from all those points where $8<\ln[\text{Mean}]<9$. The variance was then calculated and a value picked at random from the negative binomial distribution using the RAND random number function in *The SAS System 8e* (2001).

$$V = e^{\alpha \ln[M] + c + r} \qquad \text{Equation 2}$$

Where V is the variance, M the mean, α the scaling parameter, c the intercept of the log-mean and log-variance plots and r is a randomly selected residual associated with the variance-mean plot based around the magnitude of the mean.

2.2.2 Generating Alerts using simulated data

To generate Alerts the General Additive Model is fitted through the simulated data using Equation 3.

Smoothed Count = a + site + month + S(year)   Equation 3

Where *a* is a constant, *site* is a factor calculated for each site, *month* a factor calculated for each month and *S(year)* is the smooth function of year. The WeBS Alerts Report I (Atkinson *et al*. 2000) provides a full methodology.

To generate a series of indices a constant percentage change was inflicted on the index so that the species population mean declined by 50% (or the level of decline we wished to impose) over 25 years. A series of smoothed indices was calculated for a 25 year period starting with a population twice the size of the 1% international importance level for each species. The mean population size after 25 years was therefore equivalent to the level above which a site is designated for that species. We used this level as (a) we wanted to use biologically realistic population levels and (b) the species mean / variance relationship is calculated over the range of population sizes found on SPAs and predicting this relationship outside the natural boundaries was not desirable. For example, a population size of 10,000 may be appropriate for Knot on SPAs but not for a species such as Avocet.

To determine the elasticity of the Alerts methodology a series of 999 simulated indices were calculated for each different set of parameters used in the simulations. Each simulated trajectory was smoothed using a GAM and the 25 year population change calculated between the smoothed index for year 25 and year 1. Alerts were triggered when this change exceeded -50%.

2.3 Results

2.3.1 Variation between species

Figure 2.3 shows the mean coefficient of variation (CV) for the December to January WeBS counts made at all UK SPAs for which that species was designated. A mean CV was calculated for each SPA and a mean of these values is presented. This gives an impression of the degree to which monthly counts vary within the UK sites important for each species.

There are few surprises in the extremes. Turnstone showed the least variation and Black-tailed Godwit the most. The two highest extremes (Black-tailed Godwit and Avocet) may be explained as both these species occurred on most present SPAs in extremely low numbers at the beginning of the wader counts in 1969 and thus are implicitly more likely to show higher values - i.e. a change from 50 to 100 birds is a 100% change whereas a change of 1000 to 1050 is only a 5% change.

To determine the effects of count duration (number of months used in the index) and missing data, it was decided to pick three species which showed low, moderate and high amounts of inter-annual variation in counts. Respectively, these were Turnstone, Dunlin and Knot. Although for convenience, we use these species names they can be thought of as surrogates for three species that show different levels of variation. It was decided not to use Avocet or Black-tailed Godwit as these species had shown very large population increases and given changes of over 500% over a 25 year period on many of their respective SPAs, Alerts would be generated even with reasonably high amounts of missing data.

2.3.2 Effects of missing data and number of months with a simulated decline of zero over 30 years

Figure 2.4 shows the frequency distribution of 25 year change measures calculated from 999 bootstraps for Knot, Dunlin and Turnstone when no change in the population was imposed over 25 years for (a) indices calculated over three months and (b) over six months. As expected Turnstone shows relatively little variation in the change measure with no missing data. Over 90% of the change measures lie between -20 and +20% when there are no missing data. As the level of missing values increases, the bars on the graph spread out and, when 50% of data are missing, 90% of the change measures vary between -30 and +30. As expected the number of change measures that couldn't be generated also increased as all three monthly counts were missing for some of the simulations. Similar patterns were observed for Knot and Dunlin but, given the higher degree of natural variability in these counts, the spread of the change measures is greater. In fact, Knot would be expected to trigger High Alerts even with no missing data. When the proportion of missing data reaches 50%, change measures range from -80 to over +200. Species like Knot are considerably more likely to trigger False Alerts.

As Alerts deal with percentage changes, the distribution of the calculated change values is non-normal and is skewed with a longer tail towards the right hand side of the graphs. This is simply due to the fact that it is possible to get changes of over +100% but not less than -100%.

The effect of increasing the number of months over which the species is indexed was as expected (compare Figures 2.4 a and b). Given a longer sampling duration in terms of the number of months used and therefore a larger sample size, the relative variability in the total counts was reduced and thus the bars on the graphs were correspondingly closer together than when three month totals were used. The relative impact of the number of months and the amount of missing data on the standard deviation of the change measure is shown in Figure 2.5. For Turnstone the effect is small and there is an almost linear response, with small increases in the standard deviation corresponding to an increased amount of missing data and a reduction in the number of months. As natural variability increased, Dunlin showed relatively larger responses and Knot showed some large non-linear increases in the standard deviation at the highest amounts of missing data and lowest number of months. Thus, highly variable species will generate more False Alerts when the proportion of missing values is greater and when the species is indexed using fewer months.

2.3.3 Number of False Alerts generated

2.3.3.1 False Alerts triggered due to high variability in monthly counts

If a species has a simulated decline of 50% over 25 years then, given the inherent variability associated with the counts, it would be expected that an Alert would be triggered in approximately 50% of cases. As the simulated decline increases to say 60% we would expect Alerts to be triggered in a higher proportion of cases but this proportion would vary between species depending on the amount of variation associated with the monthly counts. Figure 2.6 shows that for Turnstone, Alerts would be triggered in over 90% of cases once the simulated decline was above 60% and that increasing the number of months over which the population was indexed would slightly increase the confidence one could have in an Alert. Conversely, Knot shows a much shallower response indicating that, compared with Turnstone, more Alerts would be triggered at simulated declines of less than 50% and fewer at declines of greater than 50%. The effects of the number of months used to index a site has a proportionately greater effect in Knot than Turnstone. If the population is indexed over two to three months then Alerts could be triggered when the simulated decline is zero, albeit in very small numbers. When four months were used no Alerts were triggered when the simulated decline was zero.

Dunlin showed an intermediate response and False Alerts were only triggered when the simulated decline was >20% using two months and >30% when five months or more were used.

2.3.3.2 False Alerts triggered due to missing data

Both the amount of missing data and the simulated decline in the population trend have implications for the number of False Alerts triggered (Figure 2.7). These False Alerts are those that are triggered when the Alerts status using the full dataset is different from that after having removed some counts at random, thus simulating missing data points. These include false positives (missing values cause an Alert, when the full dataset does not) as well as false negatives (missing values do not cause an Alert, when the full dataset does).

The higher the proportion of missing data, the higher the proportion of False Alerts, although at higher simulated population declines (e.g. a decline of 80%) missing values have a reduced effect. This is due, in part, to the smaller variance associated with a smaller population size i.e. as populations decline to low levels counts relatively less variation. This figure shows the decline required to be certain of generating an Alert may have to be in the order of 70-80% for species which show high variation in monthly counts and with high levels of missing data.

As expected, the highest proportion of False Alerts occur around the 50% decline where uncertainty associated with missing values is greatest. The proportion of False Alerts is approximately 20% for each species but drops away either side of this. False positives [i.e. ALL DATA: No Alert, MISSING DATA: Alert triggered] tend to be generated more often with missing data than false negatives (For example see Figure 2.8) but the number of these are drastically reduced when six, rather then three, months were used to produce the index. This errs on the side of caution when triggering Alerts and is consistent with the precautionary approach, i.e. it is better to trigger a false positive rather than a false negative.

The effect of the variation in monthly species counts also becomes more apparent with an increase in the number of months used to index the population (Figure 2.7). Turnstone generally showed <5% False Alerts at simulated declines greater than 60% or smaller than 40% when using six months of data and the zone of uncertainty is generally small. Although the likelihood of False Alerts is also reduced with an increase in the number of months used to index the population, for Knot, there is still a larger number of False Alerts generated at moderate levels of missing data compared with Turnstone. However these generally occur on fewer than 10% of occasions when declines are between 30% and 70% at moderate levels of missing data (25%).

False Alerts, both positive and negative are likely to be characterised by a series of intermittent Alerts during periods when few Alerts are raised. If a species is in true population decline Alerts are likely to be raised in consecutive years during the period of decline.

2.4 What Level of Missing Data can be Expected on Spas and How Many Months Should be Used to Produce the Smoothed Indices?

Coverage of UK SPAs has increased since the start of WeBS counts (Figure 2.9 a-h). The proportion of sites that have zero counts has declined from approximately 35% in 1966 to 5-10% in the late 1990s (Figure 2.9b). Over the same time period the number of sites with a full series of seven WeBS counts (September to March) increased from 25% to 60-70% (Figure 2.9h). The majority of sites are now fully counted and, during the 1990s, 78% of sites had coverage of six counts or more. Overall coverage has increased and in the late 1990s there were less than 20% missing values across all sites (Figure 2.10). Given various combinations of months used to index populations there is relatively little difference in terms of missing values between using two or seven months to produce the index (Figure 2.11).

From a statistical point of view, the more months that are used to produce the index, the more reliable the results. However, it is necessary to balance this with knowledge about individual species' biology. For example, many of the waterbird species may not occur at certain sites during some of the standard WeBS months. Alternatively, individuals from different bio-geographical populations may pass through or arrive on sites at different times.

This methodology may be fine for a national or even regional system as, even when all counts are missing for a particular site in a particular year, the imputed counts are based around some real data from that year - i.e. the estimated count is equal to the *year* factor (calculated from other sites in that year) plus the mean *month* factor value for the individual site and the mean *site* factor. For Site Alerts however, the year factor could not be estimated from other surrounding sites and a mean value for all three factors would be used. This is not ideal as the index value will be calculated from data collated completely outside that year. In addition, the months used to calculate the Underhill index were selected using the national dataset and individual sites may well differ in terms of the monthly pattern of occurrence.

In these simulations we have not included a month factor and have just used data from the standard three wader months used in calculating the national wader indices. We took this decision as we wanted to look at the overall principles in understanding just how elastic the Alerts methodology was rather than to concentrate on large numbers of individual species scenarios.

In terms of the future running of SPA Site Alerts there are two options. The first is a species and site-specific approach, thus selecting the months used to produce the index on a site by site and species by species basis or a blanket approach using the same series of months. For simplicity, the latter approach is preferable, especially as the way waterbird populations behave may change in time, thus rendering the first approach invalid.

The need to balance species biology and statistical reliability is therefore something that needs to be addressed, both when producing the national Underhill indices but also when calculating Alerts. The methodology used here would provide a good basis for future analyses of Underhill indexing and Alert elasticity. The general principle behind Alerts is looking at changes in *usage* and as such it needs to be decided whether it really matters if the Site Alerts system covers different populations of the same species using a site - the emphasis is on site quality rather than monitoring changes in a particular bio-geographic population which may be the aim at a regional or national scale. The separation into different populations could be performed at the interpretative phase.

In summary, we would recommend that Site Alerts are calculated using no less than three months of data, and preferably more so that peak counts on all sites will be included. When determining the length of the indexing period, site coverage, ease of sampling (i.e. using the same months for all species) and the statistical reliability of the indices generated should be taken into account.

2.5 Conclusions

2.5.1 Variance in monthly counts and the impact on Alerts

- Individual species differ in the amount of variation displayed by their monthly WeBS counts. A power law relationship between the mean and variance of these counts allowed species-specific data to be generated and the impact of this variance on the generation of WeBS Alerts to be investigated.

- Increased monthly count variance led to increased variance in the 25 year population change measures calculated from 999 Alert simulations for each species. Increased variance lead to a greater proportion of Alerts being triggered at population declines of less than 50% and fewer at declines of greater than 50%.

- Increasing the number of months from which the smoothed index was calculated reduced the variance in the change measures and, thus, the number of False Alerts.

2.5.2 Impact of missing values on SPA Alerts

- False Alerts can be generated by missing data. False positive Alerts are generated when the full dataset does not trigger an Alert but the set with missing data does. False negatives are the converse.

- Missing values increase the expected variance in the change measures calculated by the Alerts system.

- For species which show little natural variability in counts, low to moderate proportions of missing values (<50%) have a relatively small linear effect on the variance associated with the change measures. The influence of missing values on the reliability of Alerts triggered for species displaying higher levels of variability in their monthly counts is greater, with the number of False Alerts increasing in a non-linear fashion as the proportion of missing values increases.

- Increasing the number of months over which a species is indexed lessens the impact of missing data.

- The highest incidence of False Alerts occurred when populations were declining at a rate of 50% over 25 years, assuming 50% is the Alert triggering point. For these populations, when 50% of data points were missing, 20% of the Alerts generated were False Alerts. At lower levels of missing data (e.g. 10-20%) this figure was reduced to less than 10%.

2.5.3 Count coverage on SPAs important for waterbirds

- The proportion of sites with the full seven WeBS counts during a winter has increased from 25% to approximately 65-70% between 1966/1967 and 1999/2000. The proportion of sites with no counts decreased from 35% to approximately 10% over the same time period.

- Approximately 15-20% of sites are counted on six out of seven occasions per winter. Sites therefore tend to be almost or fully counted or not counted at all during the winter period. Relatively few are counted between two and five times a winter.

- Between 25-30% of the monthly values to be used for the SPA Alerts for the period 1995/1996 to 1999/2000 were missing. Projecting ahead for a further five years, this figure would reduce by approximately 5% assuming coverage does not change from current levels.

- By omitting the sites which are not counted at all, the proportion of missing values is much lower and, on average, less than 15%.

- Increasing the number of months used to produce indices will increase the number of missing values slightly. However, using September-March counts rather than December-January values will only increase the number of missing counts by approximately 5%.

- Missing values are therefore not thought to be a major problem for the generation of Alerts for the vast majority of SPAs.

2.5.4 How many months should be used to calculate indices and hence Site Alerts?

- Site Alerts are to be generated for SPAs and SSSIs with waterbird interest. Coverage is generally high, at least for SPAs. There are relatively few SSSIs with waterbird interest lying outside of the SPA network.

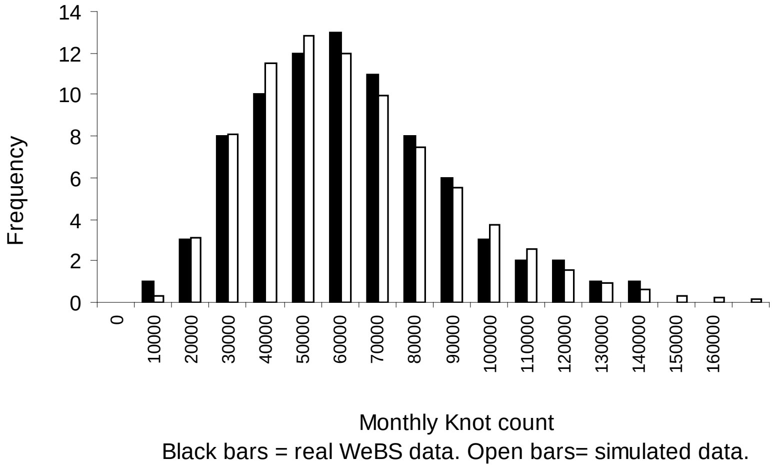
- In theory, the greater the number of months used to calculate the Alerts, the lower the variation in the change measures calculated. This reduces the probability of Alerts being generated by chance.

- Natural variation in counts of individual species will also impact on the choice of months used. For those species showing little monthly variation in counts, increasing the number of months will have little impact on the change measures and hence on the Alerts generated. For highly variable species, increasing the number of months over which populations are indexed will reduce the probability of Alerts being generated by chance.

- Species biology is also important and has not been addressed in this report. Increasing the number of months to the maximum of seven may incorporate several different biogeographic populations of a particular species of waterbird as it may cover both passage and wintering periods. For particular species or SPAs it may be necessary to tailor the months used to correspond to the occurrence of the population of interest.

- Conversely, for Site Alerts (as opposed to regional or national Alerts) it may be desirable to use a 'total usage' measure for the entire winter, irrespective of whether birds are on passage or wintering. It may be desirable to use a longer series of months for Site Alerts and use the standard Underhill months to calculate the annual indices. Further research into this topic is strongly recommended.

- However, due to the likelihood of missing values and the variability of species population trends, it is strongly recommended that **at least three months** worth of data are used to generate Site Alerts, and preferably four to six. Even using three months, some highly variable species may generate a small proportion of Alerts by chance, even when the simulated decline in the population is zero.



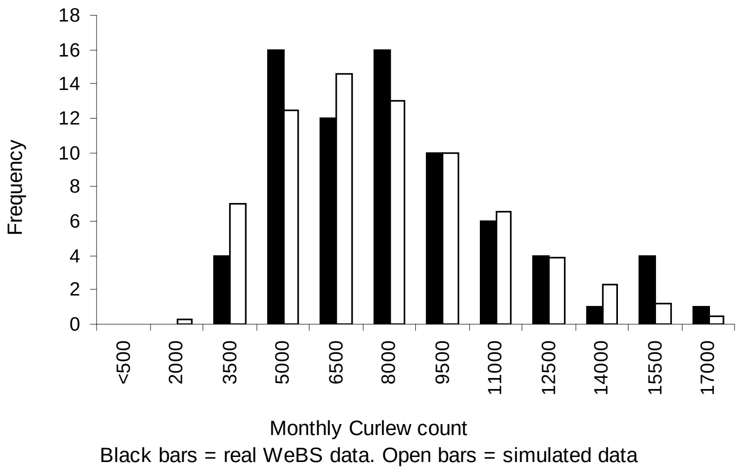


Figure 2.1 Frequency distribution of actual (black bars) and simulated WeBS counts (open bars) for December, January and February counts between the winters 1969/70 and 1999/2000. (a) Knot counted on the Wash (mean = 57,369, variance = 801,040,738, n=74) and (b) Curlew counted on Morecambe Bay (mean = 8032, variance = 10,430,445, n=81).

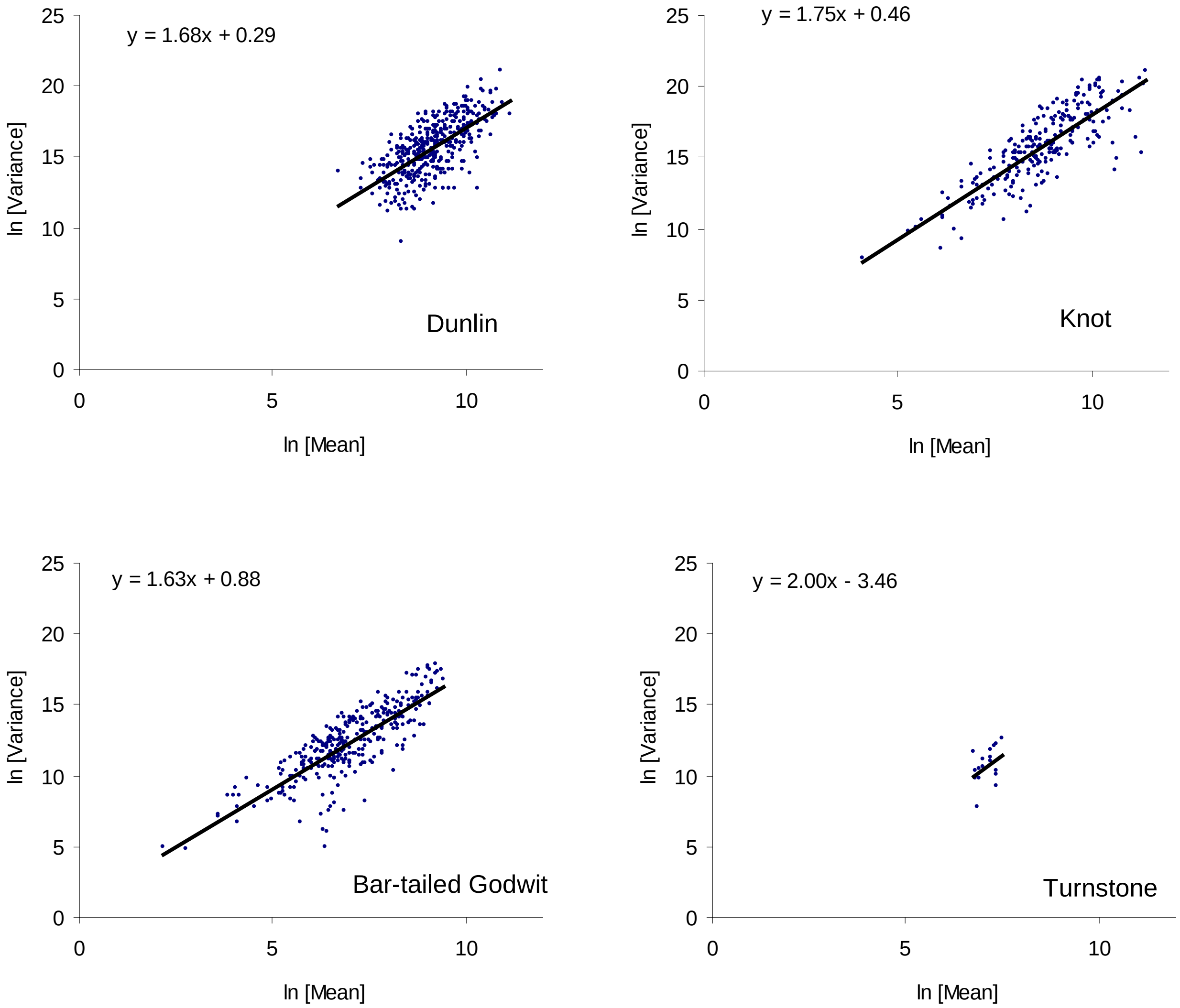


Figure 2.2 Mean-variance relationships for four selected species of waders. Each point represents the mean and variance of one winter's worth of December, January and February counts for all the SPAs for which the species has reached qualifying levels.

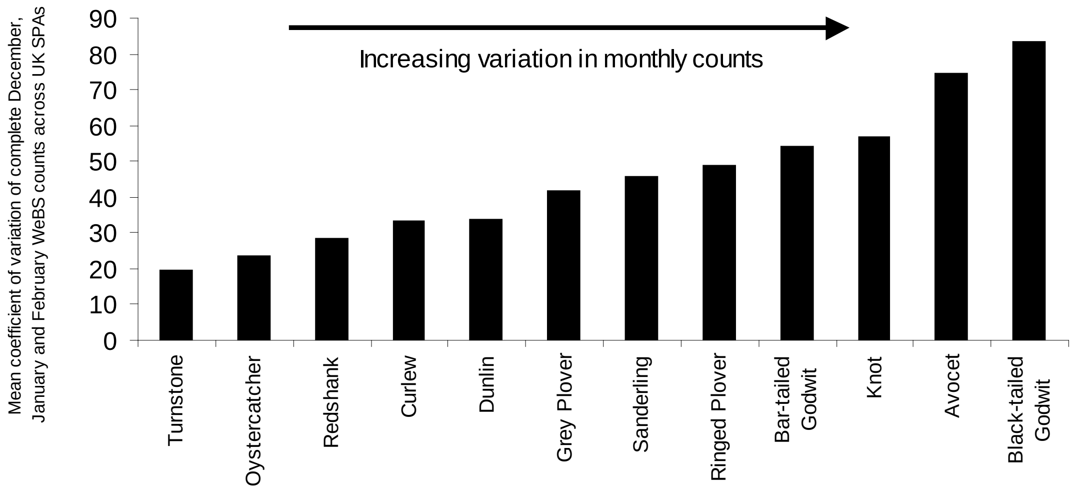


Figure 2.3 Coefficient of variation in December-January counts for the 12 regularly indexed wader species at all SPAs designated for each species. The y-axis value represents the mean of the mean values calculated for each SPA and is designed to order species along a gradient rather than for any statistical purposes.

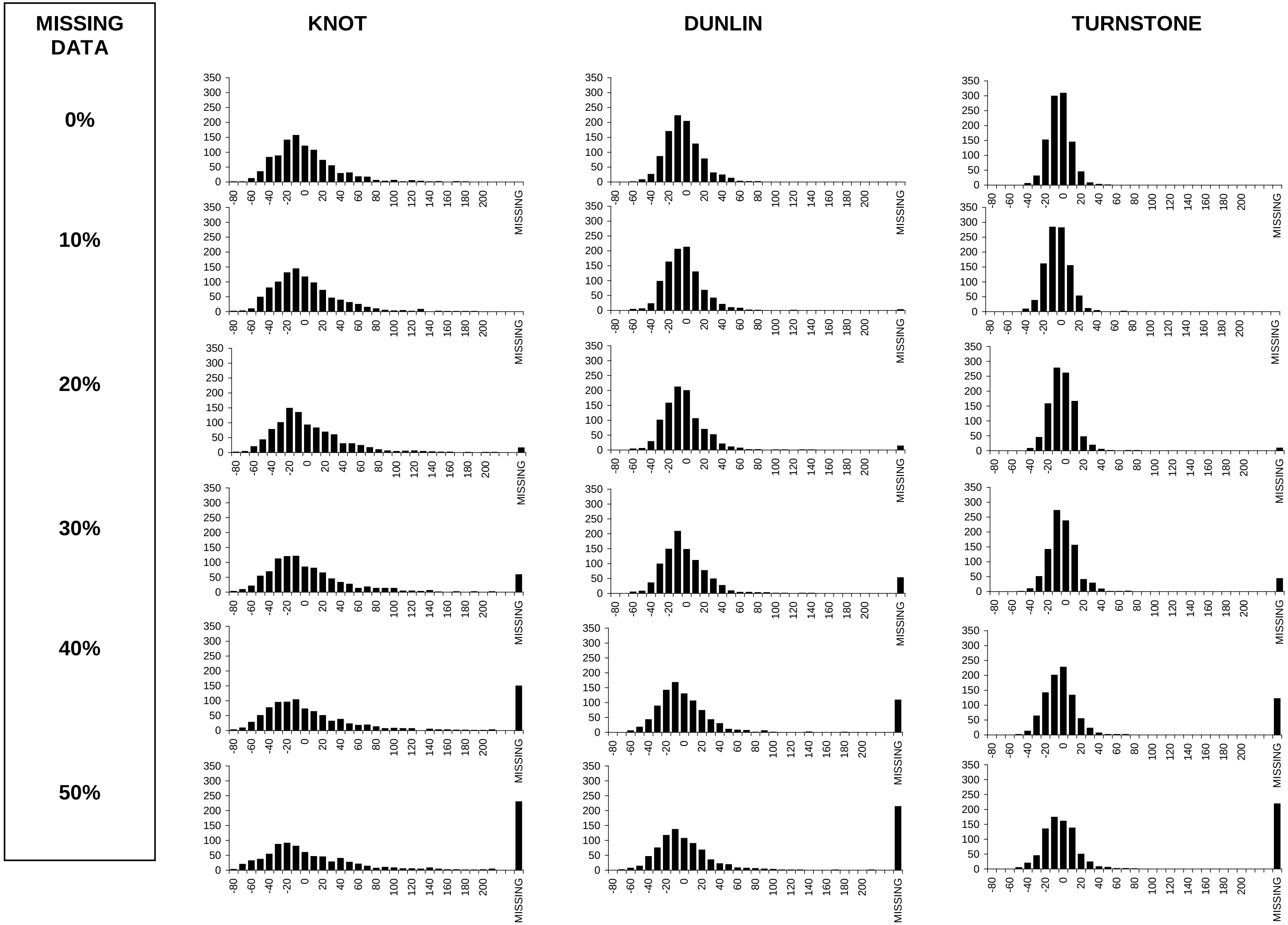


Figure 2.4 (a) Frequency distribution of the percentage changes in the smoothed index values over 25 years from 999 simulated trajectories for Knot, Dunlin and Turnstone with an overall mean of 0% change over time and with different amounts of missing data removed at random from the counts. These indices are calculated over three months. “MISSING” is the number of changes that could not be calculated as a result of missing values.

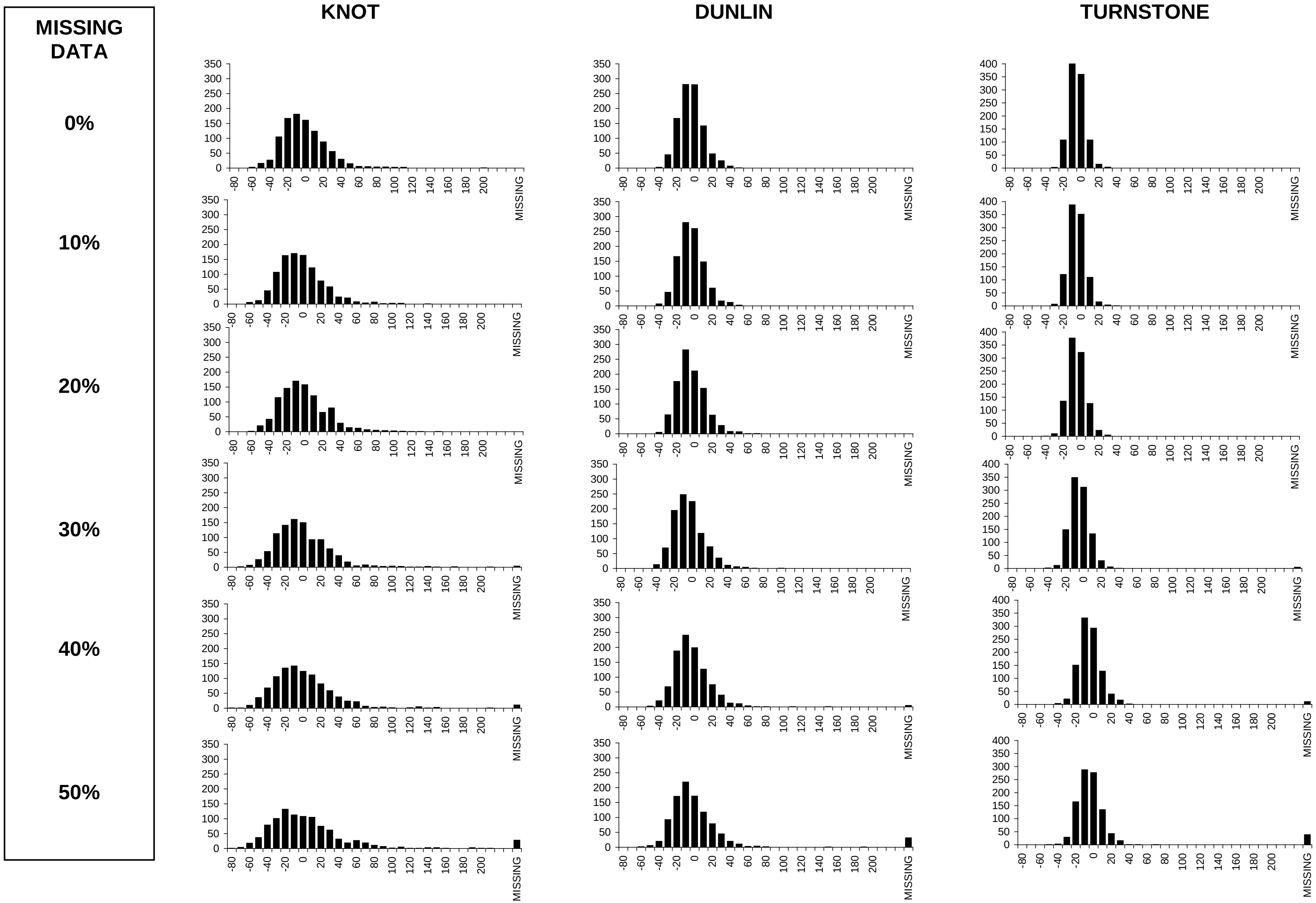
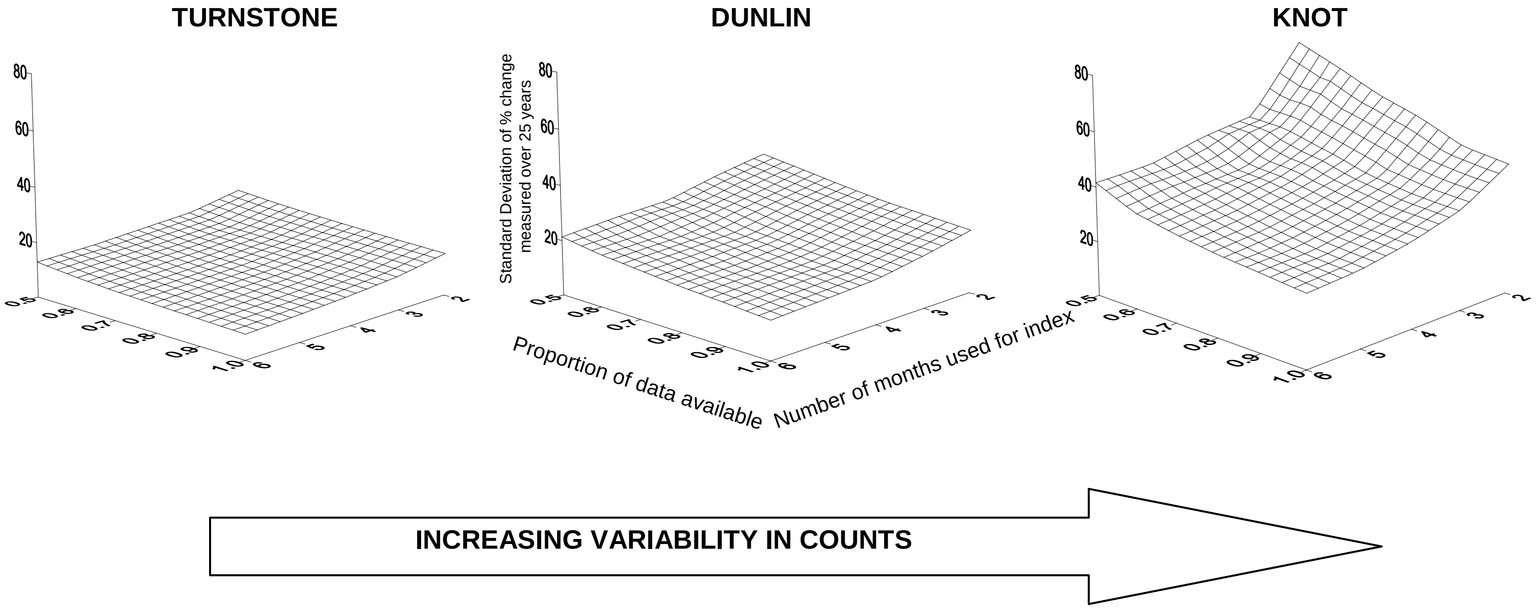


Figure 2.4 (b) Frequency distribution of percentage changes in the smoothed index values over 25 years from 999 simulated trajectories for Knot, Dunlin and Turnstone with an overall mean of 0% change over time and with different amounts of missing data removed at random from the counts. These indices are calculated over six months. “MISSING” is the number of changes that could not be calculated as a result of missing values.



Key Points: The three species have inherently different variability hence the difference in the intercepts. The more variable counts tend to be, the greater the impacts of missing counts and indexing over fewer months. With low variability in counts responses tend to be linear as in Turnstone. However, with Knot some non-linearity is observed when the proportion of missing counts is high and the species is indexed over few months.

Figure 2.5 Surface plots showing the interaction between the number of months used to produce the yearly index, the proportion of data available for indexing (missing counts: 1 = 100% of data available, 0.5–50% of data available) and the standard deviation of the change measure calculated from simulated data using standard WeBS Alerting methodology.

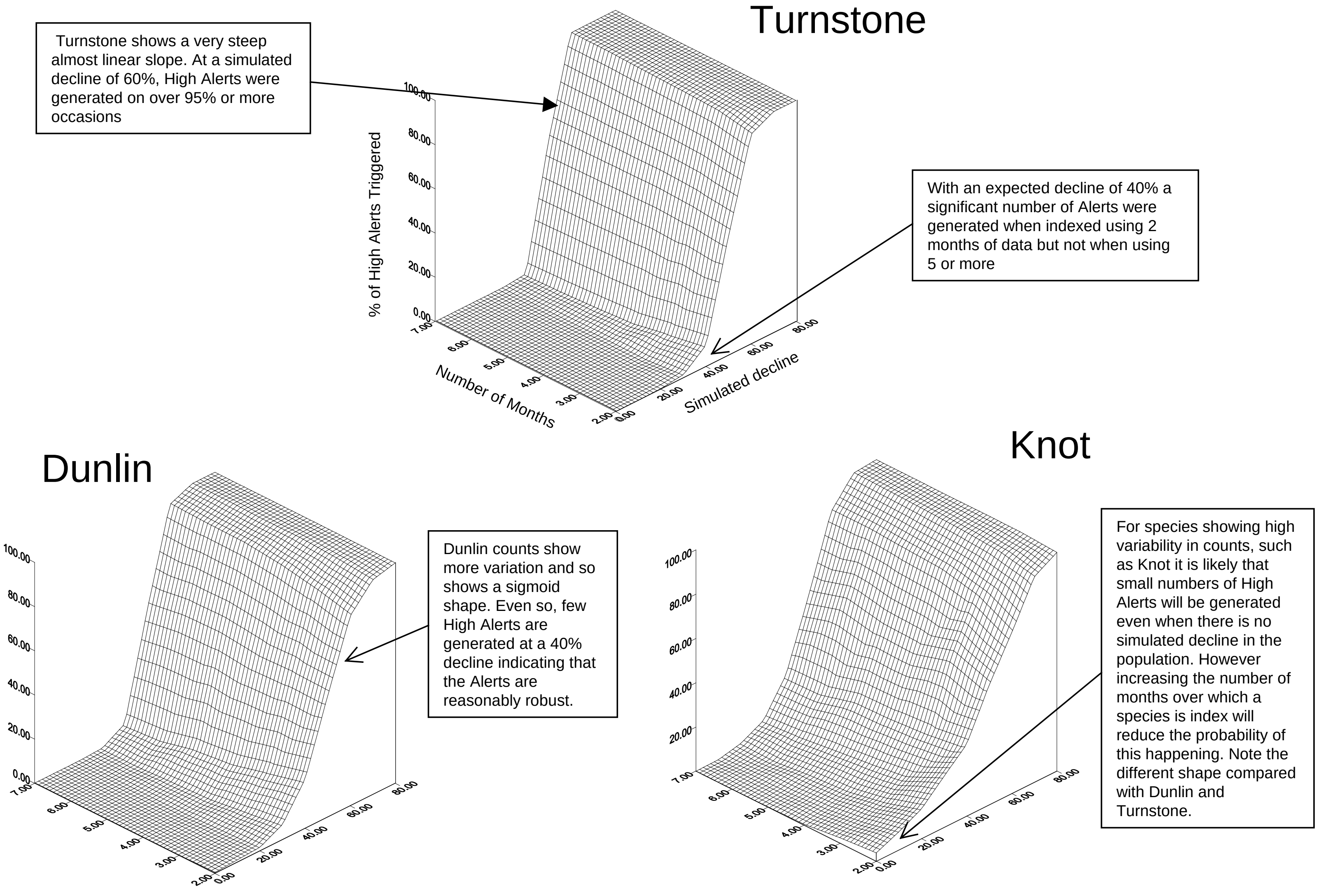


Figure 2.6 Impact of the number of months used to generate the smoothed indices on the number of False Alerts generated for species showing differing levels of variability in the monthly counts.

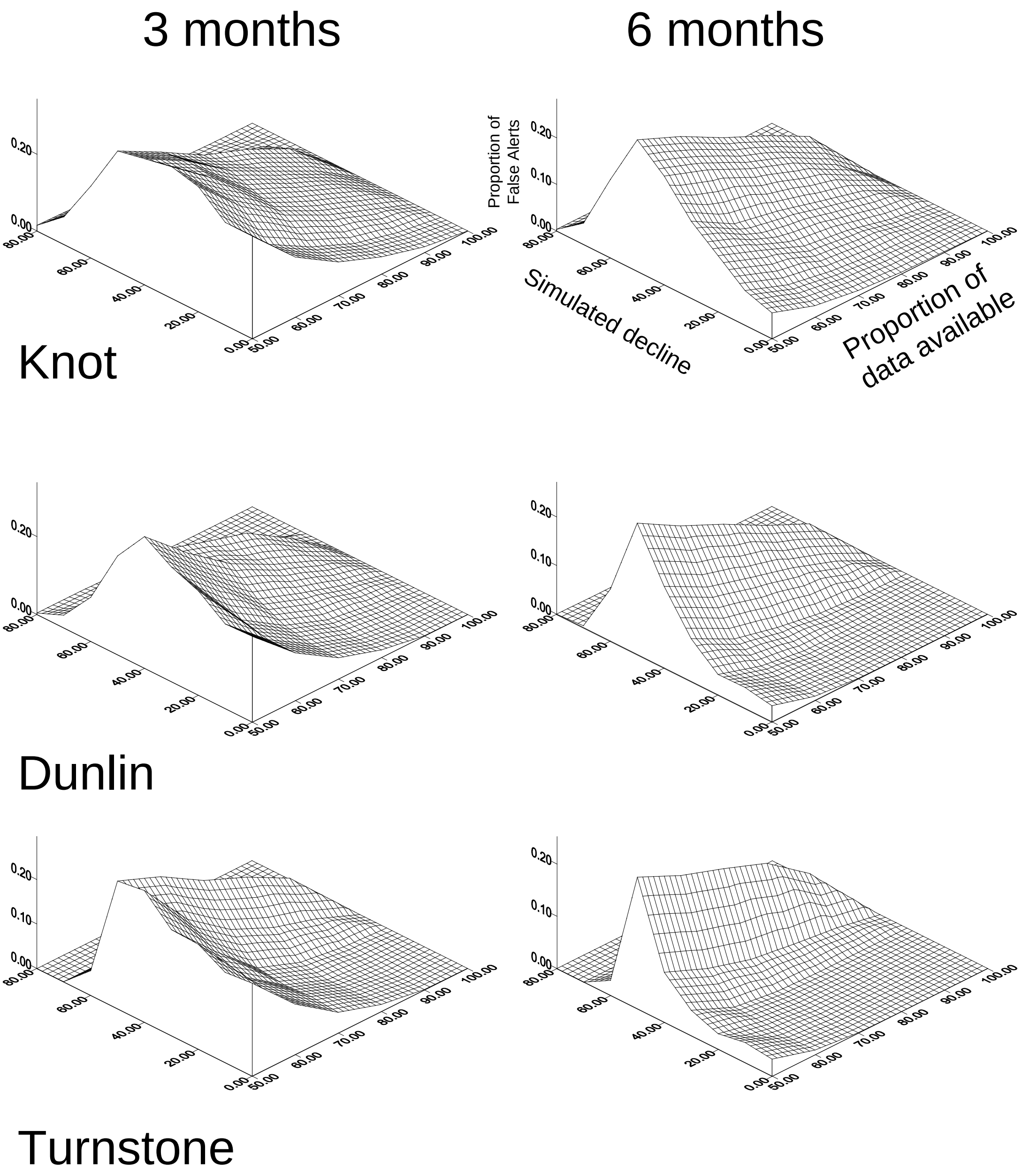


Figure 2.7 Effects of differences in the simulated decline, the number of months over which a species is indexed and the proportion of data available on generating False Alerts. At simulated declines of zero, False Alerts are likely to be triggered with increasing proportions of missing data when the number of months used in the index is small.

.

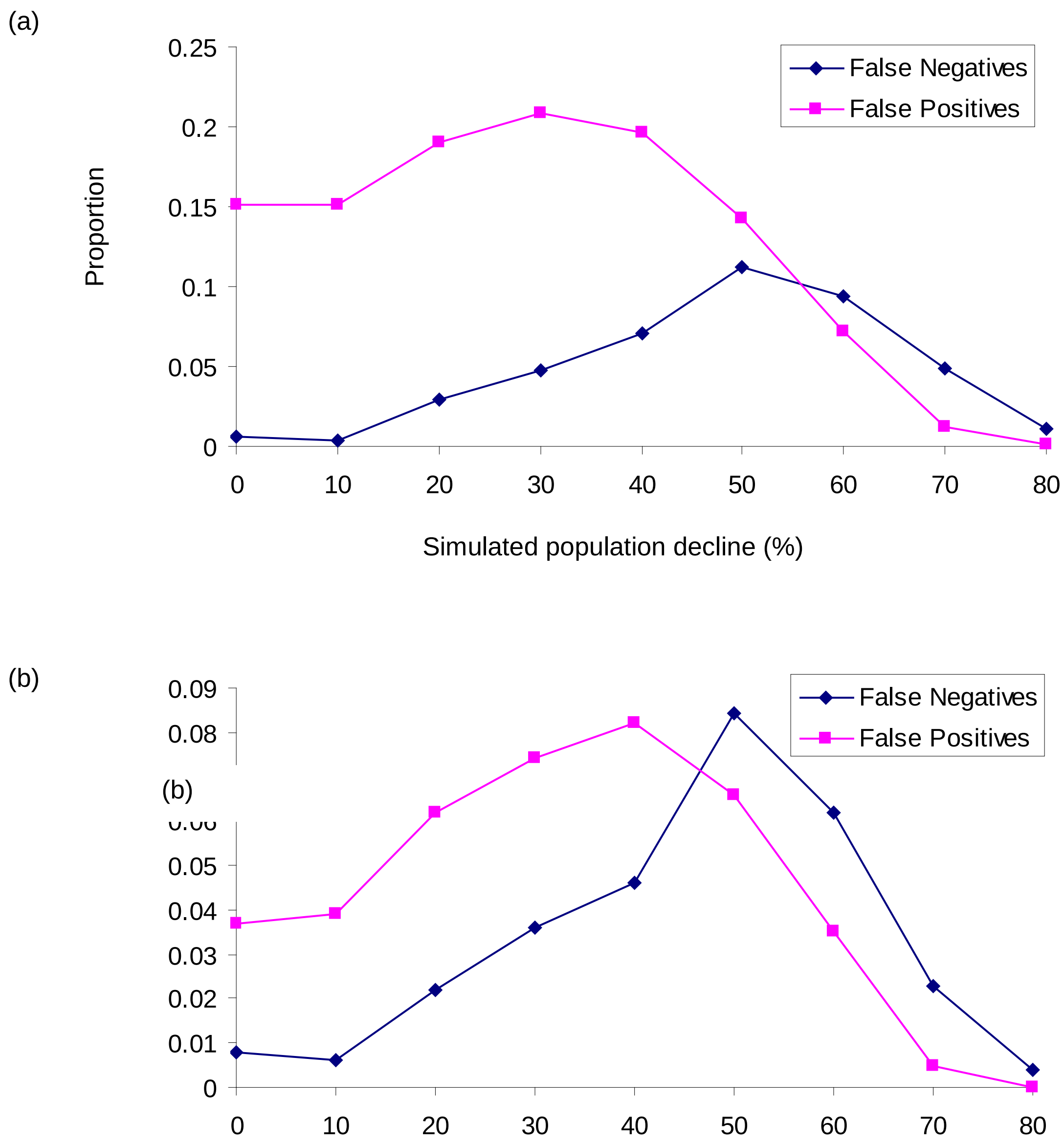


Figure 2.8 Proportion of False Alerts which might be expected for Knot with (a) 40% and (b) 20% missing data broken down by false negatives or positives. At low population declines more false positives are generated and more false negatives at higher declines

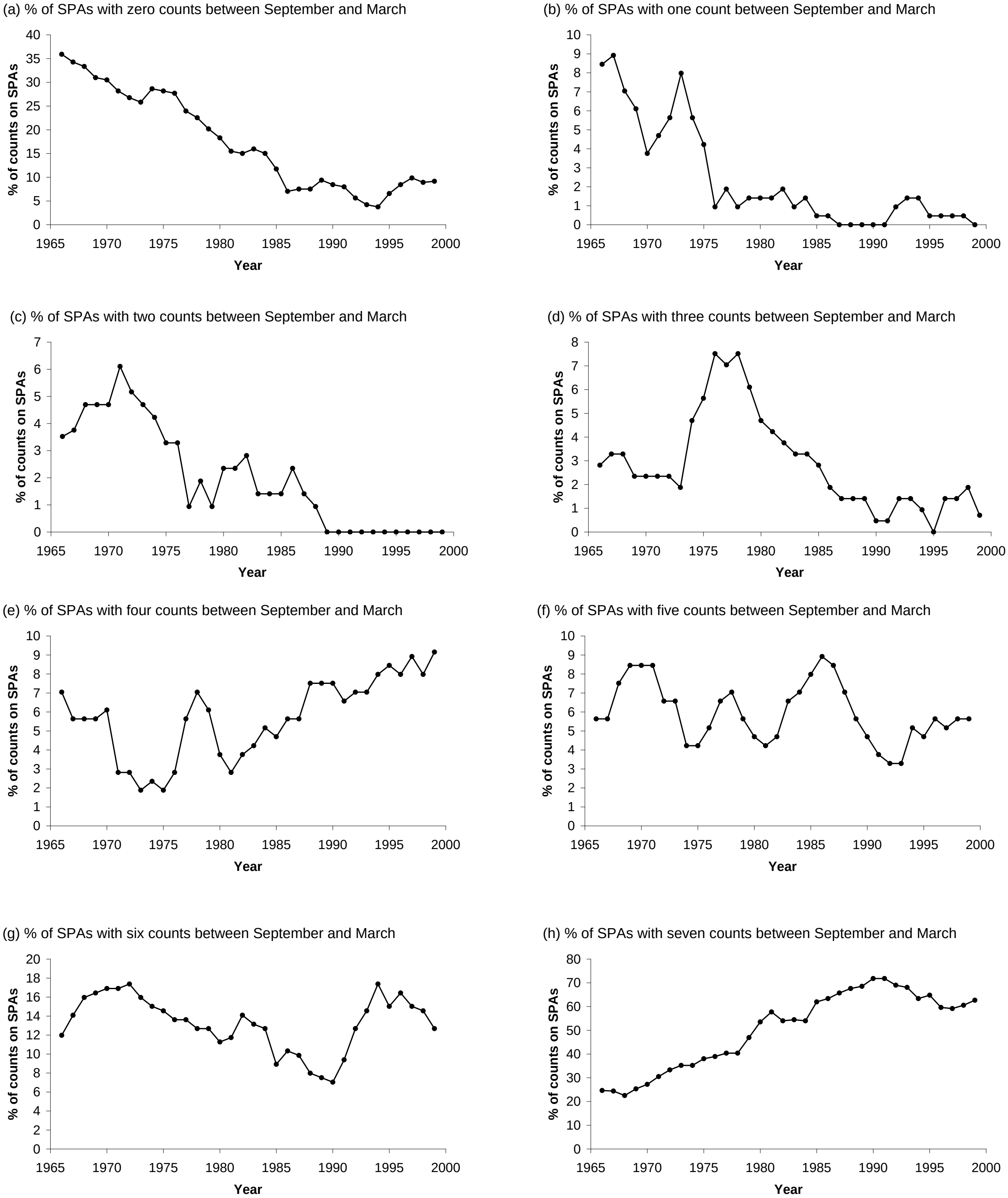


Figure 2.9 a-h WeBS core count coverage of Special Protection Areas in the United Kingdom. Each graph indicates the proportion of SPAs which were counted on zero to seven occasions between the standard WeBS core count months of September to March

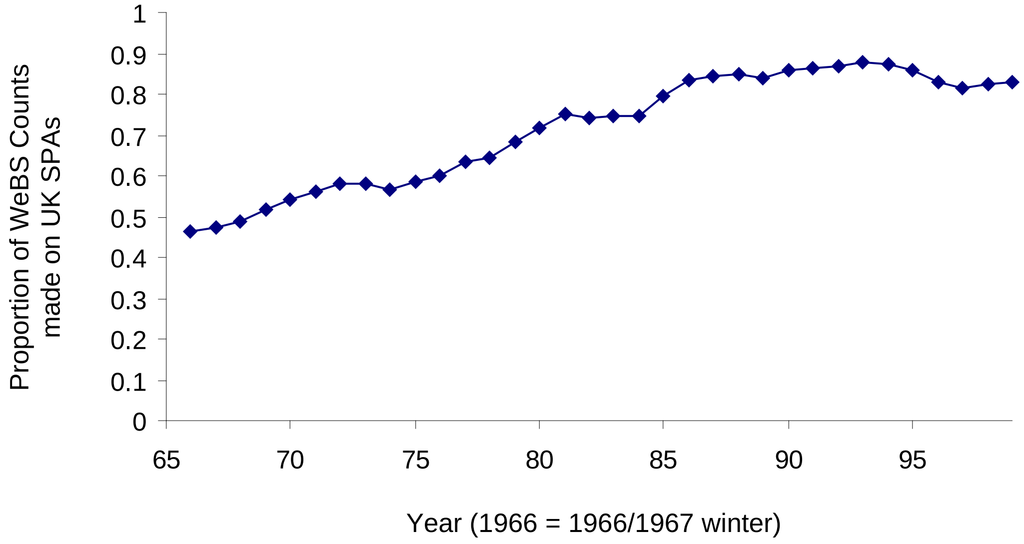


Figure 2.10 Count coverage on UK SPAs, indicated by the proportion of counts made between September to March in each winter. Coverage has increased from c. 50% in 1966 to 80-85% in the late 1990s.

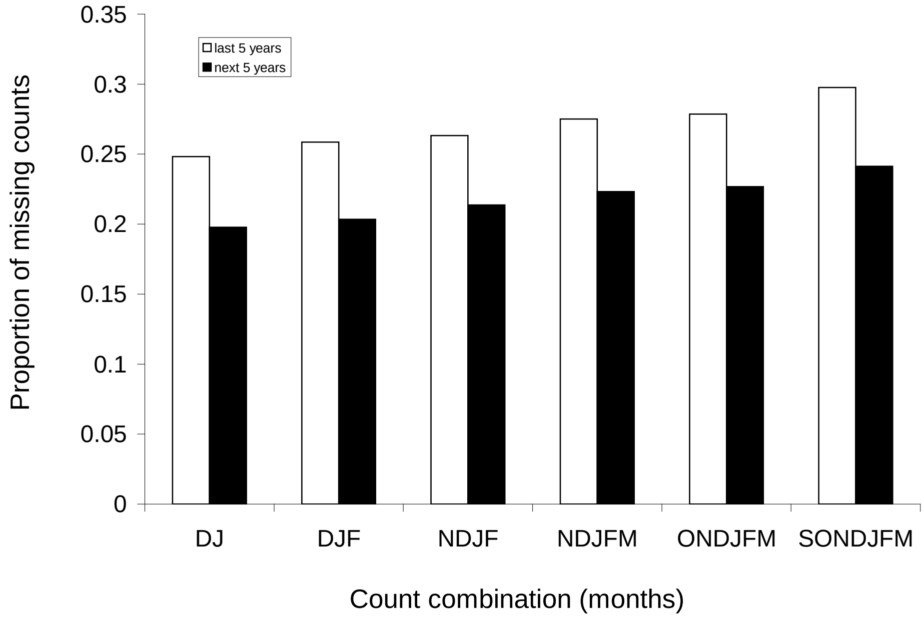


Figure 2.11 The proportion of missing counts in the WeBS Core Counts on UK Special Protection Areas for various combinations of count months. S = September, O = October, N = November, D = December, J = January, F = February, M = March. Open bars show the mean value for the period 1970/1971 - 1974/1975 to 1995/1999 - 1999/2000. Filled bars show the projected mean proportion of missing counts for five years on from 1999/2000 assuming that the coverage for the 2000-2001 - 2005/2006 remains similar to the previous five winters.

3 ASSEMBLAGE ALERTS

3.1 Introduction and General Considerations

SPAs may be designated if the number of non-breeding waterbirds exceeds a threshold level of 20,000 individuals. In theory, there is no statistical reason as to why the Alerts system could not be applied to monitoring assemblages but at present there is no standard methodology in place to calculate assemblage Alerts.

There are also several methodological problems associated with calculating assemblage Alerts. The 20,000 threshold used to designate important assemblages includes all waterbirds, including gulls and terns, which are not fully covered by WeBS, as counters have the option to count them or not. The species monitored by WeBS also differ in terms of the first year of the available data: wildfowl (1947, but 1966 is routinely used for indexing), waders (1969), Great Crested Grebe and Coot (1982), Little Grebe (1985), Cormorant and wildfowl in Northern Ireland (1986) and all other species since 1993. Calculating assemblage Alerts is thus not as simple or straightforward as might be thought.

This chapter is really a discussion document as the decision on exactly how to produce assemblage Alerts is one for the WeBS Alerts steering group. This chapter outlines several possible ways to produce the alerts and makes general recommendations as to the merits or limitations of the different methodologies.

At the outset, the steering group will need to consider the following

- Should the Alerts system even try to generate assemblage alerts given that coverage of gulls and terns is patchy and some of the more common species were not counted until 1986?

- Should assemblage Alerts be generated for just waders and wildfowl and should these be separated into ‘wader’ and ‘wildfowl’ assemblage Alerts or considered together?

- Should we ignore other species (e.g. Coot, Little Grebe, Cormorant) and treat these on an individual species basis.

- Should we just be calculating assemblage Alerts minus gulls and terns from 1986 onwards (or even 1993 but additional species will form a small part of the assemblage) but use an appropriate name for them so as not to associate them with the 20,000 waterbird threshold criterion?

3.2 Methods that Could be Used to Produce ‘Assemblage’ Alerts

There are various method that can be used to produce assemblage Alerts. The rationale behind the concept of assemblages is total numbers of waterbirds and three methods have been tailored with this in mind (Options 1-3 in Box 1). Option 4 describes a method which describes the overall rates of change in waterbird populations on SPAs and identifies periods of general increase or decline across all species.

3.3 Strengths and Weaknesses of Each Method

3.3.1 Option 1

This method takes count data from all species in a site using the standard Underhill months and imputes missing values using the Underhill method to complete the data set. A mean value is calculated for all species in each year and each yearly count is summed to provide a total count for each year. This is smoothed to produce smooth indices from which the change measures can be calculated. Optionally, the counts for each species can be scaled by the 1% international importance threshold to produce an index of change based around international importance count units. Although, this is removed from the 20,000 assemblage threshold it may be more informative when considering change in the international importance of the site.

Figure 3.1 (a,b) shows these two indices for Abberton Reservoir and only includes wildfowl and not other waterbirds. In numerical terms there was been a large increase from 1966 to 1980, then a period of stability and a slight decline since 1990. In terms of international importance count units, a slightly different patter was seen with an almost consistent increase from 1966 to 1995 and then a small decline. The two species that contribute most to this index are Shoveler and Gadwall (shown in 3.1c).

This method is probably best suited to assessing the changes in the 20,000 waterbird criterion but with the caveat of limited species coverage. In statistical terms, it is reasonable to produce them and it is also possible to devise a method to determine which species are contributing most to the changes seen.

There are two major weaknesses with this method. First, the standard Underhill months for each species (which differ between species) may not cover the months in which peak count for each species occurs and, second, dealing with average counts means that peak counts will be masked. Ideally all seven WeBS months would be used to calculate monthly totals for each winter (imputing missing counts where necessary) and then take the peak monthly count for each winter and run these values through the Alerts system.

- Should assemblage Alerts be produced using ‘average’ counts or ‘total’ counts?
- Should all seven WeBS months be used to find the peak monthly total for each winter with the caveat that the peak month may change between winters?

3.3.2 Option 2

This is similar to Option 1 but instead of imputing missing counts using the Underhill method, a smoothed count is calculated for each species by fitting a GAM through the data. The average smoothed monthly count for each species is calculated for each year and then these are summed for each year. This measure will not itself be smoothed and would be then smoothed again and change measures calculated.

We considered but rejected this method as we considered that smoothing the data twice is getting too far removed from the original data to draw sensible conclusions as to changes in waterbird numbers.

3.3.3 Option 3

This measure is used as the headline indicator for terrestrial bird species. It is calculated by producing indices for all waterbird species using the Underhill method. The log of these index values are calculated, summed across species and divided through by the number of species to provide the geometric mean. Taking the logs of each of the species' index values gives equal weighting to each species and gives an indicator of the rate of population change across all species. It would be possible to calculate this index for waterbirds, but again is a different measure to the 20,000 waterbird threshold measure.

3.3.4 Option 4 [taken from WeBS Alerts Report I]

Alerts present an opportunity to examine and measure change on a number of different scales. Site-based Alerts can be used in two ways. First, they give an effective picture of population change at individual sites over time. Individual species which are declining can be flagged up. Second, these Alerts can also be used to produce a picture of the general state of the populations on that site. To address this point, in the first Alerts report we used a simple index of change for each site which was the sum of the Alert values totalled across all evaluated/proposed SPA species where negative Alerts were given negative values (50% Alert = -2, 25% Alert = -1) and positive Alerts positive values (25% Alert = +1, 50% Alert = +2). A high negative score indicated that large numbers of nationally or internationally important waterbirds had declined over that period.

This index, or a variation on it, would be a useful measure to look at rates of population change across the assemblage. If Alerts were being raised for a number of species, this measure would provide a means of determining when periods of decline occurred.

This is a fairly crude measure and could be adapted as necessary. For example the exact percentage change, rather than Alert 'scores' could be used. In addition the percentage changes could be weighted by international count unit to look at rates of changes in the international importance of the site.

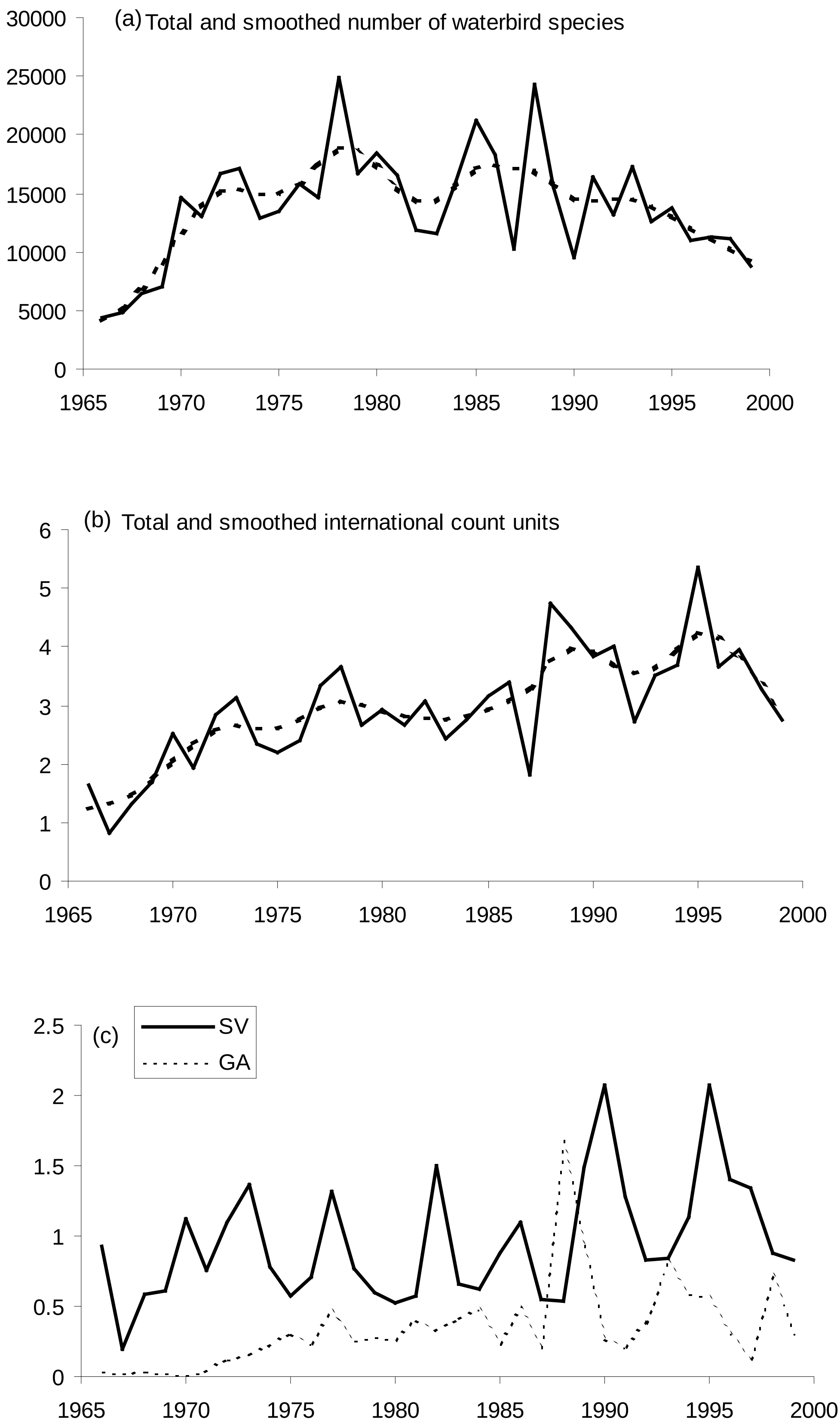


Figure 3.1 Changes in (a) the number of wintering wildfowl, (b) international count units and (c) Shoveler and Gadwall international count units at Abberton Reservoir between 1966 and 1999.

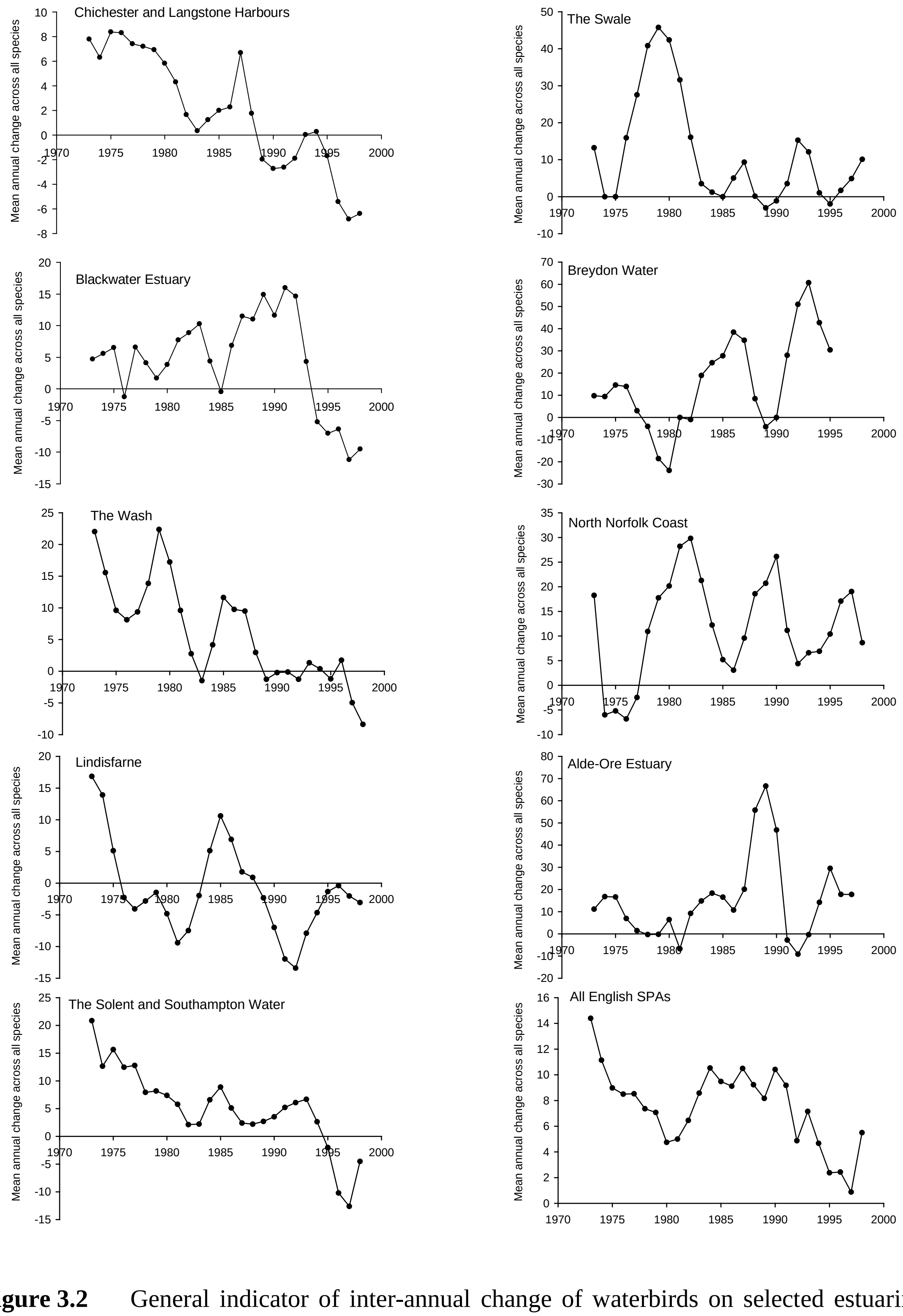


Figure 3.2 General indicator of inter-annual change of waterbirds on selected estuarine sites in England. The measure is calculated by taking the mean inter-annual change in the smoothed indices across all species evaluated for the SPA. The value for 1998 indicates the mean change in the smoothed index across all species from the winter 1997/98 to the winter of 1998/99. Sites on the left hand sides have shown recent declines across many waterbird populations, those on the right increases. Values above the x-axis indicate generally increasing populations, below declining.

OPTION 1

Impute missing values for all species for which data is available using the Underhill method.

Take an average value for the number of birds per month per year to give a single value for each year for each species.

Add up the average spp values by year to give a total across all spp (optionally weighting by International Importance thresholds)

Smooth these values and generate Alerts.

OPTION 2

Smooth indices for each species first

Work out an average value for number of birds per month of each spp per year

Sum across all spp by year (optionally weighted by International Importance)

Optionally smooth this trajectory as it will not be smooth. However smoothing twice is getting further removed from the original data.

Use this trajectory to work out % changes over 5, 10 & 25 years

OPTION 3 - method used in terrestrial headline indicators

Calculate annual indices for all species

Log all indices and sum by year across all species

Divide this index through by number of species

Take the exponential of these values and use this as the assemblage indicator

This will not be smooth but could be.

I am not keen on this one really as it gets away from any chance of weighting and treats all spp equally which may be reasonable for terrestrial spp but for waterbirds international/national count units may be more informative.

OPTION 4 - ‘rates of change’ of assemblages

This does not really have anything to do with assemblages in terms of total numbers of waterbirds but looks at the rates of change of birds averaged across all spp.

Smooth count data for each species

Work out % inter-annual change for all spp

Average this across each year, optionally weighting by international importance

Turn this into an index and use this as an indicator of how the waterbird populations are changing on SPAs.

This could be restricted to those spp the site has been designated for or for the whole suite of spp.

Box 1 Four alternative methods for incorporating waterbird assemblages in to the Alert system.

4. PRESENTATIONAL CONSIDERATIONS FOR ALERTS

Figure 4.1 South East England numbers a) 5 years b) 10 years c) 25 years.

Figure 4.2 South East England Int Units a) 5 years b) 10 years c) 25 years.

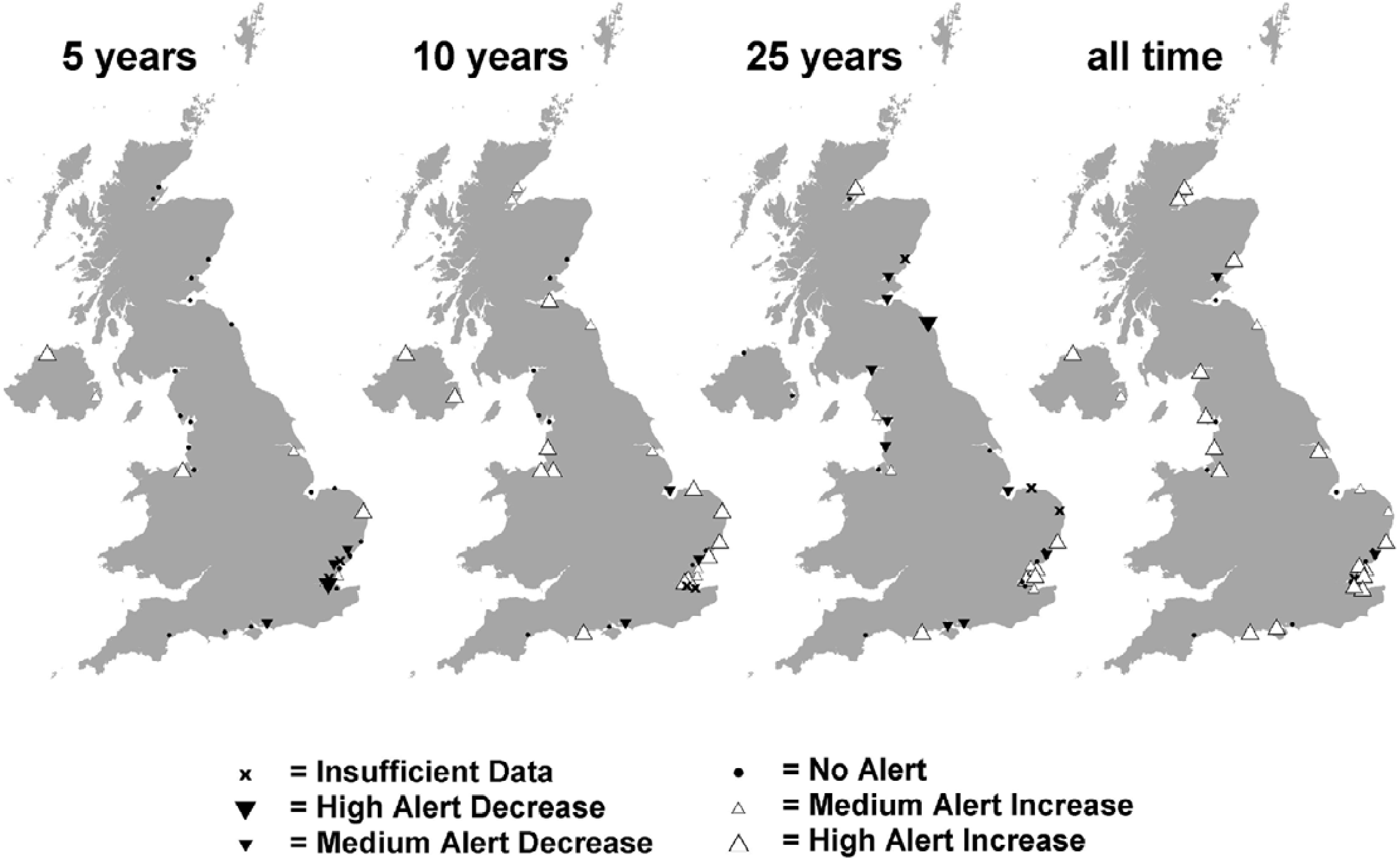


Figure 4.3 Dunlin SPA

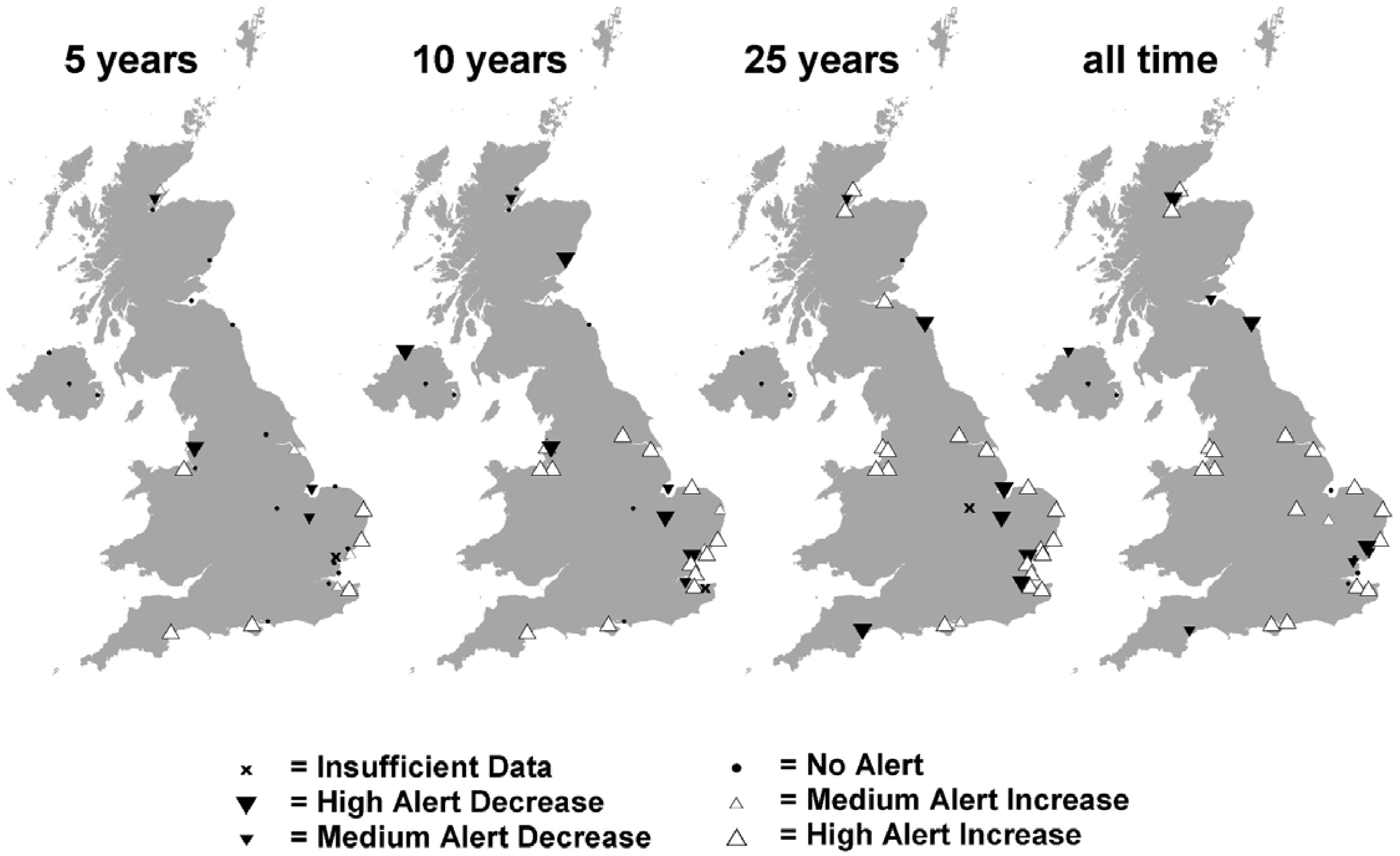


Figure 4.4 Wigeon SPA