



# ECOLOGICAL SOCIETY OF AMERICA

*Ecology/Ecological Monographs/Ecological Applications*

## PREPRINT

This preprint is a PDF of a manuscript that has been accepted for publication in an ESA journal. It is the final version that was uploaded and approved by the author(s). While the paper has been through the usual rigorous peer review process of ESA journals, it has not been copy-edited, nor have the graphics and tables been modified for final publication. Also note that the paper may refer to online Appendices and/or Supplements that are not yet available. We have posted this preliminary version of the manuscript online in the interest of making the scientific findings available for distribution and citation as quickly as possible following acceptance. However, readers should be aware that the final, published version will look different from this version and may also have some differences in content.

The doi for this manuscript and the correct format for citing the paper are given at the top of the online (html) abstract.

Once the final published version of this paper is posted online, it will replace the preliminary version at the specified doi.

Running head: List Length Analysis to detect declines

Title: Regional avian species declines estimated from volunteer-collected long-term data using List Length Analysis

Authors: Judit K SZABO<sup>1,\*</sup>, Peter A VESK<sup>2</sup>, Peter WJ BAXTER<sup>1,3</sup> and Hugh P POSSINGHAM<sup>1</sup>

<sup>1</sup> University of Queensland, Centre for Applied Environmental Decision Analysis, School of Biological Sciences, St Lucia, Queensland 4072 Australia [h.possingham@uq.edu.au](mailto:h.possingham@uq.edu.au)

<sup>2</sup> University of Melbourne, School of Botany, Parkville, Victoria, 3010 Australia, [pvesk@unimelb.edu.au](mailto:pvesk@unimelb.edu.au)

<sup>3</sup> Australian Centre of Excellence for Risk Analysis, University of Melbourne, School of Botany, Parkville, Victoria, 3010 Australia, [p.baxter@uq.edu.au](mailto:p.baxter@uq.edu.au)

Contact: E-mail [judit.szabo@cdu.edu.au](mailto:judit.szabo@cdu.edu.au), Phone +618 8946 6427; Fax +61 8 8946 7720

\* Current address: Charles Darwin University, School for Environmental Research, Darwin 0909, Northern Territory, Australia

**Abstract**

Long-term systematic population monitoring datasets are rare, but are essential in identifying changes in species abundance. In contrast, community groups and natural history organizations have collected many species lists. These represent a large, untapped source of information on changes in abundance but are generally considered of little value. The major problem with using species lists to detect population changes is that the amount of effort used to obtain the list is often uncontrolled and usually unknown. Franklin (1999 *Biol. Cons.*, **90**, 53-68) suggested using the number of species on the list, the “list length” as a measure of effort. This paper significantly extends the utility of Franklin’s approach using Bayesian logistic regression. We demonstrate the value of List Length Analysis to model changes in species prevalence (i.e. the proportion of lists on which the species occurs) using bird lists collected by a local bird club over 40 years around Brisbane, southeast Queensland, Australia. We estimate the magnitude and certainty of change for 269 bird species and calculate the probabilities that there have been declines and increases of given magnitudes.

List Length Analysis confirmed suspected species declines and increases. This method is an important complement to systematically designed intensive monitoring schemes and provides a means of utilizing data that may otherwise be deemed useless. The results of List Length Analysis can be used for targeting species of conservation concern for listing purposes or for more intensive monitoring. While Bayesian methods are not essential for List Length Analysis, they can offer more flexibility in interrogating the data and are able to provide a range of parameters that are easy to interpret and can facilitate conservation listing and prioritization.

Key-words: Avifauna; Bayesian logistic regression; Birds; Citizen science; Conservation;  
Detecting population change; Historical records; Museum data; Presence-only data; Relative  
abundance

esa

preprint

## Introduction

Detecting population declines requires long-term datasets, yet collecting high-quality data over long periods requires considerable (often prohibitive) investment and institutional support (Reid et al. 2003). High quality data are more likely to originate from intensive research projects funded for short periods (Field et al. 2007), whereas volunteer-collected records frequently cover greater areas and longer periods than more systematic surveys, providing relatively inexpensive but lower-quality data (Geissler and Noon 1981, Bart 2005). Long term datasets are especially valuable as, for example, at least 20 years of data are needed to show trends in the Breeding Bird Survey (Bart et al. 2004). Better use of inexpensive, lower quality data could expand our capacity to monitor biodiversity.

Simple lists of species are probably the most abundant form of biodiversity information available, as naturalists have long recorded the plants, birds and other taxa they have sighted or collected. However, these historic records are often compiled without using standardized methodology or rigorous sampling design. As a consequence, they have been thought to have little use in deducing population trends (Elphick 2008). Historical datasets often lack critical information on sample location, sampling effort (e.g. duration, number of observers) and survey type (e.g. point vs. transect). Analyzing large-scale avian datasets usually requires the application of sophisticated statistical methods to correct for sampling biases (Sauer et al. 1994, Sauer et al. 1995).

Recent attempts to use historical sighting or museum collection data to infer changes in species abundance and distribution show promise (Solow 1993, McCarthy 1998, Roberts and Solow 2003, Elith et al. 2006). While species surveys and atlases originally aimed to get

information about distribution patterns and species-habitat relationships, currently, focus has shifted to detecting changes in abundance, especially declines in species of conservation interest (Fuller et al. 1995) and increases in introduced or pest species (Underhill et al. 1991, Dunn and Weston 2008).

Although monitoring is not always done systematically, and often the available datasets are *ad hoc* compilations, we believe species lists have considerable potential as a biodiversity monitoring resource.

Each species list provides at least two types of information, the record of the detection of certain species, and the list length, that is, how many species were recorded in total. The list length reflects the collection effort – important information for qualifying the absence of a particular species from a list. Yet list length and collection effort are not synonymous. The length of the list also reflects local species diversity, which is influenced by many factors – a topic of great interest in ecology (Loreau 2000, Field et al. 2009). However, this has no impact on our analysis if sampling across these factors remains roughly consistent. More detailed lists provide information on location, the number of surveyors and survey date and duration. With multiple lists, prevalence can be estimated – the proportion of lists on which the species occurs – also known as reporting rate. Prevalence denotes a relative state, whereas ideally we would like to make inferences about species occurrence – an absolute state. In an important paper, Franklin (1999) introduced an example of the potential use of such lists for detecting declines in granivorous birds across a large region, northern Australia. Here we explore the potential to extend List Length Analysis and apply it to the entire avifauna of southeast Queensland, Australia, a smaller, though still sizeable region (15,000 km<sup>2</sup>), which is known to be at risk from habitat loss and fragmentation (Catterall and Kingston 1993).

In historical records there is a bias for and against the inclusion of certain species in the lists, the geographical location of surveys is changing through time and the mean search effort is changing through time (Franklin 1999). Because of violation of these assumptions associated with the use of reporting rates, Franklin (1999) suggested the use of list length as a surrogate for sampling effort, therefore enabling the detection of changes in abundance from list data. Roberts *et al.* (2007) validated this by demonstrating positive relationships between mist-net trapping effort and number of species caught. While the relationship is imperfect, because species differ in their capture probabilities, the “noise reduction” in reporting rates due to accounting for list length is considerable (Roberts *et al.* 2007). In a different manner, Link and colleagues (2006) adjust for effort using counts of all individual birds with a Bayesian hierarchical model. In this paper, we explore the potential for a wider use of List Length Analysis. This technique can use historical data to estimate long-term trends in species distributions and abundances and could facilitate a better use of community monitoring, such as atlases, which are often underused (Dunn and Weston 2008).

The aim of this paper is to develop an approach for finding useful and easily interpreted information about trends in the abundance of species derived from simple lists. We explore the performance of List Length Analysis in this respect with a regional, historical, volunteer-collected dataset with all its foibles, lacking a rigorous sampling design and with minimal metadata to investigate changes in species’ abundance. We investigate the potential of this method to answer the following questions: (a) What is the probability that a given species has declined or increased to any degree (or by a given amount, say 10%) and (b) what is the magnitude (and uncertainty) of that change? We use Bayesian statistical models for analyzing the lists as they allow great flexibility in interrogating the data.

This paper has two appendices, common names, scientific names and the number of observations for each of the 269 species are presented in Appendix 1 while details of the model selection and the effect of variation among species in prevalence and the contribution of short lists are detailed in Appendix 2.

## Methods

### *Rationale*

We assumed that as sampling effort increases so does the probability of observing each species,  $P(\text{obs})$ , and hence the expected number of species observed, which we call list length ( $L$ ). Our aim then was to model the monotonic increasing relationship between the probability of observing a particular bird species and the total number of species observed in the survey, i.e. between  $P(\text{obs})$  and  $L$ . The probability of observing each species tends to be low on the shortest lists, approaching one as the list length approaches the total number of species in all lists combined ( $S$ ). Therefore, by definition, the curve is bound between  $[0, 0]$  and  $[S, 1]$ . We can interpret  $S$  as the total number of species in the sampled region. The shape of the curve is species specific: common species appear frequently on shorter lists, so that their probability of detection rises rapidly towards one as list length increases (concave downwards). For rare species, the probability of observation stays close to zero as list length increases initially, and approaches one only on the longest lists (concave upwards).

When the relative abundance of a given species changes over time, its curve of  $P(\text{obs})$  vs  $L$  will change correspondingly. As its abundance decreases, the species will have a lower reporting rate for a given list length, which means that it has less probability of appearing on lists

of the same length and the shape of its curve approaches that of a rare species. The data are relative, i.e. as one species becomes more common and its curve shifts up and left, the remaining species become relatively less common and their curves shift down and right. The larger the species pool, the smaller the effect of any one species.

Numerous functions may fit the data to varying degrees; we investigated two reasonable candidate functions (power and logistic) and focus on the logistic form in this paper because of its superior performance (for details of the model selection see Appendix 2). We assumed that the effect of list length on  $P(\text{obs})$  was multiplicative rather than arithmetic, reflecting how many times longer one list was than another, rather than how many additional species were observed. For example, the increase in the probability of observing a species associated with comparing lists of five and 10 species is likely to be greater than the increase associated with comparing lists of 45 and 50 species. Hence we used a logarithmic transformation of list length. Apart from this functional reason, using the log of the list length was statistically superior to untransformed list length in initial comparisons. We looked at a small number of models with interactions between year and list length, but they did not perform better than simple models.

### *The nature of the data*

We used data from birdwatching outings of Birds Queensland, a non-profit organization funded by its members, and some other lists included in their database, as an example of widely available volunteer-collected data. Surveys started between 6:30 and 7:30 am (depending on the season), lasting for three to four hours. The number of birdwatchers was variable (generally 10–30 people), the locations were described in words or with low-resolution coordinates and the extent of the area covered by the surveys was not recorded (D. Niland personal communication).

Surveys were available from 1970 to 2006. Earlier surveys, from 1964 to 1970, were conducted by the local branch of the Royal Australasian Ornithologists' Union (now Birds Australia) and incorporated into the dataset by Birds Queensland. We selected an area bound by 152°30' and 153°30' longitude and -26°30' and -28°00' latitudes (six 1:100,000 scale map sheets) from southeast Queensland, Australia, as this was an area with relatively high data density from this period. We selected this dataset for its long time frame and non-standardized methodology.

Reporting rates on shorter lists are often confounded by selective searching for rare birds. However, assuming no change through time in birdwatchers' behavior, trends should be unaffected, as we are not analyzing absolute abundance. Another issue was the scarcity of very long lists (those including almost all species), as detailed in Appendix 2. Following Franklin's (1999) methods, we deleted any list with less than three species. We also eliminated data-poor species, the ones with less than five observations on those lists of at least three species, for discussion see Appendix 2. After deleting these data-poor species and the very short lists, 855 lists were left, comprising 30,773 observations of 269 species. The number of lists from each map sheet was 149 (Ipswich), 164 (Caboolture), 65 (Nambour), 165 (Beenleigh), 290 (Brisbane) and 22 (Caloundra). If our objective were the definitive analysis of these data, we would account for variation between map sheets with a random effect. However we do not, because our aim was to illustrate the approach on a large heterogeneous volunteer-collected dataset.

### *List Length Analysis*

We used Bayesian logistic regression to model the probability of a species occurring on a list as a function of the list length and the survey year (Link et al. 2002, McCarthy 2007). We compared performance of the logistic and power models and found the logistic model to be

generally better (see Appendix 2), but do not discuss this further here. Our approach could also be performed in a maximum likelihood framework and when uninformative priors are used, the parameters estimated through the two methods are equivalent. Nevertheless, the Bayesian approach has particular advantages: the flexibility of modeling allows intuitive interpretations such as the probability of a given decline and the results can be updated as further sets of lists are assembled (Clark 2005, McCarthy 2007). The statistical model for the logistic regression was:

Eq. 1. 
$$\text{Logit}(P(\text{obs})) = a_1 + a_2 \log(L) + a_3 y,$$

where  $P(\text{obs})$  = probability of observation,  $L$  = list length and  $y$  = year. The coefficients,  $a_i$ , are all assumed to be normally distributed;  $a_1$  is the intercept term and reflects the overall prevalence of the species,  $a_2$  is the term for the effect of multiplicative increase in list length and  $a_3$  is the term for change per year. If  $a_2$  is removed from the model, it collapses to an analysis of prevalence, i.e. reporting rate. An interaction between time and list length can be modeled with an additional term:  $a_4 y \log(L)$ . A term for year-to-year variation could also be fitted as a random effect. We used minimally informative priors for coefficients (means = 0, standard deviations = 10000). The distributions of input variables were centered on 0 means, however we chose not to rescale to units of standard deviations (by dividing by the standard deviation). It is easier to interpret the changes over time in units of  $\text{year}^{-1}$ , rather than units of (standard deviations of years in the dataset) $^{-1}$ .

We fitted models using WinBUGS v. 3.0.3 (Spiegelhalter et al. 2003), a statistical package for conducting Bayesian inference with Markov Chain Monte Carlo (MCMC) simulation. Three Markov chains were run for each model. Standard Bayesian diagnostics were

used to assess convergence, including plots of each of chain traces, chain auto-correlations and the Gelman-Rubin statistic for 15 species selected at random from the total of 269 species (Zuur et al. 2002). Finding that in all cases convergence was achieved within 1000 iterations, a conservative 5000 samples were used as a standard burn-in and traces of the MCMC chains were checked for all species. Posteriors were calculated on the subsequent 5000 (or greater) iterations from each of the three chains.

Model performance was compared through model fit and parsimony. Model fit was measured by the percentage of deviance for the null model that was explained by the fitted model. Parsimony was measured with the Deviance Information Criterion (DIC), which is similar in aim to AIC (Akaike's Information Criterion; Spiegelhalter et al. 2002). Among models of the same dataset, the smaller the DIC, the better the model. The parsimony of two models was compared as differences between the DIC values ( $\Delta$ DIC) for two models, in this case, the logistic and power models. The rules of thumb advocated by Spiegelhalter et al. (2002) are followed here, with  $\Delta$ DIC < 2 indicating substantial support for the inferior model,  $4 < \Delta$ DIC < 7 indicating considerably less support for the inferior model, and  $\Delta$ DIC > 10 indicating essentially no support for the inferior model.

From the basic logistic model we calculated posterior distributions of three, more intuitive parameters: first, the change in the probability of observation for a given species over 1964–2006 for the mean list length; second, the probability that this change was negative (or positive), indicating the probability that the species' prevalence declined (or increased) over the census period; and third, the decline in the probability of occurrence on a given list length as a percentage of the initial (1964) modeled occurrence, for a species list of length 45, which is near

the median length. We derived 95% credible intervals based on the observed percentiles from the MCMC replicates. The code is available from the senior author on request.

The Bayesian List Length Analysis logistic regression models were run for all 269 species. To exemplify our results, we focused on six species, representing combinations of three typical temporal trends: large, small and unchanging, and two levels of uncertainty: relatively certain and uncertain. The selected species were Speckled Warbler, Dusky Woodswallow, Forest Kingfisher, Bell Miner, Spotted Pardalote, and Satin Flycatcher. For these six species we fitted a series of models and compared them for fit (explained deviance) and parsimony (Deviance Information Criterion, Spiegelhalter et al. 2002). To verify that List Length Analysis performed best, we first compared predicted probability of observation per year for median list length for four different models, from the simplest (“Reporting Rate model”, including only intercept and random variation) to the full List Length Analysis model (see Equation 1). We also compared the List Length Analysis model to models with interactions between  $L$  and  $y$ . As the models without interaction were outperforming Reporting Rate models and were not performing worse than the ones with interaction, for simplicity, in the final runs we assumed no interaction between  $L$  and  $y$ , thus ignoring the possibility that a species’ detectability changes systematically through time. This also implies that the behavior of the birdwatchers does not change systematically through time. We were interested simply in whether there has been a change in the probability of observation on lists of equivalent length.

To investigate the performance of the candidate models and the effect of prevalence and contribution of records from short lists, we analyzed model performance more extensively for a wider set of 55 species, which included species suspected of changes and a random selection (Appendix 2).

## Results

The number of lists collected per year steadily increased though time (linear regression,  $y = 0.703x - 1380$ ,  $r^2 = 0.61$ ), however the median list length did not change substantially (linear regression,  $y = -0.106x + 258$ ,  $r^2 = 0.01$ ; Figure 1a). Moreover, the range of list lengths in any year was large on average the longest list in a year had ~7 times as many species as the shortest list in that year and in 90% of years the longest list had 46 species more than the shortest list in the corresponding year. Long lists (e.g., with  $> 70$  species) were recorded in all years after 1966, hence, the input variables were not collinear. A plot of raw reporting rates (proportion of lists for a year with the focal bird recorded) for the Speckled Warbler illustrates the problem with aggregations of unstandardized lists: some suggestion of a downward trend but with much variation (Figure 1b). We compared this raw trend to four models that predict probability of occurrence on a 45-species long list in a given year in a factorial combination of models including or excluding  $a_2$ , the effect of list length, and  $a_3$ , a change through time. The simplest model included only an intercept term ( $a_1$ ) and random year to year variation ( $\epsilon$ ; Figure 2a); variation between estimates corresponds to true variation between years, sampling variation and variation in list length (cf. Figure 1a). Some estimates are less certain because of fewer lists and have correspondingly wider credible intervals (Figure 2a) (note intervals are longest for estimates close to 0.5 and shortest for estimates near 0.0 and 1.0, on the probability scale). In the next step, we accounted for the length of the list ( $a_2$ , Figure 2b), which reduces variation in the predictions between years and uncertainty in the estimates of any particular year. Notably, accounting for list length pulls in some outlying points, such as the one for year 1964, when there was only one list, moderately long, on which Speckled Warbler was recorded. The

prediction for this year is a high reporting rate and  $P(\text{obs})$  without controlling for list length. Controlling for list length reduces the prediction for year 1996 as well, when relatively few (13) lists were collected, most of which were long (median = 71), showing a high reporting rate and high  $P(\text{obs})$  when not controlling for list length (Figure 2a). Next we modeled linear decline and year to year variation without list length (Figure 2c) and finally the most complicated model accounted for list length and included a linear year-to-year decline ( $a_3$ ). Note that 1965 and 1966 were characterized by short lists, and accounting for list length has estimated the prevalence for these years as higher and the whole decline as somewhat steeper (Figure 2d). In summary, modeling the probability of occurring on a list reduces the variation from the raw reporting rates. When controlling for list length, variation between the point estimates of prevalence and the size of the credible intervals (uncertainty about the value of the mean) are both reduced compared to modeling year-to-year variation in reporting rates only.

### *Model Comparison*

Models performed better on species with more observations and with fewer occurrences recorded from short lists as opposed to long lists (Table 1, and more detailed analyses appear in Appendix 2). Models controlling for list length fitted better and were more parsimonious than simple Reporting Rate models (Table 1, Appendix 2). Reporting Rate models produced estimates for  $a_3$ , temporal change in prevalence, that were very slightly smaller in magnitude than those produced from List Length models ( $a_3(\text{RR}) = 0.8227 * a_3(\text{LL}) + 0.0008$ ,  $R^2 = 0.94$ , Appendix 2). However, models incorporating list length explain considerably more deviance than Reporting Rate models (LL, 14% median deviance explained and 5<sup>th</sup> and 95<sup>th</sup> percentiles [0.5%, 32%]; RR, 0.5% [0.0%, 5.0%]  $n = 55$ ). List length models were mostly more parsimonious than

Reporting Rate models, with a median difference in DIC = 53.0 [-1.9, 322.5], 5<sup>th</sup> and 95<sup>th</sup> percentiles). Those ten species for which list length models were not demonstrably better are discussed further in Appendix 2.

Models that included an interaction between list length  $L$  and year  $y$  did not perform substantially better than simple linear list length models (Table 1). Although in some cases fitting an interaction term may be justified statistically, we choose to emphasize the main and effect, especially as ecological interpretation of the interaction is unclear.

Logistic regression models fitted to data for the six illustrative species are presented graphically in Figure 3 and the model coefficients are in Table 2, demonstrating the increasing probability of observation with increasing list length. Observational data are presented as reporting rates by binning into classes of at least 15 lists. The probability of observing a Speckled Warbler during a typical bird survey, relative to other species, has decreased considerably over the 40-year census period (Figure 3a). The Dusky Woodswallow has also declined, but by a lesser amount and less certainly as shown by the adjacent credible intervals (Figure 3b). Smaller declines were indicated for the Forest Kingfisher and Bell Miner (Figure 3c, d). Two species showing no trend in time were the Spotted Pardalote and Satin Flycatcher (Figure 3e, f). The response of the Satin Flycatcher to list length was very uncertain. Note that the uncertainty in the right hand panels (Figure 3b, d, f) is associated with frequent reporting rates of zero. The Forest Kingfisher exemplifies the sigmoid nature of the relationship between the probability of observation and list length (Figure 3c). The difference between the time periods is greatest at intermediate (about 40–80 species long) lists (Figure 3a, c).

Using a Bayesian approach, not only is it possible to estimate the probability of decline (or increase), but also the probability of given magnitudes of decline (or increase) (Table 3). As

an example, the Dusky Woodswallow is most likely to have declined in prevalence by about 89%, i.e. the probability of observing it in 2006 was 11% of the probability of observing it in 1964 for a list of 45 species (the median list length). For the median list length, the probability of observing it in 2006 had 95% chance of being between 52% and 98% of the probability of observation in 1964. Furthermore, there was about a 46% chance of its decline being over 90% and a 98% chance of the decline being greater than 50%. The effect of uncertainty is also evident in Table 3, for example, the Spotted Pardalote is more likely to have decreased than the Satin Flycatcher; however, on account of the lower precision estimating the extent of its decline, the Satin Flycatcher may have declined by a great deal and so is more likely to have declined by 25% than the Spotted Pardalote.

Of all 269 species, 28 had a negative  $a_2$  (estimated mean effect of list length on probability of observation) and a further 22 species had estimates with credible intervals that overlapped 0, which means that in these species the assumption of increasing observation probability with increasing list length (which we take as a proxy for effort) may not be upheld, and trends could not be deduced (see Appendix 1, Table S1). The species with negative  $a_2$  values were overwhelmingly those with few (<50) records and a high proportion (>35%) of records on short lists ( $L < 7$  species; for details see Appendix 2.)

The coefficient for the effect of year on probability of observation,  $a_3$ , ranged widely among species (Figure 4). The prevalence of some species declined considerably over 1964–2006, but that of others increased. The uncertainty can be seen to vary between species, with credible intervals of variable width in species with large and small mean changes.

Roughly 40% of the species had rates of population change ( $a_3$ ) with absolute values of their means less than twice their standard deviations, which is about equal to a 95% credible

interval (Figure 5; points shown between the dashed lines). For species whose 95% credible intervals include zero, we cannot be certain whether their rates of change are different from zero. Yet the greater their distance from the origin, the greater the potential change, therefore such species are suitable targets for further monitoring to reduce the uncertainty around potentially important changes.

The most prevalent species (those with high  $a_1$ ) tended to have more positive values of  $a_3$ , meaning that the most common species have slightly increased in prevalence (Figure 6).

## Discussion

The main contribution of this paper is to demonstrate the utility of a Bayesian List Length Analysis to model changes in prevalence over time using species lists, collected with variable effort. Historical records often comprise lists only and hence the development of statistically credible methods of detecting changes using these lists is a significant step in biodiversity monitoring. List length Analysis picks up the same trends as a simple reporting rate, but explains much more of the variation in prevalence that arises from variable effort and detectability. Thus, predictions about prevalence controlling for variation in list length have higher precision. Using reasonably robust methods, we have demonstrated that lists can provide a valuable basis to indicate changes in community composition. We have shown that we can detect declines and increases, and estimate the magnitude and certainty of those changes. Furthermore, we have demonstrated the calculation of the probability that there has been a decline of a given magnitude, e.g. an  $X\%$  chance that species  $A$  had a  $Y\%$  lower probability of being seen on a list of the same length in a later year. These measures are useful to managers who may have to

establish whether a threshold of decline has been exceeded to warrant a listing of a species as threatened, for instance.

Of the 269 species analyzed, most of the species that show substantial declines have been reported as having declined or being in decline elsewhere in Australia (Paton et al. 1994, Smith and Smith 1994, Ford et al. 2001). For the region, we have provided the first scientific validation of previous anecdotal claims, documented in unpublished data and grey literature (e.g. Walter and Walter 2007). Eleven of the 14 species showing the largest declines are insectivores, mostly inhabitants of more open woodland habitats. Their decline is not surprising, as woodlands in and around Brisbane have been replaced by suburbs to a great extent during the past half century (Catterall and Kingston 1993). Among other noteworthy decliners are rainforest species, waterbirds and two introduced species (House Sparrow and Nutmeg Mannikin). On the flip side, increasers were a mix of native parrots and introduced species, with the Common Myna showing the greatest increase. These results are also in accordance with field studies (Pell and Tidemann 1997, Oliver and Parker 2006, Shukuroglou and McCarthy 2006).

By examining model performance for a large number of species of variable prevalence we have demonstrated the reliability of List Length Analysis, within some constraints (Appendix 2). We are acutely aware of the advantages of exploring model performance on a high-quality dataset, as Roberts and colleagues (2007) did. However, there simply are no such suitable datasets at the spatial and temporal extent and of a quality that we are interested in, namely a dataset of a regional scale over decades where lists are long enough (i.e. compared to the local species pool).

The models were unreliable with rare species, contrasting with the view of Roberts *et al.* (2007), who argued that lists can provide meaningful indicators of declines in already rare

species. Roberts *et al.* (2007) used non-Bayesian logistic regression on an unusually large dataset of species lists compiled from daily mist netting in two small areas within a single 305-ha nature reserve, one of the most intensively studied nature reserves in the United Kingdom. Our results, from data collected more casually over at least 15,000 km<sup>2</sup>, suggest that insufficient records of rare species inhibit meaningful inference on their trends. Furthermore, we found that the selective reporting of some rare species led to obviously spurious results (manifested by negative  $a_2$ , i.e. probability of occurrence on a list declined as list length increased). Therefore, while unbiased sampling and reporting produce stronger correlations between abundance and reporting rate for rare versus abundant species (Bart and Klosiewski 1989, Roberts et al. 2007), our results suggest that, unfortunately, lists are generally unreliable for detecting changes in rare species, due to apparent bias in reporting. It may be possible to group rare species for analysis (e.g. Nichols et al. 2000), though this raises non-trivial questions of group membership.

*Is changing prevalence calculated from List Length Analysis merely relative?*

The primary motivation for the paper is to detect temporal changes for individual species. Ideally, we would like to infer changes in species occurrence probability. We argue that changes in species prevalence are a useful surrogate for this, subject to a set of important provisions. If, on average, species are declining such that for a given effort, fewer birds are seen and shorter lists compiled, then declines in particular species will be obscured. In theory, if the populations of every species in a community each decreased by exactly the same percentage, there would be no change in the probability of observing the bird on a list of a given length, although absolute prevalence would have changed dramatically. The data analyzed here show no substantial change in list length over time, giving confidence in our results. We suggest more work on the

effect of changes in community structure on list length analysis (see below). Nonetheless, our results indicate slightly more decreasing species (e.g. Speckled Warbler) than increasing (e.g. Common Myna) representing a considerable change in community structure. Estimating change in community structure is important in itself, and List Length Analysis provides a method that reports single species change as opposed to an aggregate metric such as community similarity as measured by Bray-Curtis, Jaccard or other similar measure (Warton 2008).

*Does List Length really reflect effort?*

List Length Analysis relies on the assumption that list length can serve as a surrogate for effort, though it may also be thought of as reflecting mean detectability (MacKenzie et al. 2003), e.g. longer lists may result when all birds are more visible. While this assumption may hold in carefully replicated surveys (Roberts et al. 2007), it is not necessarily true in all circumstances, especially for more casually-recorded lists, such as the dataset we worked on. For example, in regions with high spatial heterogeneity, longer lists could result from sampling more habitats. This may confound the interpretation of changes in observation. This is beyond the scope of the present study but we are currently investigating this using computer simulations (PWJ Baxter et al. manuscript in progress), where complex realities are created and sampled to better understand the processes behind the List Length Analysis – a virtual ecologist approach (Tyre et al. 2001).

A further assumption is that we are more likely to find each species on longer lists as opposed to shorter lists. However, some rare species may appear only on short lists as incidental records, rather than results from ordinary surveys. Species with fewer than 20 records and over 20% of records from lists shorter than seven species were indicated to be poor candidates for List Length Analysis (Appendix 2). Because we model the probability of observation against list

length, the greater the range in list length, the better the model fit. Hence, this approach will be most applicable to surveys that produce both short and long lists. By contrast, it may not be easy to analyze short standardized counts using List Length Analysis if they cover a restricted range of list lengths. Similarly, List Length Analysis would likely perform poorly in analyzing any data from assemblages with few species, e.g. for taxa or regions that are depauperate in species.

*Are declarations of decline prone to error?*

Spurious changes in prevalence could arise from a range of factors, whether variation in list length is specifically accounted for or not. It is known that detectability varies among species, habitats and environmental conditions and also depending on the observer (Allredge et al. 2007). While there are methods for accounting for these in systematic repeat surveys (e.g. MacKenzie and Kendall 2002, Simons et al. 2007, Brewster and Simons 2009), it is unclear how to apply such methods retrospectively to historic data. We note this as an area for further research. We assume that in our dataset species detection probabilities are not changing systematically through time and that any effects of changing species composition do not change the underlying relationship between effort and list length. One possible consequence of this assumption failing is as follows. If a species became more easily-detected through time, the list length for a given effort would become longer, and the inferred effect for a given focal species would be biased towards a greater decline. Different survey methods could generate different results for the same species assemblage. False negatives (missed declines) could arise if average prevalence declined, such that the effort required compiling a list of a given length,  $L$ , increased through time. False positives (declaring a decline when there is no change) could arise if it became easier to compile a list of a given length over time. If, in the face of general declines,

birdwatchers concentrated on areas of greater species richness, the incidence of false negatives would increase. Similarly, improvements to field guides and sighting optics could increase detection and reporting rates, with different effects for different species (e.g. a large, conspicuous species versus a small, elusive, brown bird). It is important that any application of List Length Analysis, or indeed reporting rate, pays due attention to these possible complications.

In this paper we have emphasized the approach and utility of List Length Analysis and drawn attention to cases where its assumptions are violated, in particular the selective recording of rare species (see Appendix 2), which distorts prevalence patterns. List Length Analysis produced informative results even after reducing the dataset by eliminating these short (but biased) lists. We do not propose List Length Analysis should replace standardized data collection as such will be crucial in detecting declines (Mac Nally et al. 2009), yet for historical and community data List Length Analysis provides real promise. Our results have also highlighted some species that require particular attention in the future. Using widely available data we quantitatively demonstrate trends previously only reported as anecdotal information.

### **Acknowledgements**

The authors would like to acknowledge the valuable input of Mick McCarthy, Yohay Carmel and Chris Wilcox and comments on the manuscript from Julian Reid and three anonymous referees. Ideas for this paper were germinated at a workshop funded by Applied Environmental Decision Analysis Centre, a Commonwealth Environmental Research Facility Hub, which also supported JKS, PWJB and HPP.

**References**

- Allredge, M. W., T. R. Simons, and K. H. Pollock. 2007. Factors affecting aural detections of songbirds. *Ecological Applications* 17:948-955.
- Bart, J. 2005. Monitoring the Abundance of Bird Populations. *The Auk* 122:15-25.
- Bart, J., K. P. Burnham, E. H. Dunn, C. M. Francis, and C. J. Ralph. 2004. Goals and strategies for estimating trends in landbird abundance. *Journal of Wildlife Management* 68:611-626.
- Bart, J., and S. P. Klosiewski. 1989. Use of presence-absence to measure changes in avian density. *Journal of Wildlife Management* 53:847-852.
- Brewster, J. P., and T. R. Simons. 2009. Testing the importance of auditory detections in avian point counts. *Journal of Field Ornithology* 80:178-182.
- Catterall, C. P., and M. Kingston. 1993. Human populations, bushland distribution in south east Queensland and the implications for birds. Pages 105-122. *in* C. P. Catterall, P. V. Driscoll, K. Hulsman, D. Muir, and A. Taplin, editors. *Birds and their habitats: Status and conservation in Queensland*. Queensland Ornithological Society Inc., St Lucia.
- Christidis, L., and W. E. Boles. 1994. *The Taxonomy and Species of Birds in Australia and its Territories*. Royal Australian Ornithologists Union.
- Clark, J. S. 2005. Why environmental scientists are becoming Bayesians. *Ecology Letters* 8:2-14.
- Dunn, A. M., and M. A. Weston. 2008. A review of terrestrial bird atlases of the world and their application. *Emu* 108:42-67.
- Elith, J., C. H. Graham, R. P. Anderson, M. Dudík, S. Ferrier, A. Guisan, R. J. Hijmans, F. Huettmann, J. R. Leathwick, A. Lehmann, J. Li, L. G. Lohmann, B. A. Loiselle, G. Manion, C. Moritz, M. Nakamura, Y. Nakazawa, J. M. Overton, A. T. Peterson, S. J.

- Phillips, K. S. Richardson, R. Scachetti-Pereira, R. E. Schapire, J. Soberon, S. Williams, M. S. Wisz, and N. E. Zimmermann. 2006. Novel methods improve prediction of species' distributions from occurrence data. *Ecography* **29**:129-151.
- Elphick, C. S. 2008. How you count counts: the importance of methods research in applied ecology. *Journal of Applied Ecology* **45**:1313-1320.
- Field, R., B. A. Hawkins, H. V. Cornell, D. J. Currie, A. F. Diniz-Filho, J.-F. Guegan, D. M. Kaufman, J. T. Kerr, G. G. Mittelbach, T. Oberdorff, E. M. O'Brien, and J. R. G. Turner. 2009. Spatial species-richness gradients across scales: a meta-analysis. *Journal of Biogeography* **36**:132-147.
- Field, S. A., P. J. O'Connor, A. J. Tyre, and H. P. Possingham. 2007. Making monitoring meaningful. *Austral Ecology* **32**:485-491.
- Ford, H. A., G. W. Barrett, D. A. Saunders, and H. F. Recher. 2001. Why have birds in the woodlands of Southern Australia declined? *Biological Conservation* **97**:71-88.
- Franklin, D. C. 1999. Evidence of disarray amongst granivorous bird assemblages in the savannas of northern Australia, a region of sparse human settlement. *Biological Conservation* **90**:53-68.
- Fuller, R. J., R. D. Gregory, D. W. Gibbons, J. H. Marchand, J. D. Wilson, S. R. Baillie, and N. Carter. 1995. Population Declines and Range Contractions among Lowland Farmland Birds in Britain. *Conservation Biology* **9**:1425-1441.
- Geissler, P. H., and B. R. Noon. 1981. Estimates of avian population trends from the North American Breeding Bird Survey. Pages 42-51. *Studies in Avian Biology*.

- Link, W. A., E. Cam, J. D. Nichols, and E. G. Cooch. 2002. Of Bugs and Birds: Markov Chain Monte Carlo for Hierarchical Modeling in Wildlife Research. *The Journal of Wildlife Management* **66**:277-291.
- Link, W. A., J. R. Sauer, and D. K. Niven. 2006. A hierarchical model for regional analysis of population change. *The Condor* **108**:13-24.
- Loreau, M. 2000. Are communities saturated? On the relationship between alpha, beta and gamma diversity. *Ecology Letters* **3**:73-76.
- Mac Nally, R., A. F. Bennett, J. R. Thomson, J. Q. Radford, G. Unmack, G. Horrocks, and P. A. Vesk. 2009. Collapse of an avifauna: climate change appears to exacerbate habitat loss and degradation. *Diversity and Distributions*:1-11.
- MacKenzie, D. I., and W. L. Kendall. 2002. How should detection probability be incorporated into estimates of relative abundance? *Ecology* **83**:2387-2393.
- MacKenzie, D. I., J. D. Nichols, J. E. Hines, M. G. Knutson, and A. B. Franklin. 2003. Estimating Site Occupancy, Colonization, and Local Extinction When a Species Is Detected Imperfectly. *Ecology* **84**:2200-2207.
- McCarthy, M. A. 1998. Identifying declining and threatened species with museum data. *Biological Conservation* **83**:9-17.
- McCarthy, M. A. 2007. *Bayesian Methods for Ecology*. Cambridge University Press, Cambridge, UK.
- Nichols, J. D., J. E. Hines, J. R. Sauer, F. W. Fallon, J. E. Fallon, and P. J. Heglund. 2000. A Double-Observer Approach for Estimating Detection Probability and Abundance from Point Counts. *The Auk* **117**:393-408.

- Oliver, D. L., and D. G. Parker. 2006. Woodland birds of the NSW central Murray catchment. Measuring outcomes of the Greening Australia fencing and tree planting program., Birds Australia Southern NSW & ACT.
- Paton, D. C., G. Carpenter, and R. G. Sinclair. 1994. A second bird atlas of the Adelaide region. Part 1: Changes in the distribution of birds: 1974-75 vs 1984-85. South Australian Ornithologist **31**:151-193.
- Pell, A. S., and C. R. Tidemann. 1997. The Ecology of the Common Myna in Urban Nature Reserves in the Australian Capital Territory. Emu **97**:141-149.
- Reid, P. C., J. M. Colebrook, J. B. L. Matthews, J. Aiken, and Continuous Plankton Recorder Team. 2003. The Continuous Plankton Recorder: concepts and history, from plankton indicator to undulating recorders. Progress in Oceanography **58**:117-173.
- Roberts, D. L., and A. R. Solow. 2003. When did the dodo become extinct? Nature **426**: 245.
- Roberts, R. L., P. F. Donald, and R. E. Green. 2007. Using simple species lists to monitor trends in animal populations: new methods and a comparison with independent data. Animal Conservation **10** 332-339.
- Sauer, J. R., G. W. Pendleton, and S. Orsillo. 1995. Mapping of Bird Distributions from Point Count Surveys. Pages 151-160. *in* C. J. Ralph, J. R. Sauer, and S. Droege, editors. Monitoring Bird Populations by Point Counts. U.S. Department of Agriculture Forest Service General Technical Report.
- Sauer, J. R., B. G. Peterjohn, and W. A. Link. 1994. Observer differences in the North American Breeding Bird Survey. The Auk **111**:50-62.
- Shukuroglou, P., and M. A. McCarthy. 2006. Modelling the occurrence of rainbow lorikeets (*Trichoglossus haematodus*) in Melbourne. Austral Ecology **31**:240-253.

- Simons, T. R., M. W. Alldredge, K. H. Pollock, and J. M. Wettroth. 2007. Experimental analysis of the auditory detection process on avian point counts. *The Auk* **124**:986-999.
- Smith, P., and J. Smith. 1994. Historical changes in the bird fauna of western New South Wales: ecological patterns and conservation implications. Pages 123-147. *in* D. Lunney, S. Hand, P. Reed, and D. Butcher, editors. *Future of the Fauna of Western New South Wales*. Royal Zoological Society of New South Wales, Sydney.
- Solow, A. R. 1993. Inferring Extinction from Sighting Data. *Ecology* **74**:962-964.
- Spiegelhalter, D. J., N. G. Best, B. P. Carlin, and A. van den Linde. 2002. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society, B* **64**:583-639.
- Spiegelhalter, D. J., A. Thomas, N. G. Best, and D. Lunn. 2003. WinBUGS user manual, Version 1.4. <http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/manual14.pdf>. Medical Research Council Biostatistics Unit, Cambridge, United Kingdom.
- Tyre, A. J., H. P. Possingham, and D. B. Lindenmayer. 2001. Inferring process from pattern: Can territory occupancy provide information about life history parameters? *Ecological Applications* **11**:1722-1737.
- Underhill, L. G., T. B. Oatley, and J. A. Harrison. 1991. The role of large scale data collection projects in the study of southern African Birds. *Ostrich* **62**:124-148.
- Walter, J., and R. Walter. 2007. Changes in the occurrence of birds and conservation of bird habitats in the Pittsworth shire, Darling Downs, since 1972. *The Sunbird* **37**:1-13.
- Warton, D. I. 2008. Raw data graphing: an informative but under-utilized tool for the analysis of multivariate abundances. *Austral Ecology* **33**:290-300.
- Zuur, G., P. H. Garthwaite, and R. J. Fryer. 2002. Practical Use of MCMC Methods: Lessons from a Case Study. *Biometrical Journal* **44**:433-455.

**Tables**

Table 1. List Length Analysis model comparison for six selected species.  $\hat{D}$  is the deviance evaluated at the posterior mean of the parameters. “Explained dev” is the percentage of the deviance explained by the fitted model, calculated as  $(\hat{D} - \hat{D}_0) / \hat{D}_0$  where  $\hat{D}_0$  is the posterior deviance of the null model (just the intercept,  $a_1$ ). Models with lower DIC values are more parsimonious, but differences of  $< 2$  are not meaningful. “N5+” denotes the number of records for the species out of 855 lists of at least five species. For scientific names of the species see Table S1. LL is the List Length model:  $\text{logit}(P(\text{obs})) = a_1 + a_2 \log L + a_3 y$ . RR designates the Reporting Rate model, i.e. no term for list length:  $\text{logit}(P(\text{obs})) = a_1 + a_3 y$ . LL:  $L.y$  designates a list length model with a term for the interaction of year and list length:  $\text{logit}(P(\text{obs})) = a_1 + a_2 \log L + a_3 y + a_4 y \log L$ .

| <i>Species</i>    | <i>Model</i> | $\hat{D}$ | <i>DIC</i> | <i>Explained</i> |            |
|-------------------|--------------|-----------|------------|------------------|------------|
|                   |              |           |            | <i>dev. (%)</i>  | <i>N5+</i> |
| Speckled Warbler  | LL           | 300.11    | 306.0      | 20.3             | 49         |
|                   | RR           | 355       | 359.1      | 5.7              |            |
|                   | LL: $L.y$    | 298.6     | 306.1      | 20.7             |            |
| Dusky             | LL           | 173.1     | 178.9      | 15.5             | 22         |
| Woodswallow       | RR           | 197.7     | 201.7      | 3.6              |            |
|                   | LL: $L.y$    | 172.1     | 179.3      | 16.1             |            |
| Forest Kingfisher | LL           | 612.3     | 618.2      | 23.9             | 152        |
|                   | RR           | 800.2     | 804.2      | 0.5              |            |
|                   | LL: $L.y$    | 611.1     | 619.1      | 24.0             |            |

|                   |                |       |       |      |     |
|-------------------|----------------|-------|-------|------|-----|
| Bell Miner        | LL             | 378   | 383.8 | 14.3 | 61  |
|                   | RR             | 436.9 | 440.9 | 1.0  |     |
|                   | LL: <i>L.y</i> | 376.7 | 384.4 | 14.6 |     |
| Spotted Pardalote | LL             | 681.9 | 687.9 | 17.4 | 159 |
|                   | RR             | 825.4 | 839.4 | 0.0  |     |
|                   | LL: <i>L.y</i> | 681.5 | 689.4 | 17.4 |     |
| Satin Flycatcher  | LL             | 125.8 | 131.4 | 6.8  | 13  |
|                   | RR             | 135.0 | 138.9 | 0.0  |     |
|                   | LL: <i>L.y</i> | 122.8 | 130.0 | 9.0  |     |

preprint

Table 2. Posterior estimates for the regression coefficients in logistic regression models for list length analysis that appear in Figure 1.  $a_1$  is the intercept term,  $a_2$  is the effect of list length and  $a_3$  is the effect of year. The lower and upper bounds on the 95% credible interval are denoted 2.5% and 97.5%, respectively. For scientific names and ID of the species see Table S1

| <i>Species</i>    | <i>Coefficient</i> | <i>Mean</i> | <i>SD</i> | <i>2.5%</i> | <i>Median</i> | <i>97.5%</i> |
|-------------------|--------------------|-------------|-----------|-------------|---------------|--------------|
| Speckled Warbler  | $a_1$              | -3.541      | 0.315     | -4.207      | -3.524        | -2.977       |
| ID 177            | $a_2$              | 1.473       | 0.278     | 0.971       | 1.460         | 2.062        |
|                   | $a_3$              | -0.068      | 0.014     | -0.096      | -0.068        | -0.041       |
| Dusky Woodswallow | $a_1$              | -4.526      | 0.487     | -5.613      | -4.480        | -3.689       |
| ID 187            | $a_2$              | 1.517       | 0.417     | 0.785       | 1.488         | 2.436        |
|                   | $a_3$              | -0.056      | 0.020     | -0.095      | -0.056        | -0.018       |
| Bell Miner        | $a_1$              | -3.205      | 0.261     | -3.742      | -3.193        | -2.729       |
| ID 208            | $a_2$              | 1.301       | 0.227     | 0.885       | 1.291         | 1.766        |
|                   | $a_3$              | -0.027      | 0.012     | -0.050      | -0.026        | -0.003       |
| Forest Kingfisher | $a_1$              | -2.439      | 0.212     | -2.876      | -2.432        | -2.049       |
| ID 114            | $a_2$              | 1.724       | 0.189     | 1.375       | 1.716         | 2.117        |
|                   | $a_3$              | -0.019      | 0.009     | -0.036      | -0.019        | -0.002       |
| Spotted Pardalote | $a_1$              | -2.088      | 0.165     | -2.428      | -2.084        | -1.777       |
| ID 193            | $a_2$              | 1.270       | 0.143     | 1.003       | 1.265         | 1.564        |
|                   | $a_3$              | -0.002      | 0.009     | -0.018      | -0.002        | 0.015        |
| Satin Flycatcher  | $a_1$              | -4.912      | 0.536     | -6.096      | -4.864        | -3.998       |
| ID 140            | $a_2$              | 1.105       | 0.432     | 0.359       | 1.068         | 2.043        |

$a_3$      -0.000     0.026     -0.050     -0.001     0.052

---

esa

preprint

1 Table 3. Estimates of change in prevalence over time.  $\delta$  denotes change in the probability of observing the species over the entire census period  
 2 (1964–2006), with negative values indicating decreases. Results of  $\delta$  are calculated at the median list length, indicating the ‘standard list’ the  
 3 difference in the probabilities of a bird appearing on a ‘standard’ list between 1964 and 2006. The estimated median percentage change in the  
 4 probability of observing the species from the start of the census is presented in the last column, note the distributions are often highly left-  
 5 skewed, hence the median is less than the mid-point of the 95% credible interval [95% CI]. For scientific names of the species see Table S1

| <i>Species</i>    | $P(\delta > 0)$ | $P(\delta < 0)$ | $P(\delta < -10\%)$ | $P(\delta < -25\%)$ | $P(\delta < -50\%)$ | $P(\delta < -90\%)$ | $\delta$ (%)            |
|-------------------|-----------------|-----------------|---------------------|---------------------|---------------------|---------------------|-------------------------|
|                   |                 |                 |                     |                     |                     |                     | [95% CI]                |
| Speckled Warbler  | 0.000           | 1.000           | 1.000               | 1.000               | 1.000               | 0.679               | -92.0<br>[-97.1, -79.0] |
| Dusky Woodswallow | 0.002           | 0.998           | 0.997               | 0.993               | 0.977               | 0.458               | -89.2<br>[-97.7, -52.0] |
| Bell Miner        | 0.013           | 0.987           | 0.976               | 0.943               | 0.75                | 0.001               | -62.9<br>[-84.1, -11.2] |
| Forest Kingfisher | 0.016           | 0.985           | 0.962               | 0.869               | 0.342               | 0.000               | -44.4<br>[-66.3, -5.4]  |

|                   |       |       |       |       |       |       |               |
|-------------------|-------|-------|-------|-------|-------|-------|---------------|
| Spotted Pardalote | 0.420 | 0.58  | 0.427 | 0.191 | 0.006 | 0.000 | -4.8          |
|                   |       |       |       |       |       |       | [-42.5, 61.8] |
| Satin Flycatcher  | 0.489 | 0.511 | 0.473 | 0.406 | 0.269 | 0.014 | -2.7          |
|                   |       |       |       |       |       |       | [-87.1, 778]  |

*esa*  
*preprint*

1 **Figure legends**

2 Figure 1. a) The number of lists (left axis, unfilled circles) and median list length (right axis,  
3 filled diamonds) in each sampling year. b) yearly reporting rates for a sample species, the  
4 Speckled Warbler.

5

6 Figure 2. Predicted probability of observation per year for median list length for four  
7 different models for a sample species, the Speckled Warbler. Models and the estimated  
8 standard deviation between years ( $\sigma_y$ ): a) Year-to-year variation in reporting rate,  
9  $\text{Logit}(P(\text{obs})) = a_1 + \varepsilon_y$ ,  $\sigma_y = 0.80$ ; b) year-to-year variation after accounting for list length,  
10  $\text{Logit}(P(\text{obs})) = a_1 + a_2 \log(L) + \varepsilon_y$ ,  $\sigma_y = 0.50$ ; c) a linear decline with year-to-year variation  
11 in reporting rate  $\text{Logit}(P(\text{obs})) = a_1 + a_3 y + \varepsilon_y$ ,  $\sigma_y = 0.40$ ; d) a linear decline with year-to-year  
12 variation after accounting for list length,  $\text{Logit}(P(\text{obs})) = a_1 + a_2 \log(L) + a_3 y + \varepsilon_y$ ,  $\sigma_y = 0.21$ .

13

14 Figure 3. Predicted probability of observation and observed reporting rates as a function of  
15 list length for old and recent surveys. Observations are reporting rates in bins of at least 15  
16 lists beginning from the shortest lists for the periods 1964–1992 (grey triangles) and 1993–  
17 2006 (black diamonds). Solid lines are predictions for median years from those periods (grey  
18 1978 and black 2000), with dashed lines as the respective 95% credible intervals. A)  
19 Speckled Warbler, B) Dusky Woodswallow, C) Forest Kingfisher, D) Bell Miner, E) Spotted  
20 Pardalote, F) Satin Flycatcher. Note the different scales on the y-axes. Note also that the lines  
21 would be further separated if we plotted predictions from the start (1964) and end (2006)  
22 years.

23

24 Figure 4. Caterpillar plot of coefficients for  $a_3$ , the effect of year on the probability of  
25 observation for all 269 species. Positive estimates indicate linear increase in logit-

1 transformed probability of observation through time. Error bars are 95% credible intervals.

2 For species ID codes see Table S1.

3

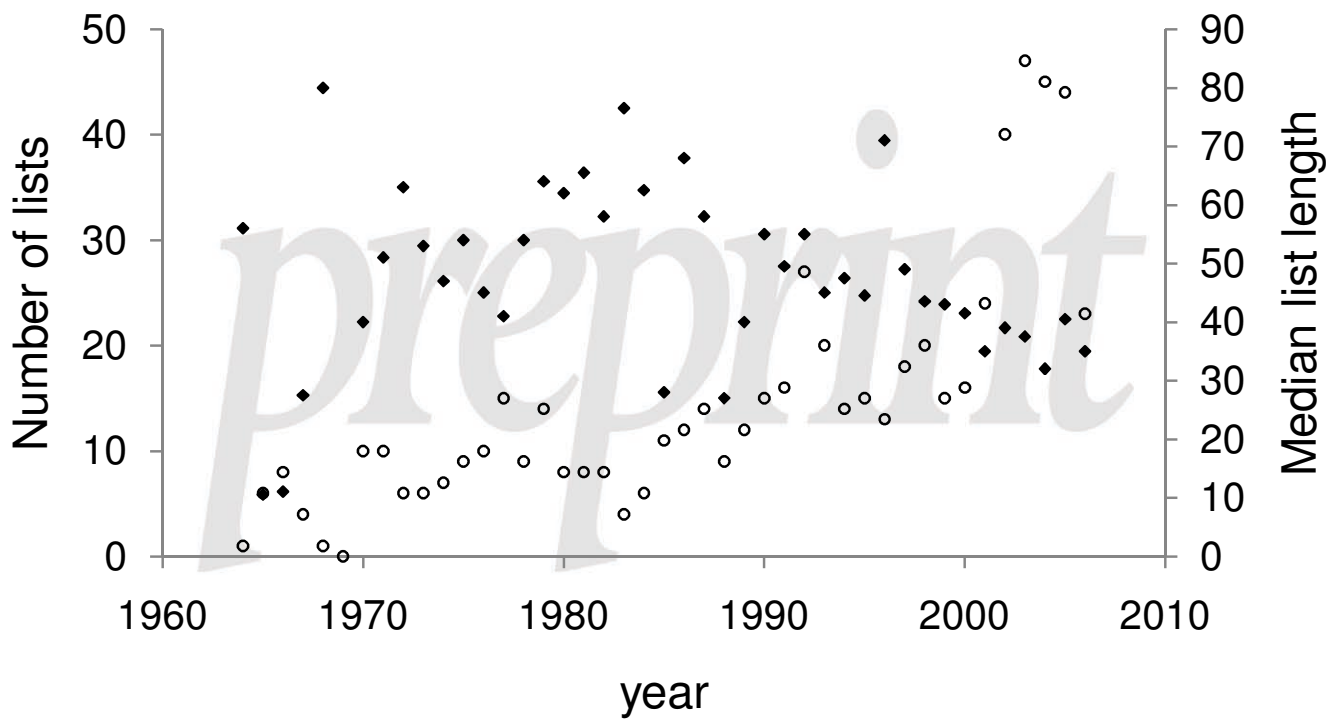
4 Figure 5. Magnitude (mean  $a_3$ ) and uncertainty (SD) of change through time. Symbols  
 5 indicate the posterior estimates for the means and standard deviation for  $a_3$ , the effect of time  
 6 on the probability of observing a given species on a standard list of 45 species. Dashed lines  
 7 indicate where the mean equals twice the standard deviation. Symbols below the dashed lines  
 8 are more likely to have changed in prevalence.

9

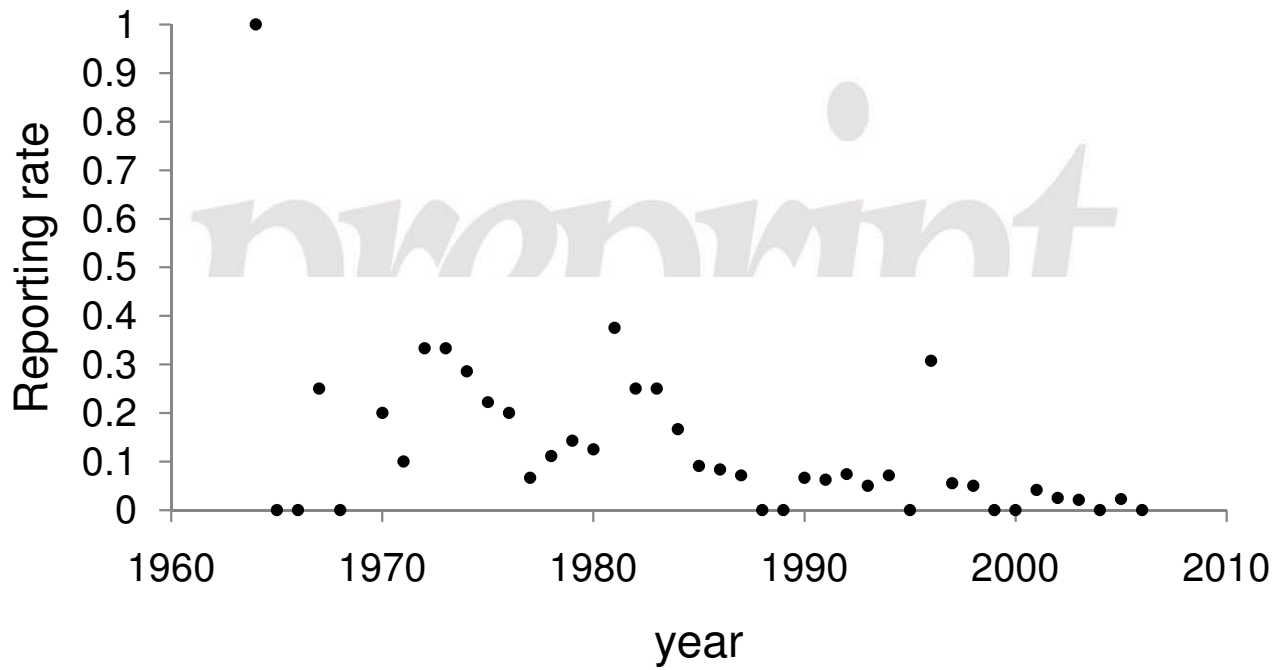
10 Figure 6. Relationship between the median of the intercept term ( $a_1$ ) and the median change-  
 11 per-year term ( $a_3$ ) from the full list length model. Data points represent species.

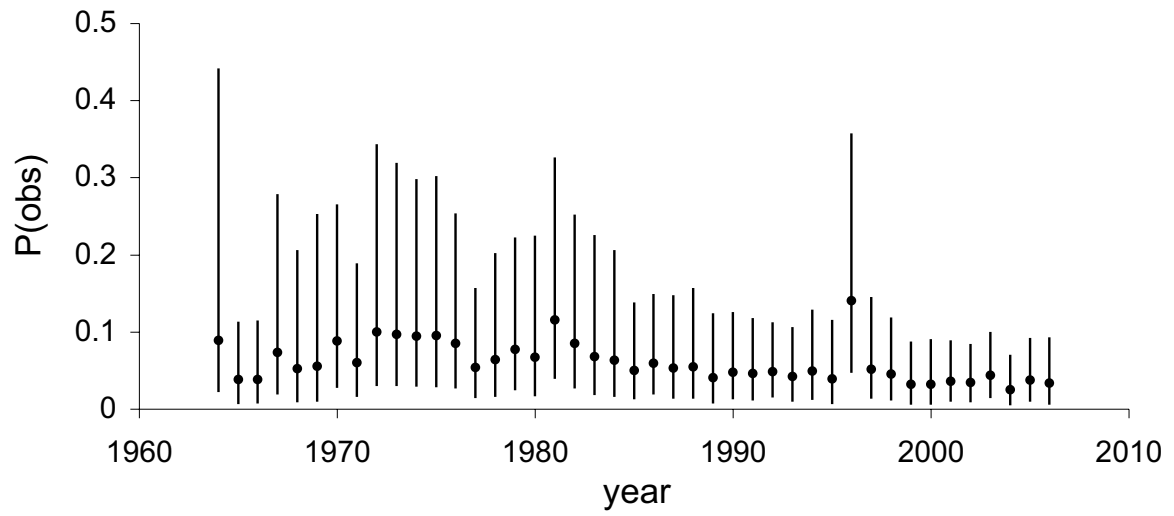
preprint

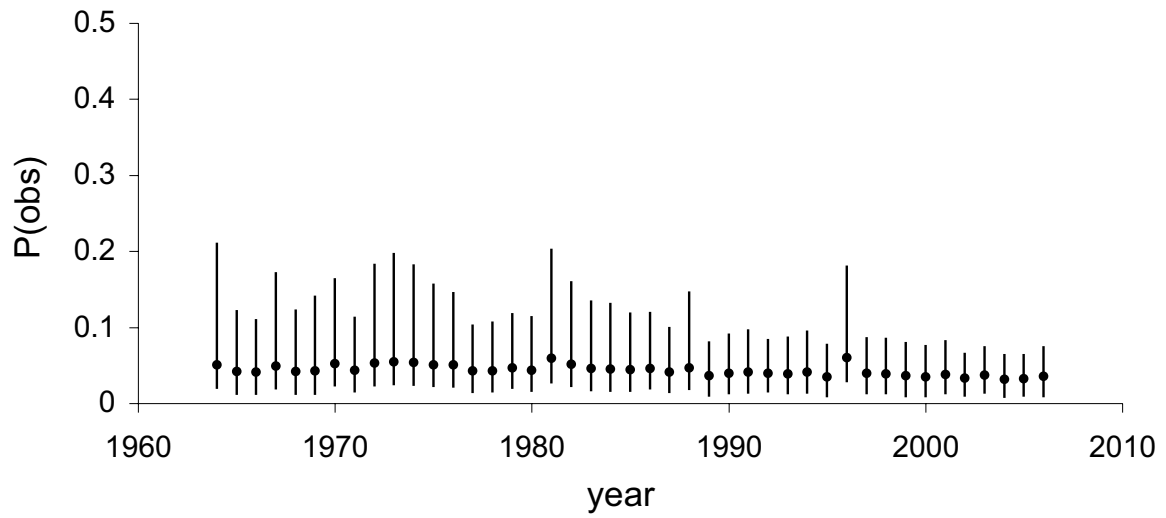
esa

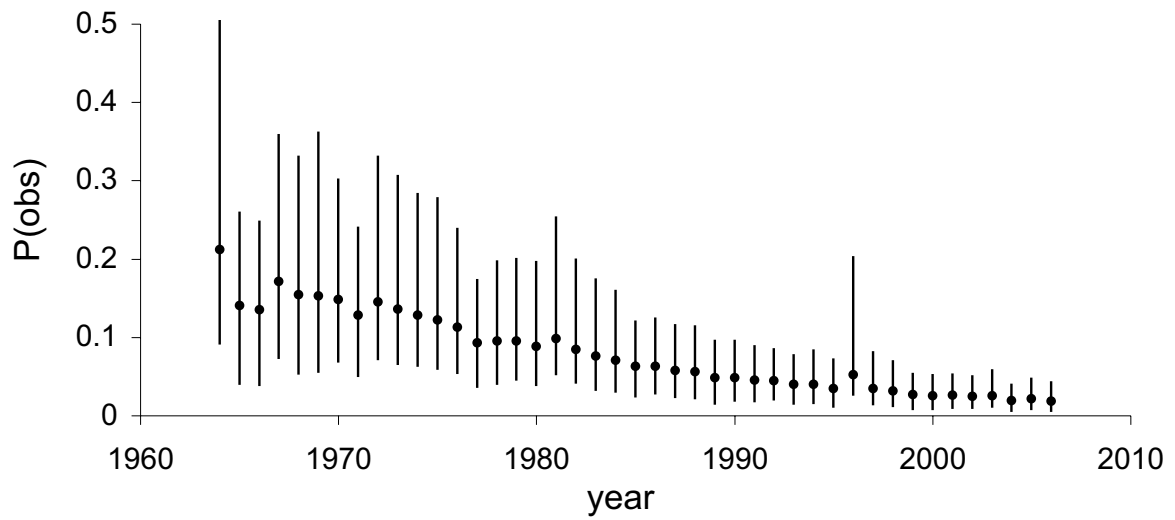


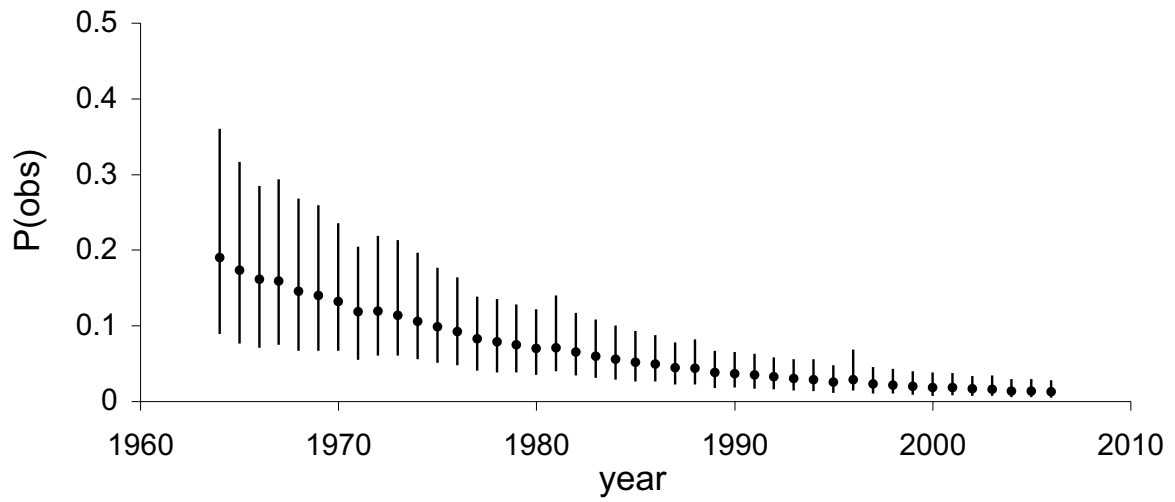
esa



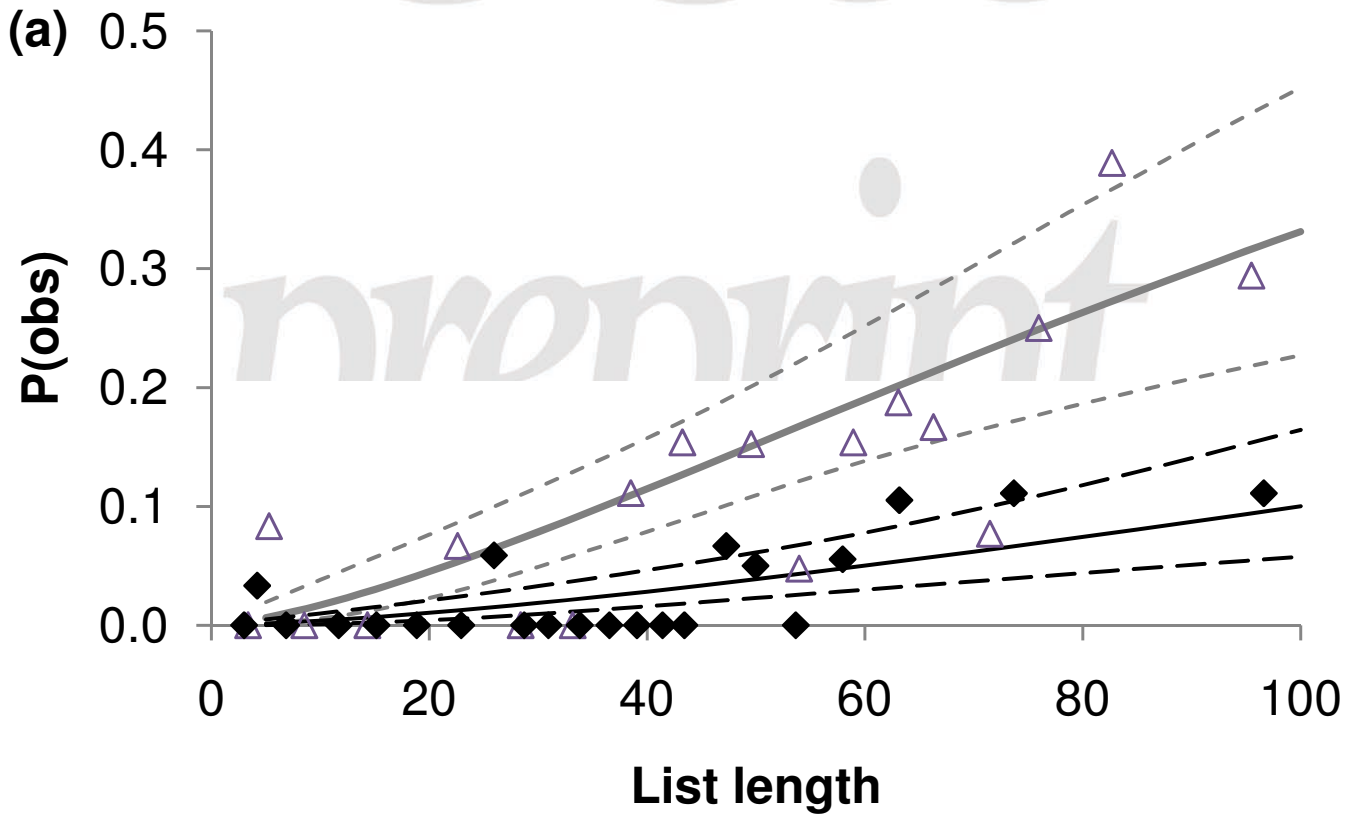




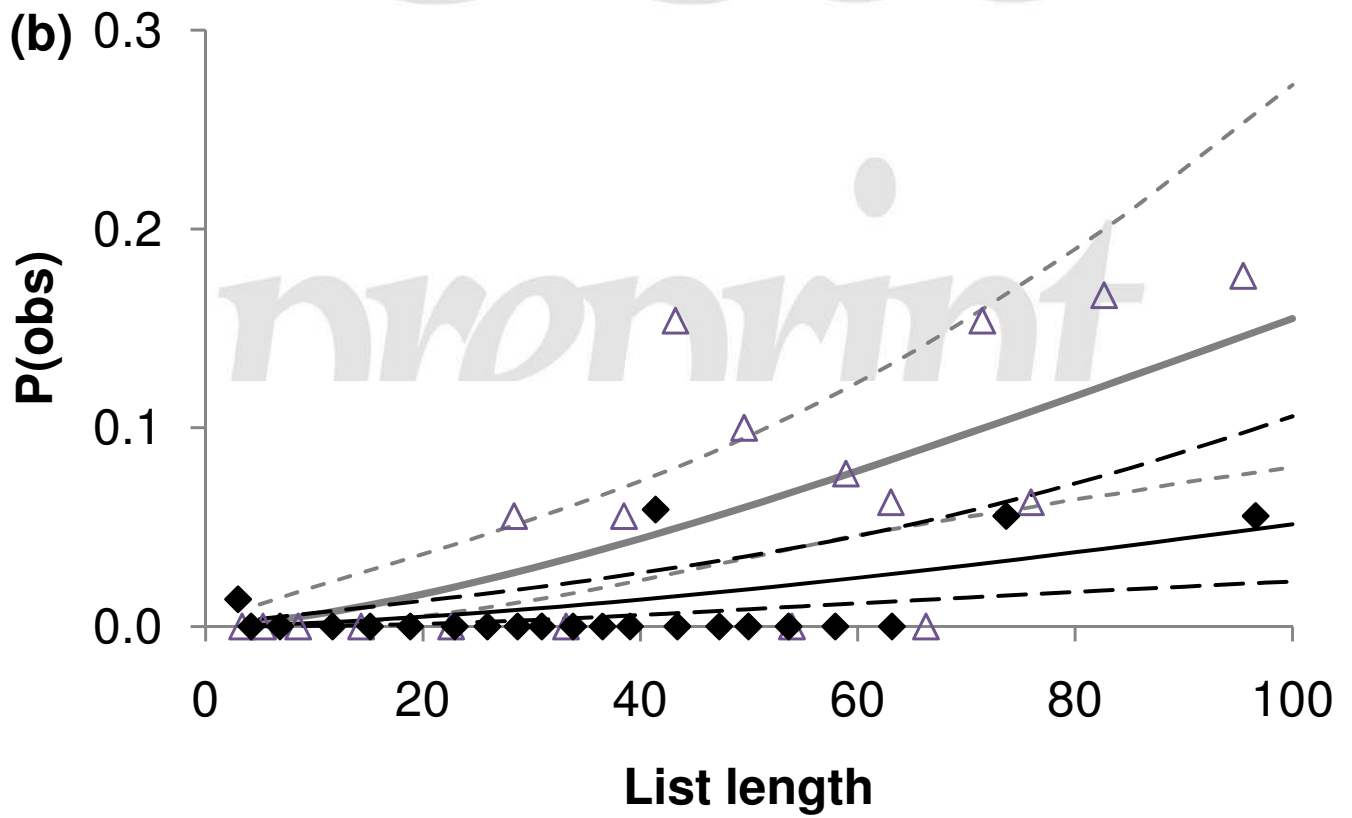




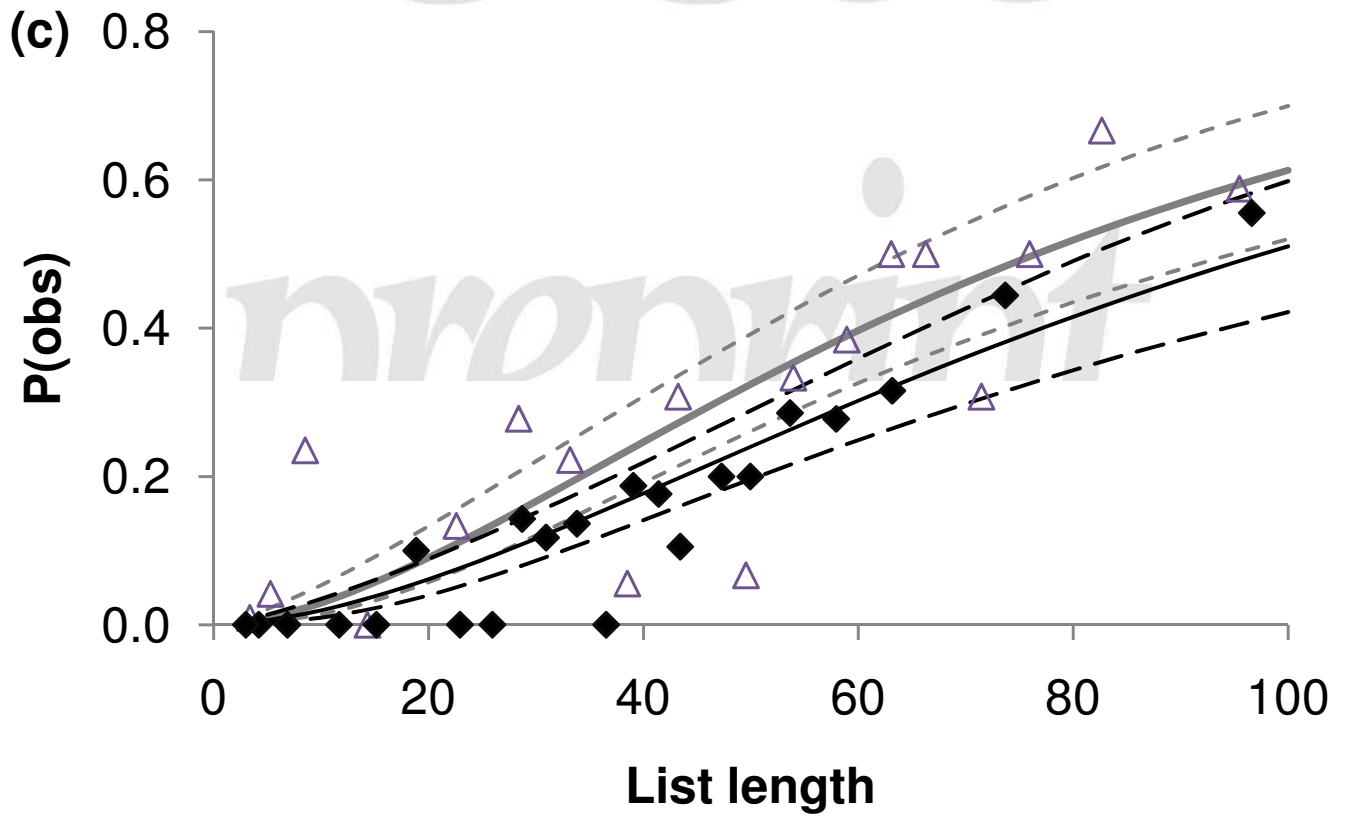
esa



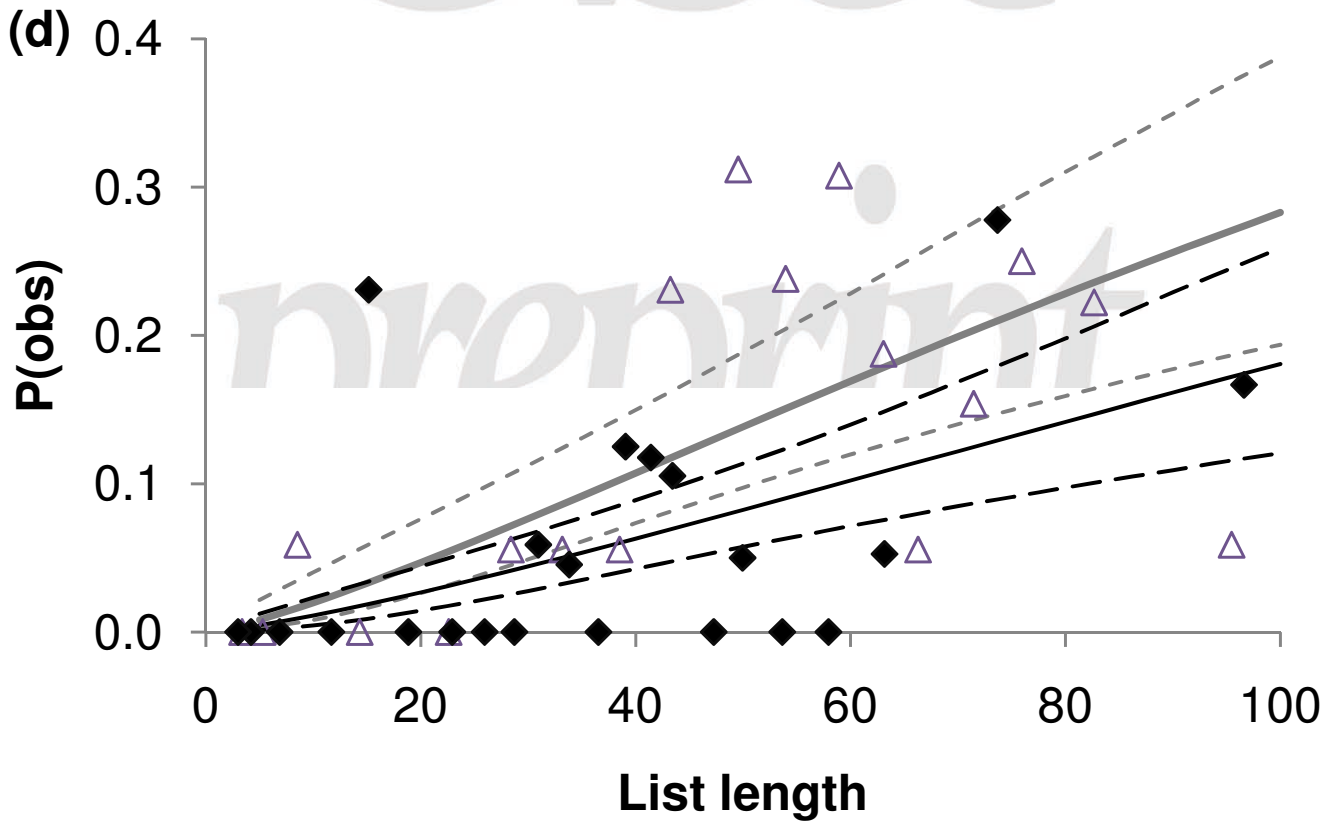
esa



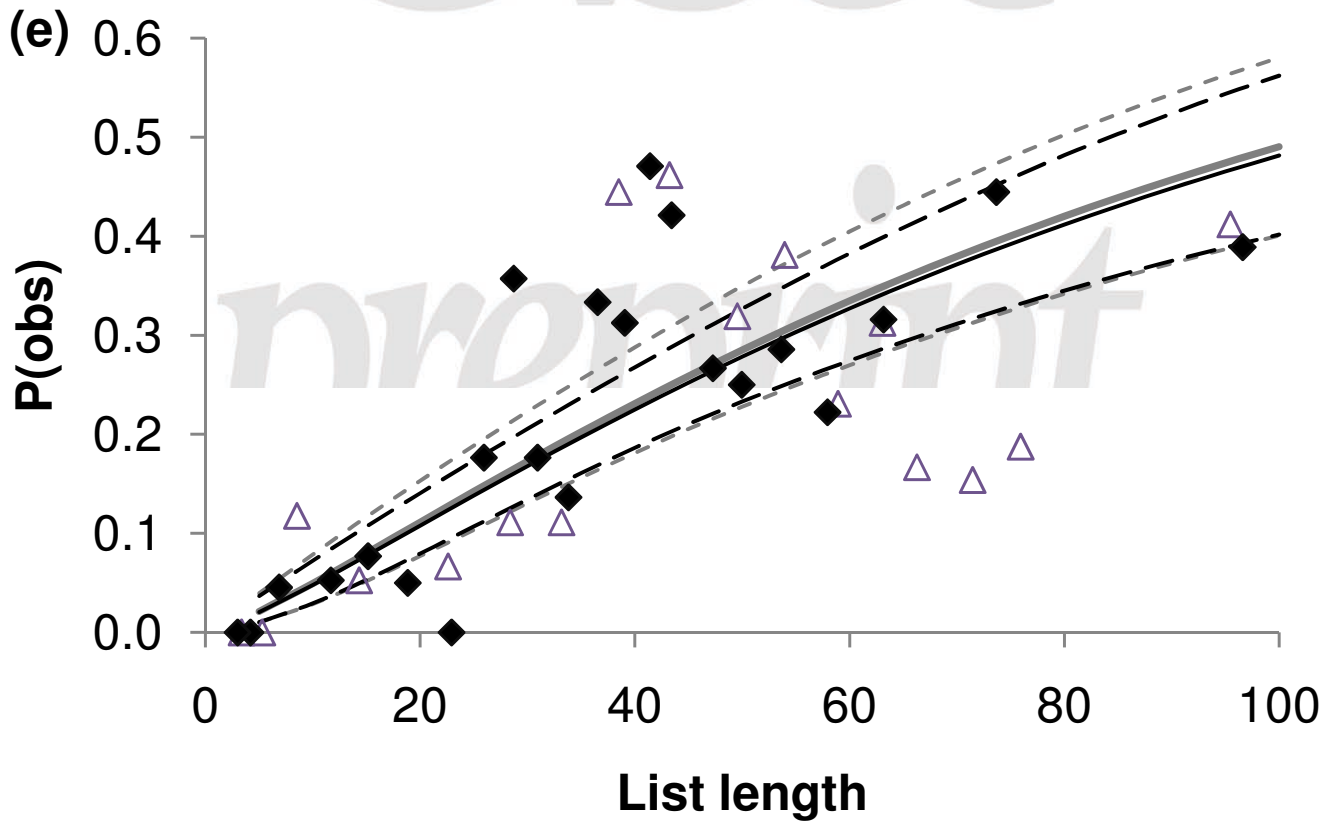
esa



esa



esa



esa

