Review of the first 25 years of the Garden Bird Survey 1981-82 to 2006-07

Martin Butterfield, Chris Davey, Michael Lenz, Nick Nicholls

Review completed September 2008 for the Canberra Ornithologists Group Inc.

Review of the first 25 years of the Garden Bird Survey- 1981-82 to 2006-07

EXECUTIVE SUMMARY

The aims of the Review of the first 25 years of the Garden Bird Survey (GBS) are outlined in the Terms of Reference provided by the Committee of the Canberra Ornithologists Group (COG). This report addresses the Terms of Reference and is to advise the general membership of COG and future GBS Coordinators of the current status of the GBS and to provide background for further analysis of the data. The Review is not a critique of any past analyses or publications.

At the start of the Garden Bird Survey in 1981 the stated aims were to 'increase our knowledge of common Canberra birds'. Since then the survey has developed a life of its own and no additional specific aims have been developed. We suggest that the present aims of the GBS are as follows:

To gain information on the distribution and abundance patterns of the birds of COG's Area of Interest and to provide a mechanism which helps members to better identify and appreciate the birds around them.

Some time has been spent examining the problems of data errors. Our conclusions are that the error rate is small and therefore unlikely to impact on past results or publications.

The Review team makes the following recommendations:

COUNTING RULES

Instead of determining whether an observer has been on site or not for the whole week in which recorded observations are made **we recommend that the wording be changed to "Include sightings made if you are present at the site for at least one day during the week and spend time observing birds".**

Of particular concern is that observers include in their counts individuals of the same species when the male and female are observed days apart. **It is recommended that the rule about counting the maximum during a single observation time should be reinforced to observers whenever possible.**

The counting rules should remain the same for all species. If information is required on movement patterns and abundance of for example honeyeaters passing through Canberra this information can only be obtained with a different survey protocol.

In any new instruction set DY (dependent young) should indicate an observation of a dependent young on site produced off site but a DF (dependent fledgling) should indicate a dependent young produced on site.

The question of a definition of dependent young (DY) has, not surprisingly, led to some confusion. For the purposes of the birds likely to be recorded within the COG area of interest the Review Team recommends the following definition:

Altricial species, that is those that are born with eyes closed, little or no down, unable to leave the nest and fed by parents, should be regarded as DY (or DF) from the time they leave the nest until they are no longer being fed by the parent or surrogate parent. Precocial species, that is those that are born with eyes open, covered with down and leave the nest after the first day or two, should be regarded as DY (or DF) until they can fly.

IMPROVING THE DESIGN OF THE DATABASE

The present database contains many inefficiencies and **this can only be corrected by a redesign of the GBS database.**

OBSERVER ERRORS

One of the outstanding problems is that there has been some inconsistency concerning the entry of abundance data when the observer is away or when the observer is at home but a comparison of data collected on the preceding and following weeks indicates that there has been an incomplete survey for that week. **We recommend that where it is indicated that the observer is away, the observer be regarded as at home if there are three or more species recorded. When it is indicated that the observer is at home we recommend that there be no change to the original data and just accept that the week was poorly surveyed.**

We note a common problem when, despite the observer being away, there is a breeding entry that has been derived from the breeding data collected during the preceding and following week. We recommend that this observation be omitted.

DATA MANAGEMENT

We recommend that:

- **Effort be put into developing a suitable (EXCEL/Visual Basic or similar) data entry system suitable for both central and distributed data entry.**
- **Any data submitted in an electronic format is backed up and retained by the COG database manager, using standard files to identify the GBS Table, collection year and site to which they refer.**
- **A Consultant be contracted to examine the problems and costs of redesigning the GBS database.**

- **The Consultant also examines the feasibility and cost of redesigning the data input system.**

- **An updated data entry program should be freely available, even though. an on-line data entry system is not recommended.**

- **A Register of issued charts be maintained by the GBS Coordinator.**

- **If consideration is given to improving the GBS data input system then this is done in conjunction with improving and updating the various input programs for the General Observations database.**

- **No changes be made to the survey protocol until a thorough analysis of the data set is completed after year 30, with any recommendations to come into effect from year 31.**

DATA PRESENTATION

It is recommended that there be a greater number of monthly and long-term graphs. In addition, trend-line and R^2 value could be included on the graphs where appropriate.

DATA ANALYSIS

We conclude that the inclusion of rural sites has little effect on the value of 'A', at least for the four species examined during this review and therefore **recommend that the practice of including all sites, in the COG area of concern when defining the level of abundance (A) should continue for the Annual Bird Report.** All sites are coded according to their location and so rural sites can easily be removed from any more detailed analysis.

- **The number of survey sites needs to be increased and COG members should be encouraged to continue to be, or to commence being, active in the GBS.**

- **Particular attention should be paid to recruiting those members of COG in Gungahlin and Tuggeranong to the survey as these urban areas are under-represented in the GBS.**

- **A detailed analysis, possibly along the lines provided by Dobbie and Welsh (unpublished) be commenced after 30 years of data collection.**

- It is recommended that **the final product should be a paper of suitable standard for submission to a quality, peer-reviewed science journal.** It is envisaged that only a suitable sample of species be used for analysis and that these species be chosen to demonstrate certain aspects of the data.

- **The Review Team does not have the required expertise to conduct such an analysis and recommends that a project brief be produced to outline the requirements of the analysis.**

- **It is recommended that enquiries be made to find a suitable person to conduct such an analysis.**

- **It is recommended that a second project brief be commissioned to examine the feasibility of redesigning the GBS database and to provide a data-input mechanism designed to capture the types of errors that have been identified in the review.**

- **It is recommended that the publication' Birds of Canberra Gardens' be updated and republished.**

Page

Executive Summary
Contents

Background 1
Terms of Reference 1

Section 1 Purpose of GBS and possible changes 4
1.1) Outcomes of the Terms of Reference 4
1.1.1) Objectives of the GBS 4
1.1.2) Stake holders 5
1.1.3) Major constraints 5
1.2) Data quality 6
1.3) Changes since start of GBS 10

Section 2 Methodology 16
2.1) Extent of current rules and definitions 16
2.1.1) Questionnaire results 16
Problems 18
Breeding codes 18
2.1.2) Database design and Analysis of errors 19
2.1.3) Adequacy of data 23
2.1.3.1) Garden Bird Survey coverage 23
2.1.3.2) Sparseness of data 26
2.1.3.3) Number of survey sites 26
2.2) Bringing the GBS and other data together 28
2.3) Changes to survey protocol 34

Section 3 Impact of collection methodology and operational constraints. 35
3.1) Known use of GBS data 35
3.2) Data analysis 37
3.2.1) Present summary statistics 37
3.2.2) Data comparison- up to 21 years and an additional four years 38
3.2.3) Improved methods to show trend lines 47
3.2.4) Regional variation in summary statistics 48

Section 4 Where to from here
4.1) Changes to operational procedures 49
4.2) Updating 'Birds of Canberra Gardens' 50

References 51

Appendix 53

Background

In 1981 the Canberra Ornithologists Group (COG) decided to undertake a survey of the birds of Canberra's gardens. This undertaking became known as the Garden Bird Survey (GBS) and the results of the GBS have been published as a companion of the Annual Bird Report (ABR) produced periodically by COG as a part of the journal Canberra Bird Notes. The history, background and methodology to the GBS have been well covered by Veerman (2002).

In 2000 the Canberra Ornithologists Group published an analysis using data for the first 17 years of the GBS covering the period July 1981 to June 1998 (Canberra Ornithologists Group 2000). The then GBS coordinator Philip Veerman undertook an analysis of the GBS using data from the first 18 years of the study (Veerman 2002). A second analysis was conducted (Veerman 2003) using the first 21 years of data and this has been subsequently updated (Veerman 2006).

By 2004 the GBS had been running for 23 years. The COG Committee decided that the GBS was an important element of COG activities and in principle should be continued and that a thorough review of the GBS should be undertaken using the first 25 years of data.

To aid in the review process a draft document outlining the Terms of Reference was provided by the Committee and Nick Nicholls was invited to conduct the review. During the 2005-06 and part of the 2006-07 periods David Rosalky was responsible for the coordination of the GBS and in November 2005 was invited by the COG committee to revitalize the review process. In January 2006 he invited long-term contributors to the GBS Henry Nix, Michael Lenz and Chris Davey to provide him with background material and ideas on how the review should progress. In July 2006 David asked Chris Davey to take over the reins of the review and this was agreed to by the COG committee at their meeting of 19th July. In January 2007 Martin Butterfield took over the role as GBS coordinator and the Review Team was then reformed; the membership comprising Martin Butterfield, Chris Davey, Michael Lenz and Nick Nicholls.

The Review Team accepted the draft Terms of Reference that had been produced sometime in 2004/05 by the then COG committee and in the light of present developments the Review Team provided a few additional terms to the draft document.

Terms of Reference

(Terms in italics added by the Review Team)

1. Consider the purposes of the GBS as defined when the survey was first established and identify areas in which changes to these purposes might be needed.
    1. The outcome of this Term of Reference should include:
        1 a broad but brief statement of the objectives of the GBS;

2 identify the major actual and potential stakeholders in the GBS; and

3 identify major constraints upon the operational aspects of the GBS.

2 This term of reference should include a description of the concepts of 'quality' used in the Review

3 Areas of change identified could include:

i. topics related to scientific knowledge;
ii. development of Canberra society;
iii. *the emergence of 'citizen science' as a research method;* and
iv. changes in technology.

2. Consider the methodology of collecting the observations for the GBS including physical processes and logical definition of data items

1 The outcomes of this Term of Reference should include a scientific review indicating the extent to which the **current rules and definitions** meet accepted scientific standards, both ornithological and statistical.

1 The review should particularly consider items for which formal definitions do not exist and recommend definitions as required;

2 Examine the types of errors and their prevalence in the GBS database;

3 To the extent possible this Term of Reference should include

1. a statistical analysis indicating the extent to which changes in the sample of sites at which observations are made should be allowed for in any form of statistical analysis;
2. specific attention should be given to the impact of the "sparseness" of the data set, examining methods of accommodating a high proportion of both zero and null cells: and
3. comment on any methods for identifying those species for which data are sufficiently reliable to permit more detailed analysis (e.g. for specific suburbs or aggregations of suburbs).

4 Where possible in this Term of Reference examples or case studies should be cited.

2 A specific recommendation should be made regarding the potential for bringing the methodology of the GBS and other data collections (especially the "General Observations" of COG) closer together.

3 Noting the value of the time series formed by current GBS data the review should carefully examine any proposals for change to ensure that the significance of breaks in the series is minimised.

3. Consider the impact of the collection methodology and the operational constraints on the ways in which the data can, could, or should be analysed and disseminated, both for the general public and for the scientific community. The outcomes of this Term of Reference should include:

    1 Commentary on experiences gained from existing uses of the GBS data set, including published works and known use in research projects.

    2 a broad set of guidelines for users of the data set (or aggregates and analyses based on it) explaining the characteristics of the data set.

    3 Guidelines recommended scientifically appropriate approaches to the dissemination of data using information for specific species or sets of species as case studies.

4. Noting the outcomes of Terms of Reference 2 and 3 make recommendations as appropriate on the data processing systems needed to transform observations of birds in the GBS into a database suitable for analysis.

    1 Specify any changes to operational procedures and definitions required to take the GBS forwards for the next 10 years in a way which satisfies the interest of the birding community.

    2 A set of recommendations for any updated book on the birds of Canberra suitable for sale to the general public[1];

[1] The compilation and production of this book is specifically excluded from the Terms of Reference of this group.

For reasons of clarity Section and Sub-section titles may have been changed but numbering remains the same as that provided by the original Terms of Reference.

1) The purpose of the Garden Bird Survey and possible changes

1.1) Outcomes of Terms of Reference

1.1.1) Objectives of the Garden Bird Survey

Hermes (1981) when announcing the start of the GBS went no further than to state that "The new Garden Bird Chart has been produced to increase our knowledge of common Canberra birds".

In the publication 'Birds of Canberra Gardens' (COG 2000) the only statement concerning the aims of the GBS was:

"In establishing (the) survey COG believed that its members could contribute to a useful bird survey from their homes, with minimal cost and effort. This would provide individual members with an opportunity to contribute in a continuing way to the bird monitoring activities of the group. It would be interesting to monitor the seasonal and long-term changes in abundance of bird species in Canberra gardens".

It would appear therefore that the original aim of the GBS was to increase knowledge of Canberra's birds although the instructions at the time were quite specific in indicating that it was only species that were actually using gardens that were to be recorded.

Originally the GBS was a tool to learn more about the birds in our gardens. Only later did it turn into a monitoring tool for Canberra's birds in general without many changes in methodology. For all participants the GBS became an important learning exercise, not all features of bird behaviour could be foreseen and existing procedures did not cover all scenarios. Through the GBS we have learnt a lot about the birds of Canberra and much of that could not necessarily have been foreseen. A prime example is the issue of dependent young (DY) and of some of our parrots begging and being fed in winter.

Since its inception in 1981 the survey has developed a life of its own and no additional specific aims having been developed. After 25 years the Review Team feels that it is time specific aims of the project were set down and we suggest the following:

The aims of the GBS are to gain information on the distribution and abundance patterns of the birds of COG's Area of Interest and to provide a mechanism which helps members to better identify and appreciate the birds around them. This definition thus includes both the settled parts of Canberra and the growing number of rural blocks. The inclusion of "abundance" in this objective is a crucial aspect of the COG GBS and adds considerable complication to any statistical analyses. Most other examples of citizen based garden bird

surveys (refer to Cannon et al. 2005) focus entirely on occurrence. It is noted that 'Birds Australia' is now collecting abundance data rather than the presence/absence data.

It is also recognized that at the start of the GBS many participants were living in places with a garden around the home. Over the years the Canberra landscape has changes significantly (in-fills, town houses, double homes on single blocks and in the new suburbs often small blocks with large homes). So, now a GBS site may or may not have a garden in the traditional sense.

There appears to be no reason to abandon abundance data in favour of presence/absence or equivalent such as reporting rates. It is always possible to convert abundance to presence/absence but not *vice-versa*. At present we can see no reason why the current measure of abundance of maximum numbers per week should generate any specific statistical issues.

1.1.2) Stakeholders

The stakeholders are determined by the broad aims of the survey and to allow information to be provided to the membership of COG, professional and amateur scientists, state and local government agencies and non-government agencies such as conservation and other environmental bodies. However, these entities are not the only users of the survey results, particularly when the potential users are considered. The following groups should also be clearly identified:

- The general public engaging in debate about issues to which knowledge about birds could contribute (e.g. global warming; urban sprawl; preservation of open spaces etc);
- The media, in both contributing to the general debate mentioned above and in their roles of reviewing aspects of society and informing the public;
- Politicians in the formulation and communication of policies. The second aim provides an opportunity for the COG membership and other contributors to the GBS to further their knowledge of the local birds and to improve their bird survey techniques.

It is also important to recognise COG as an entity as having a stake in the GBS supplementary to that of the membership. This recognises that the GBS provides COG with a vehicle which can be exploited, through publicity and contribution to debates, to furthering the broader aims of COG.

1.1.3) Major constraints

A key aspect of the GBS is that it is largely a flexible system in which people participate as much as they are able. If a person is retired they may put in several hours per week watching the birds in their site. On the other hand if they are at work for 50+ hours a week they may only be able to put in one or two hours in some weeks. If a more rigorous model (e.g.: undertake a 20 minute survey once a

week between 7 and 8 am) was employed the number of participants would probably be greatly reduced.

By basing the survey around areas where people routinely spend a lot of time (e.g. home or work or in some cases the frequently visited homes of relatives) the financial costs of participation are minimised.

The measurement systems required are again relatively simple and easy for observer/participant to employ. For example, there is no requirement to accurately determine the age and sex of the birds as is employed say for banding regimes.

It is believed – largely anecdotally – that people get satisfaction from having a physical chart to complete and display on their fridge etc. This gives some issues for consideration in data processing. In addition, the availability of a hard copy chart makes it far easier to investigate reporting errors, in contrast to provision of a spreadsheet or other electronic system.

1.2) Data quality

Four elements of data quality can be recognized:

- Relevance to user needs;
- Timeliness;
- Cost; and
- Accuracy.

Relevance

The relevance element is crucial: if the activity isn't relevant to data needs then it cannot be considered high quality. It seems that many people believe that the GBS is useful and relevant – hardly surprising or 70 observers would not participate in the collection. However, there is a close link between usefulness and accuracy: if there are 'problems' with the accuracy of the data from poor observation or data entry errors, are these sufficiently egregious to negate users' sensible requirements? The work of the Review Team suggests that by and large the data collected in the survey is certainly fit for the purposes to which it has been applied (and thus it can be regarded as relevant).

Timeliness

A key feature of most scientific data is that it must be available promptly to be of any real benefit. This is particularly so when the data may be used for monitoring purposes. To illustrate this point it is of no value to have a data set that shows a species to be threatened if the data only become available after the last breeding pair has died or been 'collected'!

It is also the case that the data only needs to be available in a time frame to suit the purposes for which it is used. In the case of the GBS the principal use of the data is to provide medium term (i.e. annual) indicators of the abundance and frequency of species of birds. It is usually the case – certainly in the past decade – that the annual data have been published in the Annual Bird Report within about 9 months of the end of the data collection year. This seems quite acceptable and it is believed that efforts to improve significantly on this performance would involve a considerable extra cost and/or a decline in accuracy, for no increase in relevance.

Cost

Two elements of cost can be recognised. The first is the cost to the organisation (Canberra Ornithologists Group Inc.) and the second is the cost to members of the organisation who either contribute data observations or who add value in one way or another to the observations.

Costs to the organisation

At the time of writing the only cost to the organisation is the cost of printing the charts approximately every three years. While the cost is not trivial – approximately $3 per sheet due to the need for a large sheet of high quality paper – it is not extreme for an organisation such as COG. In addition, the organisation generated a positive cash flow through sales of "Birds of Canberra Gardens" which was based on the results of the Survey. Potential extra costs to the organisation could be incurred if it was decided to employ professional resources to re-develop the computer system.

At present the data entry system is based upon a Microsoft Access system developed by *inter alia* Paul Fennell and Philip Veerman (Veerman 2002, 2003, 2006) and works quite well through the services of Kay Hahne and on occasions by Anne Hall and the GBS Coordinator in entering the bulk of the data. This would appear to be a case of a system not being broken and thus not requiring to be fixed.

The data repository, based upon a Microsoft Access system, developed by Graeme Deaker and Philip Veerman on behalf of COG, functions well for a system with as many variables as this survey. The size of the core database now exceeds the capacity of some alternative approaches such as Microsoft Excel and designs based upon more advanced database structures would require higher quality IT skills than held by the current GBS Coordinator.

Some additional design work and coding has been undertaken to make some approaches to interrogating the database more efficient and there could be benefits to formalising this work to make it more transferable between years. It is to be hoped that the skills to undertake this can be found, at no cost within the organisation.

Costs to members

The cost to most members is the cost of their time in recording the data. Given the large number of repeat recorders it seems clear that this cost is currently seen by those members as being balanced by their enjoyment in recording their observations and/or

contributing to the objectives of the organisation. However, it must be recognised that any change to the GBS which either dilutes the pleasure or adds to the time required to participate may work to undermine this balance. Any significant change to the GBS must be carefully tested to ensure that it is satisfactory to the observers.

Those members involved in operating the GBS also seem to derive sufficient pleasure from their activities to justify the investment of time. Again this must be borne in mind in formulating any proposals for change. Authors of the Annual Bird Reports also invest significant time in analyzing the GBS data and care should be taken not to impose further burdens upon them.

Accuracy

Errors may be defined as impossibilities or improbabilities.
Impossibilities, also known as fatal or physical errors, occur when there has been a breach of the rules of collection. In the case of the GBS this can occur:

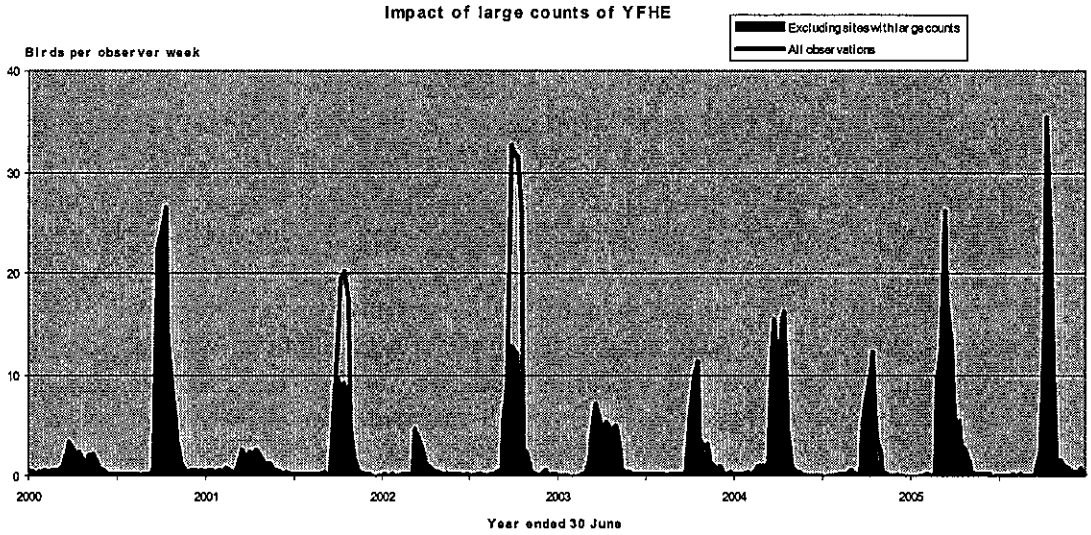
- where birds are reported on the chart when the observers have indicated that they are away. This is a relatively common error and presumably arises because an observer is away for most of the week but on returning home notes something of interest. Another example of this type of error is where an observer enters an abundance figure on a specific week and then draws an arrow across many other weeks irrespective of whether the site is being observed or not.
- The opposite can also occur when an incorrect week has been crossed out and no birds are entered onto the database despite the observer being at home.
- If there is a breeding record there should be an associated abundance record.
- Another impossibility, which would arise during data entry or processing rather than observation, is that there should be no duplicate records.
- Ever since the GBS commenced some observers have followed the rule of counting the maximum number of birds seen at a point in time while others have – at least from some species - accumulated the maximum number seen in the course of a week. It was said to at least one member of the Review Team when commencing their first Chart that the choice between these was optional.

This has been a particular problem for the Yellow-faced Honeyeater during the autumn migration when the maximum numbers at a point in time is the size of one migration flock (perhaps one or two hundred) while the maximum number of different birds passing through in a week (i.e. the sum of all the various flocks) can be several thousand.

Veerman resolved this problem in the database by arbitrarily setting the maximum number of birds at a point in time to 500 (Veerman 2003, p 21) and it was not possible to use the original data in the current investigation, although data are still available from the original GBS charts.

In later years most observers seemed to have followed the rules, although two sites have continued to accumulate flocks. To illustrate the difference between the two measures the chart below shows the number of birds per site per week for all active sites as a line,

with an area filled in to show the number of birds per site per week excluding the two accumulating sites.



The difference is very pronounced. Further, it also suggests a difference in interpretation, in that a large drop off in the autumn migration in 2003 could be interpreted as showing the impact of the January 2003 bushfires in the ranges, whereas excluding the accumulating sites suggests that the major impact on the migration through urban Canberra might well have been the December 2001 fires which destroyed some of the pine forests to the North of the Cotter Road.

In processing Year 26 it was apparent that only one site was accumulating the numbers and in discussion with the observer an alternate set of values representing the maximum flock size was adopted. It thus seems probable that all observers will report following the rules in the future. To assist in achieving this outcome it is suggested that the instructions on the GBS chart be emphasized and that all observers be advised to count the maximum number of birds in a site at a point in time.

- A further source of impossibilities is when an observer does not report an actual number of birds but merely ticks to indicate presence or absence of the species from a site during the week. In the course of reviewing the data set for this Review it was discovered that one Chart, from year 1, only included ticks. The Chart was annotated to indicate that one of the Coordinators had used G values to estimate the average number of birds present. The Review Team decided that this data was not in accordance with the basic concept of the Survey and deleted the chart.

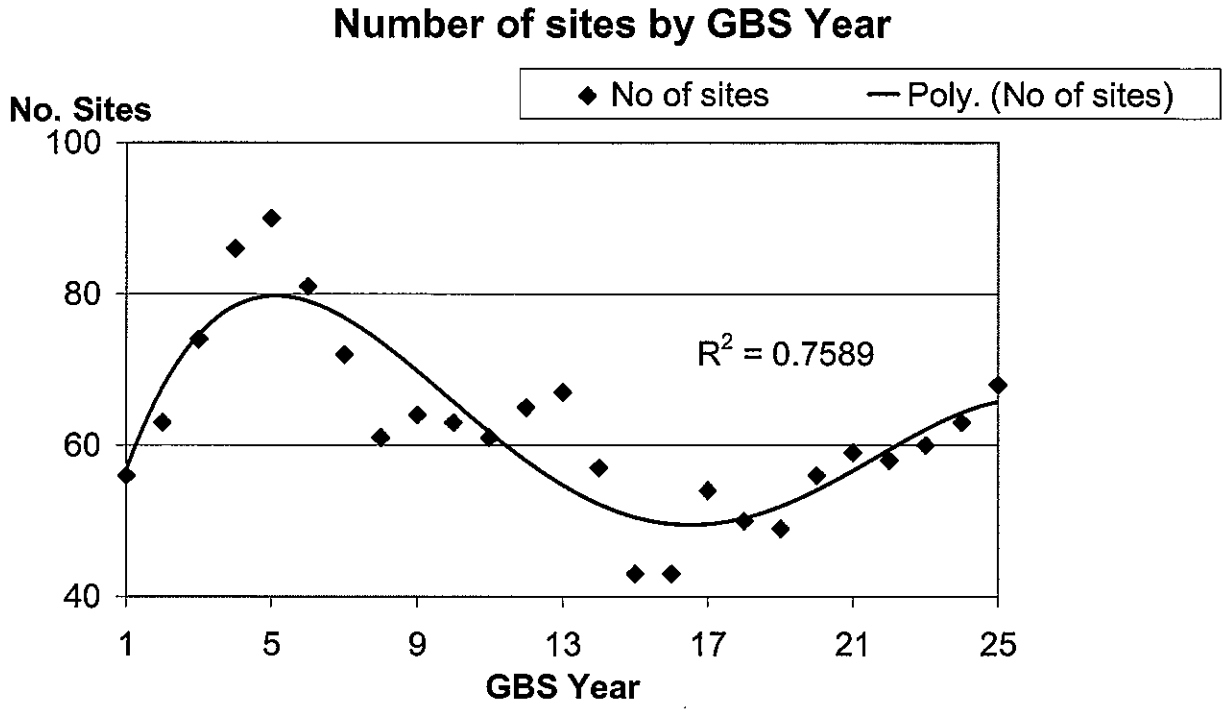
During the course of the Review the Team found a few instances of "impossible" data due to data entry errors of one sort or another. With the exception of the invention of data from ticks on a Chart in Year 1, the action of the Review Team was to set the data in the database to that reported on the Chart.

Improbabilities also known as non-fatal or logical errors occur when the data are incorrect. Such errors can be picked up by the Coordinator's quality control checks or by scrutiny by authors of species accounts.
An obvious form of improbability is where the observer misidentifies a bird and includes this in the Chart. In many (most?) cases this is impossible to pick up but where the bird is reported out of season (for example, a Common Koel in July) or in unusual numbers (for example 50 Australian Ravens) this can be detected and queried back to the observer. Reports of unusual birds (for example Pink Robin) can also be picked up and queried where not endorsed by the Rarities Panel. Where data errors are confined to a typographical error it could be considered an impossibility error but because the error can only be picked up *post hoc* by a logical process we prefer to regard this as logical. Two examples that have been detected involved a typographical error of 335 House Sparrows instead of 35 and 9 Pallid Cuckoos where an observer had noted in the body of the Chart that a single Pallid Cuckoo had been first recorded on the 9th August.

Typically, the impossibilities are seen as more important than the improbabilities because they can affect the operation of the entire system. They should also be more amenable to identification, especially when the data set has relatively few constraints. The improbabilities are likely to only affect a single species and by and large to be difficult to identify.

1.3) Changes since the start of the GBS

Since 1981-82 the population of the ACT has increased from 229,000 (source ABS Bulletin 3105.065.001) to 339,000 at June 2007 (source ABS Bulletin 3301.0). Development has continued in the Woden Valley and now includes the southern suburbs of Tuggeranong. In addition, the development of the township of Gungahlin is well underway. But, despite a 48% population increase there has not been a proportional increase in the number of sites being surveyed.



From an analysis of the distribution of the survey sites over the past number of years it is obvious that the survey does not collect data from a random selection of sites of different ages neither does it represent a random set of locations. The lack of sites within the newly developed areas of Canberra is responsible for the uneven distribution in location. In addition, a change to the instructions set down for collecting data has led to a fundamental change in the data collected, see p. 14.

Analysis of site frequency

There are considerable differences in summarising data using means calculated for levels of a single factor (for example months or years across many years) and the calculation of means for the combinations of levels for more than one factor at a time. For example, monthly values for year 1 followed by monthly values for year 2 etc. The former practice is easier but hides the variability present across the second and subsequent factors. It is well understood that the significance of a factor can change dramatically when a second and subsequent factor is added to the analysis. One of the current shortcomings of the analysis of the GBS data has been the use of marginal summaries (that is summarising the data for levels of one factor at a time) without exploring the potential for the influence of other factors. The question of the source and magnitude of the variability of the data to date has largely been ignored with the exception of Dobbie and Welsh (2001). For example, for 'A' values we have site, week and year as sources of potential variability. Modern statistical analytical techniques have the potential to accommodate these different sources of variability. The failure to take them into account can have significant

impacts on the interpretation of the responses and the inferences drawn from any analysis.

In addition, there are two fundamental problems that need to be overcome when describing data such as that produced by the GBS. The first is the question of zero's in the data set, that is where an observer, for all intents and purposes, records a bird as having not been observed and therefore by inference not present. This question was raised in the Dobbie and Welsh (undated) report and in their analyses they attempted to overcome the problem. Unfortunately, the GBS data they were supplied with did not permit them to separate between zero when no individuals were observed in a week and zeros when no observations were made in a week.

The second is the question of 'sparseness', which is the creation of an incomplete data set because there were never any observations in the first place, and is for instance produced by sites no longer being observed or by inconsistent observations at a site. Much has been written on this issue and it is difficult to comment on the implications of the level of sparseness and requires a level of statistical expertise beyond that of the Review Team.

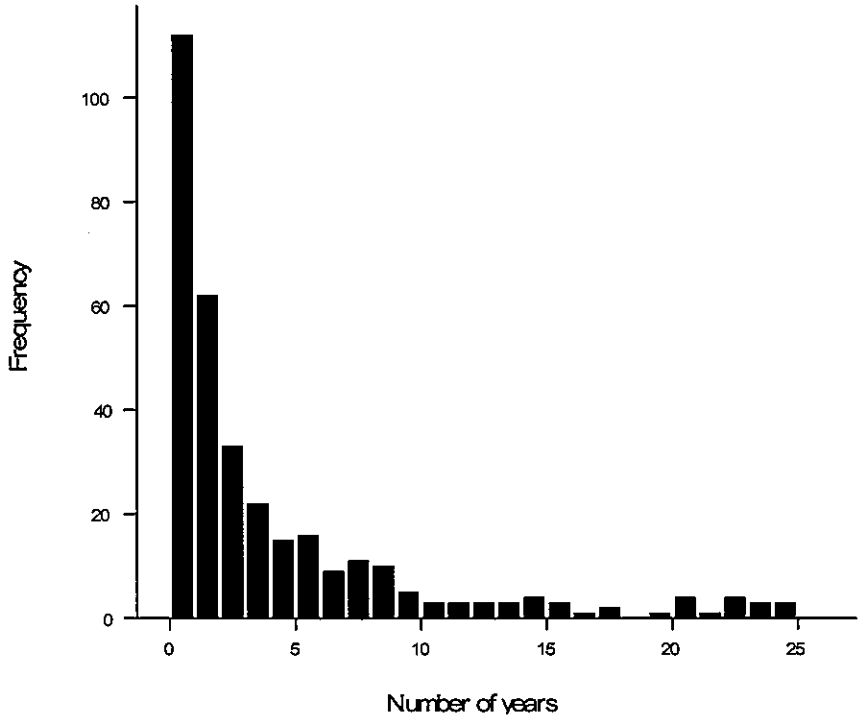
A total of 333 sites contributed 1561 GBS sheets over the 25 years. Of the maximum number of possible site year combinations (333 sites by 25 years) the available sheets represent 18.75% of the combinations. A more realistic estimate of the sparseness of the data is by estimating the proportion of cells from the start for each site that is filled. The number of such cells is 5603 and the proportion represented by charts is 27.9%. This low rate of site year monitoring is potentially important as most models that are likely to be fitted to the data will involve both year and site effects. There are some 50 sites that have breaks in the record from a single year to up to 17 years. A few sites have multiple breaks in the record.

Of the 333 sites only 15 (or 4.5% of the sites) recorded 21 or more years. At the other extreme a total of 174 (or 52.3% of the sites) were recorded once or twice only. This means that 22.1% of the sheets come from just 15 sites and 15.1% of the sheets from 174 sites or the fifty most surveyed sites contribute approximately 50% of the sheets whilst the remaining 280 sheets contribute the remainder.

It is worth noting that the 15 sites monitored for 21 or more years are wide spread across Canberra except for Tuggeranong other than Kambah.

GBS Site Number	Number of years	Suburb	Region	Observer
38	25	Aranda	Belconnen	Hahne
46	25	Holt	Belconnen	Davey
15	25	Ainslie	North Canberra	Lenz
79	24	Melba	Belconnen	Moncur
77	24	Weetangera	Belconnen	Wood
62	24	Campbell	North Canberra	Cleland
32	23	Chapman	Weston Creek	Holland
54	23	Kambah	Tuggeranong	Wilson
22	23	Hughes	Woden	Anderson
60	23	Yarralumla	South Canberra	Cannon
126	22	Cook	Belconnen	Brookfield
152	21	Kambah	Tuggeranong	Veerman
13	21	Ainslie	North Canberra	Barr
28	21	Farrer	Woden	Rentz
143	21	Weston	Weston Creek	Bounds

Details of sites that have been monitored for 21 or more years.



The number of sites recorded for specific number of years. Note that there were no sites recorded for 19 years out of the 25 years.

Changes to the instruction set

There have been three major changes to the design of the GBS chart, all occurred with the introduction of Version 3 in July 1993- Year 13, (Veerman 2002, 2003, 2006). For the first two versions of the chart the instructions were as follows:

> *Enter the maximum number of birds of each species seen at any one time within 100m of your house for each 7 day period.*
>
> *If birds are not dependent on your area (i.e. are not feeding, perching or hawking) or if they are seen flying over or are outside your area then underline the record.*

In Version 3 the instructions were changed to:

> *Decide on, and keep to, a survey area equivalent to a 100 metre radius (3.1 hectares) from house/work and include only observations from within or over this area. There is no altitude limit and no requirement for the birds to be 'using' the area.*

From Version 3 the site did not have to be a garden, (although this was never really the case as 61 of the 836 Charts completed before year 13 came from clearly non-residential addresses), a house did not have to be the centre of the site and birds not 'using' the area could no longer be identified from the data. These changes imply that the number recorded no longer reflects an estimate of species using gardens but rather provide an index of abundance within the urban area. The emphasis of the survey was therefore directed away from the garden *per se* to a survey using sites that captured data about birds using a suburb. This means that it is now not possible, without seeking further data from respondents, to examine the data in terms of garden type or garden age but it is possible to examine the data in terms of suburb location, suburb age and to examine questions concerning changes in general distribution and changes in seasonality and abundance.

There has recently been an increase in the awareness of the importance of 'citizen science' (Greenwood, 2007) and the Canberra Ornithologists Group GBS has been recognized as an important survey in the area of this type of work. Citizen science based projects have the advantage of being able to capture data from a large geographical area and over a long time frame. Citizen-science has become increasingly important because of the difficulty for institutions to generate long-term data sets due to funding shortages. Further, the general public is more and more aware of environmental problems and of the need to collect such data.

One problem with surveys that collect information obtained by the general public is that the aims and objectives of the survey must be simple, easy to understand and the data to be collected should be unequivocal. Any analysis of the data must bear this in mind.

Citizen based surveys must therefore be based on a set of simple questions with the overriding value being the long-term nature of the data set. The Canberra Ornithologists Group's GBS is a citizen-based survey but taking data from a limited number of observers at a limited number of sites.

One important aspect of the data set from this type of survey is that the introduction of errors can have a much greater influence on the results than from a data set collected by a much larger number of contributors. Further, few atypical sites may well over-contribute and therefore decrease the level of confidence in the data. This Review has spent some time examining data quality.

In addition to providing data that may be of value in improving our understanding of the birds of Canberra, the survey also provides an excellent education tool for COG members. Unfortunately, any analysis must take into account that the two major aims of the survey may not be compatible. Adopting a wide approach to recruiting observers will give an increased possibility of introducing errors associated with misidentification, over and/or underestimating abundance and misinterpreting the way in which the data are to be collected. That is, participants have not been able to satisfy the counting rules.

2) Methodology

2.1) Extent of current rules and definitions

2.1.1) Questionnaire results

The Review Team recognises various problems with the current survey methodology. To obtain baseline information and assess which rules of collection the GBS contributors had difficulty with, a questionnaire was produced and mailed out to all members in the May 2007 Gang-gang. From a possible membership of 360 approximately 65 COG members contribute to the GBS each year. Some of these are long-term contributors whilst others contribute for a short time only. Forty-seven of the questionnaires were returned.

Apart from obtaining background information, the questionnaire was designed to investigate how contributors assess the instructions concerning recording the maximum number of individuals occurring in the area at any one time and assessing how they interpret instruction concerning species flying over and/or using the site, how flock size is estimated, how to count birds seen and/or heard and how to define the shape and size of the site. In addition the questionnaire was designed to assess problems associated with when and when not to record species and to determine which species observers have the greatest difficulty with when assessing abundance. For results of the survey see Appendix.

With over half the respondents having contributed to the GBS for more than 10 years it is possible that the questionnaire results do not represent a cross-section of the contributors. Those who have contributed for a short time may possibly feel that they were not in a position to comment. Most (87%) had contributed during the past 5 years.

Whilst 53% had contributed for more than 10 years, 51% had contributed from their present site for less than 6 years indicating the mobility of the respondents. This may help to explain why there are so few sites that have been surveyed for a long time.

The rule indicating that 'there is no altitude limit and no requirements for the birds to be 'using' the area' appears to be adhered to with a positive response of 83% and 98% respectively. This confirms the change in emphasis of the GBS since July 1993 as noted previously.

There are obvious problems associated with defining when a person is recording for the week, see question 6. Although the instruction states that if away for a week then cross out the week number and do not enter any species, respondents did not interpret the rules consistently with 45% correctly filling in species if they have been on site for only one day but the remainder do not whilst 51% indicate that they have not been recording yet have been on site for one day.

If for any one week you are away for 6 days do you cross out the week number at the head of the column?	If for any one week you are away for 6 days do you record the species seen on the one day at home?	
	Yes	No
Yes	2	21
No	19	2

The cross-tabulation of question 6 (see above) indicates that that the majority of those who don't mark the week off, still report if they spend one day at home, whilst those who do mark the week off don't report any observations from the one day they are at home.

The evenness of this result suggests that there has not been sufficient guidance offered to observers regarding the application of this rule. Many observers are in the workforce and thus likely to concentrate their observation, particularly in winter, in the weekend. It seems that reporting the birds seen when only home for one day is not going to introduce great bias to the results.

As indicated elsewhere in this review, the Review Team has adopted a rule that if only one or two species are seen it is assumed that the observer did not spend time observing birds and the data was revised to indicate that the site was not active during that week. The question is not whether an observer was on site recording for a week but whether, despite the number of days observing, enough effort has been given to provide a true account of the activity on the site for the week. We examined this question in detail and for a person recorded as being away for the week the number of species either was similar to the weeks on either side or the number recorded was much less, usually less than four. Therefore, for the weeks in which an observer was recorded as being away yet three or more species were recorded the observer has been deemed to be on site and the appropriate entry in the database changed.

Instead of determining whether an observer has been on site or not for the week **we recommend that the wording be changed to "Include sightings made if you are present at the site for at least one day during the week and spend time observing birds".**

Replies to question 8 indicate that there is some confusion in interpreting the counting rule that "the number of individual species to be recorded must be the maximum occurring in the area at any one time".

Taking the Australian King-Parrot as an example, of particular concern is that 49% of observers would record two individuals when the male and female were observed days apart. It **is recommended that the rule about counting the maximum during one such observation time should be reinforced to observers whenever possible.**

Respondents listed 37 species where they had trouble estimating species abundance. Not surprisingly, the most common difficulties were with the small, active, flocking species.

The Review Team discussed the problems of collecting meaningful data for migrating honeyeaters and other migratory species (woodswallows etc.) with the GBS counting rules. All indications are that for those species other counting rules may be required in order to get more meaningful data. At the 30 year completion point and after further analysis this issue will have to be re-visited. For the time being, the Review Team therefore **recommends that the counting rules should remain the same for all species and if information is required on the movement patterns and abundance, for example, of honeyeaters passing through Canberra this information can only be obtained with a different survey protocol.**

Problems

Breeding codes

There was discussion by the Review Team on the merits of some of the breeding codes, how to define them and how to add nest contents to the abundance record for weekly totals. It would appear that there is a need to define the term 'breeding event'. In the view of the Review Team this term should not infer any degree of breeding success yet something needs to be produced whether it be eggs or young. Some terms tell us about the occurrence of a breeding event (NE, NY) at a site, some tell us about an intention to breed (IH, CO and displays) whilst others tell us about a possible breeding event (DY CF) at a site. The term NB is difficult to define because of the possibility of the building of nests for roosting.

The question is 'are we interested in the intention to breed and the possible breeding'? However, because it is hard enough to get any information on breeding we should retain as much information as possible. There is every possibility that DY indicates that the birds bred somewhere in the local area so the information should not be disregarded. In addition there is value in collecting data that provides an indication that the species has an intention to breed in the area. After all, it is most unlikely that a species will exhibit such behaviour in an area that it never breeds in.

At the same time it would be beneficial to record whether the observer felt that the DY arose from the site rather than from outside it. **We therefore suggest that in any new instruction set DY (dependent young) indicates an observation of a dependent young on site produced off site but a DF (dependent fledgling) indicates a dependent young produced on site** and so GPS coordinates can be assigned to it.

The question of a definition of DY has, not surprisingly, led to some confusion. For the purposes of the birds likely to be recorded within the COG area of interest the Review Team recommends the following definition:

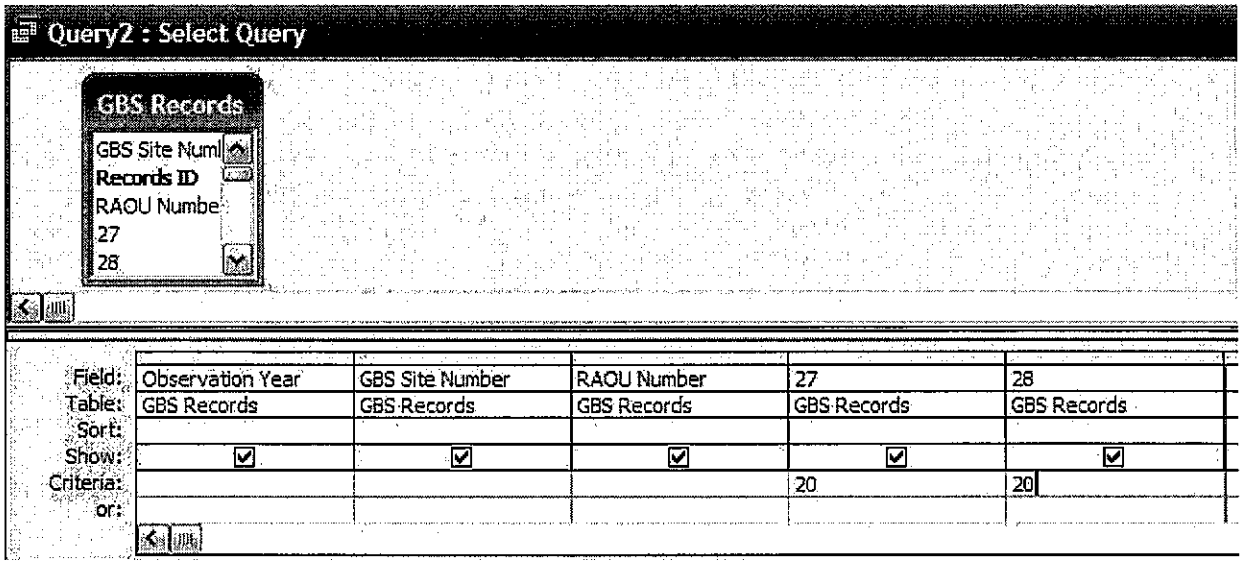
Altricial species, that is those that are born with eyes closed, little or no down, unable to leave the nest and fed by parents, should be regarded as DY (or DF) from

the time they leave the nest until they are no longer being fed by the parent or surrogate parent. Precocial species, that is those that are born with eyes open, covered with down and leave the nest after the first day or two, should be regarded as DY (or DF) until they can fly. The exception being the Emu in which case young are DY if accompanied by an adult

2.1.2) Database design and analysis of errors

Examining the occurrence of errors has been difficult and time consuming due to a fundamental design approach in the database and has involved checking data against the original charts. A record (defined as a single unit of data) consists of the annual observations from a particular site of a particular species. It is therefore not a simple process to obtain the abundance of a particular species in a particular year for a particular week. To obtain these data requires a specific query that needs to be altered for each week. This problem has arisen because 'week' is not a specific identifiable variable. Although it is possible to export queried data from ACCESS into EXCEL for subsequent manipulation, in many cases this would be unnecessary in an appropriately designed database.

The key practical issue is that fields within a table (or select query) are treated as independent for the purposes of querying the data within that field or table. As an example in validating the database the Coordinator wished to check 'large' values (say, flocks of >19 birds). If this was done simply by designing a query along the following lines

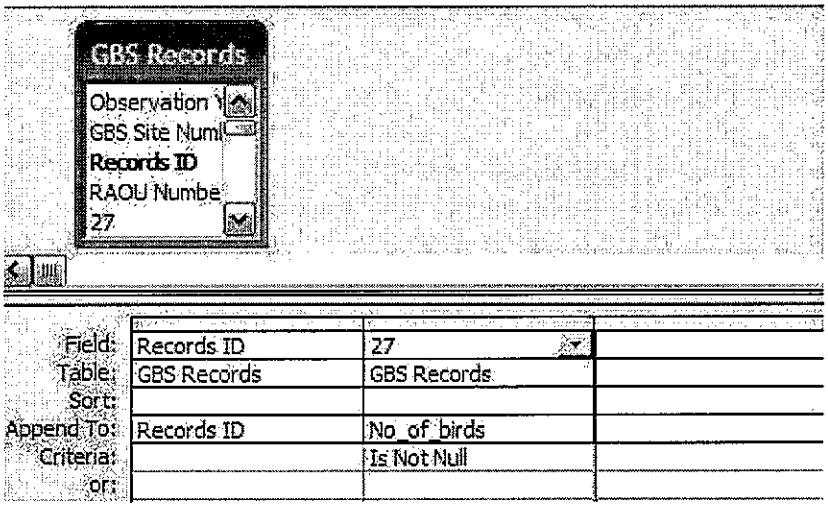


The result would be to identify those cases in which 20 birds were seen in <week 27 **and** week 28> rather than <week 27 **or** week 28>. Obviously extending the example to all 52 weeks merely makes the problem more acute since using the query facilities of ACCESS

requires 52 separate queries to be required. (While it might be possible to write SQL queries to cover all fields from one statement the current Coordinator is not familiar with SQL.) The approach adopted was to create, a new table containing:

- the record number from the table GBS Records (to act as a relational link to the rest of the system);
- week number (generated by a specific query of the table GBS Records); and
- the number of birds reported in that week for that record.

When the skeleton table was designed the default value for "week number" was set to 27. An Append Query was then run selecting value for the field '27' in the table GBS Records. To avoid the problems of ACCESS not summing fields containing null values any empty cells were excluded from the Append Query.



This process was iterated 52 times, updating the default value in the new table by 1 and selecting the next "week number" field. The final restructured table was large, comprising 1.04 million records.

While it is somewhat time consuming to create (and great care must be taken not to duplicate or omit weeks in the iterations) the saving in time, and reduction in potential for errors in addressing questions such as "which records contain more than 19 birds?"; "how many Crimson Rosellas were seen in Year 23?"; In which week were Pallid Cuckoos first seen in each year?"; more than compensate for this once-off effort. Ideally an SQL module should be written to perform the iteration but to date no-one with the required skills has been identified.

Despite using the relational database ACCESS, the design of the GBS database is not relational. Because of this there is much duplication of data but more importantly analysis is conducted from various tables derived from queries that have no relationship to each other. The consequence of this is that all tables produced from the original queries must be updated if an error is detected. This is a recipe for errors and is due to

the poor original design of the database. **The views of the Review Team are that this can only be corrected by a redesign of the GBS database.**

Errors in the GBS database may arise in a variety of ways. They can be due to an input error either because of a typing error or because the data has been read incorrectly from the chart. These are referred to as input errors and are improbability errors. A second source of error can be due to the entry of data that has been incorrectly entered on to the GBS chart. These are referred to as observer errors and can be either improbability or impossibility errors. Without doubt there are further observer errors that remain undetected in the database

Data are read from a chart and entered into one of four ACCESS database tables. Abundance data are entered to tbl_GBS Records (3,368,820 cells), breeding data to tbl_GBS Breeding Records (306,000 cells) and data indicating whether the observer is at the site or not for the week is entered to tbl_GBS Sheet (81,276 cells). A fourth tbl_GBS Breeding Records Sheet (1094 cells) also exists. Errors may occur in any of these tables. Veerman (pers. comm.) indicated that Figure 21 and 22 and Appendix 2 of his publications shows the specific use of this Table.

There may be interactions of errors between the three tables. For instance an incorrect entry in tbl_GBS Sheet may affect errors in both tbl_GBS Records and tbl_GBS Breeding Records.

There may be two sources of error in the abundance data, where either tbl_GBS Records or tbl_GBS Sheet has a data input error or an observer error. There may also be two sources of error on the breeding data, where again either tbl-GBS Breeding Records or tbl_GBS Sheet has an input or an observer error.

An error is counted when a mistake is found concerning a particular site on a particular year and in a particular week. The error will then affect one or usually more species depending on the number of observations recorded for the particular week. Therefore, the number of errors in any analysis will far exceed the number of errors in the database.

We examined the prevalence of impossibility errors in two ways. First we examined the number of instances where for any one week in any one year an observer recorded the abundance of a species yet the data in Table_GBS Sheet indicated that the observer was away from the site for the week. Because the abundance value 'A' is determined by the total abundance of a species divided by the total number of weeks that observers are at home any mis-match will result in an incorrect value for 'A' for all species that were observed during the particular week when the observer was away. Second we examined the prevalence of incidence where there was a breeding observation during a week when no abundance was recorded for that particular species. This typically happened when observers 'filled in' breeding observations that they did not observe, e.g. week 27- dy, week 28- away, week 29- dy, yet week 28 is given a breeding observation of dy.

The numbers of errors found in the database to date were as follows:

Abundance data		
Tbl_GBS Records:	Input errors	14
	Observer errors	26
Tbl_GBS Sheet:	Input errors	135
	Observer errors	0
Breeding data		
Tbl_GBS Breeding Records:		
	Input errors	12
	Observer errors	39
Tbl_GBS Sheet:	Input errors	44
	Observer errors	0

On average there are only 1.16 errors for every 10 charts, of which 0.4 consist of observation errors and 0.76 occur as input errors.
The most common input error arises from an incorrect entry into tbl_GBS Sheet (e.g. indicating a person is at home when they are away) or typing abundance data into the wrong week. The most common observer error is reporting the abundance of species when the observer has indicated that he/she is away.

One of the outstanding problems is that there has been some inconsistency concerning the entry of abundance data when the observer is away or when the observer is at home but a comparison of data collected on the preceding and following weeks indicates that there has been an incomplete survey for that week. **For the first case we recommend that the observer be regarded as at home if there are three or more species recorded. For the second case we recommend that there be no change to the original data and just accept that the week was poorly surveyed.**

We note a common problem when despite the observer being away, there is a breeding entry that has been derived from the breeding data collected during the preceding and following week. We recommend that this observation be omitted.

A comparison of the abundance index 'A' over 25 years before and after the detected errors were corrected was made for four species. The Crested Pigeon represented a species that is rapidly changing in abundance, Rufous Whistler represented a migrant that does not flock, Yellow-faced Honeyeater represented flocking migrants and Australian Magpie represented common resident birds. Overall the values of A differed by less than 1% for each species. A relatively small number of values for species/weeks change by >1% (but none by >2%). The only negative change (for these species) was for the Australian Magpie - which probably reflects the correction of a typographical error from 115 to 15.

Up to and including Year 25 all analyses have included these errors. The errors, though small, will be reflected in the values published in the Annual Bird Reports, in Canberra Ornithologist Group, 2000 and in Veerman, 2002, 2003, and 2006. To preserve this data set we have therefore made no corrections to the original database (Version 1) and have made a copy of the original to which we have made corrections. Version 2 contains corrected data for the first 25 years of the survey whilst Version 3 contains corrected data for all years and up to the present. Copies of all are being retained by the Database Manager and the GBS Coordinator.

By definition, improbability errors are difficult to detect and can only be detected by a *post hoc* comparison of data in the database and on original charts. Because impossibility errors are caused by a breach of rules they can be picked up by an appropriately defined data input program. Similarly it should be possible to reduce many of the improbability errors by providing stopping rules to the data input program. Through the use of an appropriately designed 'look-up' table and defining for each species an abundance figure that is unlikely to occur, any value input at or above the value will be queried by the system. As an example, a value of more than or equal to three Pallid Cuckoo for any one week should be queried.

In conclusion, we have found that the small number of errors discovered in the database up to now will have little impact on the published observations. For most years there have been somewhat more than 2,400 weeks of activity. It would be surprising if the changes made are more than 0.8% of the total. So for a very common species with an A value of (say) 5 and a change to total weeks of -20 weeks the A value would drop by 0.04. Only those species with an A value >1.25 would show a change in the second decimal place of A. .For a bird seen once in a year with 2400 observations reduced by 20 to 2380 the value of A would increase from .0004167 to .0004202 which would not be shown in the ABR as a difference.

2.1.3) Adequacy of data

2.1.3.1) Garden Bird Survey coverage

The GBS has, on average, been conducted over 65 sites ranging from 43 to 90 sites. The question arises whether a citizen-based project is able to produce meaningful results from such a small sample base. Over the past 25 years all sites have been analysed with no allowance being made for site location. We note that in the analysis by Veerman (2002, 2003, 2006) he has been aware of this problem and has not included those sites that he regards as rural rather than urban sites.

An analysis on the abundance value 'A' was conducted to examine the difference by including or excluding rural sites. Four wide ranging species typical of both urban and rural sites were chosen and the impact of removing the rural sites for year 24, 25 and 26 was examined.

Impact of Rural sites

Species used: 273 Galah, 322 Laughing Kookaburra, 401 Rufous Whistler, 638 Red Wattlebird

Rural Sites: *246 Wamboin (Year 26); 309 Gundaroo (Years 24, 25, 26); 312 Hoskintown (Years 24, 25, 26); 313 Crace (Years 24, 25); 314 Fyshwick (year 24), 347 Carwoola (pt year 26) 349 Murrumbateman (pt year 26)*

Ainslie Hackett sites: *15, 239, 278, 281, 295.*

Rural A values

Year	273	322	401	638
24	12.32	0.73	0.33	2.23
25	10.95	1.52	0.38	3.76
26	10.19	0.69	0.31	1.26

Total A values

Year	273	322	401	638
24	8.66	0.27	0.04	3.25
25	7.41	0.29	0.03	2.99
26	8.58	0.32	0.04	2.67

Non rural A values

Year	273	322	401	638
24	8.40	0.24	0.02	3.32
25	7.27	0.24	0.02	2.96
26	8.49	0.30	0.03	2.75

% Difference in A values (R -NR)

Year	273	322	401	638
24	31.80	67.53	93.89	-48.66
25	33.61	84.16	95.91	21.28
26	16.66	56.39	91.79	-118.17

% Difference in A values (T -NR)

Year	273	322	401	638
24	2.97	12.03	50.26	-2.20
25	1.93	17.11	47.67	1.04
26	1.00	6.14	36.13	-2.82

There was a surprising difference, see above, but the level of difference attributable to rural versus other regions could just be a factor of 5 - 6 sites versus 65 - 70 sites. A separate analysis was therefore conducted where the same four species were examined for years 24, 25 and 26 with five suburban sites excluded. The sites chosen were from the suburbs of Hackett and Ainslie.

Comparison of Ainslie Hackett with other sites

Ainslie Hackett A values

Year	273	322	401	638
24	10.61	0.92	0.01	2.87
25	8.26	0.91	0.03	2.65
26	10.46	0.88	0.01	2.40

Total A values

Year	273	322	401	638
24	8.66	0.27	0.04	3.25
25	7.41	0.29	0.03	2.99
26	8.58	0.32	0.04	2.67

Non Ainslie Hackett A values

Year	273	322	401	638
24	8.51	0.22	0.04	3.28
25	7.36	0.25	0.03	3.01
26	8.45	0.28	0.04	2.69

% Difference in A values (AH -NAH)

Year	273	322	401	638
24	19.74	75.94	-292.43	-14.42
25	10.91	72.53	-5.96	-13.51
26	19.27	68.00	-732.07	-11.89

% Difference in A values (T -NAH)

Year	273	322	401	638
24	1.69	18.07	-5.49	-0.89
25	0.74	13.87	-0.34	-0.73
26	1.54	12.24	-6.13	-0.70

Comparison of comparisons

% Difference in A values (T -NR)

Year	273	322	401	638
24	2.97	12.03	50.26	-2.20
25	1.93	17.11	47.67	1.04
26	1.00	6.14	36.13	-2.82

% Difference in A values (T -NAH)

Year	273	322	401	638
24	1.69	18.07	-5.49	-0.89
25	0.74	13.87	-0.34	-0.73
26	1.54	12.24	-6.13	-0.70

Again, there was a surprising difference, see above. Without having a rigorous test available it appears that the effect of including/excluding the sites in Ainslie and Hackett is, at least for the 4 species chosen, similar to the effect for including/excluding the rural sites. The apparent differences are therefore due to the differences in the number of sites rather than to any difference in including or excluding the rural sites. We conclude that the inclusion of rural sites has little effect on the value of 'A', at least for the four species examined and **recommend that the practice of including all sites, irrespective of location when defining the level of abundance (A) should continue for the Annual Bird Report.** All sites are coded according to their location and so rural sites can easily be removed from any more detailed analysis

2.1.3.2) The Review Team did not have the expertise to examine the problems associated with the 'sparseness' of the data set due to the high number of zero and null cells, see p 12 for further comments.

2.1.3.3) The number of years with active sites for each of nine regions is presented below. Six regions had active sites for all 25 years, Gungahlin and Queanbeyan had active sites for only 13 years.

Nine regions have been defined; see below listed in the key to the figure below. Most of the regions are self explanatory, but some have outliers. Gungahlin includes Hall and Crace but not Mitchell. North Canberra includes Pialligo and Mitchell. South Canberra includes Fyshwick.

Generally the region with the largest number of sites is Belconnen. Regions Queanbeyan, NSW Rural and Gungahlin all have too few sites to draw inferences about changes over time but the other six regions probably have sufficient sites to test if the pattern of increases and decreases are similar. It should be noted that the sites are not evenly distributed within the broad regions; the more southerly and eastern parts of

Tuggeranong are poorly represented as are the densely settled suburbs of Phillip and Kingston.

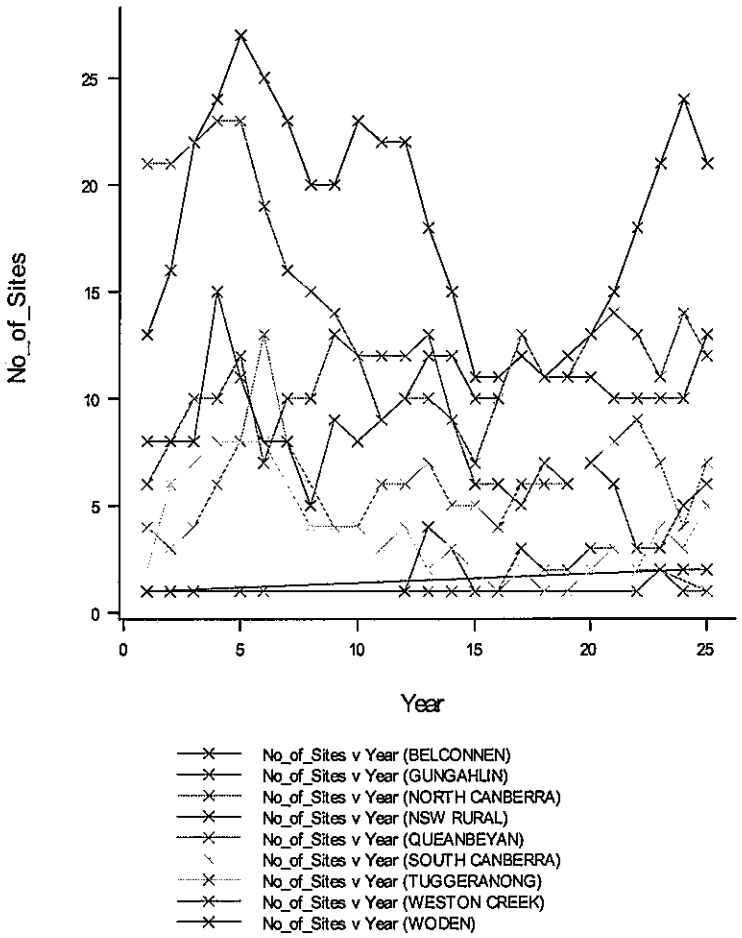
A log-linear model was fitted to the number of sites recorded in each year for the six regions and the interaction between the classifying variables, Region and Year, tested. The probability of getting the observed spread of number of sites across years within regions was p=0.10 (χ^2 = 100.8 with 120 degrees of freedom). This suggests that the gain or loss of sites over time is proportionally similar across the six regions.

Recommendations:

- **To increase the number of survey sites COG members should be encouraged to continue to be, or to commence being, active in the GBS.**
- **Particular attention should be paid to recruiting those members of COG in Gungahlin and Tuggeranong to the Survey as these urban areas are under represented in the GBS.**

Number of years with sites recorded for each of the nine regions participating in the GBS.

Region	No. of Years
Belconnen	25
Gungahlin	13
North Canberra	25
NSW Rural	3
Queanbeyan	13
South Canberra	25
Tuggeranong	25
Weston Creek	25
Woden	25



Number of active sites recorded in each Region for the period 1981/82 to 2005/06

2.2) Bringing the GBS and other data collections together.

Improving the processing system

The Garden Bird Survey cannot be undertaken in anything like its current form without a computer processing system. The current system was designed in 1995 and has been incrementally updated since then (largely an event driven process to meet the needs and desires of the Coordinators and their stakeholders).

The basis of the system is Microsoft ACCESS. It currently contains four databases:

GBSData: As its name suggests, this contains the data about birds and sites. It is the heart of the system. It also includes the ephemeral queries used by the current Coordinator to investigate topics of current interest.

GBS: In essence this contains the data entry system, and includes a range of linked tables and some tables required for the data entry system to function. It

also includes a very large number of queries developed by a previous Coordinator for his publications (Veerman 2002, 2003, 2006).

GBSStats: This is the home of the basic aggregate data used to compile the Annual Bird Report.

GBSAnnRep: This is primarily concerned with compiling and holding the reports used in each years Annual Bird Report.

A key point to note is that the data entry system is currently two programs. The version used by the Coordinator and one of the data entry workers uses a full version of ACCESS. However the other data entry worker does not have a full version of ACCESS on her computer and thus uses a 'run-time version" of the software. Unfortunately that version no longer appears portable to more recent versions of Windows. The COG database manager has explained that the run time version of the data entry system utilises an out of date version of ACCESS and would need to be completely rewritten under a more up to date version of that program. Thus if any further data entry volunteers come on board they will have to have a full version of ACCESS.

It is important to note that the data is transferred from the data entry workers to the Coordinator as EXCEL spreadsheets since it is difficult simply to export an ACCESS table without creating it as a separate database. However this process does not cause any problems since the files concerned exactly match the format of the 'official' database.

In addition to this system some observers have at various times provided material as EXCEL spreadsheets, which are imported to the system. Some of these are satisfactory since they are designed to replicate the database tables. In other cases observers have:

- employed special fields;
- added comments fields;
- recorded more than one value in a field;
- included alpha characters in numeric fields; and even
- renumbered the fields so that GBS week 27 was recorded as field 1.

All of these added significantly to the difficulty of loading the data to the database.

A major problem has been that at least four Coordinators have been responsible for the system and each has added their own items to it. A particular difficulty is where the objects (tables; queries; reports etc have not been designated with an indicator of the nature (e.g. T or TBL for Table, Q for Query, R for Report etc). Not only is this in defiance of established good practice but it makes problem solving very difficult: if an expression fails, the Coordinator has to spend time trying to establish whether the data assessed by the query is in a table or a query before addressing the cause of the problem.

There is an amount of redundant data in the system, largely due to:

- successive Coordinators finding 'better' ways of doing things (or adding additional data fields; in conjunction with
- limited documentation (and the issue of being uncertain what comes from where) making deletion of existing material a high risk option.

By way of example, an historic table, Suburbs and Regions, coded suburbs to Zones and Regions since that was used as a presentational array for some early Reports. However the current Coordinator has not had need to use this table but did need one, TBL Suburb characteristics, which indicated the time since development started in the suburb and its population density. Both tables are in the database whereas in a 'better' system suburb would appear once with all required attributes.

A very significant flaw of the system as it exists is that the element "week" is not defined as a variable of any core table. This appears to have been done as a matter of getting the data entry screen to "look like" the chart (and possibly programming efficiency). However it adds very significantly to the complexity of writing efficient queries and is a major blockage to efficient use of the data. The current Coordinator has developed a work-around but it is no more than that.

The current Coordinator is also concerned that the current practise of using a system of weeks in which the first week of an observation year (starting on July 3) is coded 27 is intrinsically illogical. The use of a year ended 30 June suits analysis of the behaviour of most species of birds quite well but having calendar year 2005 as weeks 1 – 26 of year 24 and weeks 27 – 52 of year 25 is far less intuitive and thus prone to error than having it represented by year 24, weeks 27 – 52 and year 25, 1:26. This complexity is most significant when the data are tabulated having been sorted by year and week or the data plotted.

Due note should also be taken of the very significant workload that will be required if the anticipated many changes to the taxonomy occur. If, as expected, this leads to a number of lumps and splits there will need to be considerable work to recalculate the historic data - if in fact that can be done – as well as recoding the BIRDLIST table.

With regard to the data entry, the current system has some good features. However, the fact that only 12 weeks of data appear on the screen at once and the marginal entries disappear as one scrolls across the screen causes a very significant amount of risk to arise. Also the system does not undertake a range of checks that would be helpful (for example, seeking confirmation of large numbers of birds or unusual species.

The size of the main table GBS Records, which holds the abundance data is now beyond the number of rows which can be handled by Microsoft EXCEL. The benefits of using a relational database to **manage and manipulate data** also argue against using a spreadsheet system as the core of the GBS data holding system. From the view of the current Coordinator continuing to base the system around ACCESS is quite satisfactory. This does not necessarily mean that the data entry system also has to continue to be based around ACCESS.

Possible solutions for a data entry system

There are two possible methods of recording bird observations, either on a paper chart (present method) or electronically. For electronic entry there needs to be a program available to each observer, either as a .exe file distributed to every user or as online data

entry via the internet. There is generally a compelling logic for an online data entry system as an alternative to distributed .exe entry programs. In both cases each observer does his or her own data entry work but, the problem of upgrading and distributing .exe files to all users should not be under-estimated.

However, there are several factors of the GBS that make online data entry the least preferred option. The data set is complex, the survey period is a week long, it is difficult to get a very easy to use, robust online entry system on a limited budget, and even computer-literate people would still find the chart easier to use and with less work than the online system.

At present the data entry system is based around observers maintaining a hard copy Chart which is provided to the Coordinator at the end of each year and the data entered by a small pool of data entry staff. The strengths of this system are:

- Many observers get a positive feel from having the Chart available to look at;
- The data can be entered instantly (or at least as soon as a pen can be found);
- The Charts form a very safe back-up to the electronic system;
- There is no requirement for observers to have a computer (or any specific software on a computer);
- The small number of observers provides great assurance that consistent practices are followed.

The problems with the system are:

- It is proving difficult to print the Charts at the required size, on good quality paper for a reasonable price;
- Some observers would like to have a digital copy of their data; and
- The approach relies heavily upon the continuing ability of a small number of workers to undertake the data entry.

However, the present entry system now used must be improved, and distributed more widely to willing volunteers and to those observers who would prefer to use an entry system rather then the Chart. There is therefore a case to be made for developing an .exe data entry program for distribution

For the reasons above we do not recommend an online data entry system, given that only a small number of users would use it. Overall we support a simple .exe input system that could be distributed to those who wish to use it, provided the number is not more than 10 users. The key issue is whether input would occur on a weekly or an annual basis. The current input programs are meant for input of an entire chart at a time at the end of the year. This should be the intention of the future input program if it is distributed to individuals. The program could also be used on a weekly basis, but not necessarily designed for that purpose.

The pros and cons of different approaches are briefly outlined below:

1	Runtime Access System	This is possible, but someone would have to design it using up to date Access	Compatible with Access so data transfer easy PF could do it, but needs new software	System may become incompatible and require updating every few years (as is case now)
2	Visual Basic enabled Runtime Excel System	This is possible, but someone (a programmer) would have to design it	Will produce a transferable data set familiar to users with experience in Excel	Some minor problems with transferring data to Access, but solvable. Will need a programmer to troubleshoot and update in future.
3	Visual Basic entry module distributed to users .exe file	This would collect the data through forms similar to online systems of Access/Excel front-end programs	Can be made fairly robust and produce data in csv files for ready assimilation by Access (similar to Tony Harding's programs)	Needs a programmer to make, troubleshoot and update Problems with version security and updating a number of programs Data could come in in old formats and cause confusion May be difficult for some users to come to grips with.
4	Online system	Similar to 3, but online.	Single authoritative data source Readily available Update on one site	As for 3. Slightly more expensive to maintain website

It is considered that on balance, and in the absence of any compelling reason to the contrary, the current practice of completing hard copy charts should continue. This implies that there remains a need for a central data entry system.

This could possibly be done effectively using EXCEL as the basis of the system, supported by Visual Basic code to build in functionality to provide quality assurance. By way of example, if the observer indicates that the site is inactive in a week that column of the spreadsheet could be locked (preventing any data entry) as opposed to the warning message currently generated. Warning messages could be generated if a large number of birds was entered (say >50 birds) and a warning to submit a rare bird report if an unusual species was entered. By use of look-up lists it would be possible to ensure that data was reported under standard names rather than the "folk idioms" employed for some species. Of course it should be noted that a careful balance is needed between building in checks and keeping the system working at a suitable speed.

It is recommended that effort be put into developing a suitable (EXCEL/Visual Basic or similar) data entry system suitable for both central and distributed data entry. Obviously the views of the current data entry workers and the observers must be sought at the specification stage and as comment on work-in-progress.

A radical alternative that could be investigated would be the use of a commercial data entry facility. The problems of accessibility to current software are then unlikely to be an

issue. However, it is likely to add significantly to the cost of the exercise, particularly since the basic data entry document (the Chart) is a non-standard size and thus unlikely to be accommodated within the production-line systems of a commercial facility.

A major obstacle to the use of spreadsheets for data entry is that the original data are not available as a check against the database. There is no policy of retaining the spreadsheets that were submitted. **We therefore recommend that any data submitted in an electronic format is backed up and retained by the COG database manager, using standard files to identify the GBS Table, collection year and site to which they refer.**

Possible solutions for a data holding/management system

When maintaining the ACCESS system there is a clear need to adopt a logical analysis of the various objects in the system, and their inter-relationships. Without attempting to formally define the objects, or their attributes, the following offers a starting point. Underlining indicates an attribute which can be used to form a relationship between tables

1. Birds (species code, species name(s)[2], scientific names, systematic order and ABR author code)
2. Site (site reference number, observer code, street address, latitude, longitude, [map grid reference[3]], suburb/locality name, type of site, habitat notes)
3. Suburb (suburb/locality name, population size; density, date of initial development)
4. Observer (observer code, name, telephone number, email address)
5. Observation data (site reference number, species code, year number, week number, number of birds, breeding event code).

Having undertaken this analysis and identified the desirable set of objects and attributes a set of formal naming conventions should be implemented[4] so that the objects in the database can be referred to in a consistent manner.
The final steps in the formal redevelopment would be:

- rewriting the current set of (approximately) 15 queries required to calculate the summary table for the Annual Bird Report using the conventional names; and
- writing a user's manual so that all people involved in the operation of the system, now and in the future, can readily see what needs to be done.
- At the start of each GBS year a Register be maintained by the GBS Coordinator of all recipients of a GBS chart. The Register will include postal address and, if different, address of site. This will allow the Coordinator to follow up on any charts not produced at the end of the year.

---

[2] Some species currently have several names recorded under the official names allocated at various times. Species name as current 'Birds Australia' nomenclature.

[3] It is possible to automate the calculation of map grid from latitude and longitude)

[4] It is highly likely that there are professional standards for this.

The Review Team recommends that:

- **A Consultant be contracted to examine the problems and costs of redesigning the GBS database**
- **The Consultant also examines the feasibility and cost of redesigning the data input system**
- **An updated data entry program should be freely available, but an on-line data entry system is not recommended**
- **A Register of issued charts be maintained by the GBS Coordinator**

Although it is not envisaged that the data input system be based on-line, it is envisaged that those contributors who wish to input their own observations have assess to the program on their own computer. Even so, we consider that all original observations must be entered onto the GBS Chart.

At present the COG database consists of separate databases. One contains the information received from the GBS with input via a data input program. The second database contains information received from general observations submitted through the COG data sheet or through the on-line data entry program. In addition, data are received from Birds Australia, Greening Australia and from various specific surveys such as the COG woodland survey. All require various forms of data manipulation and some have their own specific data input programs and all with varying degrees of error checking. Though the basic structure of the General Observations database is simple (Paul Fennell *pers. comm.*) there is a need to upgrade and refine the data input procedures. This also applies to the GBS database. **We therefore recommend that if consideration is given to improving the GBS data input system then this is done in conjunction with improving and updating the various input programs for the General Observations database.**

2.3) Changes to survey protocol

Apart from the inadvertent changes in the aims of the GBS brought about by changes to the instruction set in 1993, the basic methodology has remained the same throughout the past 25 years. **We recommend that no changes be made to the survey protocol until a thorough analysis of the data set is completed,** see Section 4.

3) Impact of collection methodology and operational constraints.

3.1) Known usage of the GBS data

The GBS results are summarised each year in the Annual Bird Report (ABR) published by COG and it is not possible to indicate in detail what use is made of that data. However, examples are identified from time to time including use of the ABR data on some migrant species by Beaumont (2006) in research into arrival dates and subsequently quoted in Wingspan (Chambers, 2007) in an article on global warming. In addition, the results of the GBS are used from time to time in the preparation of articles published in the COG journal "Canberra Bird Notes (see for example Butterfield, 2004).

In terms of the volume of data utilised and the profile of the products that resulted the two largest users of the Garden Bird Survey have been Veerman's reports on the first years of the Garden Bird Survey (Veerman, 2002, 2003, 2006) and the COG publication "Birds of Canberra Gardens" (COG, 2000). As "Birds of Canberra Gardens" is also included on the COG website it will be assisting birders worldwide to understand the birds of Canberra's gardens.

Information about the numbers of Common Mynas was provided to Dr Chris Tidemann of the Australian National University to assist in development of his research into this species. He indicated that the data was very helpful in providing a baseline for comparison with his own more detailed studies. Dr Tidemann is also intending to use the data on Common Myna, Crested Pigeon and Spotted Turtle-Dove to research the spread and growth of these invading species.

Information about Australian Magpie was provided to Heather Gibbs of Deakin University for a research project into the relationship between climate change and breeding of Australian Magpie, (Gibbs, 2007).). Documentation of the spread of Yellow-tailed Black Cockatoos after the 2003 bushfires is mention in Life in the Suburbs: Urban Habitat Guidelines for the ACT, (Mitchell, 2006).

Dobbie and Welch (undated) utilised the data in development of some statistical models of the distribution of three species of birds in Canberra.

In 2007 aggregate information on a number of prey species was provided to Dr J. Olsen of Canberra University to set benchmarks for a study comparing prey of Peregrine Falcons and Australian Hobby from two periods 10 years apart. (Olsen, 2007 *pers comm.*).

Finally, extracts of the data are used each month by the Coordinator to illustrate brief articles written in the COG Newsletter Gang-gang.

The results of a computer generated literature search are shown below:

- Many citations of Veerman's reports
- *Zoological Record* 1978 to 2007 database - the only hit is an unnamed source
- *SORA: Searchable Ornithological Research Archive*- no hits
- Ornithological Worldwide Literature – no hits
- BioOne – no hits

- *Mitigation of the Impacts of Mynas on Biodiversity and Public Amenity*', Dr Christopher R. Tidemann, 2001, http://sres.anu.edu.au/associated/myna/minimise_files/Myna_Mitigation_PhaseI.pdf
- Google Scholar has a dead link to *Mitigation of the impact of mynas on biodiversity and public amenity in the ACT,* Dr Christopher R. Tidemann, 2003
- 'Modelling Correlated Zero-inflated Count Data', Authors: Dobbie, Melissa J.1; Welsh, A.H.1, Source: *Australian & New Zealand Journal of Statistics*, Volume 43, Number 4, December 2001 , pp. 431-444(14), http://www.ingentaconnect.com/content/bpl/anzs/2001/00000043/00000004/art00006
- This is probably Dobbie's thesis covering the same ground: http://www.cmis.csiro.au/melissa.dobbie/PhDthesis2001.pdf
- Predation by House Cats, *Felis catus* (L.), in Canberra, Australia. I. Prey Composition and Preference, *D. G. Barratt, Wildlife Research, 1997, 24, 263–277*
- 'Predation by house cats, *Felis catus* (L.), in Canberra, Australia. II. actors affecting the amount of prey caught and estimates of the impact on wildlife' *D. G. Barratt* Applied Ecology Research Group, University of Canberra, PO Box 1, Belconnen, ACT 2616, Australia. http://www.publish.csiro.au/?act=view_file&file_id=WR97026.pdf
- The influence of remnant bushland on the composition of suburban bird assemblages in Australia H. Parsons, K. French and R. E. Major Austral Ecology 2006 Volume 31, 217-227.
- Temporal variation in bird assemblages: How representative is a one-year snapshot? * MARTINE MARON, ALAN LILL, DAVID M. WATSON AND RALPH MAC NALLY Austral Ecology Volume 30 Issue 4 Page 383-394, June 2005
- The Ecology of the Common Myna in Urban Nature Reserves in the Australian Capital Territory A.S. Pell and C.R. Tidemann EMU Vol. 97, 141-149, 1997
- The impact of two exotic hollow-nesting birds on two native parrots in savannah and woodland in eastern Australia A. S. Pell and C. R. Tidemann *Biological Conservation* Volume 79, Issues 2-3, February-March 1997, Pages 145-153
- LINDA J BEAUMONT, IAN AW McALLAN, LESLEY HUGHES (2006) A matter of timing: changes in the first date of arrival and last date of departure of Australian migratory birds Global Change Biology 12 (7), 1339–1354.
- Climate change and its impact on Aust avifauna, Chambers, Hughes & Weston Emu 2005 1-20
- COG's Birds of Canberra Gardens
- Many species accounts in Handbook of Australian, New Zealand and Antarctic Birds.

3.2) Data analysis

3.2.1) Present summary statistics

To date the data from the GBS has been summarized annually in the form of various statistics which are the mean values of 'A', 'F', 'W', 'R' and 'G' from all sites irrespective of location, see Annual Bird reports for further details. In the case of 'A' the value is obtained by summing the abundance for any one species across all sites and dividing by the number of observer weeks. It is not calculated by determining the individual values of 'A' for each site and then determining the average value. If it were then some measure of variance could be attributed to 'A'.

The values of 'A', 'F', 'W', 'R' and 'G' are published in the COG Annual Bird Report for each 12 month period. These annual values have subsequently been graphed and published in Canberra Ornithologists Group (2000). In addition Veerman (2002, 2003, 2006) has removed the non-urban sites, recalculated the summary statistics, graphed the results and for 'A' has, in many cases, provided a linear trend line on the graphed values.

A comment was received indicating that there could be problems in the calculation of the long-tem average values of the summary statistics as a simple average of the annual values of the statistics. The Review Team agrees that taking an 'average of averages' is an area of risk since it could ignore the impact of greater observer effort on the values of the summary statistics. A redesigned GBS database could include procedures to replace this simple average with a true value of the overall value of the summary statistic in question across all previous years.

However, an investigation by the Review Team compared the simple average values of 'A' for 26 years against the overall value for three species with a range of patterns of 'A'. In each case the difference between the simple value and the overall value was non-significant, (Australian Magpie 2.87 v 2.84, Common Myna 2.95 v 2.81 and Yellow-tailed Black-Cockatoo (0.31 v 0.30). Thus, the Review Team concluded that there was no practical problem, in terms of the current published values, of long term averages being grossly misleading.

Peter Woodall (2002) in his review of Veerman 2002 noted that 'Annual variation is shown by line graphs and linear regression lines are fitted to some of these graphs. In some cases the trends do appear linear but in other cases some other form of analysis (curvilinear, GLM, etc) would probably have been more appropriate. In any case there is no indication of whether the regressions are significant or not'.

Nicholls has previously critiqued Veerman's 2003 publication and commented on the manner of analysis and interpretation of results especially in the failure to take into account the spatial and temporal dependences that exists within the data and warns that any inference drawn by the author needs to be treated with caution, (Nicholls 2003).

In addition, a more sophisticated trend analysis was undertaken on the Golden Whistler, Noisy Friarbird and Australian Magpie by Dobbie and Welsh (undated) in which they analysed the counts to determine whether there had been changes in an index of abundance of birds temporally (within a year as well as over years) and/or spatially (in location) and to quantify and possibly explain any such changes. It was noted that the high percentage of zero counts in the data meant that standard statistical discrete distribution, such as the Poisson and Negative binomial were unsuitable for modeling the counts and instead a conditional model (Welsh *et al.* 1996) consisting of two components was adopted. The first component was based on modeling the probability of recording a species based on the recoded count data to be coded zero for not recorded or one for all positive counts recorded. The second component was a truncated Poisson model and was based on only those observations for which a positive count was recorded and thus is a subset of the previous data. The two models were combined to give estimates of the mean abundance of the species as a function of the predictor variables used in the two models. One of the strengths of this analysis is that the individual data for each week are used rather than values calculated to reflect levels of a single covariate. The model can take into account the fact that there are groups of weekly observations that are from the same site in the one year. This permits the exploration of the level of serial correlation to be undertaken. Where there is significant correlation, as was found by Dobbie and Welsh, the confidence intervals for the model parameters are modified. This changes the inferences about changes over time (between years or within years) or space that might otherwise be drawn. We attempted to replicate their analysis but the computer platform (UNIX) on which this analysis had been performed was unavailable to us and we were therefore unable to replicate or expand on their analysis.

3.2.2) Data comparison- up to 21 years and an additional four years.

Given that Veerman (2003, 2006) has analysed and published the first 21 years of data from the GBS we have examined what difference an additional four years has made to the top 14 species as measured by long term average value of 'A' and value of 'A' in 2004-05. The difference between these two measures is that only the former includes Common Blackbird and only the latter included Crested Pigeon. The analysis plotted trends for each of the chosen species for two time periods and in two ways.

The two time periods are: up to and including year 21 which enables a comparison with Veerman's comments ; and including data up to year 25 to enable an assessment as to whether and how things have changed since Veerman undertook his analysis.

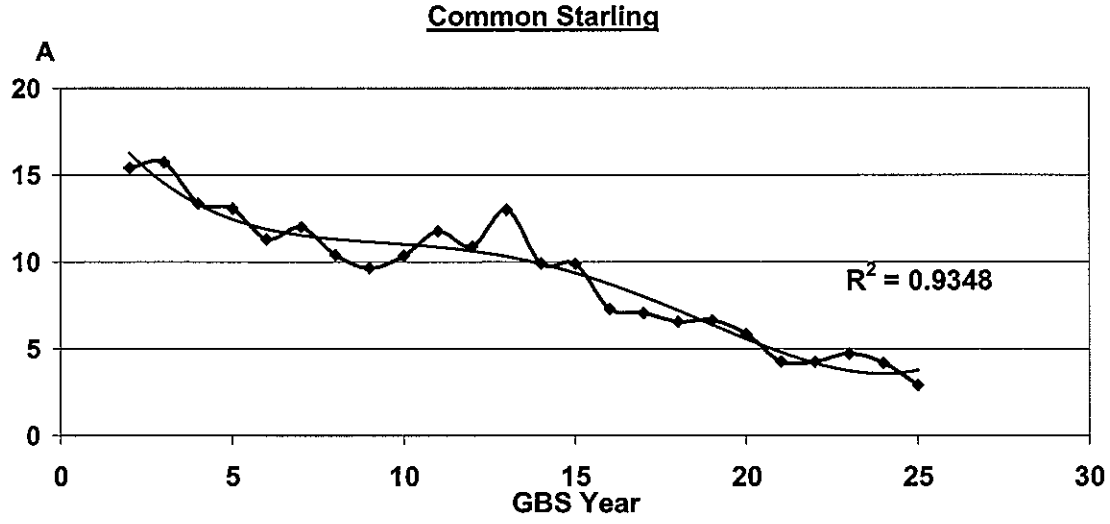
Common Starling, code 999, Long term rank 1

Veerman comment: Clear that the Canberra population has declined significantly since then (i.e. 1990). Numbers have decreased over the first 9 years then a slight resurgence for the next 4 years. It then showed a sharper decline.

Review Team comment:

21 years: Ignoring Year 1 (for Veerman's reasons) decline appears steady except for strange value in year 13. (This is due to very high numbers 300-500 per week in spring reported for site 230 for that, but no other, year. From a query to the observer, this is because the roost was abandoned after that year. It is unclear why these observations were not deleted from analysis as for year 1?).

25 years: Continues in same pattern



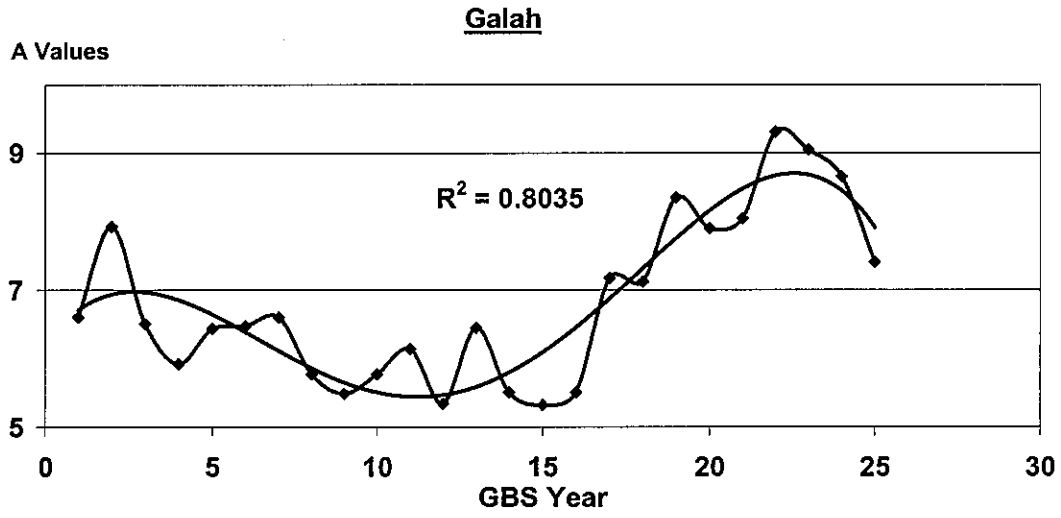
Galah, code 273, Long term rank 2

Veerman comments: Numbers over the years have been mostly stable, with an increase over the last six years,

Review team comments:

21 years: May be better to say an initial decline followed by a strong increase (using 4th power trend). No linear trend should be graphed.

25 years: Adding in last 4 years shows a decline from peak in year 21. It looks to be almost cyclic?



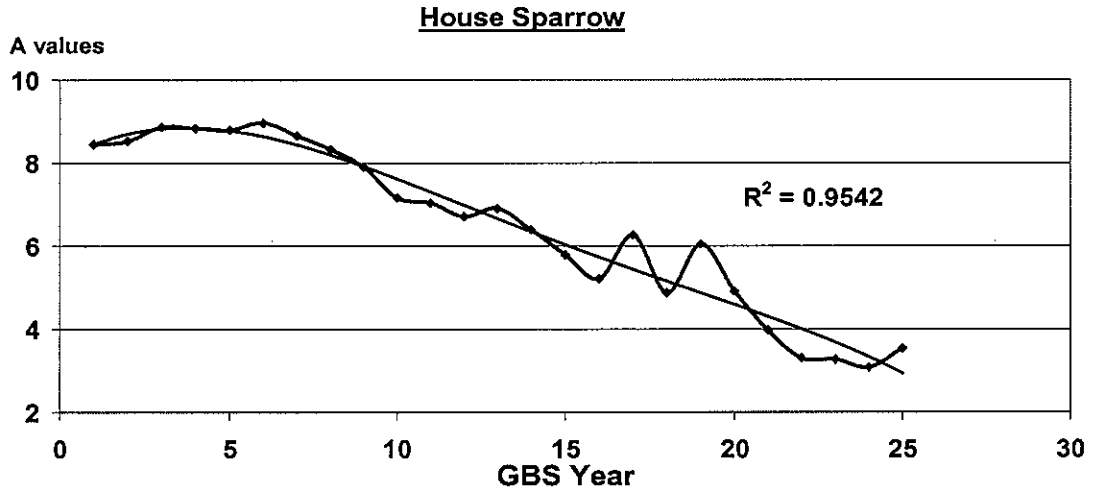
House Sparrow, 995 code, Long term rank 3

Veerman comments: The abundance has declined smoothly and markedly from year 6. There could be a long term undulation as the species was increasing from year 1 to 6.

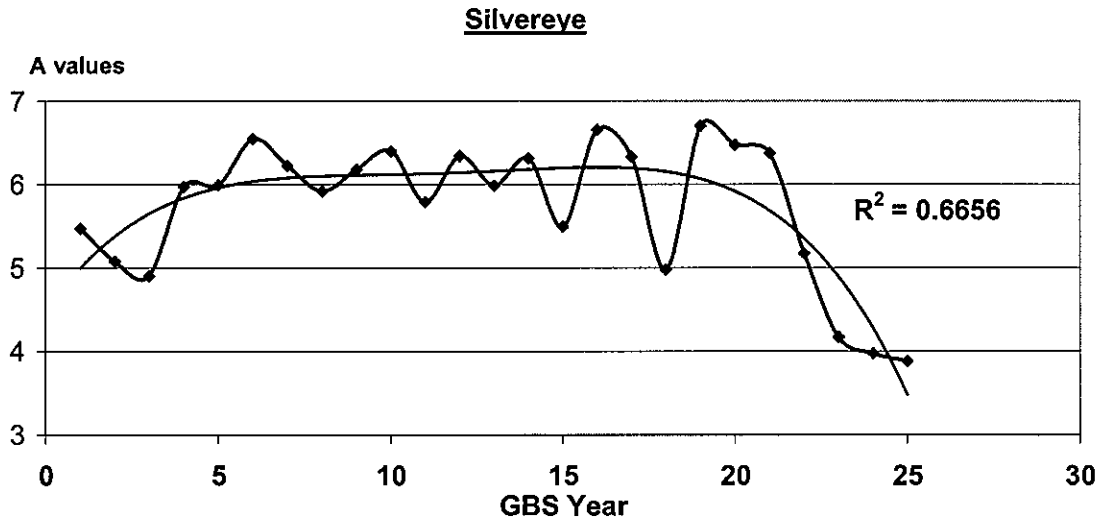
Review team comments:

21 years: Agree.

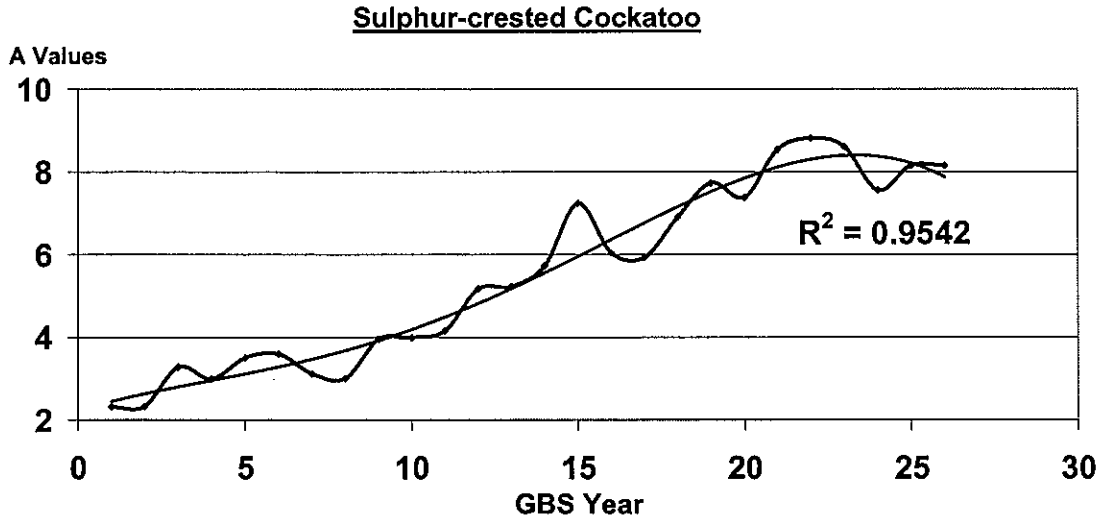
25 years: Trend continues.



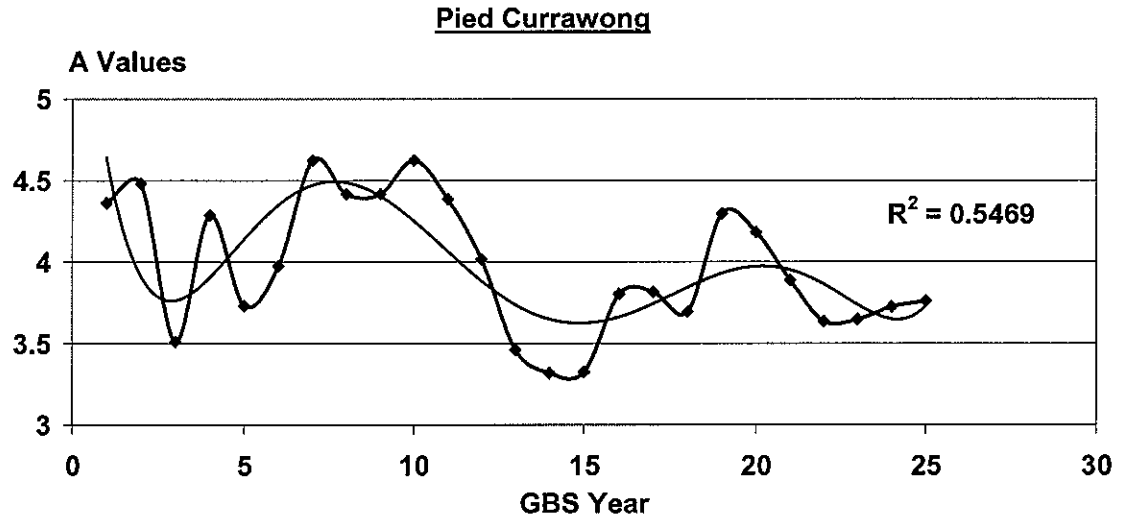
Silvereye, code 574, Long term rank 4
Veerman comments: One of our most regular species.
Review team comments:
21 years: Agree. No trend line should be shown (r2= 0.19).
25 years: Non-linear trends show a marked decline since year 21. $R^2$is not great (0.59) but this may be due to fluctuations earlier in the series.



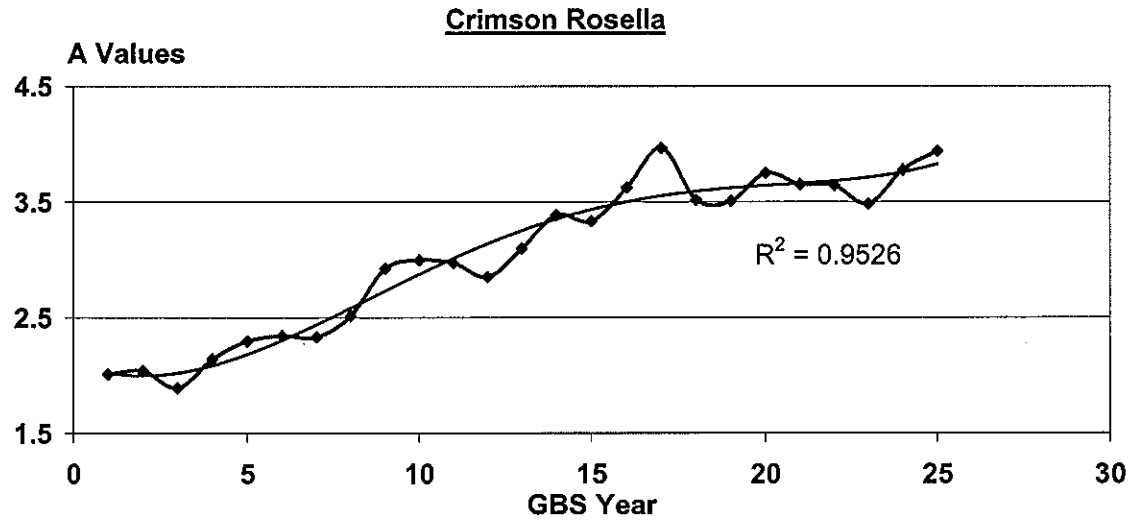
Sulphur-crested Cockatoo, code 269, Long term rank 5
Veerman comments: Numbers have risen significantly.
Review team comments:
21 years: Agree.
25 years: Possible plateau after year 21?



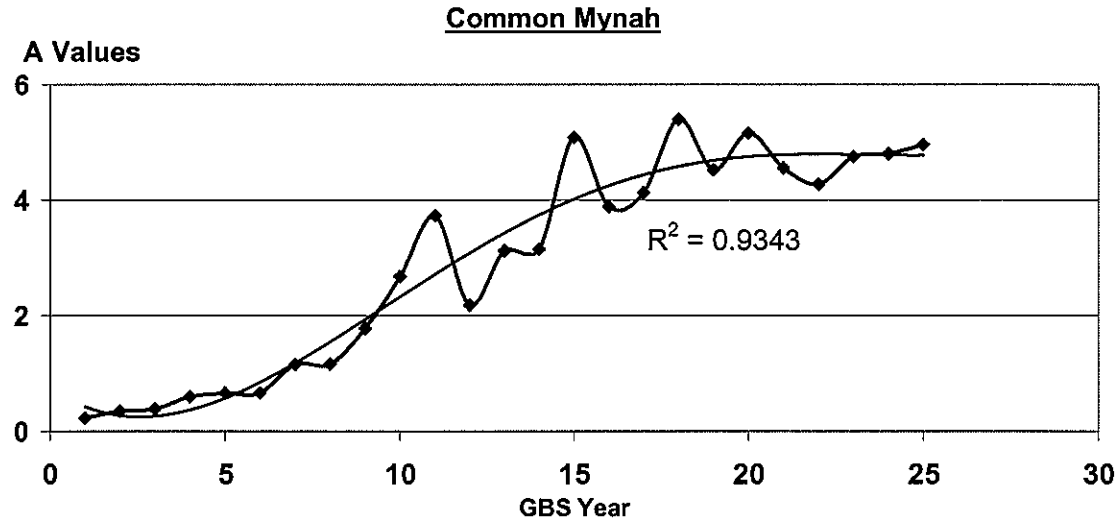
Pied Currawong, code 694, Long term rank 6
Veerman comments: *Comments focused on summer abundance rather that year to year trend.*
Review team comments:
21 years: Year values show a lot of fluctuation
25 years: No clear trend with possible complex behavioural changes as the species becomes more urbanized.



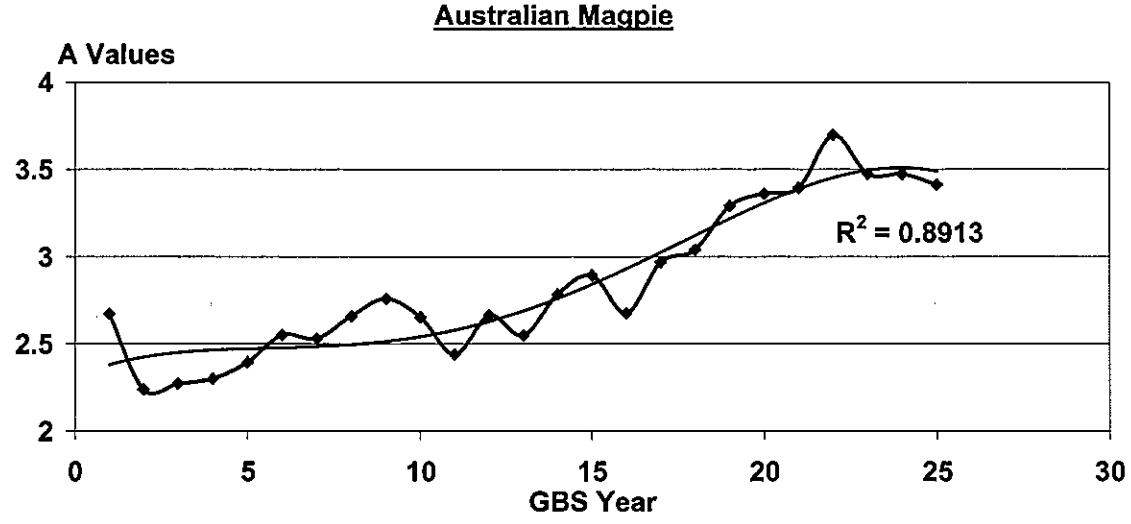
Crimson Rosella, code 282, Long term rank 7
Veerman comments: This species has increased having doubled in numbers.
Review team comments:
21 years: Agree. R^2= 0.9259
25 years: 4th power trend suggests a plateau after year 16.



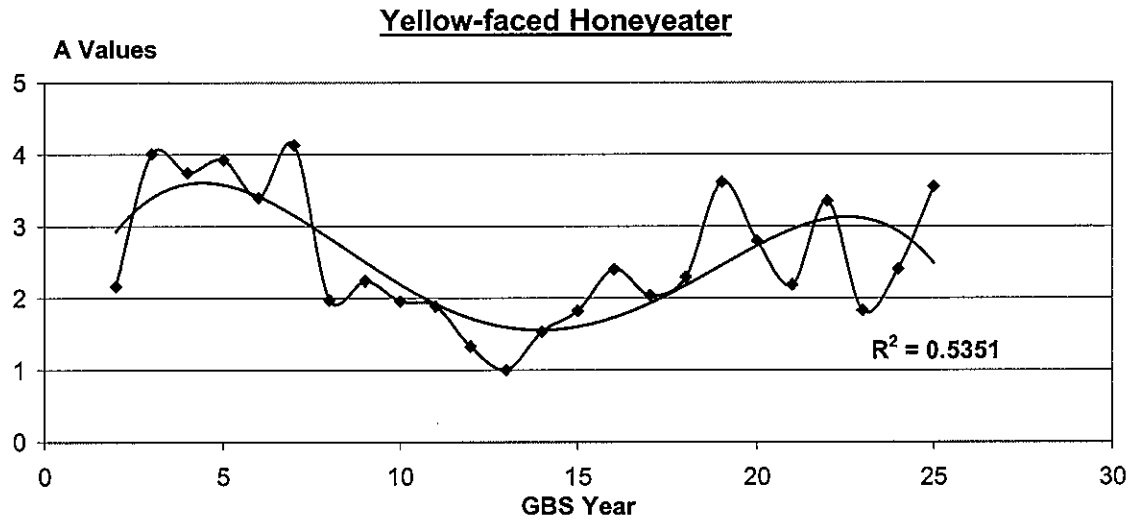
Common Myna, code 998, Long term rank 8
Veerman comments: Dramatic rise in numbers. ...steady population increase..
Review team comments:
21 years: Agree: Linear R^2 is 0.90.
25 years: Linear trend does not fit so well (R^2 drops to 0.87). A 4th power trend suggests a plateauing.



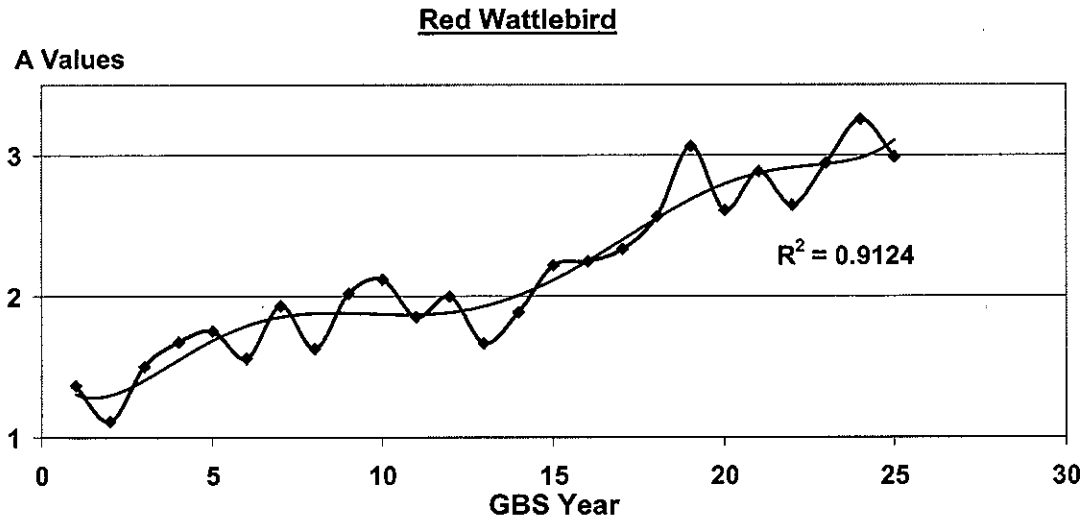
Australian Magpie, code 705, Long term rank 9
Veerman comments: The long term abundance has barely perceptible change although overall a small increase.
Review team comments:
21 years: Compared to many charts the upward linear trend is quite strong (R^2= 0.74). A 4th order polynomial shows an increasingly upward trend (R^2= 0.86).
25 years: Upward trend continued. Local peak in year 22 may be worth investigating.



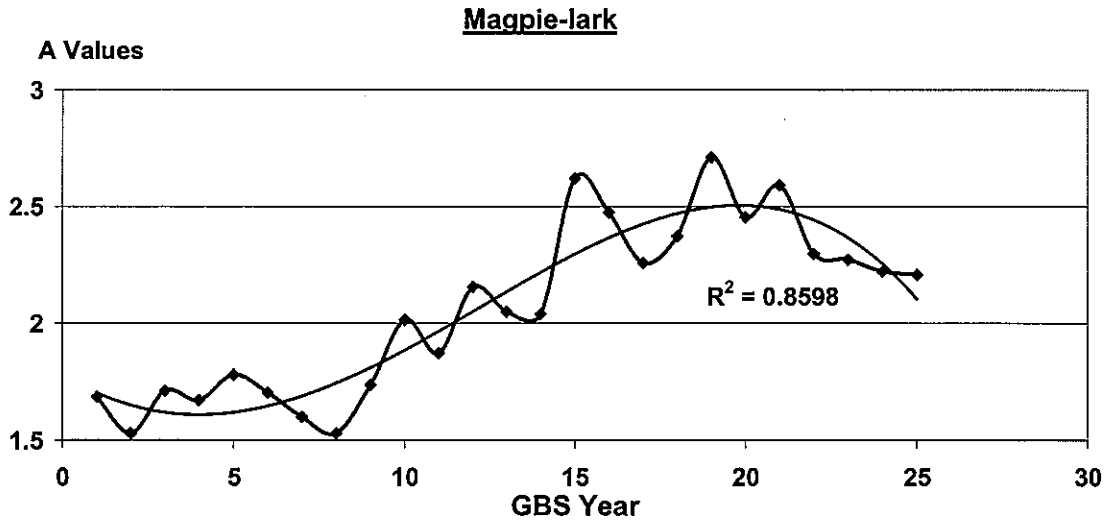
Yellow-faced Honeyeater, code 614, Long term rank 10
Veerman comments: Even with the problems of reliably counting the birds there are marked changes in the birds' abundance over the years.
Review team comments:
21 years: Linear trend has R^2 of 0.25 and shouldn't be shown. Deleting year 1 (as an extreme) gives a 4th order sine wave with an R^2 of 0.81.
25 years: No attempts give a clear trend. The last few years appear particularly random..



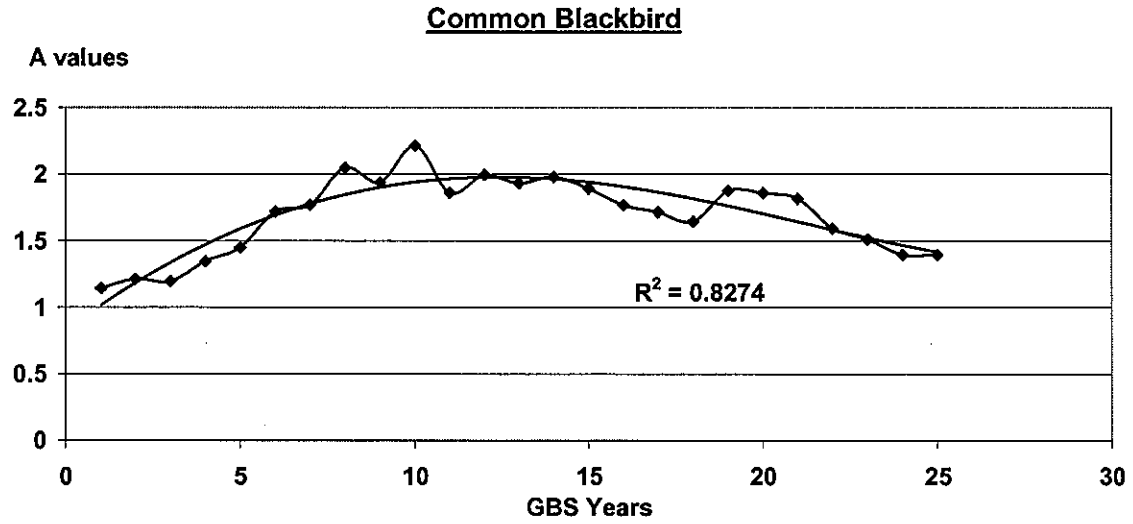
Red Wattlebird, code 638, Long term rank 11
Veerman comments: Long term the species has had a steady increase, its abundance has doubled.
Review team comments:
21 years: Agree.
25 years: The trend continues with an R^2 close to 0.90.



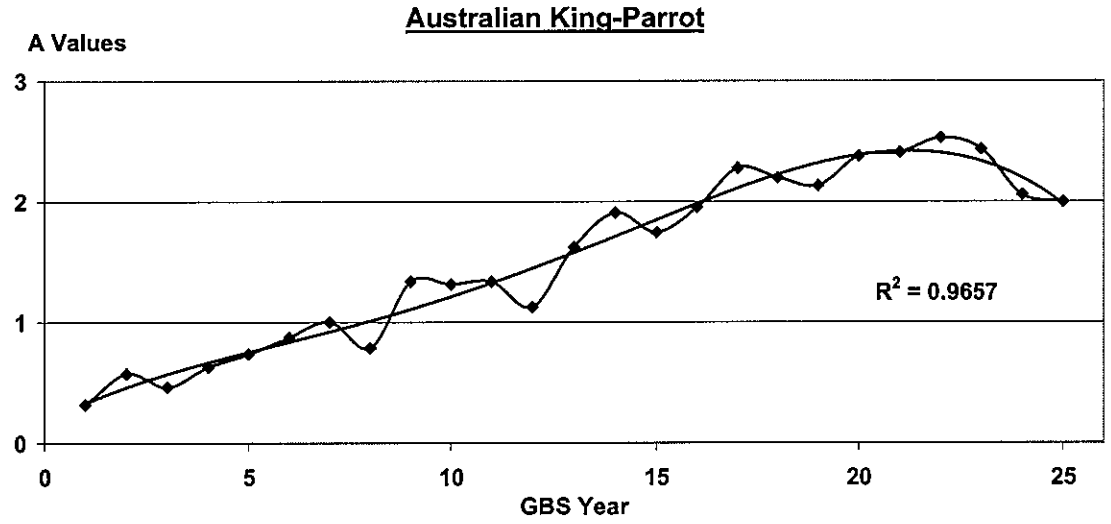
Magpie-lark, code 415, Long term rank 12
Veerman comments: Long term numbers have had a smooth to undulating increase.
Review team comments:
21 years: A fair comment.
25 years: A major decline from year 19. Looks like a cycle appearing??



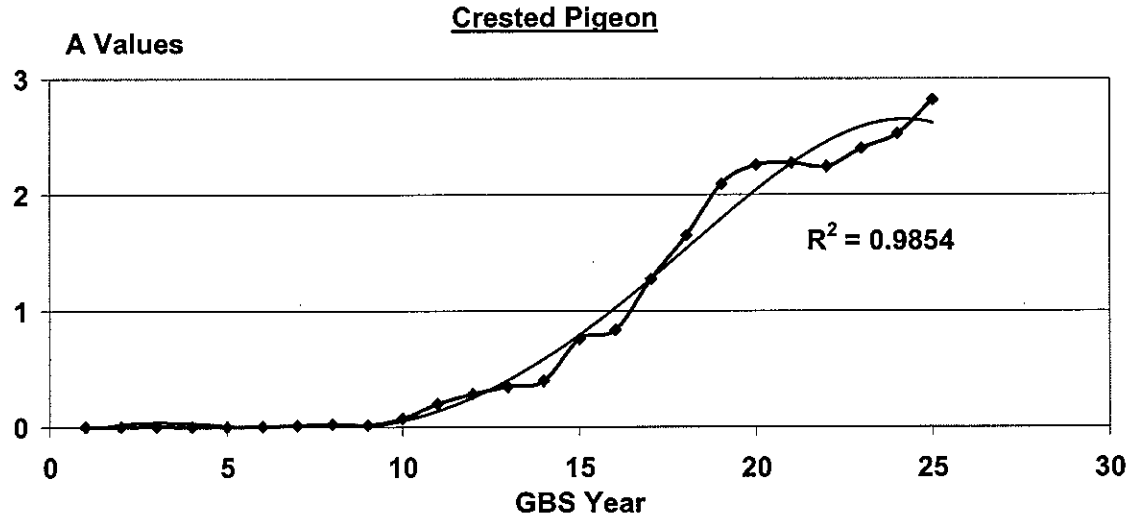
Common Blackbird, code 991, Long term rank 13
Veerman comments: Increased smoothly to year 10 and then marginally declined.
Review team comments:
21 years: Actually increases in last three years.
25 years: Shows peak around year 10 with fairly steady decline since then with local peak between years 19-21.



Australian King-Parrot, code 281, Long term rank 14
Veerman comments: Numbers have increased dramatically in a smooth manner.
Review team comments:
21 years: Agree.
25 years: A drop-off in abundance may be apparent after year 22: possibly an effect of the bushfires on breeding habitat.



Crested Pigeon, code 43, Long term rank 19 (year 24= 11)
Veerman comments: Numbers have risen dramatically. ..it appears to be stable and evening off now.
Review team comments:
21 years: Veerman doesn't show a linear trend. A 4th order trend has an excellent R^2. In essence his comment is sound.
25 years: After a local plateau for years 20-22 the strong upward trend continues.



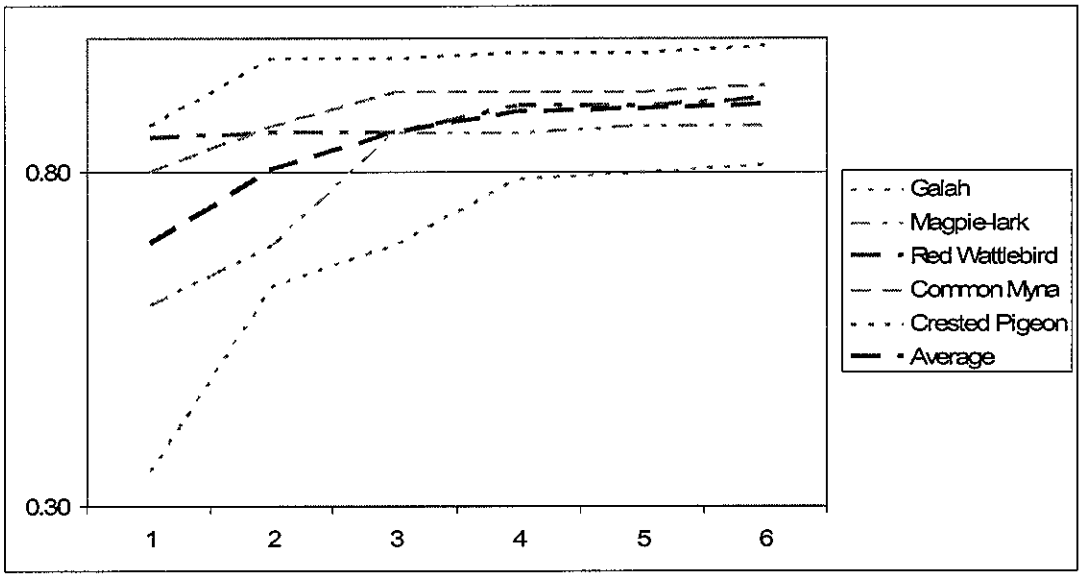
In most cases there is no disagreement with the textual comments made. With regard to the graphs, a measure to indicate the strength of the trend should be shown as a matter of course (together with advice for non-statisticians about the interpretation of that measure). In some cases no trend-line should be shown as there is no reliable trend.

3.2.3) Improved method to show trend-lines

Adding in the extra 4 years seems to suggest some changes to the comments Veerman made, as would be expected. It does, however, suggest that there is a need for updating the graphs in some way. Graphs used to be a feature of the ABR and a small set for the commonest species, as measured by A, has been supplied to the authors of species accounts for year 26. **It is recommended that, as in the past, there be a greater number of monthly and long-term graphs. In addition, trend-line and R^2 value could be included on the graphs where appropriate.**

EXCEL was used to plot the charts which give a range of options for plotting trendlines including linear trend and fitting a polynomial. Using R^2 (available as a ready-made option in EXCEL) as a measure of goodness of fit, we found that in several cases the linear trend gave a very low value of R^2 but using a 4^{th} order polynomial gave quite good results. The shape of the polynomial trend gave an indication of potential plateaux for several species, although exactly how these lines should be interpreted is not clear (and EXCEL does not provide any diagnostics about the significance of the individual coefficients. This does suggest that undue reliance on packages which are restricted to linear trends may not be totally productive. Investigations into the use of the package R, which gives measures of dispersion about the mean for trend-lines should be undertaken.

We have found that in those cases where a linear trend is not appropriate due to a low value of R^2 a 4^{th} order polynomial appears to be more appropriate, see below. From a low goodness of fit value produced by a linear trend (1^{st} Order Polynomial) the R^2 value asymptotes at around the 4^{th} Order.



Increase in the goodness of fit (R^2) for trend line for selected species with increasing order of polynomial.

3.2.4) Regional variation in summary statistics

The summary statistics of 'A', 'F', 'W', 'R' and 'G' provide an overall index with which to compare general population parameters from year to year but they are not associated with any comment on the reliability of the estimates. The 'A' summary statistic provides information on abundance for instance but no information is provided on distribution. This could be provided with an overall value of 'A' followed by a value of 'A for sites located within Belconnen, North Canberra, South Canberra, Woden/Tuggeranong and Rural. This could involve considerable work, due to the current design of the database.

The Review Team appreciates the importance of providing as much summary information as possible to GBS contributors. This is done via the Annual Bird Report but in addition the present contributions provided by the GBS Coordinator (Martin Butterfield) to the monthly newsletter Gang-gang goes a long way to encourage further support for the GBS. Even so, additional information can be provided that may be of use to the Gang-gang newsletter contribution or to the Annual Bird Report.

4) Where to from here?

4.1) Changes to operational procedures

The Review Team's efforts have been directed to determine whether the present survey protocols are providing data that can be directed towards satisfying the aims of the survey. We have examined the error rate within the database and have made suggestions as to how this rate can and is being reduced. We now have three versions of the GBS database, one of which contains corrected data for the first 25 years of the survey and to the best of our knowledge should now be free of the improbability errors. Without doubt there are still some impossibility errors, within the database but these may never be detected due to the nature of the errors although some may come to light with further analysis.
In 2002 and again in 2003 and 2006 Veerman used the summary statistics to provide an analysis of long term trends and seasonality of the birds in Canberra's gardens. The Review Team felt that an addition of a further four years of data is unlikely to add any significant understanding to the literature. The COG woodland survey indicates a decline in many species but at this stage it is unclear whether this is a long term feature due to habitat change or whether there is a significant drought effect. The same may well apply to the GBS.

We therefore recommend that a detailed analysis, possibly along the lines provided by Dobbie and Welsh be commenced after 30 years in the hope that any possible effects of the present dry conditions can be identified.

The Review Team does not have the required expertise to conduct such an analysis and recommends that a project brief be produced outlining the requirements of the analysis.

The Review Team then recommends that enquiries be made to find a suitable person to conduct such an analysis. This person will work in close cooperation with the COG Database and Records Management sub-committee and to ensure an appropriate analysis that is conducted within a timely fashion the project should be commissioned under contract with appropriate remuneration. It is recommended that **the final product should be a paper of suitable standard for submission to a quality, peer-reviewed science journal.** It is therefore envisaged that only a suitable sample of species be used for analysis and that these species be chosen to demonstrate certain aspects of the data. The Review Team agrees that COG is the caretaker of a unique data set and its detailed analysis must be subject to the rigors of a peer-reviewed process.

In the meantime **the Review Team recommends that a second project brief be commissioned to examine the feasibility of redesigning the GBS database and to provide a data-input mechanism designed to capture the types of errors that have been identified in the review.** At the same time it may be appropriate to commission a review of the General Observations database and its associated data input programs. Initially the brief should commission an appropriate person to provide a report that details possible ways with associated costs to improve the system. For the GBS, consideration

will need to be provided on the costs and benefits associated with an on-line data entry module.

4.2) Recommendations for an updated book on the birds of Canberra

It is recommended that the well received publication 'Birds of Canberra Gardens' be updated with the 9-10 years of additional data and republished. A point to consider will be whether to use all available data and thus a need for a change of title or use data from Canberra gardens only.

References

ABS Bulletin 3105.065.001.

ABS Bulletin 3301.0.

Beaumont, L.J., McAllan, I.A.W and Hughes, L. (2006). A matter of timing: changes in the first date of arrival and last departure of Australian migratory birds. Global Change Biology 12, 1339-1354.

Butterfield, M. (2004). Reporting of breeding activity in the Garden Bird Survey. Canberra Bird Notes 29, 99-106.

Canberra Ornithologists Group. (2000). Birds of Canberra Gardens. Canberra Ornithologists Group and ACT Department Urban Services, Canberra.

Cannon, A.R., Chamberlain, D.E., Toms, M.P., Hatchwell, B.J. and Gaston, K.J. (2005). Trends in the use of private gardens by wild birds in Great Britain 1995-2002. J. Appl.Ecol. 42, 659-671.

Chambers, L. Observed climate change impacts on birds. In: The State of Australia's Birds 2007. Suppl. Wingspan 14, No 4.

Dobbie, M.J. and Welsh, A.H. (undated and unpublished). The Canberra Garden Bird Survey: analysis of counts of the Golden Whistler (*Pachycephala pectoralis*), Noisy Friarbird (*Philemon corniculatus*) and Australian Magpie (*Gymnorhina tibicen*)

Dobbie, M.J. and Welsh, A.H. (2001). Modelling Correlated Zero-inflated Count Data. Australian and New Zealand Journal of Statistics 43, 431-444.

Gibbs H 2007 "Climatic variation and breeding, in Australian Magpies". Emu v107, vide http://www.publish.csiro.au/?paper=MU07022.

Greenwood, J.J.D. (2007). Citizen science and bird conservation. J. Ornithology, **148**, Suppl 1, S77-S124.

Hermes, N. (1981). Garden Bird Chart and Survey. Canberra Bird Notes **6,** 112-114.

Mitchell, B. (2006). Life in the Suburbs: Urban habitat guidelines for the ACT. ANUgreen, Australian National University.

Nicholls, A.O. (2003) Review of *Canberra Birds: a report on the First 21 Years of the Garden Bird Survey* by Philip A. Veerman, Wingspan, December, p. 39.

Veerman, P.A. (2002). Canberra Birds: a report on the first 18 years of the Garden Bird Survey. Published by the author.

Veerman, P.A. (2003). Canberra Birds: a report on the first 21 years of the Garden Bird Survey. Published by the author.

Veerman, P.A. (2006). Canberra Birds: a report on the first 21 years of the Garden Bird Survey. 2nd Edition. Published by the author.

Welsh, A.H., Cunningham, R.B., Donnelly, C.F. and Lindenmayer, D.B. (1996). Modelling the abundance of rare species: statistical models for counts with extra zeros. Ecological Modelling 88, 297-308.

Woodall, P. (2002). A review of 'Canberra Birds: a report on the first 18 years of the Garden Bird Survey' by Philip A. Veerman. Canberra Bird Notes 27, 82-83.

APPENDIX

1.1	How many years, in total, have you contributed to the GBS?	1-2 10	3-5 5	5-10 7	>10 25
1.2	Have you contributed over the past 5 years?	Yes 41	No 6		
1.3	How many years have you contributed at your most recent/present site?	0-5 24	6-10 5	11-15 5	>15 13
2	Have you changed the shape of your survey area at your most recent/present GBS site?	Yes 3	No 44		
3	Do you count species flying >100m above the site?	Yes 39	No 7		
4	Do you count species that you have heard and not seen?	Yes 43	No 4		
5	Do you count species passing through but not using the site?	Yes 46	No 1		
6.1	If for any one week you are away for 6 days do you record the species seen on the one day at home?	Yes 21	No 25		
6.2	If for any one week you are away for 6 days do you cross out the week number at the head of the column?	Yes 24	No 20		
7	For your most recent/present GBS site, how did you assess the size of the site?	By eye 32	Paced 10	Off map 5	
8	How many individuals would you record if you saw an adult male and female Australian King-Parrot not together but: Within 30 seconds of each other? Within 5 minutes of each other? On different days within the GBS week?	1 1 16 24	2 46 31 23		
9	Please list the five bird species which present the greatest difficulty in assessing the numbers of birds on your site at any one time	Thornbills/Weebills Silvereyes Pardalotes House sparrows	29 29 15 10		

10	Any comments on GBS?	