Importance of accuracy in co-ordination and integration of terrestrial vertebrate fauna survey databases in Western Australia

R Teale

Biota Environmental Sciences, Suite 2 First Floor, 186 Scarborough Beach Road,
Mt Hawthorn WA 6016
email: biota@biota.net.au

Abstract

Fauna data collected at present typically comprise species identity, abundance, morphometric and meristic measurements, and some basic information on life history. Any database incorporating such data must be designed with its end use in mind. End uses may include determining species distributions, abundances, average weights, *etc*. In the context of this workshop, a key outcome is to facilitate the documentation of Western Australia's biodiversity. Accurate species identifications are imperative in a database designed to document biodiversity, and the only clear mechanism to ensure this is vouchering. The obvious repository for this information is thus FaunaBase, which is held by the Western Australian Museum. It is after all the responsibility of the WA Museum to review and update the taxonomy of the fauna of Western Australia, which surely is the first step in documenting biodiversity.

On a parity with accuracy is precision of data, particularly in relation to morphometrics and meristics. There are clear examples of where differences between observers in measuring techniques, *etc* have generated different interpretations. Similarly, abundance values are contingent on the methodologies used to collect the data and the experience of field practitioners, particularly in relation to opportunistic collecting techniques such as head-torching, raking, *etc*.

There is clearly a "need for better coordination and integration of terrestrial fauna survey data" from surveys across the State. However, there are numerous issues inherent with an observational database, including differing data collection methodologies and experience of field practitioners, which reduce its scientific value. I would argue that much of the underlying infrastructure required to achieve the necessary coordination and integration is already in place in the WA Museum FaunaBase, but that it is currently under-utilised and perhaps under-resourced.

Keywords: database, fauna survey, voucher specimen

Introduction

The intention of this workshop is to investigate the development of a database to house all relevant data from all fauna surveys across the state. The rationale behind the development is "If data from surveys were coordinated and integrated it would provide an enhanced opportunity to understand and interpret the biodiversity and ecosystems of WA." In this case the definition of biodiversity follows that of the Environmental Protection Authority (EPA):

> *It is the variety and variability of all life forms, all plants, animals and micro-organisms, the genes they contain and the ecosystems they form.*

It is worth noting at this stage that the proposed database would largely be an observational database, in contrast to say the Western Australian Museum FaunaBase that only stores information on vouchered specimens. A key advantage of the latter is that through the vouchered specimen we can be more confident of the accuracy of the dataset. If there are any questions or taxonomic revisions, then we can go back to the original specimens and check the records again.

This paper was presented in similar form at the Workshop on the Co-ordination and Integration of Terrestrial Vertebrate Fauna Survey Database for Western Australia, held on April 5, 2002, at the University of Western Australia

Stimulus for workshop

The stimulus for this workshop seems to have been recommendations made within the Environmental Protection Authority Position Statement No. 3 "General Requirements for Terrestrial Biological Surveys for Environmental Impact Assessment in WA". This Position Statement highlights, with respect to databases:

> "The need for both a consolidated database and for data to be collected by the proponent or their consultants in a format to allow ease of assessment at the local, regional and national levels, and to facilitate transfer to State biological databases"

The above EPA recommendation covers a broad range of topics. It first mentions a need for a consolidated database, secondly that data be collected in a format to allow ease of assessment, and thirdly that it can be transferred into State biological databases.

Compatibility with State databases

I think that the issue of compatibility is a key component of the EPA's statement and leads us to ask ourselves some fundamental questions:

- can existing databases give us insights into understanding and interpreting biodiversity in WA?
- recognising the cost associated with maintaining databases, can we use those resources to improve on existing databases in preference to embarking

on the development of the proposed new observational database?

- would any such improvement negate or partially offset the need to develop a new database?

In my role as a consultant, I use on a regular basis a number of State databases and some private ones including our own. One of the most useful resources is provided by the Western Australian Museum (WAM) collection, which is essentially the foundation of their database FaunaBase. The value of the collection is that it gives us an assurance of accuracy, provides us with a means to review taxonomy and is readily updateable in response to taxonomic review. Interestingly, the WAM is looking towards coordinating data storage/retrieval not only at a State level but also at a National level with other museums across Australia (P Berry, WAM, personal communication).

Fundamentally, the checks and balances required to ensure the highest accuracy of data in any new, largely observational database would also enhance the value of FaunaBase. Experience indicates that databases are costly to establish, populate and maintain. Therefore any resources we might have to put towards developing integrated data management would perhaps be better channelled into existing databases to further enhance their functionality and value, particularly those databases that underpin our understanding of biodiversity.

Accuracy and vouchering

So what are the 'checks and balances' alluded to above? Basically they are the process of ensuring accuracy through vouchering. In a sense vouchering could be considered as peer review of field identifications, which are otherwise often carried out in isolation, in an analogous way to peer review of the resulting papers. This is particularly important where the work is carried out by inexperienced field practitioners or where experienced practitioners are working in new regions. Where the focus of a paper is a species inventory or where species identification is critical to the conclusions of the study, then it is crucial that the field identifications are correct.

Similarly, the worth of a database lies in the accuracy of data that it contains *i.e.* "put rubbish in, get rubbish out". Accuracy of data is acknowledged as one of the key issues requiring discussion at this workshop and encompasses the following sub-topics in relation to development a new observational database:

- Mis-identification
- Who checks
- Who changes
- Nomenclature
- Voucher specimens
- Submissions, additions and corrections to data

Observational databases already exist for avifauna records and provide a template to address the above subtopics. For example, in the Bird Atlas program (Birds Australia) records can come from any registered individual. These records are vetted by a panel of experts and questioned if the records are of species at the edge of, or beyond, their known range of habitat preferences or distribution, or involve uncommon or rare species. However, these unusual records are often exactly the types of records that the scientific community is interested in, particularly many ornithologists. Observational datasets are largely unavoidable for birds because vouchering of this group is impossible for the typical amateur bird observer. In addition, very few bird studies, particularly those carried out during inventory surveys, use mist netting or other trapping techniques to record birds, as it is much simpler to observe and record. However, inexperienced bird watchers do make mistakes and often only record the common species (Saffer 2002), and as such these sorts of datasets are limited in their use.

Are similar identification errors made when identifying non-avian fauna? Clearly the answer is yes. Species within the *Neobatrachus*, *Heleioporus*, *Ctenotus*, *Pseudomys* and *Sminthopsis* genera often provide challenges for experienced field biologists, let alone inexperienced ones. Indeed, experienced field biologists working in one region of WA may not necessarily be familiar with species in a different region. Furthermore, and perhaps more importantly, the taxonomy of many of Western Australia's herpetofauna and mammalian fauna (unlike our birds) is incomplete (*e.g. Gehyra* spp, *Cryptoblepharus plagiocephalus/carnabyi*, *Lerista muelleri* complex, *Menetia greyii* complex, Aplin & Smith 2001; B Maryan, WAM, personal communication; G Harold, personal communication), *Planigale* spp, *Sminthopsis macroura* complex (N Cooper, WAM, personal communication). Quite the opposite of birds, reptiles and mammals are largely recorded from trapping and rarely through observational data (with the exception of large macropods, possums, varanids, *etc*) and this gives us the opportunity to confirm preliminary field identifications through subsequent vouchering.

This is not to say that every record obtained during inventory and survey needs to be supported by an accession number, nor that vouchering should be indiscriminate. Rather, each taxon needs to be supported by at least one vouchered specimen (qualified below) and possibly more, depending on the geographical spread of the study sites and the taxonomic status of the species. Vouchering also needs to take into consideration the fragmentation and size of habitat, historical collections from the locality, and the conservation status and distribution of the species.

However, it is not just a question of getting the identification right; it is also a matter of resolving taxonomy, being able to describe new taxa and enhancing the States biodiversity knowledge base. Taxonomy is constantly under revision and without voucher specimens we cannot describe new taxa. Surely taxonomy is a first step to documenting biodiversity, a key aim of this workshop.

Crucially, without vouchers we may not be certain of the identity of records for recently resolved taxa identified from within large species complexes (*e.g. Lerista muelleri)*. LA Smith (WAM, personal communication) has recognised approximately 17 species in the *L. muelleri* complex, all of which were previously registered as *Lerista muelleri*. Such a process of revision for a large and well collected species complex, with constituent taxa demonstrating some sympatry, renders universal changes in observational databases nearly

impossible and may in fact invalidate large amounts of data *e.g.* nearly all *L. muelleri* records will need to be eliminated. There are many other examples of such species complexes in WA (Aplin & Smith 2001).

Whilst the resources are not yet available to undertake every required taxonomic review, a database supported by voucher specimens will not suffer from problems associated with invalidated data. Taxonomists can undertake reviews at a later stage and then make the changes to records of the vouchered specimens.

Once voucher specimens are lodged with the WAM, the responsibility of the tasks identified in the above bullet points becomes obvious:

- Mis-identification / Who checks — The WAM staff, using the extensive collection and their specialist knowledge, can compare newly vouchered specimens against the collection to check field identifications.
- Who changes — If necessary, WAM staff can then make the changes, or alternatively the vouchered specimens remain in the WAM until relevant experts undertake a revision.
- Nomenclature — The WAM is responsible for determining which nomenclature practitioners in WA should be using, *e.g.* 'Checklists of the vertebrates of Western Australia' published in 2001 in Records of the Western Australian Museum, Supplement 63.
- Submissions — WAM staff enter the information for vouchered specimens into FaunaBase, so there is no need to duplicate this information in another database.

Vouchering of specimens in WA seems to be limited to a handful of individuals (N Cooper, WAM, personal communication). This lack of vouchering seems to be central to the issue of the perceived need to develop a new database infrastructure.

If as a group, field biologists are interested in gaining a better understanding of biodiversity (as defined by the EPA), then we first need to ensure accuracy. It is my belief that the best mechanism for ensuring accuracy is through vouchering. Furthermore, this provision of specimens also enables taxonomists to carry out taxonomic revisions and identify new taxa, another key component to understanding biodiversity. In addition if field biologists voucher more diligently, then FaunaBase itself becomes a much more useful tool for other types of data interrogation.

Ramifications of an increase in vouchering

If as a first step to integrating our data, we need to ensure that accuracy is assured through vouchering, then we will obviously need to consider the significant resource/funding constraints that this will place on the WAM. By way of example, a total of 1300 herpetofauna records encompassing 79 taxa were made, and approximately 330 specimens were vouchered, from a recent survey (Biota Environmental Sciences 2002 Proposed Hope Downs Rail Corridor from Weeli Wolli Siding to Port Hedland – Vertebrate Fauna Survey; unpublished report for Hope Downs Management Services) involving over 5000 trap nights spread along a 330 km transect. From the same survey, there were 480 records of non-volant mammals, with 109 individuals vouchered.

What are the approximate resource demands on the Western Australian Museum from this level of vouchering? If the animal is alive, the process of euthenasing, extracting tissue for DNA, preserving, databasing, labelling and putting in the collection takes at least 45 mins per specimen, longer for mammals (1.5 hours) if the skull has to be prepared. So, 330 herpetofauna would take 245.7 person hours or 31 person days, whilst 109 mammals would take 163.5 person hours or 21 person days (N Cooper, WAM, personal communication).

If we have time and energy/resources to assist with developing a new database, I think that these would be more appropriately channelled (at least initially) into ensuring that the foundations of the database (*i.e.* accuracy and taxonomy) are adequately addressed. The feedback mechanism to the field biologists also provides important training.

Misplaced conservation

Misplaced conservation (after M Craig, personal communication) refers to the focus of biologists and others at the level of individual animals, rather than at the population level, when considering species conservation. In large intact areas of vegetation, considered vouchering would not in all likelihood cause a detectable impact to population levels. In the above example, the 330 vouchered individuals from in excess of 70 taxa were from 33 sites, each spread on average 10 km apart. This equates to ten individuals per site. The Biota (2002) survey recorded on average 7.8 ± 5.2 species per site, thus in most cases only one or two individuals of any one species were taken from each site. Clearly this level of collecting would not have a detectable impact on local populations. In contrast, in highly fragmented and small remnants, vouchering of some species may hasten the demise of local populations. As stressed above, vouchering should not be indiscriminate.

In the above example, vouchered specimens included one species possibly new to science (*Ctenotus* aff *uber johnstonei*), a recently "re-discovered" species that is poorly collected (*Ctenotus* aff *robustus*), and another poorly collected taxon, *Vermicella snelli.* Other specimens that were vouchered included taxa (*Diplodactylus stenodactylus, Lerista bipes, Lerista muelleri* and *Menetia greyii*) that are known to belong to species complexes, and some that are just extremely difficult to accurately key out in the field (*e.g.* members of the genus *Ramphotyphlops*).

Limitations of observational data

What about the remaining observational data that were collected in the above example (*i.e.* the other 1000 herpetofauna records) — what are their value? My response to this question is, what is the intended end use of the data? Fauna data collected typically comprises species identity, abundance, morphometrics and meristics, and some basic information on life history. End uses may include determining species distribution, abundances and average lengths, weights, *etc.*

If the end use is species distributions, then the Western Australian Museum FaunaBase is already established and can incorporate representative vouchers from the study site. For rare species, where observational data is important due to restrictions on collecting threatened taxa, the Department of Conservation and Land Management Rare Fauna Data Base is already established.

Abundance data are contingent upon the methodologies used for collection, for example, Elliott trapping, pit fall trapping (buckets or PVC tubes), head torching, raking, *etc*, seasonal timing, and survey duration. Undoubtedly the experience of the practitioner also markedly affects abundance values, especially from opportunistic collections and in selecting trapping sites.

Morphometric data are highly variable between different observers; even within observers some characters cannot be scored reliably (Humphreys 1990). Clearly the amalgamation of meristic and/or morphometric datasets from a large number of different observers is of little value for any detailed descriptions.

Exceptions

There are some extremely good practitioners whose observational datasets would be of huge value if used correctly. For example, Greg Harold (a very experienced field biologist) has recorded every single herpetofauna he has seen since the mid 1970s. This would be a fantastic dataset and one that would probably have no more errors in it than FaunaBase. These data could thus reliably be used to obtain a better idea of species distributions. Similarly, where datasets are supported by adequate vouchering, location details from non-vouchered records can add to our knowledge of species distribution.

Conclusions

Any new database would largely store observational data. An observational database would not meet the primary objective of the rationale behind its development, that is:

"If data from surveys were coordinated and integrated it would provide an enhanced opportunity to understand and interpret the biodiversity and ecosystems of WA."

Clearly there are many limitations on collective data placed into an observational database by many observers, which end users may not always be aware of or may not consider. These include different methodologies, experience of the field biologists, weather, fire history, *etc.* When combined with questions of accuracy and precision, these can significantly undermine the value of an observational dataset.

We cannot understand nor interpret the States biodiversity (as defined by the EPA) without ensuring accuracy of our field identifications. The need to voucher to ensure accuracy means that FaunaBase will support more records and become a more useful tool for interrogation.

To summarise, we should initially examine existing databases before embarking on the development of a new database. FaunaBase is an obvious choice as it resides with the Western Australian Museum, whose staff can confirm initial field identifications, are responsible for nomenclature in the State and regularly undertake taxonomic reviews. FaunaBase is already operational, has the necessary IT infrastructure in place, has a clearly identified custodian and is widely accessible to the scientific and general community. This addresses many of the issues that would undoubtedly arise in the delivery of any new database.

Implications are that the increased level of vouchering that would be required to ensure the accuracy of any data that we would place into a proposed observational database would place considerable strains on the resources at the Western Australian Museum. However at the same time the value of FaunaBase would be enhanced so that it may offset the need to develop a new database. Rather than investing in a new database, what I believe is required - at least initially - is better support, funding, resourcing and in particular USE of existing databases such as FaunaBase. If through this process we became more confident in the accuracy of our data, then we could re-examine an integrated observational database. As indicated above, there are also other databases available that we need to consider, for example the CALM Rare Fauna Database and the Pilbara Biological Reporting Database (see Biota 2001).

References

Aplin K P & LA Smith 2001 Checklist of the frogs and reptiles of Western Australia. Records of the Western Australian Museum Supplement 63:51-74.

Biota 2001 Pilbara Biological Survey Review. Unpublished Report for the Department of Resource Develpment. Biota Environmental Sciences, Perth.

Humphreys G 1990 Geographic variation in King's Skink, *Egernia kingii* (Scincidae); island populations and genetic differentiation. Honours Thesis, Department of Zoology, University of Western Australia, Perth.

Saffer V M 2002 Community involvement in the assessment of the health of selected remnants in south-western Australia. Western Australian Naturalist 23:237-246.