COLLATION AND DIGITISATION OF FRESHWATER AQUATIC INVERTEBRATE DATA FOR SOUTH AFRICA

Report to the
Water Research Commission

by

Jenny A. Day[1] & Belinda E. Day[2]

[1]Biological Sciences, University of Cape Town
[2]Cape Waters

WRC REPORT No. KV 315/13

JULY 2013

Obtainable from

Water Research Commission
Private Bag X03
GEZINA, 0031

orders@wrc.org.za or download from www.wrc.org.za

The publication of this report emanates from a project titled *Collation and digitisation of freshwater aquatic invertebrate data for South Africa* (WRC Project No. K8/998).

EXECUTIVE SUMMARY

1. The objective of this project was to initiate a digital inventory of the freshwater aquatic invertebrates of South Africa by developing a database containing taxonomic and distributional data for South African aquatic invertebrates. This has been achieved in the form of South African Freshwater Invertebrate Database (SAFID).

2. Taking into account the various types of platform available, we decided to use the very simple Excel spreadsheet program because most software for housing biodiversity data is record-based and therefore unsuitable for our purposes.

3. The database consists of 42 files, each consisting of a species list with distributional data, plus a sheet explaining geographical codes used in the species list and a third sheet detailing sources of data. An additional file contains this report together with an extensive list of the references used in compilation of the checklist.

4. We have attempted to include all species known for southern Africa and the Afrotropical region and also to record distribution ranges, even those beyond the shores of Africa, of all species included in the checklists.

5. The checklists are based largely, if not entirely, on the records in the 9-volume series *Guides to the freshwater invertebrates of southern Africa* published by the South African Water Research Commission between 2000 and 2007.

6. Other sources of data include original species descriptions in the taxonomic literature; papers by taxonomic experts detailing additional records; existing databases at museums and other organisations; species lists in theses and papers reporting on ecological studies; and lists of species in private collections.

7. Various institutions, particularly museums, hold taxonomic and distributional data that are already warehoused. Contact details are provided for people wishing to access these existing databases.

8. It is recommended that the SAFID checklists are warehoused and curated by the South African Environmental Observation Network (SAEON), at least until the South African Encyclopaedia of Life (SAEOL) has been developed by SANBI.

9. While taxonomic and systematic studies on terrestrial plants and fishes are relatively well funded (at any rate there are institutions that employ systematists in these fields), we urgently need to fund continued work on the taxonomy and systematics of freshwater invertebrates, especially of important and under-studied groups.

ACKNOWLEDGEMENTS

We thank the Water Research Commission for providing the funds, and WRC Research Manager Bonani Madikizela for help and encouragement throughout the course of the project. We also thank the following people for providing information about access to museum records: Amy Kirk-Spriggs of the National Museum, Bloemfontein; Ferdy de Moor of the Albany Museum in Grahamstown (also for a list of trichopterans); Mikhail Mostowski of the KwaZulu-Natal Museum in Pietermaritzburg and Simon van Noort of the Iziko Museum in Cape Town. We also gratefully acknowledge the compilers, too many to mention individually, of the valuable data provided in the published papers and unpublished reports, theses, and checklists used in this project. Each is individually acknowledged in the references to the relevant checklists, however.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

SABIF	South African Biodiversity Information Facility
SAEOL	South African Encyclopaedia of Life
SAEON	South African Environmental Observation Network
SAFID	South African Freshwater Invertebrate Database
SAIAB	South African Institute for Aquatic Biodiversity
SANBI	South African National Biodiversity Institute
WRC	Water Research Commission

1. INTRODUCTION

The objective of this project is to initiate a digital inventory of the freshwater aquatic invertebrates of South Africa. The rationale is that, although numerous papers have been published on the taxonomy and ecology of these organisms, we still do not have a checklist of species, let alone accompanying data on their distribution both within South Africa and further afield. Such a list is an important contribution to the requirements of the Convention on Biological Diversity, to which South Africa is a signatory. Indeed, the South African National Environmental Management: Biodiversity Act (no. 10 of 2004) requires the provision of Biodiversity Management Plans but these cannot be adequately developed if we do not know how many species we have, or how they are distributed through the landscape. Furthermore, if we want to be able to examine the biogeography of invertebrate taxa, particularly in the light of climate change, collation of existing data is necessary. When this has been done it will also be possible to investigate the biogeographical distributions of, and produce Red Data lists for, some of the better-studied taxa; to inform State of the Environment and State of Biodiversity reports; and to assist wetland managers in determining Reserves, especially in biodiversity 'hotspots'. Furthermore, it will be possible for the first time to address the issue of "dark diversity" - the taxa that have NOT been recorded for particular areas. Biodiversity scientists want to know whether a lack of records for individual taxa is a matter of inadequate sampling or if there are real gaps in distribution. Real gaps may indicate something of interest about the management of rivers and wetlands to water managers.

Invertebrates are targeted in this project because checklists of fishes are already available from the South African Institute for Aquatic Biodiversity (SAIAB). Information on wetland plants has been collated by personnel at the South African National Biodiversity Institute (SANBI) and indications are (Rene Glen, SANBI, pers. comm.) that it is likely to be published soon. (Furthermore, a guide to identification of wetland plants has very recently been published by the WRC: van Ginkel *et al.,* 2010). This leaves the invertebrates as the major group for which an inventory/checklist is required.

1.1 Objectives of the project

1.1.1. To decide on and develop an appropriate database;

1.1.2. To locate existing databases and provide details of access to organisations holding them;

1.1.3. To digitise literature records.

2. RESULTS

2.1. The database: the South African Freshwater Invertebrate Database: SAFID

The outcome of this project is the South African Freshwater Invertebrate Database, SAFID. The database consists of a series of 42 checklists, each containing the species and distribution data to be found in the various chapters of the *Guides to the Freshwater Invertebrates of southern Africa* (Day *et al* 1999; Day *et al* 2001a, 2001b; Day & de Moor 2002a, 2002b; de Moor & Day 2003a, 2003b; Stals & de Moor 2007), together with files containing the important literature on each of the groups. Some additional data from personal records and the literature have been added to the checklists but this process is not yet complete. Below is a description of the approach used in developing the database.

The first step in this project was to decide on which electronic platform to use. Thereafter it was necessary to decide on the structure of the database and the types of data to be entered.

2.1.1 Choice of database

Initially we had anticipated using one of the relatively sophisticated databases (such as PRECIS or Specify: http://specifysoftware.org/) commonly used by museums and other repositories of taxonomic data. The advantage of such a system is its common use in museums, and this is an obvious benefit. Discussions with museum and SAIAB (South African Institute for Aquatic Biodiversity) personnel indicated that these data repositories are specimen- or record-based. This means that the 'unit' is a record rather than a species, requiring repeated entry of identical data for each record. Where numerous records exist for a species, this process can be tedious and time consuming. Furthermore, relatively sophisticated databases of this sort often require purchase (although Specify is free Open Source software) and customising, both of which can be expensive. For our purposes, then, we decided to use the very simple Excel spreadsheet program. It is easy to use, it is available on virtually all personal computers, and data can be migrated easily from it to more sophisticated systems if need be. Furthermore, PowerPivot, an add-on to Excel, allows interactions between numerous individual spreadsheets. Thus it is possible to ask questions such as, "which species have been recorded in all provinces?" or "which species have distribution ranges confined to South Africa?"

2.1.2 Structure of the database

The database consists of 42 files, each consisting of a species list with distributional data, plus a sheet explaining geographical codes used in the species list and a third sheet detailing sources of data. An additional file contains this report together with an extensive list of the

references used in compilation of the checklist.

The primary component of the database is a list of orders, families, genera and species. Where these are easily available (e.g. in some of the Invertebrate Guides), the names and dates of the original authors of the species are included. Note that authors' names in parentheses indicate that the species was originally described by her/him under another name.

The other major component of the checklists is distributional data. This is entered as follows:

- name of sampling site and collector: in this column one can enter specific records of species occurrence obtained from the taxonomic literature, from field collections recorded in the ecological literature, from theses, personal collections, and so on; geo-coordinates can be listed here;

- general area: the country, province or district in which the record was made, or the "hydrobiologial regions" of Harrison (1959);

- other records: records outside of southern Africa

- comments: notes on name changes, habitat preferences, or other features of interest.

Note that records of unpublished species (e.g. *Chironomus* sp.) are not included because such "species" are taxonomically invalid.

2.1.3 Geographic coverage

With the exception of the south-western Cape endemics, very few species found within South Africa are confined to her borders. Many are distributed through the Afrotropical (= Ethiopian) faunal region, which extends throughout sub-Saharan Africa. Some occur throughout the continent or even on other continents. Furthermore, the *Invertebrate Guides* (referred to in section 2.1.4 below) cover southern Africa, defined here as "south of and including the Cunene River catchment in the west and the Zambezi River catchment in the east". We have thus attempted to include all species known for southern Africa and the Afrotropical region and also to record the entire distribution ranges, including those beyond the shores of Africa, of all species included in the checklists. The reason for making the checklists as all-encompassing as possible is that these additional data will be important for future biogeographical analyses.

As with any biogeographical exercise, the distributions of many taxa have to be inferred from very few collecting sites. It is not possible to be sure, from a few scattered records, whether

an absence of records is because the species really does not occur in an area, or whether there has been little or no collecting effort in the region. What is more, collecting effort is almost always concentrated along main roads, near university towns, and in scenic areas. The only way to get detailed data from further afield is by physically visiting areas of interest. While such visits would provide valuable distributional data, they are unlikely to happen except inadvertently during sampling for ecological studies or deliberately by individual taxonomists in search of particular taxa.

2.2 Sources of data

In most cases the list is based largely, if not entirely, on the records in the 9-volume series *Guides to the freshwater invertebrates of southern Africa* published by the South African Water Research Commission between 2000 and 2007 (Day *et al.*, 1999; Day *et al.*, 2001a, b; Day & de Moor 2002a, b; de Moor *et al.*, 2003a, b; Day *et al.,* 2003; Stals & de Moor, 2007). The hydracarine and chironomid checklists were compiled by the late Arthur Harrison and the trichopteran checklist by Dr Ferdy de Moor.

For various reasons, the contents of the different chapters in the *Invertebrate Guides* vary greatly in the depth of coverage. A few taxa are very well known but many have never been seriously studied. Some taxa are currently under investigation by taxonomists, and others are not. Some taxa are very large and so the authors decided to cover genera but not species, while other chapters include all valid species. Distributional data also vary from taxon to taxon. At one extreme, some authors provide world-wide distributional data for each species; some provide distributional data only within Africa or within southern Africa, and some provide almost none at all. These constraints mean that the *Invertebrate Guides*, and the checklists on which they are based, are the best available but are not complete. The situation is unlikely to improve until more emphasis is placed on (i.e. more funding is available for) taxonomic studies in South Africa and the rest of the continent.

The task of compiling all of the data from the *Invertebrate Guides* was far more time-consuming than we had expected and cut down on time available for entering other smaller data sets (see section 2.2.3 below).

Other sources of data include

- original species descriptions in the taxonomic literature and papers by taxonomic experts detailing additional records
- existing databases at museums and other organisations
- species lists in theses and papers reporting on ecological studies
- lists of species in private collections.

Details of each are provided in the next section.

2.2.1 Original species descriptions and additional records

The *Invertebrate Guides* are based on original descriptions so this aspect is well covered. Of course new species are constantly being described, and new records published, so lists go out of date unless a concerted effort is made to keep up with the emerging literature. Taxonomists usually do this as a matter of course. In Africa, however, the lack of taxonomists working on most groups means that few new species are described and few new records published except on groups being studied by taxonomists from elsewhere, usually Europe or North America.

2.2.2 Existing databases

Various institutions hold taxonomic and distributional data, as indicated below. It is not necessary to capture these data but a means of quick access should be developed at a later stage, perhaps within a data repository at the South African Earth Observation Network (SAEON).

Note that no attempt has been made to contact overseas institutions, even if they are known to curate southern African material.

Museums

Albany Museum, Grahamstown: http://www.ru.ac.za/albanymuseum/

> National Collection of Freshwater Organisms (1.5 million specimens), including the Omer-Coopers' collection of aquatic beetles.
> Curator Ferdy de Moor, F.deMoor@ru.ac.za

Iziko Museum, Cape Town: www.iziko.org.za/static/landing/natural-history-collections

> Formerly the South African Museum; mostly Phyllopoda and K.H. Barnard's other arthropod collections, which include specimens in various freshwater insect and crustacean orders. Digitised records (captured using Specify software) are all publically available on the South African Biodiversity Information Facility (SABIF) portal: http://www.sabif.ac.za/; Barnard's wet collection has not yet been digitised. Staff of the Museum intend to make their data available on the South African Encyclopaedia of Life (SAEOL) website, which is currently under construction by personnel from the South African National Biodiversity Institute (SANBI).
> Curator Simon van Noort svannoort@iziko.org.za

KwaZulu-Natal Museum, Pietermaritzburg: www.nmsa.org.za

Formerly the Natal Museum; some aquatic taxa, including some molluscs, dipterans (including the dipteran collection - mainly culicids and ceratopogonids- from the former South African Institute for Medical Research), heteropterans and oligochaetes. Curator Mikhail Mostowski: mmostovski@nmsa.org.za

National Museum, Bloemfontein: www.nasmus.co.za

Not many aquatic invertebrates in the collections; Diptera, Megaloptera and Neuroptera are currently being digitised; the intention is to complete digitisation of the Coleoptera by the end of 2013; data are being checked by curators, after which species lists will be made available, again towards the end of 2013.
Curator Ashley Kirk-Spriggs: ashley.kirk-spriggs@nasmus.co.za

Ditsong National Museum of Natural History (previously the Transvaal Museum), Pretoria: www.ditsong.org.za/naturalhistory

Large collections of beetles, mostly terrestrial; Pinhey's & Balinsky's collections of odonates; CK Brain's collection of rotifers and microcrustaceans
Contact person Klaas Manamela: manamela@ditsong.org.za

The South African Institute for Aquatic Biodiversity (SAIAB): www.saiab.org.za

SAIAB has an extensive collection of fishes but, despite its name, does not house invertebrate material.

Universities, Research Institutes, etc.

The National Freshwater Snail Collection, Dept of Biological Science, University of the North-West

A large collection of freshwater molluscs, mostly of medical importance as vectors of bilharzia and other parasitic diseases, and mostly from the now-defunct South African Institute for Medical Research.
Contact person Prof. Kenne de Kock kenne.dekock@nwu.ac.za

The Biobase, Freshwater Research Unit, University of Cape Town

Collated data on South African rivers up to the mid-1990s; includes most of the significant ecological studies, both published and unpublished, for which water chemistry and species data are provided; available on CD from the Water Research Commission.
Contact person Dr Helen Dallas, Helen.Dallas@uct.ac.za

The Rivers Database, Dept of Water Affairs, Pretoria

Collated data from the River Health Programme; covers much of the country; biodiversity data to family only; coverage variable across the country and over time; archived on the RHP website: http://www.dwa.gov.za/iwqs/rhp/database.htm; queries to rivers@dwa.gov.za

The River Classification Programme, Freshwater Research Unit, University of Cape Town, and the Dept of Environmental Affairs

A fairly extensive set of records of ephemeropterans, plecopterans, trichopterans and simuliids mostly extracted from Albany Museum records; on aged software and probably overtaken by the *Invertebrate Guides*; available from Prof Jenny Day, Jenny.Day@uct.ac.za

2.2.3 Species lists in the scientific literature: papers and theses

Data captured and/or transferred to the database

Records for all taxa from the 9-volume *Guides to the Freshwater Invertebrates of Southern Africa* (Day *et al.*, 1999; Day *et al.*, 2001a, b; Day & de Moor 2002a, b; de Moor *et al.*, 2003a, b; Day *et al.,* 2003; Stals & de Moor, 2007) - all captured

Records scattered in JA Day's collection of reprints of published papers on the taxonomy of southern African invertebrates: papers have been sorted and those with relevant data identified. As predicted, this task could not be completed in the time available but we will continue to add data to the checklists from time to time.

Personal records requested from various taxonomists, ecologists and consultants and are added to the checklists as they become available.

Records from undergraduate and Honours student projects, theses and personal collections, Freshwater Research Unit, UCT, have been collected and some have been entered into the database. Details are available in the files for each taxon.

Usable records in all available papers and reports on the ecology of SA rivers and wetlands (some of which will overlap with data in theses listed below) have been examined and where they represent new distributional records they have largely been included in the checklists.

3. CURATION OF THE DATABASE

It will be necessary to find a 'home' where the SAFID checklists will be properly curated and from where they can be made available to the public. The Water Research Commission (WRC) has been in talks with the South African Environmental Observation Network (SAEON) about the storage and curation of WRC data and it would seem to the authors of the present report that SAEON would indeed be a suitable host.

Other alternatives are the South African Institute for Aquatic Biodiversity (SAIAB), and/or SABIF and SAEOL within SANBI. Although responsible for aspects of the aquatic biodiversity of South Africa, SAIAB seems an unsuitable home for the SAFID database because it has traditionally been involved specifically with fishes. Furthermore, SAIAB is funded not by SANBI but by the Department of Science and Technology. Should this department's interest in SAIAB decline, the SAFID database might lose its home. The other alternative is SANBI, and in the long run that is where, in our opinion, all biodiversity data of this kind should be housed. SANBI arose, however, from the National Botanical Institute and is still biased in favour of terrestrial plant data. SABIF, which is SANBI's South African Biodiversity Information Facility, has extensive and valuable holdings of terrestrial plant data but animal data are only now being included, and slowly at that. SAEOL (the South African Encyclopaedia of Life) is still under construction. Probably this will be the most natural warehouse for SAFID when once SAEOL is functional.

In summary, SAEON is probably the best home for the SAFID database, at least in the short term. At a later date it should also be provided to the SAEOL website.

4. FURTHER WORK

As far as the database itself is concerned, we hope that the contents will be used to develop a better understanding of the biogeography of our freshwater invertebrates, particularly in the light of the effects of climate change on distribution patterns. Such analyses will also identify areas where collecting has been inadequate.

While taxonomic and systematic studies on terrestrial plants and fishes are relatively well funded (at any rate there are institutions that employ systematists in these fields), we urgently need to fund continued work on the taxonomy and systematics of freshwater invertebrates, especially of important and under-studied groups.

5 REFERENCES

NOTE that the following references were cited in this Final Report. The extensive references lists that form part of the SAFID database are to be found on the CD that houses the checklists.

Day, J.A., B.A. Stewart, I.J. de Moor & A.E. Louw (eds), 1999. *Crustacea I: Notostraca, Anostraca, Conchostraca and Cladocera.* Guides to the freshwater invertebrates of southern Africa, Vol 2. Report TT121/00, Water Research Commission, Pretoria.

Day, J.A., B.A. Stewart, I.J. de Moor & A.E. Louw (eds), 2001a. *Crustacea II: Ostracoda, Copepoda and Branchiura.* Guides to the freshwater invertebrates of southern Africa. *Vol 3*: Report TT 148/01. Water Research Commission, Pretoria.

Day, J.A., B.A. Stewart, I.J. de Moor & A.E. Louw (eds), 2001b. *Crustacea III. Bathynellacea, Amphipoda, Isopoda, Speleogriphacea, Tanaidacea and Decapoda.* Guides to the freshwater invertebrates of southern Africa. Vol 4: Report TT141/01, Water Research Commission, Pretoria.

Day, J.A. & I.J. de Moor (eds), 2002a. *Vol 5: Non-Arthropods*. Guides to the freshwater invertebrates of southern Africa. Report TT 167/02. Water Research Commission, Pretoria.

Day, J.A. & I.J. de Moor. 2002b. *Vol 6: Molluscs and Arachnids*. Guides to the freshwater invertebrates of southern Africa. Report TT182/02, Water Research Commission, Pretoria.

De Moor, I.J., J.A. Day & F.C. de Moor (eds), 2003a. *Vol 7: Insecta I. Ephemeroptera, Odonata and Plecoptera.* Guides to the freshwater invertebrates of southern Africa. Report TT 207/03, Water Research Commission, Pretoria.

De Moor, I.J., J.A. Day & F.C. de Moor (eds). 2003b. *Vol 8: Insecta II. Hemoptera, Megaloptera, Neuroptera, Trichoptera & Lepidoptera.* Guides to the freshwater invertebrates of southern Africa. Report TT 214/03, Water Research Commission, Pretoria.

Day, J.A., A.D. Harrison & I.J. de Moor (eds), 2003c. *Vol 9: Diptera.* Guides to the freshwater invertebrates of southern Africa. Report TT 201/02, Water Research Commission, Pretoria.

Harrison, A.D. 1959. *General statement on South African Hydrobiological Regions.* Report no. 1, Project 6.8H. Internal Report, National Institute for Water Research, Council for

Scientific and Industrial Research, Pretoria

Stals, R. & I,J, de Moor (eds) 2007. *Vol 10: Coleoptera.* Guides to the freshwater invertebrates of southern Africa. Report TT 320/07, Water Research Commission, Pretoria.

van Ginkel, C.E., R.P. Glen, K.D. Gordon-Gray, M. Muasya & P.P. van Deventer. 2011. *Easy identification of some South African wetland plants*. Report No.TT 479/01, Water Research Commission, Pretoria.

6 USEFUL WEBSITES

South African Biodiversity Information Facility (SABIF): http://www.sabif.ac.za/

Albany Museum, Grahamstown: http://www.ru.ac.za/albanymuseum/

SANBI (South African National Biodoversity Institute): www.sanbi.org

Iziko Natural History Museum(formerly the South African Museum), Cape Town: www.iziko.org.za/static/landing/natural-history-collections

National Museum, Bloemfontein: www.nasmus.co.za
KwaZulu-Natal Museum (formerly the Natal Museum): http:// www.nmsa.org.za

Ditsong National Museum of Natural History (formerly the Transvaal Museum): www.ditsong.org.za/naturalhistory_exhibits.htm

South African Earth Observation Network (SAEON) data portal access: www.saeon.ac.za/data-portal-access