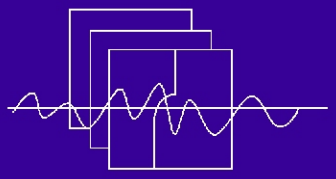
Vol 14 No 2 May 2003

SADCO SADSO



Southern African Data Centre for Oceanography
P O Box 320, Stellenbosch 7599
South Africa

Email: mgrundli@csir.co.za

Website: http://fred.csir.co.za/ematek/sadco

SADCO is sponsored by ...

Department of Environmental Affairs & Tourism
SA Navy
CSIR Environmentek
NRF (SA Universities)
Namibian Ministry for Fisheries & Marine Resources

Highlights of 2002/2003

A brief overview from the SADCO Annual Report

SADCO draws up an Annual Report every year, and this is presented to the Steering Committee at its meeting in May. After obtaining feedback from Committee, the report is finalised and bound, and a copy provided to each member (and the sponsor organisations).

To allow wider insight into the activities, the following provides a brief overview of some of the major achievements of the data centre during 2002/3.

In terms of web-enablement of the marine database, extraction routines have been completed, including a useful conversion to ODV format. (The VOS data base has been fully web-enabled in 2001, while the Current Meter and Weather Station data bases still need some final aspects to be dealt with see article on *Status of web-enablement* this issue).

A total of 5134 CTD stations were loaded. Of these, the majority was received from Namibia, but some of these need checking and editing.

The process of preparing and transferring all (existing) underway ADCP data from MCM (Marine and Coastal Management) to SADCO has been started. An inventory of the data has been established, and the first 6 (of 142) cruises submitted (but not loaded).

Off-line requests comprised 28% of the total budget (15% last year). This is a record for the first time since the request load has been monitored (1990).

The inventory was kept up to date with surveys and track charts (except for current meter and automatic weather station data).

Contact with the user community (and specifically BCLME) and promotion of the Data Centre was maintained through Newsletters and visits.

There were no system crashes in the year, and no data was lost (backups are done every day).

ISSN 1017-2361

Priority Activities for 2003/4

At the SADCO Steering Committee meeting of May every year, priority (development) activities for the following year are identified.

SADCO is experiencing considerable growth in terms of the infrastructure that is being established (e.g. web-enabling the data bases), data that is being loaded and even new data sets that are being identified and prepared for loading (e.g. the underway ADCP data from Marine and Coastal Management).

A number of activities are therefore in progress and need completion, or need to be started (and hopefully completed in this year). Some activities will obviously extend over more than one year, or only start in following years. The following list provides some insight into the activities of this year:

The **marine load program** will be completed (*this program loads data into the Oracle database. It is the last component of the web-enablement of the Marine database, and also the last item that links SADCO to its previous Database Management System, Informix.*)

The **ADCP database** system needs to be defined, and loading, extraction and product software created *(this is required for the impending storage and retrieval of the Acoustic Doppler Current Profile data from MCM)*

The **VOS data** set needs to be cleaned up (*the data from Voluntary Observing Ships seems to contain anomalous observations, that need to be identified and edited*)

Data loading:

- *Backlog of "underway" ADCP data (MCM have more than 140 cruises where ADCP data has been collected) needs to be loaded. It is planned that somebody will be appointed at MCM to assist with the processing and submission of the data, and this process will extend into 2004.*
- *The loading of backlog weather station data (MCM) should be completed*

Internet access to the current meter database should be established (*although the database has been web-enabled, it is still not accessible to on-line users*)

Internet access to weather station database should be established

Current meter and weather station entries need to be added to the inventory (*although a large amount of Current Meter and Automatic Weather Station has been loaded onto SADCO, the Inventory has not been updated with these entries yet*)

Discrete CTD data loading (MCM)

Some activities are provisionally foreseen to fall outside the 2003/4 budget, and may have to stand over for the following year.

Loading WOD2001 data (*we have kindly received an updated version of the hydrographic data holdings from WDC, and should consider whether/when this should be loaded*)

Finding and loading XBT data (from ships of opportunity)

Moored ADCP data (*these are ADCP moorings deployed by MCM, but it has to be clarified first whether/when this data will be made available to SADCO*)

Continuous coastal temperature data (*MCM have sites around the southern African coast where sea surface temperature is continuously monitored. It should be investigated whether this data should be stored by SADCO*)

In summary it seems that a rather full year is foreseen for SADCO.

The SADCO Team

We thought that we will show (remind) you what the SADCO "team" looks like, and what they do.

The "SADCO team"

Seated L t r: Sharifa Engel, Ursula von St Ange.
Standing: Mario August, Steven Pietersen, Marten Gründlingh.
Insets: Louise Watt (left) and Tammy Morris (right).

As a reminder: *SADCO does not employ anybody full time*. All the staff that provide a service to SADCO do so for only a small part of their year (roughly varying between 5 and 30%). This has been a significant aspect of SADCO's ability to provide a cost-efficient service over the past 13 years (i.e. having full-time access to suitably qualified staff, but only funding actual time spent on the data centre activities).

The roles of the team members are:

Marten Gründlingh
Management, including user liaison, requests, newsletters, reporting, scouting
Ursula von St Ange and Mario August
Programming and data base design and maintenance
Louise Watt and Steven Pietersen
Data handling, editing, loading, extraction
Sharifa Engel
Secretarial, mailing list
Tammy Morris
Data processing of MCM data (and located at MCM)
Magdel vd Merwe
Newsletter layout

Sharifa forms part of the CSIR infrastructural support (not paid by SADCO), which also includes Franco Oosthuizen (finances) and Diane Grobler (computing hardware and network)

The "SADCO team" also includes staff at the sponsor organisations who spend time and effort to prepare and check data, or assist with transfer to the data centre.

Last but not least, the members of the Steering Committee help to guide, support and encourage SADCO, and are also part of the team.

Status of web-enablement

The web-enablement of the data bases maintained by SADCO has been a major undertaking, but forms part of SADCO's philosophy of having a modern, albeit small, data management system. The arrival of web-based systems, allowing direct access via the web, was a natural development in the data management domain over the past years, and SADCO needed to keep up with technology or run the risk of falling behind (and becoming obsolete).

For interest sake, SADCO has moved over various platforms over the years. In the 1980s, they keyword for all organisations' computing requirements for data management was “mainframe”. The readers may recall that, when visiting another institute, they were soon confronted with two questions: (a) whether you would like to have some coffee, and (b) whether you would like to see the MAINFRAME.

Much of this changed with the advent of smaller platforms, and SADCO took the bold step (almost unheard of at the time) in 1994 to move its entire operation from a mainframe 1500 km away onto a PC (albeit a jacked-up version) 10 m away and onto a user-friendly Database Management System (Informix). The ease of access, and control over hardware, software upgrades, inspections, etc was a remarkable improvement, and led to a significant reduction in real computing costs.

At the end of the 1990s, SADCO prepared for, and started, the process of web-enablement. It was planned that for each of the databases in the Data Centre, the following steps would be taken

a. **Transfer** of data from Informix to Oracle
b. Create programs to handle data **loading** and access
c. Create programs to handle data **extractions**
d. Create programs to handle **products**

The progress achieved thus far, is:

The **VOS database** (weather and sea state observations) was completely web- enabled at the end of 2001.

The **Marine database** (hydrographic stations) web-enablement has not been completed yet. After completing Tasks a) and d) in 2001, Task c) was completed in 2002. Additional functionality has been added in the extraction programme to convert data into the ODV (Ocean Data View) format. The loading program is quite complex and still needs some work, and will be completed early in 2003/4.

After the loading programme (b) is completed the final transfer of the database from Informix to Oracle will also occur.

It is now e.g. possible to locate and extract cruises on-line, reformat the data (e.g. into ODV), create products and graphs. Text products can be viewed instantaneously on screen, while the HPGL graphs can be downloaded for local printing or easy insertion into a word processor file. The system will nevertheless be improved as its use is expanded.

Current meter database. Apart from a few deployments, all the current meter data submitted to SADCO (CSIR and MCM data), as well as the data from current meters deployed by overseas researchers, was loaded during 2001. The web-enablement of the database has been done during 2002, but is still not fully accessible via internet (foreseen for 2003).

Weather station database. Similar to the current database (above) the weather database has been web-enabled but is not accessible on-line yet. Similar alterations are planned to establish full accessibility via SADCO (2003).

Additional effort was spent on aspects of web maintenance, as well as to install a time-out procedure on the web access to improve the security of the system.

It is foreseen that the majority of the web-enablement activities will be finalised this year, although loading backlog data will continue for some time.

BCLME Workshop in Swakopmund

The Benguela Current Large Marine Ecosystem (BCLME) is quite unique in the Southern African context. It is the first time that such a geographically large, multidisciplinary and multi-national experiment has been tackled for the Benguela Current (or for any current off southern Africa). The requirements for the successful execution of the project are challenging. Not least of all is the requirement for efficient data management, which will also ensure that the legacy in terms of data collected during the experiment, is maintained.

(a) Cape Town

SADCO has been invited to address meetings of the BCLME on two occasions: During the first one, in the Breakwater Lodge in Cape Town, November 2002, SADCO provided some insight into its data holdings, capabilities and present/planned developments.

Feedback from the meeting was discussed by the SADCO Steering Committee after the BCLME meeting, and the decisions that were taken were

> SADCO would be prepared and keen to undertake the data management for BCLME
>
> SADCO could expand its data handling expertise in the field of **plankton data**, by establishing a plankton database.

For such a work load, additional funding would be required above the present budget of the Data Centre.

(b) Swakopmund

The SADCO manager was again invited to attend a discussion of BCLME, specifically around the data needs of Angola. The meeting took place in Swakopmund, in April 2003.

The opportunity was used to give an indication of the amount of data that SADCO has on its data bases, or to which it has access, to participate in discussions about data needs, and obtain some insight into data inventory in Angola.

Recommendations of the meeting included an identification of the role of SADCO in the planned data management, but this still has to be formulated and agreed upon.

Organising and executing a multinational experiment like this, has considerable logistic and other challenges. Some issues that come to mind, are the distance (making physical interaction between participants costly and difficult), and infrastructure (what may be the accepted norm in one country, may not be the same in other countries). There is nevertheless a significant degree of mutual support between the countries, and between the various groups/organisations participating in the BCLME.

BCLME is planning to start a web site and newsletter, which will be used to provide information to the marine community.

Data Collector of Yesteryear

All organisations have individuals that have left some sort of mark on activities of the organisations. In terms of oceanographic data, I often recognise names and data sets where an individual's influence is still visible. For the period of the mid-sixties to the mid-seventies of oceanography in the CSIR off the KZN coast, one of those individuals was Alan Pearce. (It may be mentioned that SADCO plans to load current velocities that were measured by Alan about 30 years ago, this year.)

Alan is a mechanical engineer by training, and was located at the CSIR in Durban in the 1960's before leaving for Australia in 1977.

In his initial involvement with marine matters, he studied the flow in pipelines and submerged jets and nozzles, did tank model studies on this topic and received a masters degree in engineering (fluid mechanics) in 1965. Many of the formulas and concepts used today in marine effluent disposal are still recognisable in his thesis.

Alan's joining the physical oceanography group coincided with the acquisition of the RV *Meiring Naudé*, a state-of-the-art research vessel and which opened the way for a large number of projects. Of the data Alan collected on the KZN coast, the experiments off Richards Bay (in preparation for the development of the port of Richards Bay), Durban and Port Edward (to investigate the characteristics of the Agulhas Current, and the meteorological forcing of the Current), remain high quality, insightful baseline sets that are still valid today. This data was used for his second M Sc (1977).

In the middle seventies, Alan became the Head of the Oceanography Division of the National Research Institute for Oceanology of the CSIR.

This was a time in South Africa when imported scientific items were very exceptional, and in most cases equipment had to be developed inhouse. Understandably, innovation was the order of the day. Alan participated in the establishment of accurate navigational technology for the RV *Meiring Naudé*, and suitable processing methods for surface and underwater data. The meticulousness with which individual measurements were quality controlled, would be unheard of today. This was quite necessary given that the research was conducted in a rather foreign ocean environment, and because oceanography in South Africa was in its infancy and many of the data handling routines and systems still had to be developed.

In a career spanning about 40 years, Alan is still active in oceanography in Australia, having moved to Perth and more into providing satellite data support for the fishing industry off the coast of Western Australia. His emigration 25 years ago left a distinct gap in the oceanographic community of KZN, and one often wonders how things would have turned out if he had continued his role in the local oceanography.

Alan with his wife, Valerie, and two children Steven and Catherine (see also Alan's poetry on the back page)

What's all this fuss about historic (environmental) data? (or, Data Management 101)

Some time ago, somebody asked me why we should look after old data, and also why we seem to "discover" historic data sets all the time. The impression has namely been created that, for every data set we know about and have a copy on a database, there is another data set that is "still out there" (in a drawer or local PC).

These questions seem to be somewhat philosophical, and also touch on the issue of why we collect data at all. **Don't we, or won't we, get to a point where we have sufficient data to answer all existing (and future) questions?**

The answer seems to be No, but why not?

Environmental data has been collected tacitly since prehistoric times on this planet, and used by early man to determine the hospitability of the environment (safety and security), the characteristic and location of food, etc.

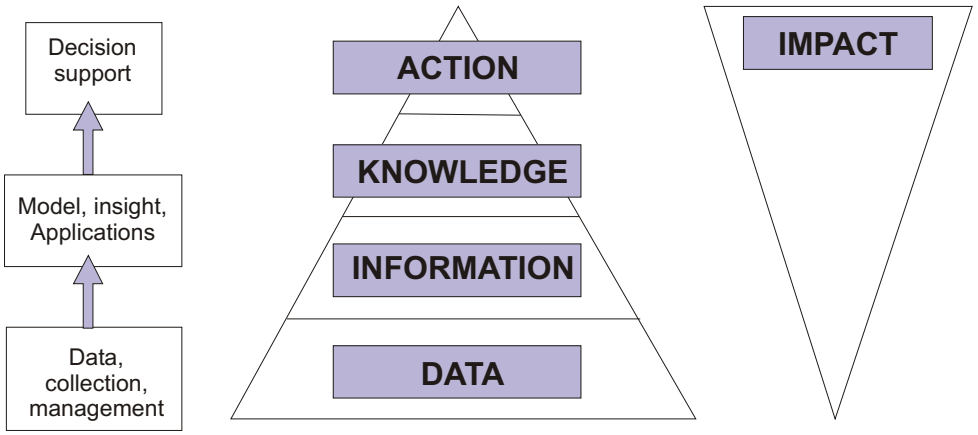
Much, much later (in more recent times), mankind started moving from being not only a user or beneficiary of the environment to also becoming a custodian or benefactor of the environment. This is (not only) due to man's wider sense of responsibility, but most probably also to preserve his own life and lifestyle (as manifested by the move toward higher levels on Mazlow's expectation pyramid). In the present era, man seems to be living out a very fine balance between these two roles (beneficiary and benefactor), only using the environment in a sustainable mode.

For both of these roles, sound decisions (for the environment as well as mankind) are based upon our growing understanding of the function of, and interaction between environmental entities. In turn, this understanding relies on the scientific process of hypotheses and testing, and this "testing" requires observations.
Why do we need such large amounts of data?

The enclosed diagram portrays the value chain of data and information. A large amount of **data** is needed to produce a smaller amount of **information**, which in turn creates an even smaller amount of **knowledge**, etc. In the end, a limited number of actions (this can take the form of a policy) can be taken, based originally on a large amount of data.

While the quantifiable magnitude of the value chain reduces upward, the corresponding **impact** increases.

So, to have the maximum impact, large amounts of data (which probably have negligible impact) first have to be converted into smaller amounts of knowledge and a limited number of actions, for the latter to maximise the impact.



What's all this fuss about historic (environmental) data? (or, Data Management 101)

(Continued)

This is fine, but **why don't we just measure more rapidly, thus shortening the time required?**

The problem is that a lot of the applications of data are based on real (= clock) time. E.g., marine engineers need to provide design conditions for maritime structures, and these specifications are based on the probability of occurrence of environmental extremes in real time. For these calculations, as long a time series as possible is required. Similarly, if climate variations are studied, the time series of valid data needs to be as long as possible (in real time) to ensure that valid conclusions are drawn.

So, in order to predict the probability of a given circumstance occurring further into the future, we often need to be able to look further into the past.

In a southern African concept, the historical marine record is not very long (here we are excluding the historical record in the fields of geology, palaeontology or geophysics, where information is available on the movement of continents, reversal of the earth's magnetic field, fossils, etc). In comparison, SADCO has data only going back to the middle of the 19th century!

To be continued

Ode to an Eddy

By Alan F Pearce, 1975

The mighty stream Agulhas
Still wends its wand'ring way
Down the coast of Afrik
And on past Durban Bay

Twas there we found *The Eddy*,
Majestic and serene,
Cyclonically rotating:
The best we've ever seen.
It's heart was cool yet shallow
It's motion slow and steady
It's structure well-defined
Ah, yes, *that* was an eddy!

We know not of its birthplace,
Nor when it first saw day,
But Grundling has a theory:
They're born in Richards Bay.
The coastline at St. Lucia
Deflects the Mighty Flow,
And southwards in its wake
The eddies form and grow.

Yet that's not *all* the answer,
For other data shows
The inshore flow is governed
By the weather-bringing "lows".
These cells of falling pressure
Arrive from west-sou-west
And bring the dreaded "busters"
Which the mariners detest.

And years of observation
Along the eastern coast
Suggest that eddies form
When pressure's down the most.
It therefore seems quite likely
That instead of from the *north*
It's really from the *south*
That the eddies issue forth.

Alas, our single section
(repeat occupation)
did not show the direction
of eddy propagation.
Is it to north or south
The eddies set their face?
Or maybe they are born
And die in the same place?

We're planning now a programme
To show them in their glory.
We'll seek them find them -
track them -
Until we know their story.

Published in Nature 258,(5535), p 486

Layout & Design: Magdel van der Merwe
CSIR