Information management

G8

Large quantities of data are generated in an MPA, from records of visitor numbers and patrols, to the results of monitoring and research. It is essential that this is organised, stored, analysed and made available appropriately to provide the information that is needed for management. This sheet provides advice on how this can be achieved, including basic principles for setting up databases.

The data collected within an MPA provide a vital source of information by which the effectiveness of the MPA can be assessed and the best management decisions taken. Data entry, storage, analysis and write up takes as much time as data collection, and this is often overlooked in work plans. As a result, data are often never used and monitoring programmes fail to be useful because data are lost or never analysed. A good information management system can help to overcome this. Types of data may include:

- Textual or qualitative - e.g. words, sentences;
- Graphical - e.g. map, photo;
- Numerical or quantitative - e.g. areas, units, a ranking score.

An information management system may involve electronic and/or hard copy files. Software for electronic systems include Excel-type spreadsheets or database programmes such as MS Access. Spreadsheets are easy to set up and view and analysis is easy on a one-off basis, but the regular analysis and reporting required in most MPAs generally requires a more rigid system. Furthermore, as the quantity of data increases, so do the limitations of spreadsheets, and they are less secure as automated tasks and values in cells can be changed by any user.

Electronic, packaged databases are preferable because they can record changes over time more easily, take up less space and can be duplicated, and allow for efficient, accurate data entry and retrieval, safe storage, and greater accessibility. Relational database programmes such as Oracle, Microsoft SQL Server, or Microsoft Access for PC users, or FileMaker Pro for MAC users, are particularly efficient and powerful. These store data by dividing the information up into tables containing different fields. Queries can be set up to do analytical tasks in a consistent way, and standardised reports generated. For example, a query can be written to ask how many people visited the MPA over a period and the results can be printed out as a report in a format designed by the user.

Databases must be kept up-to-date which requires good maintenance, especially as software is regularly revised, and it is best if someone is made responsible for this. The following procedures are involved in management of electronic data:

DATA COLLECTION

Agree on the terms, format and abbreviations before data are collected (i.e. create a data dictionary), and use them consistently; always indicate measurement units, and be clear about how dates are to be recorded. Maintain a logbook as a back up. Fill in all fields on data sheets to show that no data are missing and note any problems or irregularities. Transcribe data on to clean datasheets after returning from the field if necessary, and make photocopies so that the originals can be stored.

DESIGNING THE DATABASE

This should be done jointly by the staff responsible for the monitoring, research or management programmes and those responsible for information technology. An external advisor or consultant in database design is invariably necessary. Close links, either in-house or with partners, should be developed between the database programmer, scientists with experience in analysis, and managers who know the questions to be answered.

A management-oriented database must have data entry, verification and analysis pages designed for easy use by non-specialist staff. Focus on what is relevant or essential for the analysis so that the required outputs are obtained. Numerical data fields are preferable for analysis; comments can be added in text fields.

Other people or institutions may need access to summarised data, and the database may need to be compatible or harmonised with international or regional databases, such as COREMO II and III (developed by the Indian Ocean Commission), CORDIO, the WIO Fisheries

Database (developed with assistance from IUCN), ReefBase and FishBase). CORDIO has developed databases for coral reef, fisheries catch and socio-economic monitoring data which, as open-source databases in MS Access, can be continually upgraded and improved.

A database may need to include a security protocol, and this is already available in most good software packages. While it is advisable to allow only certain individuals to enter or make changes to the data and structure, some form of access is essential for those responsible for analysis and preparing reports.

DATABASE DOCUMENTATION

This is best done by creating word-processing files that describe how and why the data were collected (including any known problems or data gaps) and the data within each table, and any analyses performed.

DATA ENTRY

A key aspect of data entry is quality control. The following procedures are recommended:

- Enter data as soon as possible after collection; it is best if the data collector does this or at least is available for consultation;
- Enter raw data. These can be aggregated later to produce summaries (e.g. daily averages, site totals), but it is generally impossible to extract raw data from a summary;
- Be consistent, as abbreviations, misspellings and data entered in a different format will not be recognised and risk being lost.

Customised data-entry forms assist by:

- allowing (or requiring) users to select entries from a list (e.g. species, fishing gears) which makes data entry quicker and ensures that the same terms are used every time;
- standardising formats (e.g. the user has to enter dates as dd-mm-yy) and preventing entry of text into numerical fields;
- automatically filling in data fields from entries made in other fields, which speeds up data entry and provides additional checks.

Where data-entry is done by someone who may not know whether the data are correct or not, validation rules can be set up that indicate values that are unusually low or high and need checking, and that certain fields are filled before the user can move on.

DATA VERIFICATION

Summary analyses of data should be carried out regularly to check that the data being collected are what is required and that data entry is accurate and complete. The queries tool can be used for this by, for example:

- Counting the records to check that there are enough for statistical validity;
- Looking for data entry errors, e.g. unusually high or low values;
- Counting the records to see if they match the number of data sheets.

DATA ARCHIVING

Data must be archived for future users, and backed up in case of damage or loss. Back-ups are short-term copies of current work. An archive remains in storage as a record of a database at a particular time, and should be conducted regularly, but perhaps only monthly or 6 monthly. A backup is done much more frequently (e.g. weekly) and the new back-up is written over the old one. An archive should include the final version of all database files and data document files in a clearly labelled folder, saved on a CD-R (non re-writeable) and stored in a central office, with the original raw datasheets and the print out of the data tables. The version on the computer, as well as photocopies of the data sheets, are back-ups and should be labelled accordingly and stored separately.

KEY POINTS FOR THE MPA

- ❑ **If possible, appoint a data manager and provide appropriate training. Alternatively, ensure that all personnel are trained in data collection, management and analysis for the areas for which they are responsible.**
- ❑ **Make sure that data management and the equipment needed are put into the budget for the MPA.**
- ❑ **Ensure that information on visitor numbers, ticket revenues, sightings of rare or endemic species, and other management or occasional events is collected and stored, as well as data from monitoring programmes and research.**
- ❑ **Maintain close contact with other programmes in the region to exchange experiences, and help in the development and maintenance of a database, especially if funds and staff capacity are low.**

Sources of further information

Bainbridge, S.J. & Baker, V.J. 1994. Database design and operation. pp. 313-328. *In*: English, S. Wilkinson, C. & Baker, V.J. (eds.) *Survey Manual for Tropical Marine Resources.* ASEAN-AIMS Project, Townsville. Queensland 4810, Australia.

Pomeroy, R.S., Parks, J.E. & Watson, L.M. 2003. *How is your MPA doing? A Guidebook on Natural and Social Indicators for Evaluating Marine Protected Area Management Effectiveness.* IUCN, Gland, Switzerland and Cambridge, UK. xv + 230pp. (section 3-3 on 'Manage Collected Data')

Samoilys, M. & Giles, B. 2003. *Data management. Project Seahorse Technical report No. 5.* Version. 1.0. Project Seahorse, Fisheries Centre, University of British Columbia. 8pp. **www.projectseahorse.org**

Samoilys, M.A. (ed.) 1997. *Manual for assessing fish stocks on Pacific coral reefs*. Queensland Department of Primary Industries, Brisbane. 78pp.

Wilkinson, C., *et al*. 2003. *Monitoring Coral Reef Marine Protected Areas: a practical guide on how monitoring can support effective management of MPAs.* Australian Institute of Marine Sciences and IUCN, Switzerland. 68pp.

Reefbase: **www.reefbase.org**

Fishbase: **www.fishbase.org**

WIO Fisheries Database **www.wiofish.org**

CORDIO **www.cordio.org**