South East Water Quality Project
Water Quality Data Collection and Management Protocols

Prepared For: SEWQ Project Management Committee

Prepared By: WBM Oceanics Australia

Offices

Brisbane
Denver
Karratha
Melbourne
Morwell
Newcastle
Sydney
Vancouver

DOCUMENT CONTROL SHEET

WBM Oceanics Australia Newcastle Office: 126 Belford Street BROADMEADOW NSW 2292 Australia PO Box 266 Broadmeadow NSW 2292 Telephone (02) 4940 8882 Facsimile (02) 4940 8887 www.wbmpl.com.au ACN 010 830 421	***Document:***	R.N0938.002.01.protocols.doc
	Title:	SEWQP: Water Quality Data Collection and Management Protocols
	Project Manager:	Philip Haines
	Author:	David Wainwright/Mark Wainwright/Philip Haines
	Client:	SEWQ Project Management Committee
	Client Contact:	Debby Lenson
	Client Reference:	
	Synopsis:	This report documents the work undertaken by WBM Oceanics Australia in reviewing, auditing, modifying and extending the southeast water quality database.

REVISION/CHECKING HISTORY

REVISION NUMBER	REVISION DESCRIPTION	DATE	CHECKED BY		ISSUED BY	
0	Draft	26/11/04	PEH/MEW		DJW	
1	Final	26/9/05	PEH		PEH	

DISTRIBUTION

DESTINATION	REVISION										
	0	1	2	3	4	5	6	7	8	9	10
SEWQ Project Management Committee	1-Elec	4									
WBM File		1									
WBM Library											

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

1.1 Background

This report documents the initial work of WBM Oceanics Australia in assisting with the review and modification of the South East Water Quality (SEWQ) Database and providing recommendations regarding the protocols and management of water quality data for the SEWQ Project.

1.2 Protocols

Sections 2 and 3 of the report detail protocols for the establishment of water quality monitoring sites and protocols for water quality data collection, respectively. Through necessity, they are generic in nature and are essentially a summary of the water quality monitoring approach outlined in the joint Australian and New Zealand Environment Conservation Council (ANZECC) and Agriculture and Resource Management Council of Australia and New Zealand (ARMCANZ) guideline document *Australian Guidelines for Water Quality Monitoring and Reporting.*

More specific recommendations associated with these protocols will be developed following analysis of the data contained within the database.

1.3 Database

Section 4 includes a description of modifications made to the original database provided by the SEWQ committee. Section 5 describes the way in which data requests are generated from, and the received data entered into, the database.

2 PROTOCOL FOR ESTABLISHING MONITORING SITES

2.1 Overview

The selection of suitable monitoring site locations will be significantly dependent on the objectives of the monitoring program or study. In addition to the overlying objectives, there will be a number of other issues that will require consideration across different spatial scales. Key regional, catchment and site considerations for establishing monitoring sites are summarised in the following sections.

2.2 Regional considerations

- Select sites that provide appropriate spatial data to enable comparison of the measured parameter results with the objectives. Typically if a specific issue is identified, the location of the monitoring sites will largely be determined by the location of the issue. If the objectives require broader reporting on a particular parameter across a region, then the sites would be more distributed.
- Consider the location of existing or planned infrastructure that can be utilised to support sampling (e.g. weirs, stream gauges, rainfall stations).
- Existing sites where water quality monitoring data have previously been gathered would assist with providing historical data for comparative purposes.

2.3 Catchment considerations

- Sites for monitoring ecological impacts are typically assessed at a specific site/s where potential impacts from localised disturbances have occurred.
- Sites to assess physical, chemical and biological parameters are typically established following pilot studies to confirm suitable locations.
- Determine whether water quality is only to be monitored along major streams and receiving waters, or also along minor tributaries connected to these waters.
- Locate downstream and upstream of inflows from areas with differing land use characteristics (e.g. urban areas, sewage treatment plants, DEC licenced premises, industrial sites, stream confluences, forested/agricultural land).
- Locate near catchment 'hot spots'.
- Sites should not be spaced too closely to avoid invalidating statistical assumptions that sites are independent (where this is relevant to the objectives and parameter under consideration).
- The smallest differences or changes that must be detected determine the number of sampling stations, spatial distribution of the stations and temporal distribution of the sampling.
- Groundwater quality is affected by the local geology and by pollution from surface land use; sufficient sampling stations are needed to account for the variability that these factors can introduce.

2.4 Site considerations

- Access to the site should be feasible and safe during all weather conditions. If vehicles are required to transport monitoring equipment, the access should be designed for vehicle loading and the alignment of the access planned accordingly.
- Monitoring sites should be positioned away from highly accessible and visible areas to minimise the potential for unrepresentative impacts on water quality caused by human activities in a localised areas surrounding the monitoring station (i.e. locate sites away from roads, bridges and jetties).
- Where it is important that free flowing water is sampled, sites should also be positioned away from the hydraulic influence of structures that alter the flow and chemical conditions of the water in the stream (i.e. weirs, culverts and bridges).
- Where equipment is to be located on-site to support monitoring, the sites selected should ideally not be highly visible to minimise the potential for vandalism. Where the utilisation of a highly visible site is important to achieving the monitoring objectives, appropriate security measures should be implemented to protect the monitoring equipment from vandalism. This may include housing the equipment within lockable steel cabinets and protecting sensors from damage.
- Sites should be readily identifiable and locatable. GPS will assist with this task, particularly for offshore monitoring stations.

3 PROTOCOL FOR WATER QUALITY DATA COLLECTION

3.1 Planning for Water Quality Data Collection

Planning of a water quality data collection program identifies what the specific data requirements are, and outlines whether the parameters are to be measured in the field or in the laboratory. Planning is required to ensure that the sampling falls within the budget, without compromising to a detrimental extent on the statistical power of the findings. The water quality data collection program outlines the protocol to be followed in sample collection, preservation, storage and transportation.

Specific sampling protocols will be required for each water quality parameter monitored against the established objectives. The specific protocols would nominate the sample collection device, type of storage container, preservation procedures and types/numbers of quality control samples to be taken.

3.2 Selection of Water Quality Parameters

The water quality parameters selected will depend on the environmental values of the water body, and therefore on the objectives of the study. The water quality parameters can be separated into physical, chemical and biological parameters. Physical and chemical parameters essentially reflect the causes, and biological parameters, the effects. Normally these collectively will describe the state or health of the waterbody to be monitored. The attributes listed in **Table 3-1** should be considered.

Table 3-1 Checklist for selection of measurement parameters (ANZECC, 2000)

Relevance	Does the measurement parameter reflect directly on the issue of concern?
Validity	Does the measurement parameter respond to changes in the environment and have some explanatory power?
Diagnostic value	The measurement parameter must be able to detect changes and trends in conditions for the specified period. Can the amount of change be assessed quantitatively or qualitatively?
Responsiveness	Does the measurement parameter detect changes early enough to permit a management response, and will it reflect changes due to the manipulation by management?
Reliability	The measurement parameter should be measurable in a reliable, reproducible and cost-effective way.
Appropriateness	Is the measurement parameter appropriate for the time and spatial scales of the study?

Physical parameters include flow, temperature, electrical conductivity, suspended solids, turbidity, and colour. Chemical parameters include a range of general measurements such as pH, alkalinity, hardness, salinity, biochemical oxygen demand, dissolved and total organic carbon. These are generally considered routine measurements in most investigations. In addition, other major controls on water chemistry include main anions and cations, and nutrient species (phosphate, nitrate, nitrite, ammonia, silica). These, together with the physical measurement parameters, will determine the stability, chemical forms and bioavailability of a range of minor and trace contaminants or toxicants such as metals, metalloids, and specific organic compounds.

Biological measurement parameters consist of ecotoxicological and ecological measurements and is non-specific in that it responds to the sum of the contaminants in the system. Taxonomic groups that have been used or proposed as indicators of ecosystem health include macroinvertebrates, macrophytes, fish, algae, bacteria, protozoa and fungi. Some taxa are particularly susceptible to certain contaminants and so provide a sensitive tool for early warning, via bioassays. Biomarkers (measurable levels of biochemical products in an organism's body) can show that it has been exposed to a contaminant.

A whole-environment approach can be used, focusing on biological monitoring (effects) with measurements of physical and chemical data (causes) to aid interpretation of the biological data. This applies especially when the objective is to protect aquatic ecosystems or to assess ecosystem health or some change from a reference condition.

Ecological assessment aims primarily to measure the structure and function of biological communities. It principally involves field-based measurements of the abundance and diversity of species, community structure and function, and how these are altered as a consequence of known or unknown stressors and their modifiers in both waters and sediments. Macroinvertebrates, fish, and algae appear to be actually or potentially the most useful groups for the bioassessment of water quality. Bacteria, protozoa and fungi have not been widely used in ecosystem health studies, but bacteria and protozoa have been used extensively to test that waters are safe for human use. Before choosing a particular taxonomic group as a measurement parameter of water quality or ecosystem health, it should be checked that the taxonomic group fulfils the following criteria:

- the response measured reflects the ecological condition or integrity of the site, catchment or region to be monitored;
- approaches to sampling and data analysis can be highly standardised;
- the response can be measured rapidly, cheaply and reliably; and
- the response has some diagnostic value.

3.3 Flow measurement

If concentration measurements are being used to calculate loads, it will be important to decide how to relate flow and concentrations, and on what time basis. In the majority of Australian rivers, most (70–90%) of the annual flow and constituents are discharged under high flow or event conditions even though these may prevail for only 1–10% of the time. Under these conditions, the dominant water quality processes are the transport and deposition of discharged material during the flow event, followed by in-stream remobilisation of deposited material in the 10–30 days following the event.

To solve the difficulty of sampling at all flow regimes, a range of robust and reliable automatic sampling devices can now be obtained. Auto samplers can be used to trigger sampling when a pre-determined condition is reached, e.g. a flow level in a stream.

- Size of the catchment and storms to be sampled will determine how many sample bottles are required.

- Refrigeration of samples may be required for a number of parameters as significant changes in the samples can occur over short periods of time. Ideally samples should be preserved soon after sampling has occurred.

Automated stream gauges can also be utilised to assess flow rates and volumes, and are particularly useful for confirming flow volumes estimated utilising auto samplers. Ideally stream gauges should be sited where-:

- The site is accessible both during dry and wet weather;
- It is upstream of a hydraulic control where a suitable rating curve can be derived (e.g. weir); and
- The geomorphology of the stream should be stable and relatively straight.

3.4 Sampling approaches

In stratified random sampling, different numbers of samples can be taken in each layer or section of the waterbody, in proportion to the variance of that stratum or its size.

In systematic sampling, samples are collected at regular spacings or intervals of time.

Only rarely will sampling be random, but when it is, the number of sites and the extent of homogeneous areas in which they may be located can be determined from a pilot study. Multivariate classification procedures can be used for grouping sites, to define homogenous areas.

3.5 Sampling Frequency and Timing

If a water quality parameter has a predictable pattern (e.g. deoxygenation during thermal stratification, or migration at a certain time of day), the monitoring program must sample this parameter at a frequency that suits this pattern. For sampling in tidal waters, tidal charts should be referenced to ensure that samples are collected at appropriate stages necessary for comparing the testing results against objectives established for the site.

If a water quality parameter has a highly variable pattern (e.g. stormwater discharges), the monitoring program must be flexible enough to enable sampling over variable time scales.

When the monitoring program aims to compare test data against particular guidelines or standards, it is important to sample very often to note the possibly brief occasions when the guideline is exceeded.

3.6 Cost Effectiveness

Most water quality monitoring programs will require some form of compromise to remain cost effective. It is preferable for the cost of sampling programs to be as small as possible while still meeting the stated objectives of the monitoring study. Cost-effectiveness considerations involve trade-offs between loss of statistical 'power' (i.e. the capacity of a program to discriminate between various hypotheses) and the cost of data acquisition. It is necessary to determine all the resources and associated costs required, thereby ensuring the study can be carried out.

Costs of data acquisition are determined by:

- the number of sampling stations, sampling occasions and replicates;

- the cost of collecting samples (staff, transport, consumables);
- the cost of analysis;
- the cost of data handling, and
- the cost of analysis and reporting.

Cost-savings can result from collaborative monitoring, for example when local councils pool resources with other water managers to comprehensively monitor a particular waterbody.

3.7 Sampling Methods

Selection of a sampling method, whether for biological or physical or chemical parameters, should be guided by the objectives, the local conditions, the safety of the field staff, the acceptability of the method, and commonsense.

Sampling protocols should be followed carefully. An alert person or team wearing plastic disposable gloves can avoid sample contamination by using plenty of plastic sheet to cover work areas and wrap equipment. Care is required with sampling to prevent contamination of the samples by dust, powder, skin and hair. Care is particularly required when boats and helicopters are used to assist with sampling to ensure that these forms of transport do not lead to contamination of the samples.

Basic precautions for avoiding contamination should be considered: e.g. use decontaminated containers to store reagents for use in the field; pre-clean all field equipment to the same standard as the containers; if containers were filled with water as part of the preparation protocol, empty them well away from and downstream of the sampling location before rinsing them with the sample and refilling.

The sample collection methods are selected based on the objectives of the monitoring program, parameter to be measured and the sampling frequency and timing. The methods include hand sampling, auto sampling, integrated sampling, automated real time sampling, field measurements, remote sensing and field observation.

Measurement parameters can vary from place to place within a site, randomly or in strata. When measurement parameters are being sampled in the water column, it is sometimes assumed that the water is well mixed and that a mid-water or mid-stream sample will be sufficiently representative, although this will not always be the case. Even if the monitoring goal is just to measure the average concentration of a chemical in the water at a site, the sampling process must be planned so that the within-site variation is included in the estimate.

It is important to recognise that stratification in the measurement parameter will affect the data being obtained. There are three options for dealing with such strata:

- restrict the scope of the inference to a particular stratum; for example, sample only one type of substrate but make the stakeholders aware that the inferences drawn are applicable only to that sort of substrate and cannot be generalised to unsampled strata within the sites; or
- divide the sampling effort among the strata; here the goal is to estimate the value of the measurement parameter for each site as a whole rather than for a stratum within the site; or
- make separate estimates for each stratum (if this is consistent with the study objectives).

Groundwater quality is almost always stratified vertically, and there can also be significant lateral variation in quality. There is much less dispersion of contaminants in groundwater than in surface waters, and so its natural spatial variability is potentially much greater than in surface waters.

Observations and characteristics of the site should also be recorded at the time of sampling on a standardised field sheet.

3.8 Sampling equipment

The equipment used for sampling of surface waters includes bottle samplers, pumping systems, depth samplers, automatic samplers and integrated samplers. Groundwater sampling utilises a number of pumping methods. All methods and equipments are required to meet relevant Australian or ISO standards.

The sample containers are selected to minimise the potential for adsorption or contamination of the sample. Glass and plastic containers each have potential limitations and methods to consider for preparing the containers for sampling. Refrigeration or preservation may be required to prevent changes to the sample during the period between sampling and analysis. Preferred sample containers include-:

- For sampling metals – bottles made of fluorocarbon polymers, PTFE (Teflon) or FEP, and high density polyethylene. High quality bottles are recommended, e.g. Nalgene, because these have good closures that prevent sample leakage. For Selenium, bottles made of polycarbonate and some types of polyethylene are not suitable.
- For sampling nutrients – bottles made of low or high density polyethylene are the most favoured type. Glass is not favoured because there can be high concentrations of trace metals in the glass and it has the potential to adsorb ions, e.g. phosphate.

The sampling device should not significantly disturb the environment being sampled or alter the samples taken; if it does, the samples will not reflect the actual conditions. This is a particular problem in sediment and groundwater sampling. Contaminant-enriched fine particles of sediments can be lost as grab samplers are pulled to the surface, for example; and the redox state of the sediments (oxic or anoxic) can change irreversibly on contact with air. Also, the construction of a bore and the effects of sampling procedures inevitably disturb the environment from which groundwater samples are taken.

If the sampling device comes into contact with media other than the sample of interest, sampling errors may be caused. For example, to collect sub-surface water samples for hydrocarbon analysis, the device must enter the water closed or it will pick up hydrocarbons from the water surface microlayer; and when sampling shallow water, the device should not stir up bottom sediment.

The sampling of waters for trace and ultratrace contaminants is increasingly a requirement for monitoring studies, especially for comparison with regulatory standards. Non-contaminating equipment is essential for these measurement parameters, and it should be cleaned with acids for sampling metals or cleaned with detergents and solvents for sampling organic compounds.

Before going into the field, sampling staff should check with their analytical laboratory to ensure bottles have been appropriately cleaned and prepared.

3.9 Specific Sampling Considerations

3.9.1 Surface water

Surface water can be sampled using bottle samplers for shallow waters, pumping systems for surface to medium (10 m) depths, depth samplers (50 m to >100 m depending on design), automatic samplers, or integrating samplers. Automatic samplers may not be appropriate for sampling bacteria, DO, pH or other variables that are likely to change significantly between the time of collection by the automatic sampler and retrieval from the field for analysis.

Membrane-based passive samplers are effective for the time-integrated sampling of hydrophobic contaminants. They partition the sampled contaminant between water and a lipophilic solvent enclosed in a semi-permeable polymeric membrane, thereby indicating its bioavailability.

3.9.2 Groundwater

Groundwater can be sampled after the construction of a bore or other access hole. Possible equipment includes displacement pumps, submersible pumps, suction pumps, down-the-hole grab samplers, and balers. Groundwater sampling should generally be carried out by experienced field staff or in close consultation with experts, to ensure sample integrity. To retrieve a representative sample, the following principles should be considered:

- the sampling equipment should not change the water quality in any way; particular effort should be made to avoid cross-contamination between bores and sampling equipment;
- sufficient water should be removed to ensure that the sample is newly derived from the aquifer itself rather than from water that has sat in the bore; and
- the methods of collection and storage in bottles and transportation to the laboratory should be suitable for the type of analysis required.

3.9.3 Sediments

Sediments often are surveyed to determine the composition and concentration of contaminants in them, as well as the numbers of organisms located at various depths. There are two broad-based sediment classifications; suspended sediments and bottom sediments.

Suspended sediments are generally dealt with as part of the water column, although specialised sampling techniques are required to obtain representative samples. The benthic organisms in bottom sediments are investigated as measures of aquatic health, pollution or contamination, and as part of the ecology of aquatic systems. They must be removed from samples that have been taken to assess sediment and pore water only.

Sampling equipment includes sediment corers, grab samplers or dredges. The redox state of the sediment must not change in sampling and storage. Particle size distribution is another important factor that will affect contaminant distribution.

3.9.4 Aquatic organisms

The aquatic organisms typically sampled comprise plankton, bacteria, periphyton, protozoa, algae, fungi, macrophytes, macroinvertebrates, benthic macroinvertebrates and algae, bivalves and fish. Methods include grab sampling, netting, trapping, scraping and electrofishing with nets, traps, brushes, and other suitable equipment.

3.10 Sample preservation

It is usually necessary to preserve the samples to retard biological, chemical and physical changes. Matters for consideration to ensure successful preservation and storage include selection and decontamination of sample containers, selection of a preservation technique and the time lapse acceptable between sample collection and analysis. Choices available will depend on the parameter being measured.

The preservation time expected before the samples can be analysed needs to be confirmed before samples are collected, and samples must be analysed before a significant change in composition occurs. For samples that must be collected in the field and then analysed in a laboratory, fixative, preservatives and cold storage during transport can minimise changes.

Complete and unequivocal preservation of samples is a practical impossibility. At best, preservation techniques only retard chemical and biological changes that inevitably take place after sample collection. Chemical preservatives should be avoided, if possible, because they may contaminate samples or interfere in chemical or biological analysis. If preservatives are used they should also be taken into account in the analysis of blanks.

3.11 Sample transportation

Samples should be rapidly transported to the laboratory to minimise reactions within the sample bottles. Ideally samples should be transported in cool and dark conditions where possible. Disturbance to the samples should be minimised during storing, loading and transportation to the laboratory. If transportation delays are unavoidable, consideration should be given to preserving the samples using methods appropriate for the parameter being tested.

3.12 Quality Assurance

A quality assurance/control system is required for field sampling to control sampling errors and manage the samples following collection. Sampling errors can be minimised by ensuring that correct procedures have been followed during the field sampling, transport and storage. Sampling procedures need to be clearly written and adhered to and include detailed descriptions for collecting, labelling, transporting and storing the samples and necessary ancillary field data specific to each parameter.

During sampling or field measurements, it is important to record the samples taken, their labels and other details. Clear and distinctive sample labelling is important. After collection, it is important to maintain the integrity of each sample and to ensure that it does not become contaminated, or change between collection and analysis.

Comprehensive field records are vital, and the records must note the position of a sampling site; the condition of the water body and the weather during measurement and sampling; the time of measurement; descriptions of the measurements and any associated samples taken, their labels and other details about them, including who took them; all field data and instrument calibration data; and any incidents. Video or photographic records are highly desirable for future reference. Observations or information on the conditions at the time of sampling may explain unusual data that otherwise might be attributed to problems in measurement, sampling or analysis. All field records must be completed before leaving a sampling station.

If it is possible that contamination could occur during the sampling process, an appropriate procedure for taking blank samples should be devised to detect and measure the contaminant. Field blanks involve taking extra containers with suitable contents to the site. There, the container is opened and closed and the contents are handled just as if this were a real sample during transfer and storage. Other types of blanks are filter blanks which are filtered in the field, container blanks which determine the contamination from the container, equipment blanks which measure contamination introduced through contact with sampling equipment or sampler, and trip blanks that assess gross cross-contamination of samples during transport and storage.

Besides blanks, duplicate samples in the form of sub-samples and replicates are useful for comparative analyses to check methods or practices. Another alternative is to ‘spike' sub-samples in the field to detect change. Samples to check the quality of data should be labelled in such a way that they are not distinguishable from other samples in the batch.

Transfer of results from the field to a database should be automated where possible, and the printout of the entry should be checked against the field sampling sheet and the laboratory register. Entries can be validated by electronic screening against the expected range and against other analytes for the same site and sampling date, and field measurements.

Quality control also relates to data security and backup. With respect to security, those personnel who have read or write access to the data must be specified. Data backup is always essential in case of system or file failures.

3.13 Occupational Health and Safety

Field staff will need to be trained and be competent at undertaking the sampling tasks. The hazards likely to be encountered by the sampling staff will need to be identified and addressed in a risk management plan to ensure that these hazards are minimised.

Sampling staff should be trained to use sampling equipment and should anticipate problems that may occur in the field (e.g. loss of sample containers, low volumes of sample, occurrence of foreign objects, impossible conditions for sampling a site). Before sampling staff are permitted to do reportable work, they should demonstrate competence in field procedures. As a minimum this would include being able to adhere to protocols, being able to avoid contaminating samples, and being able to calibrate field instruments and make field observations. Some field staff may also need vehicle handling or bush skills.

All staff must be appropriately trained as part of the formal risk minimisation strategy. Training should include familiarisation with environmental hazards that may be encountered, sampling protocols (sampling procedures, chain of custody considerations, etc.), use of sampling equipment, and safety procedures. Staff must be qualified to drive appropriate vehicles, e.g. off-road 4-wheel-drive vehicles or boats, and to administer first aid.

Some key risk minimisation strategies during sampling include-:

- limit continuous driving;
- choose safe sites with safe access;
- wear appropriate clothing, e.g. for the expected weather;
- take appropriate safety gear and a first aid kit;
- maintain contact with help and never sample alone;
- never go into deep water;
- avoid contact with contaminated water.

Professional practice requires sampling staff to obtain approvals and permits as required. Also, individual sampling staff have a duty of care to other field personnel (helping each other where necessary, not discriminating, respecting privacy) and to the environment with regard to such matters as littering, fire, removal of human wastes, keeping to tracks, etc. Staff should never work alone, and junior staff will need appropriate supervision. For some procedures, such as in sampling of groundwaters and sediments, it is important that experienced staff are involved.

3.14 Field Analysis

Only a fraction of the parameters can be accurately measured in the field (eg temperature, flow), whilst others are susceptible to change in the sample after collection (eg dissolved oxygen, pH). Consideration should be given to the feasibility of utilising highly reliable sensors that are capable of accurately recording these types of parameters in the field.

The reliable sensors that are readily available make it often convenient to measure many parameters in the field, obtaining on-the-spot values that can be checked immediately.

3.15 Laboratory Analysis

The objective of the laboratory analysis is to obtain accurate and precise data in a safe environment. The selection of appropriate methods for analysing the physical, chemical or biological parameters will be dependent on the objectives of the monitoring programme, available budget, laboratory resources, speed of analysis, matrix type and contamination potential. The choice of an appropriate method is based on four primary considerations-:

- The range of concentrations of the analytes that need to be detected (detection limits are method specific);
- The accuracy and precision required (all results are only estimates, higher accuracy and precision increases the costs);
- The period between sampling and analysis (field analysis may be required); and
- Familiarity with a particular method when more than one suitable method is available.

Laboratory analyses also need to consider a range of issues including-:

- Data management (data storage and reporting);
- Quality assurance/control; and
- OH&S (identification of hazards, risk minimisation plans, education).

3.16 Data Analysis and Interpretation

A number of common statistical methods are used for the analysis of water quality data. The data initially needs to be summarised from the field/laboratory into a form that is suitable for analysis, and preliminary checks made for erroneous, missing or below detection data. Suitable approaches for dealing with these types of data will need to be considered.

The checked data is then reduced and summarised using a combination of commonly used statistical tools, including graphs, tables and statistics (e.g. mean, standard deviation). The objective of summarising the data is to present essential information contained in the data set concisely, and to assist in clearly identifying outliers. Advances in computer software have enabled highly sophisticated graphics to be readily accessible.

A number of more advanced statistical methods are then typically used to assess data relationships. Methods including transformations, checking distributional assumptions, trend detection and smoothing can be undertaken. It is then important to subject the data to a process of statistical inference to assist with determining characteristics of the data set, enabling testing of hypothesis and comparison of the data with water quality guideline or trigger values. The relationship between pairs of water quality variables can be evaluated using correlation and regression analysis.

Following the data analysis, the results should be summarised concisely, and compared with the monitoring programme objectives to determine if the original questions are answered by the results.

3.17 Reporting

The reporting of data from the water quality monitoring program requires careful consideration of the end user to ensure that the technical nature of the topic is presented to the intended audience in an appropriate format. An example template for reporting is outlined below-:

- Executive summary – summarises technical findings for managers;
- Introduction – Outlining study objectives, study location and review of previous/related studies;
- Methodology – Outline of study design, sampling and analysis methods;
- Results – Summarised results;
- Discussion – Data interpretation and implications for management;
- Conclusions
- Recommendations
- References
- Appendices – laboratory reports, data tables

The report findings can be disseminated in a number of ways depending on the intended audience. Methods for disseminating the data include publications, meetings, internet web pages, film and video presentations, and media reporting.

4 DATABASE MODIFICATIONS

4.1 Quality Analysis of the Existing Database

4.1.1 Introduction

A number of quality tests were performed on the existing database to determine:

- The presence of duplicate water quality data records;
- The presence of any duplicate sites;
- The presence of empty records;
- The presence of outlying records; and
- The presence of any unused parameters.

The results of these quality tests are summarised in the following sections.

4.1.2 Duplicate water quality data records

The entire database was scanned for corresponding records where the testing results were identical for each parameter measured. Further evaluation was only undertaken if records included values for three or more parameters. This ensured that a positive match was not returned for the numerous records where only zero concentrations were recorded (e.g. coliforms). There were a number of duplicate records found in the database. Closer examination of these records suggested that a data entry or reporting error was likely to have occurred.

4.1.3 Presence of Duplicate Sites

To determine the presence of duplicate sites, the recorded longitude and latitude for all sites were compared. Sites with a longitude and latitude within 10 m were assumed to be the same site.

A number of duplicate sites were identified. These fell broadly into three categories:

- Locations where multiple sites were created to differentiate sampling depths. Multiple sites were merged into a single site and a depth field was added to the 'rec' table to differentiate between sampling depths.
- Locations where the same coordinates were entered for a number of closely spaced sites at a definable geographical feature (e.g. Lilli Pilli Beach). The SEWQ Project Management Committee provided new coordinates for these sites.
- Genuine duplicate sites where two sites have inadvertently been established at nearly the same location. There was only one instance of this and it was determined that one of the sites in question could be removed from the database as no records had been assigned to that site.

4.1.4 Presence of Empty Records

The database was scanned for records that contained no water quality data. Nine records were found and removed from the database.

4.1.5 Presence of Outlying Records

Statistical analyses were undertaken to determine the presence of outliers. The following steps were undertaken:

- The maximum value was determined for each parameter. Where the maximum exceeded the values that would normally be expected for that parameter, that parameter was highlighted for further review.

- A coarse statistical filter was applied to all parameters to extract the values that were positioned more than 3 times the standard deviation away from the mean. Although the statistical distribution of most water quality parameters is not normal, this test managed to identify most of the suspected outlier records. The resulting list of nearly 500 suspected outliers was examined further, considering the value recorded for each parameter and the location of the site (e.g. STP readings are expected to have poor water quality). This process identified for further review, a number of records that are considered to be outliers.

- Reference was made back to the maximum values of all parameters. Where no potential outliers had been selected for a parameter based on the coarse filter applied in the second step, that parameter was further analysed to extract a list of the top 10 values recorded. These were then reviewed to identify outliers.

Two parameters were analysed further based on these initial assessments, as there were numerous values that appeared to be outliers:

- Salinity %: Some records contained salinity readings with very high values (13 individual values > 100 %) that were clearly in error. It was also observed that there was a cluster of records around 25 – 40 %, which suggested that these records were incorrectly entered as a percentage rather than in parts per thousand. The majority of values are within the expected range (around 3% or 30ppt).

- Chlorophyll – a: Numerous records were observed to include very high readings for chlorophyll. The high chlorophyll values were compared with corresponding dissolved oxygen and total nutrient measurements (where available) to determine any indications of potential algal blooms.

Based on these analyses and consultation with members of the SEWQ Project Committee, the following approaches were considered:

- Removal of a parameter record from the database;

- Modification of the record (e.g. where a pH of 683 was entered, this was altered to 6.83 to be consistent with the normal range of readings).

- Assignment of an unknown, or poor quality code (5).

4.1.6 Presence of Unused Parameters

The database was scanned for parameters for which no records existed. Twenty parameters were identified and removed from the database following consultation with the SEWQ Project Management Committee.

4.2 Quality Coding

It was initially considered that quality codes for data existing on the database would be adopted from codes that the organisation collecting the data had previously established. It was also initially assumed that all data for a particularly site were collected by the same organisation. However, it was subsequently determined that there was no common system for quality code assignment for the existing data. Accordingly, a common system of quality code assignment was derived following consultation with members of the SEWQ Project Management Committee. This consultation resulted in the definition of the following quality codes:

- 1 – Professionally sampled and analysed in accordance with NATA certification and / or established field measurement protocols exist;
- 2 – Sampled and / or analysed by professional staff without NATA certification using laboratory based methods;
- 3 – Sampled and / or analysed by Council staff;
- 4 – Sampled and / or analysed by community groups using field based kits; and
- 5 – Quality of data unknown.

Within the existing database, it could not be confirmed whether any of the data had been analysed in accordance with NATA certification and established field measurement protocols. Accordingly, a quality code value of 1 was not assigned to any of the values in the database.

Where data comprised parameters measured in the laboratory, a quality code of 2 was assigned. Where data measurements were taken in the field, a quality code value of 3 was adopted.

No community data was present on the database, so a quality code of 4 was not used for data that already existed within the database.

Where the data was subject to modification for quality reasons as described in Section 4.1, and it was deemed appropriate, a quality code of 5 be assigned to those data items.

As part of the overall project, there was a need to collate the geographical coordinates (longitude and latitude) to enable linking to a Geographical Information Systems (GIS) application. Coordinates were available for the majority of sites, although the coordinates for some sites could not be confirmed from available data and knowledge. For those sites where the location was unconfirmed, a quality code of 5 was assigned to every record on the database, which pertained to those sites.

In order to store the quality codes, an additional table was added to the database (the "qual" table). This table has exactly the same structure as the 'rec' table. However, in place of the actual water quality values, the quality code is entered to the table.

4.3 Site Coding

As a component of the protocol development, uniform and consistent naming conventions were developed for the sites on the database. Prior to developing these conventions, a variety of systems have been used by different water quality collection organisations in the region.

Following a review of existing database systems (Australian Water Resources Council (AWRC) and Department of Infrastructure, Planning and Natural Resources (DIPNR)) and considering a future objective to facilitate linking with those systems, it was determined that the site code should be formatted as follows:

"AAA_00000"

The proposed site code comprises a three letter identifier, followed by an underscore, followed by a five digit integer. The three letter identifier represents the AWRC basin, of which there are six within the SEWQP area. These six basins and their associated three letter codes are as follows:

- Clyde (CLY);
- Moruya (MOR);
- Tuross (TUR);
- Bega (BEG);
- Towamba (TOW); and
- Genoa (GEN).

The five digit integer is used to differentiate sites within the same AWRC region.

4.4 Incorporation of Additional Datasets

As part of the project, significant effort was undertaken to add additional data sets to the database. Contact was made with a variety of different organisations and a significant amount of additional data was obtained. A summary of the sources contacted and the information gained is provided in Table 4.1.

Table 4.1 Sources of Additional Data

Organisation	Comment	Additional Data Values Added	Sites Added
Australian National University	No water quality data were available for the study area.	N/A	N/A
Clyde River Shellfish Quality Assurance Program	Water quality database is currently under construction and was not available for entry into this database within the project timeframe.	N/A	N/A
Department of Environment and Conservation (formerly EPA)	Water quality data obtained from Geoff Coade.	11097	82
Department of Environment and Heritage – Water Watch NSW	David Leviston from the DEH indicated that the water quality data should already be in Council's existing water quality database.	N/A	N/A
Department of Infrastructure Planning and Natural Resources	Water quality data obtained from Tim Cooney via the TRITON water quality database.	57766	181
Department of Primary Industries	Water quality data obtained from James Haddy on behalf of DPI.	66	2
Sydney Catchment Authority	Martin Krogh of the SCA indicated that there is no water quality data available for the Eurobodalla and Bega Valley LGAs.	N/A	N/A
State Forests	Steve Dodd of State Forests indicated that water quality data is available for the study area in a hard copy format. State Forests also indicated that the possibility of obtaining this data within the project timeframe was not possible.	N/A	N/A
University of Canberra	Water quality data obtained from Daniel Spooner. Additional water quality data is available from Professor Bill Maher in a hard copy format.	10312	11
University of Wollongong	Contact did not respond.	N/A	N/A
University of New South Wales	Ross Matthews of the Water Research Laboratory (WRL) indicated that no water quality data were available within the study area.	N/A	N/A
University of Sydney	Contact did not respond.	N/A	N/A

The data was examined, appropriate quality codes were assigned and data was converted into the standard template format (see Section 5.6) and imported to the database. In conjunction with the additional sites and records added to the database, 14 additional parameters were added to the database.

4.5 Other Modifications

Throughout the course of the database review, there were a number of other modifications made to the database. A summary of the modifications is provided below:

- Merging of site separated depth profiling records into a single site;
- Delete outlying records;
- Included new column for changed site codes (old site codes retained for backwards compatibility);
- Merging of some parameters at the request of Stephen Hure from Ecowise;
- Included an allowance of a column in the 'rec' and 'qual' tables for the addition of a restricted data flag, following a request from the University of Canberra;
- Modification and removal of some sites for which no useful data exists at the request of Norm Lenehan and Debby Lenson from Eurobodalla Shire Council on 27/10/04;
- Included the capability to enter quality codes via the existing database entry routines;
- Included an option to read in template formatted data (see Section 5.6);
- Included an option to partly automate requests for data collection;
- Included an option to automatically assign valid site codes based on the protocol outlined in Section 4.3;and
- Modification of a number of forms within the database to refer to the newly assigned site codes.

5 USE OF THE DATABASE

5.1 Management of the Database

It is presumed that the database will be managed by specifically trained personnel with permissions to perform tasks as outlined in Table 5-1.

Table 5-1 Database Permissions

Personnel	Assigned tasks
Database Administrator	• Modifications requiring changes to the underlying database structure. • Maintenance and support. • Addition of new parameters.
Database Entry Operator	• Importation of new data. • Quality checking. • Manual entry of data. • Addition of new sites

At the time of writing, a number of issues relating to the day-to-day maintenance of the database were yet to be resolved. Issues to be resolved are:

- **How will data be synchronised if different individuals are entering data onto different versions of the database (i.e. if both Eurobodalla and Bega Valley Councils are able to assign data).**
- **How the system security is to be managed. At present some security has been implemented on the database with a requirement for the entry of passwords on manual data entry routines.**

These are issues we consider will need to be resolved among the various stakeholders of the SEWQ Project.

5.2 Data Requests, Receipt and Entry

It is considered that either of two main approaches would be adopted for gathering and entering data into the database:

1. Via a formal request to an external organisation to collect the data (refer Section 5.5). The data is subsequently returned in a specified file format and imported to the database using automated checking and entry routines (Section 5.6).

2. In a more generic fashion. For example, where additional data is obtained opportunistically, or internally by the database managing organisation. In this instance, the data is entered using the manual entry routines on the database, as described in Section 5.7.

5.3 Database Installation

Additional directories should be created for proper functioning of some of the new routines on the database. The steps required are as follows:

- Create a directory in “C:\Program Files” called “SEWQ”
- Within the “SEWQ” directory, create three directories, named “Database”, “Imports” and “Exports”.

The database (*.mdb) file must be saved in the “Database” directory for SEWQ MAP to execute correctly (C:\Program Files\SEWQ\Database). Further information relating to the structure of SEWQ MAP and the location of all relevant files used by the program is contained within the SEWQ MAP Description and Reference Manual.

5.4 The Database Opening Screen

Upon loading the database, the main menu displays (refer Figure 5.1).

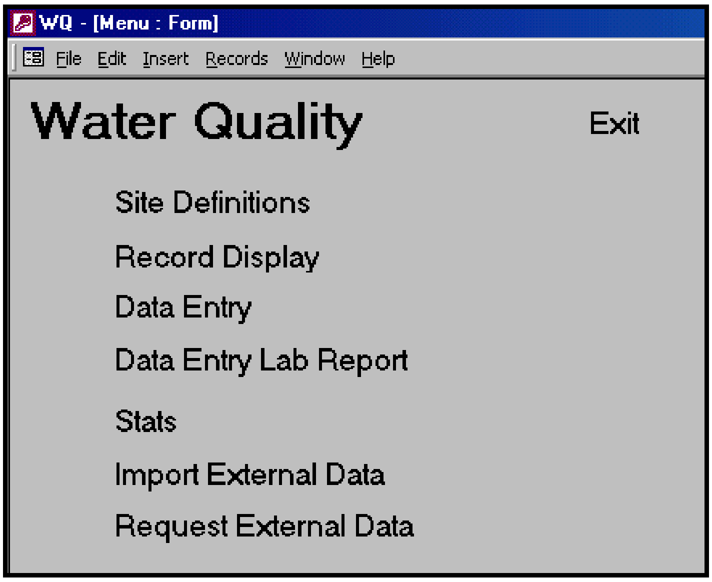

Figure 5.1 Database Main Menu

Changes have been made as part of database improvements undertaken by WBM. These changes, and the way they impact upon the function of the database are described in the following sections. The discussion is organised as follows:

- Section 5.5 describes the function of the “Request External Data” item, which has been added to enable rapid development of requests based on the database;
- Section 5.6 describes the process of importing external data following receipt of external data in the required format.
- Section 5.7 describes changes to the manual data entry routines which provide for the addition of quality codes.

- Section 5.8 describes the process of adding new sites and generating appropriate site codes.
- Section 5.9 describes the process of adding new parameters.

These sections primarily focus on recent modifications to the database. Items that have not substantially changed, such as the "Record Display" and "Stats" functions are not discussed here. It is assumed that these are either understood at present, or are documented elsewhere.

5.5 Requesting Additional Data

5.5.1 Standard File Format

The standard water quality file format for the SEWQ project is a .csv file format. It is envisaged that this file format will be generated from spreadsheet software by saving the file as .csv. Figure 5.2 shows an example from Microsoft Excel to illustrate the file format.

	A	B	C	D	E	F	G	H	I
1	Instructions: Collect Field Measurements of Temperature, DO and Salinity on a monthly basis from March to August.								
2	Site Code	Date	Time	Depth	Parameter ID	Value	Quality	Collecting Organisation	Restrictions
3	EG1_00001	15/11/2004	12:00:00	1.00	L1	23.4	3	DPI (James Haddy)	NO
4	EG1_00001	15/11/2004	12:00:00	1.00	L2	10.7	3	DPI (James Haddy)	NO
5	EG1_00001	15/11/2004	12:00:00	1.00	L3	40.0	3	DPI (James Haddy)	NO
6	EG1_00002	15/11/2004	12:00:00	1.00	L1	25.6	3	DPI (James Haddy)	NO
7	EG1_00002	15/11/2004	12:00:00	1.00	L2	9.7	3	DPI (James Haddy)	NO
8	EG1_00002	15/11/2004	12:00:00	1.00	L3	35.0	3	DPI (James Haddy)	NO

Figure 5.2 Standard Water Quality File Format

The first line of the standard water quality file is reserved for instructions or comments for the collecting organisation. This line is used by the commissioning organisation to provide any additional instructions or comment to the collecting organisation. The second line of the csv file is a header line (refer Table 5.2) and describes the purpose of each column. The first column of the csv file contains the site code. The next six columns of the csv file (See columns B, C, D,…) contain the date of the record, the time of the record, the depth at which the measurement was collected, the measured parameter ID, the measured value and the quality code to describe the reliability of the record. The last two columns (H and I) contain additional data relating to the organisation that collected the data, and whether there should be any restriction on access to the record.

The quality code shall be selected from the five categories outlined in Section 4.2.

Each row of the standard water quality file below the header line should contain water quality data for only one water quality parameter and site location. The standard water quality file can contain data for multiple sites. Each standard water quality file may contain a maximum of 65535 water quality records corresponding to the maximum number of rows available in Microsoft Excel.

Water quality data should be entered into the standard water quality file in the same manner as shown in Figure 5.2. In this example standard water quality file, the first row of data below the Header data (Line 3) corresponds to a measurement of Faecal Coliforms (cfu/100mL) at a depth of 1m for site EG1_00001 (Site 00001 in fictitious river basin EG1) on the 15/11/2004 at 12pm. The collecting organisation in this example record was DPI and there are no restrictions on the data. The next two lines of the water quality file are also for site EG1_00001 with the same data and time, however, the measured water quality parameters are Enterococci (cfu/100mL) and Total Coliforms (cfu/100mL) respectively. The next line of the water quality file, Line 6, corresponds to water quality data for site code EG1_00002. The water quality data for this site location has been entered in the exact same manner as the previous site.

Table 5.2 Description of the standard file format

Column	Parameter	Format	Information
A	Site Code*	Basin ID_Site Number (e.g. EG1_00001)	SEWQ Database site code for the data within each line. A site code must be entered on each line of the file.
B	Date*	(DD/MM/YYYY)	Date samples were recorded
C	Time	(HH:MM:SS)	Time samples were recorded in 24-hour format.
D	Depth	Numeric value in metres	Depth at which sample was recorded in metres.
E	Parameter ID*	Text	SEWQ database code representing the parameter (e.g. F1 for Temperature (oC), F2 for Dissolved Oxygen – DO (mg/L))
F	Value*	Text	Data readings corresponding to the parameter described in Column E.
G	Quality Code*	Integer (1 to 5)	SEWQ Database quality code for the data within each line.
H	Collecting Organisation*	Text	Name of the Collecting Organisation
I	Restrictions*	YES or NO	“YES” or “NO” indicating whether the data is clear for public release.

*Indicates field is compulsory.

5.5.2 “Requesting External Data” Menu Item

Selecting the “Request External Data” menu item results in the opening of the form displayed as Figure 5.3. To generate the template file and instructions for the collecting organisation, perform the following steps.

1. Enter a brief description of the required data collection to be commissioned.
2. Select the sites for which monitoring is required.
3. Select the parameters which are to be monitored.
4. Click on the “Generate Request” Button

A dialog will appear asking where you would like the resulting data request template to be saved.

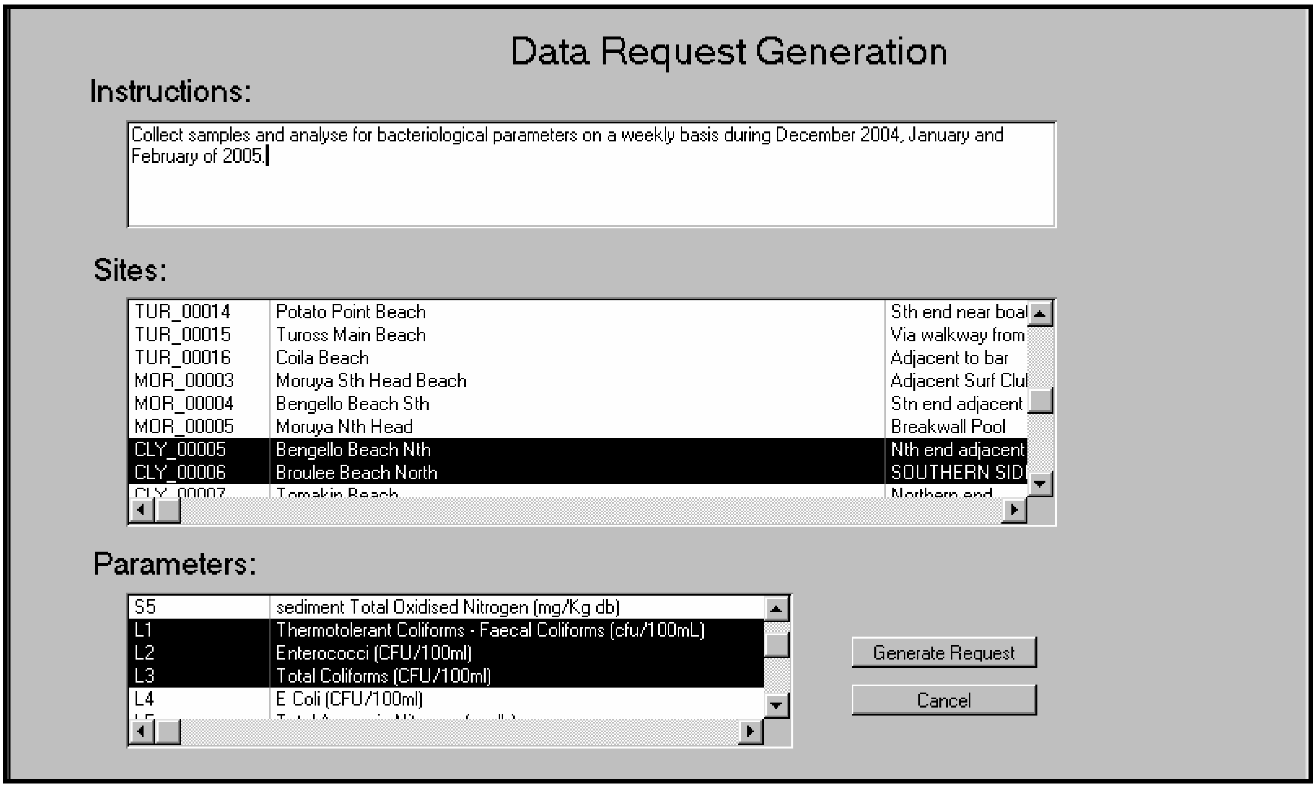

Figure 5.3 Form for Automating Request of New Data

When the request is generated, two files will be saved.

- An empty template file for data collection as requested.
- A descriptive file which provides more detail on the parameters (Names, units) and sites (Description, latitude and longitude values) to guide the collecting organisation on exactly what is to be collected. This file is saved in the same location as the template file, but with a .txt extension. The text file can be opened in Microsoft Word, or other word processor / text editing software.

5.6 Importing Data Using Template Formatted Files

The “Import External Data” option from the opening menu opens the SEWQ file import form as shown in Figure 5.4.

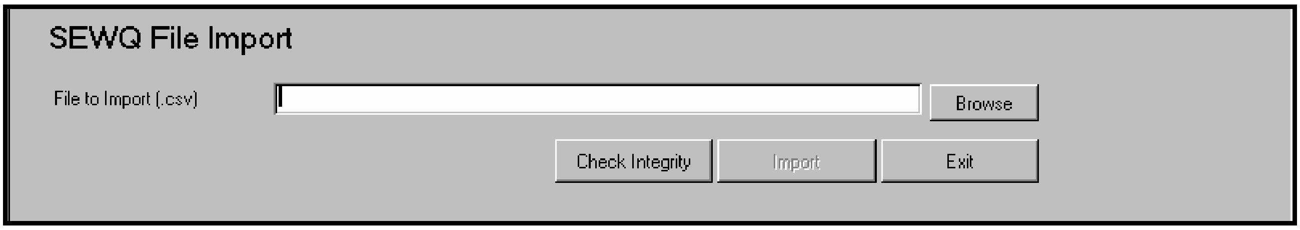

Figure 5.4 SEWQ File Import Form

The import requires the following steps:

- Locate the file requiring import using the browse button (file should be formatted in standard SEWQ import format – See Section 5.5.1);
- Check the integrity of the file by clicking on the "Check Integrity" button. When this button is pressed, a "tracking file" named "*_track.csv" is saved to the "C:\Program Files\SEWQ\Imports" directory (where * represents the name of the file being imported). This tracking file records the results of the integrity check.
- The routine then reads each record in the file and performs a number of checks on the integrity of the data being supplied. Where problems are identified, the user is prompted to provide an appropriate course of action. The integrity checks performed for each record include the following:
    - Checks that parameters on the file exist on the database. If not, the record is set to be ignored in the tracking file.
    - Checks that sites on the file exist on the database. If not, the record is set to be ignored in the tracking file.
    - Checks whether coordinates exist for the sites, if not a quality value of 5 is assigned.
    - Checks to make sure that a quality value has been provided. If not, the record is set to be ignored in the tracking file.
    - Checks whether an existing water quality record exists on the database. If so, the record is set to be ignored in the tracking file.
    - Checks whether the value for the water quality parameter violates the maximum or minimum conditions specified in the "SEWQLimits.csv" file (refer Section 5.3).
- Following the integrity check, the data can be written to the database using the "Import" button. If there were substantial issues identified during the integrity check it may be worthwhile to check the tracking file (openable in a text editor or spreadsheet software) to determine whether the import should proceed. If it is decided to abandon the import, click on the "Exit" button.

The integrity checks provided by the automated routine do not ensure that all data being entered is reliable. It is strongly recommended that the formatted data file be inspected by a qualified scientist or equivalent to provide an additional quality check of the data. Furthermore, it is noted that the limits specified in the limits file are, in some cases, based on scant data and may need revision.

This item may require implementation of a security password to prevent just anyone from adding data to the database.

5.7 Manual Entry of Data

Manual entry of data is accomplished via forms opened following selection of either the "Data Entry" or "Data Entry Lab Report" options from the main menu. Changes made to these two forms include the following:

- Data can now only be entered for sites with a valid SEWQ formatted Site Code.
- A quality code between 1 and 5 must be entered for each parameter for which data has been collected. Descriptions of the quality codes are provided on the forms. If the relevant quality code is not known, a value of 5 should be adopted.

5.8 Adding New Sites

The addition of new sites requires knowledge of where those sites are located. Prior to attempting to add the site to the database, it is essential that the data entry operator knows which AWRC region the site is located (refer Section 4.3). This task can be achieved via GIS, with a prior knowledge of the site location and a GIS data set showing the location of the AWRC region boundaries.

Once the AWRC region has been determined, the process for adding a new site can be undertaken. This involves the following steps:

- Select "Site Definitions" from the Main Menu.
- Click on the "Create New Site" button, and enter the required password, the "Edit Sites" form opens (refer Figure 5.5). This form can be used to directly add a new site, however it is recommended that a valid code be generated as follows:
    - Click on the "New Site" Button. The "Create New Site Code" form appears (refer Figure 5.6).
    - Select the correct AWRC region and click on "Next Valid Code". After a short delay, a new site code will be determined and displayed.
    - Note down the new code, and, if you want to proceed, click on the "Add to Database" button. The "Edit Sites" form will reappear.
    - By using the record selector, the user can now scroll through to the required location (refer Figure 5.7) of the new site code (they are in alphabetical order) and fill in the details for that site.

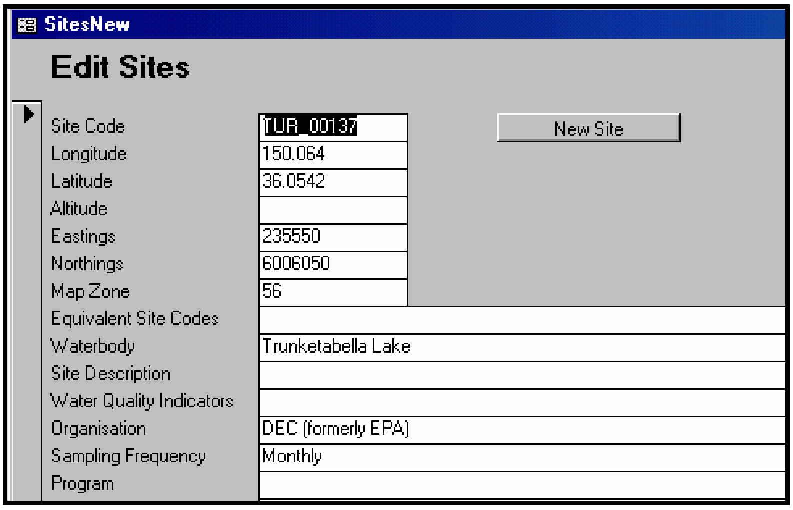

Figure 5.5 Edit Sites Screen

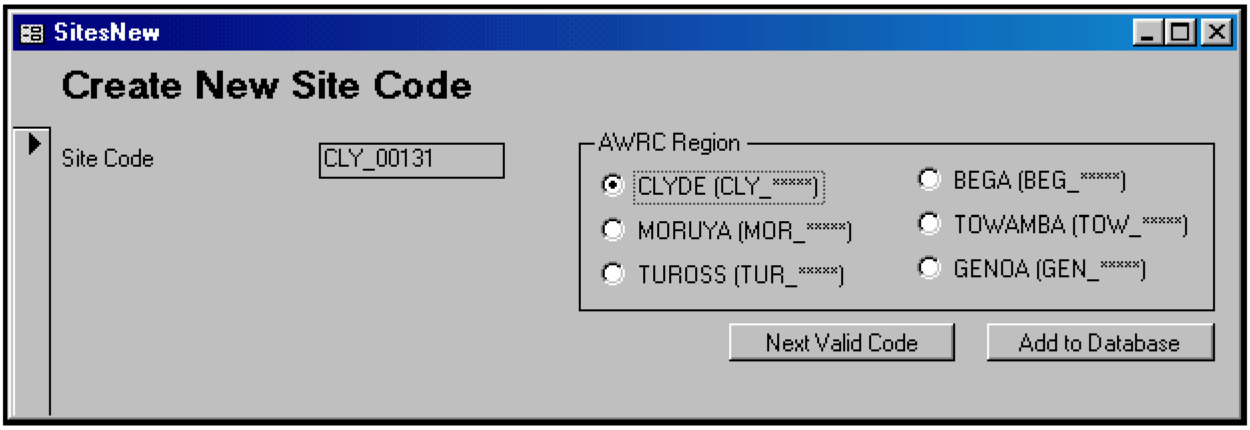

Figure 5.6 Create New Site Code Form

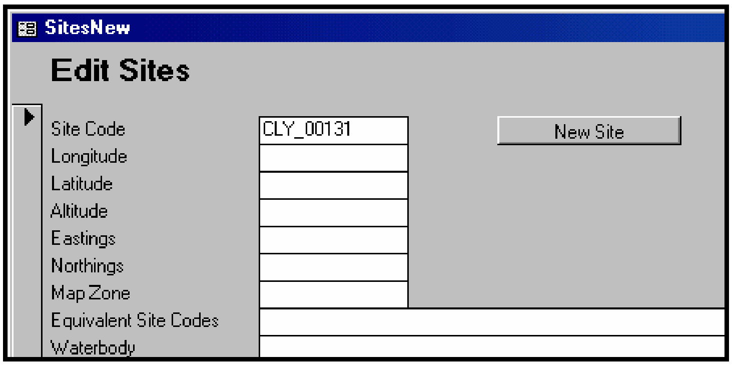

Figure 5.7 Editing New Site Details

Even though latitude values will represent sites that are all located south of the equator within the project area, the values should be entered as positive.

5.9 Adding New Parameters

The addition of a new parameter requires structural changes to the database. As there are limits to the number of columns that can be held within the 'rec' and 'qual' tables of the database, this is not a change that should be made without due consideration.

The steps required to update the database include the following:

- Addition of a new site to the 'sites' table;
- Addition of corresponding columns to the 'rec' and 'qual' tables
- Addition of values to the "SEWQLimits.csv" file (found in the "C:\Program Files\SEWQ\Control" directory.

Care needs to be taken to ensure that changes are consistent.