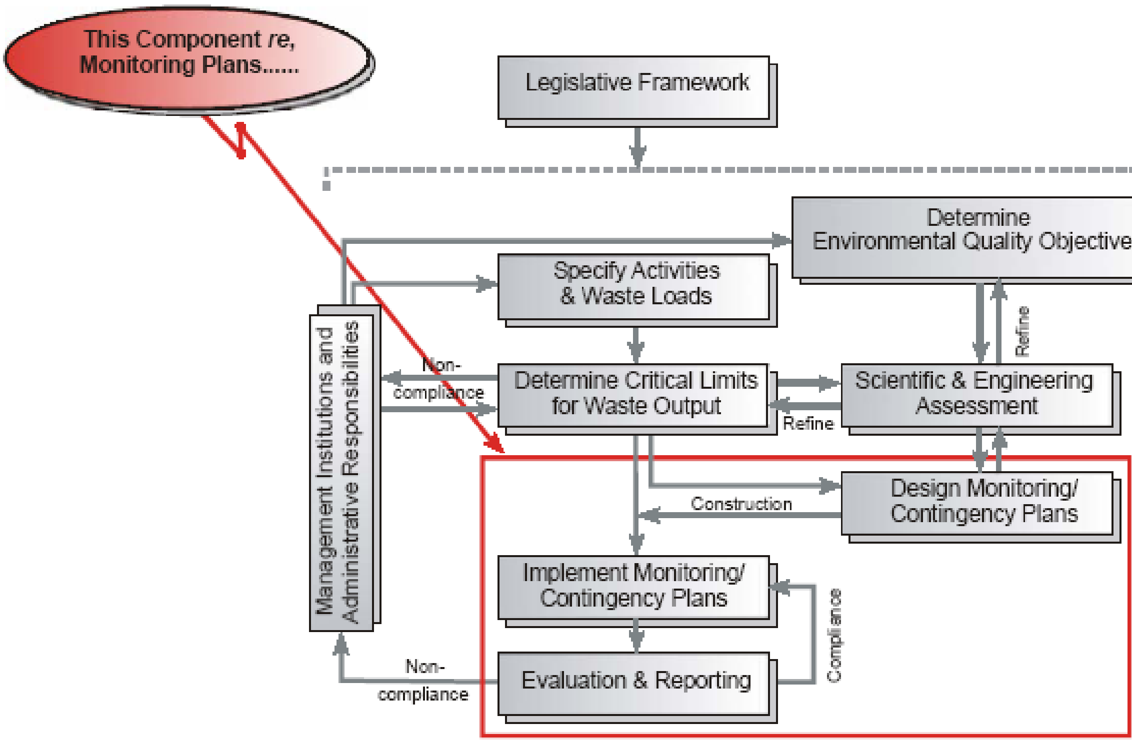
SECTION 7:
MONITORING

SECTION 7: MONITORING



PURPOSE:

In the context of this operational policy, this component refers to the longer-term monitoring programmes that are undertaken once marine disposal activities become operational. Such monitoring programmes typically become part of the licence agreement issued by the DWAF for a particular discharge, under section 21 of the National Water Act. In essence, the purpose of these long-term monitoring programmes is to provide data to continuously evaluate:

- ***The effectiveness of management strategies and actions to ensure compliance with the licence conditions and design criteria (Compliance monitoring and System Performance monitoring)***
- ***The trends and status of changes in the environment in terms of the health of important ecosystems and designated beneficial uses in order to respond. Also, to evaluate whether the actual environmental responses match those predicted during the assessment process (Environmental monitoring).***

Such information is required to provide feedback to management, authorities and other interested and affected parties on the performance of the disposal practice, and to enable informed debate on operations, and operations modifications, where appropriate.

These monitoring programmes are designed and implemented at the cost of the authority or industry responsible for the management of the waste disposal system (following the Polluter Pays Principle).

In order to be useful from a management perspective, monitoring data must be evaluated against pre-determined objectives. Results need to be presented in a clear format, providing the appointed management institution/s with the scientific and engineering information for effective decision making.

Long-term monitoring programmes are typically only designed as part of a <u>Detailed Investigation</u>.

Long-term monitoring programmes related to a systems discharging wastewater to the marine environment consist of two main components, namely:

- *Compliance monitoring*, to determine the effectiveness of management strategies and actions to ensure compliance with licence conditions, e.g. the limits set for the volume and composition of the wastewater

- *System Performance monitoring,* to determine the effectiveness of management strategies and actions to ensure performance in accordance with the design criteria, such as the hydraulic performance of the discharge system

- *Environmental monitoring*, to determine the trends and status of changes in the receiving marine environment, in terms of the health of important ecosystems and designated beneficial uses. Also, to evaluate whether the actual environmental responses, match those predicted during the assessment process. This evaluation is necessary in order to respond, where appropriate, in good time to potentially negative impacts, including cumulative effects.

If compliance monitoring shows non-compliance, the licensee can be prosecuted even if the system performance monitoring and environmental monitoring indicates compliance with the design criteria and environmental quality objectives. If there is compliance with licence conditions and/or the system performs within the design criteria, but the environmental quality objectives are not being met, the licence needs to be reviewed to set stricter limits or the objectives need to be re-evaluated.

7.1 COMPLIANCE MONITORING

Once a wastewater discharge to the marine environment is approved by the DWAF, a licence is issued, which among other aspects, specifies the volume and composition of the wastewater. The aim of this monitoring component is to assess continuously whether these licence requirements are adhered to. Parameters to be monitored include:

- *Flow*: The sampling frequency need to be sufficient to resolve the actual *variability* in the wastewater volume.

- *Composition of wastewater*: The list of constituents to be monitored will depend on the composition of the wastewater, while the frequency of monitoring needs to reflect the actual *variability* in wastewater composition.

  Urban/Municipal wastewater discharges, consisting mainly of domestic sewage have a characteristic wastewater composition. Key constituents that need to be included in the monitoring programme of discharges to the marine environment are:

  - Biochemical oxygen demand/Chemical oxygen demand
  - Total suspended solids
  - Particulate organic carbon and nitrogen
  - Inorganic nitrate and nitrite
  - Total ammonia
  - Dissolved reactive phosphate.

  In the case of industrial wastewater discharges, or where industrial wastewater discharges entering a municipal WWTW, the constituents included in the monitoring programme will depend on the constituents present in the wastewater and their potential to impact negatively on the receiving marine environment and its designated beneficial uses (refer to Section 3).

NOTE:

An industry, discharging to a WWTW or directly to the marine environment (or if applying for a licence to do so), will be required to provide a detailed description of the waste stream both in terms of volume (quantity) and quality (i.e. listing all substances present and their concentrations). Where industries discharge into WWTW, the WWTW authority is responsible for obtaining this information from the industry concerned (in accordance with Ground Rule 20 in the* Operational Policy for the treatment and disposal of land-derived water containing waste to the marine environment of South Africa *(RSA DWAF Water Quality Management Sub-Series No. 13.2).

The sampling frequency of the composition of the wastewater will depend on the actual *variability* in wastewater composition.

Sample analyses must be conducted by an analytical laboratory accredited to analyse of the relevant constituents in the wastewater.

- *Toxicity testing:* It is imperative that wastewater streams are routinely subject to toxicity testing (refer to Section 4.5 for further details). Such tests have been routinely used in South Africa for the past 15 years for monitoring the performance of wastewater discharges to the offshore marine environment. The frequency of toxicity testing of the wastewater will depend on the actual *variability* in the wastewater composition.

7.2 SYSTEM PERFORMANCE MONITORING

Monitoring of the performance of the wastewater disposal system comprises two main components:

- Physical inspections of the outfall system (for marine outfalls)
- Hydraulic performance (this typically applies to marine outfalls, i.e. wastewater discharges to the offshore marine environment).

7.2.1 Physical inspections

<u>Head works and land line</u>. The Standard Operating Procedure for the head works and treatment plant should include specified and scheduled monitoring procedures. These include daily routine observations and longer-term mechanical, electrical and hydraulic testing. These records form an integral part of the maintenance programmes and service contracts for the plant and specific components.

<u>Underwater section</u>. The stability of the pipeline and the diffuser of a wastewater discharge system should be checked regularly, especially after major storms, to ascertain that no undesired displacements or damage have occurred.

An underwater video recording of exposed sections of the pipeline, the diffusers and ports is used to check the structural/physical condition of an outfall as well as its operational efficiency. These inspections should be conducted annually and after major storms.

7.2.2 Hydraulic performance

After the commissioning of a new marine outfall, the hydraulic/dilution performance should be confirmed. This field test(s) should include:

- Controlled injection of a conservative tracer material, such as Rhodamine-B dye, into the wastewater at the head works
- Continuous sampling downstream of the injection point and at one or more of the ports (depending on the length of the diffuser)
- Spatial sampling in the initial mixing zone ('boil')
- Accurate recording of wastewater flow and physical conditions at the discharge location (currents and seawater density throughout the water column)
- Sample analysis (e.g. by using a calibrated Turner Design Fluorometer or similar device)
- Statistical analysis of the distribution of concentrations in the boil to determine the achievable dilutions
- Comparison of measured achievable dilution with the theoretically predicted dilutions.

The performance test should also be conducted at any stage during the lifetime of the outfall when physical changes or alterations, which may have an effect on the hydraulic characteristics, are introduced or when there is a substantial change to the wastewater quantity or composition.

During this performance test, the sampling can be extended to the far field to confirm the estimated achievable secondary dilutions.

7.3 ENVIRONMENTAL MONITORING

In the context of this operational policy, the purpose of establishing monitoring programmes in the receiving marine environment is to continuously provide data for the evaluation of the status of the receiving environment in terms of the health of important ecosystems and designated beneficial uses. This evaluation enables a response, where appropriate, in good time to potentially negative impacts, including cumulative effects.

The requirements for monitoring in the receiving environment are usually site-specific and depend on the type of wastewater discharge and the variability in its waste loads, as well as the site-specific physical, biogeochemical and ecological characteristics of the receiving environment and the variability thereof. As a result, this section is not intended to be prescriptive, but rather sets out the approach to follow when formulating long-term monitoring programmes associated with wastewater disposal activities to the marine environment.

Key elements of a successful monitoring programme include (ANZECC, 2000b; US-EPA 2003):

- Setting clear monitoring objectives
- Design and implementation of a cost-effective programme.

NOTE:

Useful documents to consult in the design of monitoring programmes are NATIONAL RESEARCH COUNCIL (1990), ANZECC (2000a), ANZECC (2000b), NZWERF (2002), US-EPA (1994) and US-EPA (2003).

It is also important to remember that any long-term monitoring programme is a dynamic, iterative process that needs to be adjusted continuously to incorporate new knowledge, thereby supporting the principle of adaptive management.

NOTE:

The Methodology for the Determination of the Preliminary Ecological Reserve for Estuaries *(RSA DWAF, 2004), issued under the National Water Act, through the Directorate: Resource Directed Measures, also requires that Resource Monitoring Programmes be implemented for estuaries. To ensure alignment with existing initiatives on estuaries, those methods (or future updates thereof) need to be consulted when formulating monitoring programmes for estuaries as part of this operational policy. Proposed procedures for the design of long-term environmental monitoring programmes for estuaries, as part of Resource Directed Measures, are provided in a WRC report entitled:* Resource Monitoring Procedures for Estuaries for Application in the Ecological Reserve Determination and Implementation Process *(Taljaard et al., 2003).*

7.3.1 Monitoring objectives

Measurable site-specific monitoring objectives are a key component of a sound monitoring programme. Such clear objectives make it possible to design a focused and cost-effective monitoring programme. These objectives can also be translated into hypotheses that could be proved statistically.

The monitoring objectives are distilled from the environmental quality objectives previously specified for the study area and, in turn, are based on the requirements of the important marine ecosystems and the designated beneficial uses (refer to Section 3).

NOTE:

Where a monitoring programme is designed for an existing wastewater discharge to the marine environment and where the study area boundaries, as well as the environmental quality objectives of important marine ecosystems and other beneficial uses, have not yet been established, these will have to be addressed prior to setting monitoring objectives.

EXAMPLE…

Monitoring objectives for a discharge from a municipal WWTW typically include:

- Determine whether *E. coli* levels measured at designated recreational beaches exceed the recommended target values for contact recreation, as set out in the *South African Water Quality Guidelines for Coastal Marine Waters*
- Determine whether particulate organic matter discharged from the WWTW accumulates at depositional areas, thus creating a ‘sink’ for toxins such as trace metals
- Determine whether trace metals are present in the wastewater discharged from the WWTW accumulates at depositional areas and whether these exceed the limits set for such constituents under the London Convention
- Determine whether toxins are accumulating in biological tissue (e.g. that of filter feeders such as mussels and oysters) at concentrations exceeding the environmental quality objectives set, for example, for human consumption of these organisms or the protection of organism health
- Determine whether the composition of the biotic community in the study area is being altered as a result of the wastewater discharge.

Note that monitoring objectives do not specify details on temporal and spatial scales since these are addressed as part of the monitoring programme design (ANZECC, 2000b).

7.3.2 Programme design and implementation

A key component of the design of a focused and cost-effective monitoring programme is an understanding of dominant physical, biogeochemical and ecological processes that govern the 'cause-and-effect' linkages between the receiving environment and the wastewater discharge. This design should also take into account modifications to such processes resulting from existing human activities. Such information is collated as part of the *Scientific and Engineering assessment* component (refer to Section 6).

NOTE:

Where a monitoring programme is designed for an <u>existing</u> wastewater discharge, it is crucial that an understanding of the dominant physical, biogeochemical and ecological processes be obtained prior to implementing any such monitoring initiative, even if a only conceptual model can be formulated. The procedures to be followed in such circumstances will be similar to those described for the characterisation of physical, biogeochemical and ecological processes of the study area as part of the **Scientific and Engineering assessment** ***component (refer to Section 6).***

In areas in which the marine environment receives waste inputs from multiple sources, monitoring programme design efforts need to be combined and integrated to ensure cost-effective usage of both human and financial resources. For example, in Saldanha Bay, the authorities and industries disposing of waste to the bay established the Saldanha Bay Water Quality Forum Trust, to enable each authority/industry to contribute towards an integrated monitoring programme (Taljaard & Monteiro, 2002).

Setting spatial boundaries for a monitoring programme is important because inappropriate boundaries might focus efforts away from driving or consequential factors (ANZECC, 2000b). The anticipated influence of the proposed discharge, therefore, needs to be taken into account when selecting the spatial boundaries. This selection, in turn depends on the transport and fate of the wastewater plume, both in the near and far field, as well as potential synergistic effects associated with other anthropogenic activities occurring in the area.

Numerical modelling has proven to be very useful in enhancing the design of monitoring programmes and improving the interpretation of the results of monitoring. Such numerical models provide the process links that enhance the ability to diagnose problem areas as well as to anticipate problems through their predictive capacity. The benefits of numerical modelling in the design of long-term monitoring programmes include:

- Definition of the most critical space- and time-scales of impact in the system: Important insights are provided by the combination of the synthesis of the existing understanding of the key processes and the model assumptions and inputs.

- Improve interpretation and understanding of the monitoring results in the context of a dynamic environment that determines the transport and fate of pollutants.

The aim, therefore, is to use the capability of numerical models (set up as part of the *Scientific and Engineering Assessment* component – Section 6) to reduce uncertainties in relation to system variability, key processes and, how these influence the transport and fate of contaminants. Because this increased understanding provides greater confidence in the predicted outcomes, allowing the investment in the monitoring to be limited to only a number of critical parameters measured on critical time and spatial scales.

i. Measurement Parameters

The selection of measurement parameters is site-specific. A key determining factor in the selection of such parameters is the composition of the proposed wastewater discharges as well as anticipated effects on the physical, biogeochemical and/or biological characteristics of the receiving environment.

Biogeochemical parameters (e.g. pH, dissolved oxygen, turbidity, particulate organic carbon and nitrogen, dissolved nutrients, toxin concentrations and microbiological parameters) can be measured in the water column, and/or sediments, including interstitial waters.

NOTE:

Geochemical ratios of trace metals can be used to determine whether the trace metals are of natural or anthropogenic origin. It is possible for conditions to arise in which the total trace metal concentration in the sediment is high (particularly in depositional areas) but completely linked to the natural structure of clay minerals, in which case the trace metals will not be bio-available. This condition would be characterised by geochemical ratios very similar to those of unpolluted sediments typical of the area. The geochemical ratio of each trace metal relative to aluminium (TM [µg/g]: Al [%]) is used, usually allowing a conservative two-fold natural variation in the geochemical ratios. Natural geochemcial ratios are site-specific for different geographical regions and need to be sourced from the literature (Monteiro & Scott, 2000).

Stable isotope ratios of nitrogen ($\delta^{15}N$) and carbon ($\delta^{13}C$) have also been used successfully in tracing the origin of organic matter in marine sediments, e.g. sewage-derived organic matter versus organic matter from natural, marine origin. The use of stable isotope ratios in this manner is discussed in more detail in Sweeney et al., 1980, Vivian, 1986 and Monteiro, 1997.

Toxins (e.g. trace metals and hydrocarbons) can also be detected through the *biotic component* of the receiving environment. For example, filter feeders (such as mussels and oysters) are internationally recognised as suitable bio-indicators for trace metal and hydrocarbon accumulation in the marine environment (Cantillo, 1998). These organisms filter food from the water in which they live and tend to retain contaminants that often accumulate to high concentrations in their tissues, thus reflecting changes in water quality over time. Their sedentary nature also prevents confusion about where a filter feeder might have accumulated a particular chemical compound.

NOTE:

A range of specific organisms can be used as indicators for the likely presence of pathogens and viruses in the receiving marine environment (Taljaard et al., 2000). In South Africa, faecal coliforms (**E. coli*****) are used as indicator organisms. However, new developments in health monitoring propose the expansion of monitoring parameters to include a selection of indicator and/or pathogenic organisms rather than a single indicator organism. The extended surveys need not be conducted frequently, but, for example, could be conducted every three months. In doing so, the assessment of the extended list of parameters assists in detecting anomalies which may otherwise pass undetected when assessing only a single group of organisms. For example, sampling at quarterly intervals for a range of micro-organisms such as*** **E. coli, Salmonellae, Salmonella typhi*****, pathogenic staphylococci, shigellae, helminth parasites and faecal streptococci has proved useful for auditing purposes to detect problems which may have passed unnoticed in routine*** **E. coli** ***assessments. Measuring more than one indicator for microbiological contamination, together with salinity (as a physical parameter for assessing the dilution or impairment of pristine seawater), has proved valuable for interpretation purposes (Rathbone et al., 1998; CSIR 1999).***

Other studies that assessed the suitability of indicators for the likely presence of pathogens and viruses in the receiving marine environment along the South African coast include Grabow **et al*****. (1989), Idema and Kfir (1990), Genthe (1996) and Grabow*** **et al*****. (1999).***

Measurement parameters can also include biota for ecological evaluations. In essence, biological measurements involve field measurements on the relative abundance, species diversity and community structure and composition of the biological communities in the study area to establish any change as a result of the wastewater discharge/s. Changes need to be measured against a pre-impact baseline data set and/or appropriate control sites.

As it is often very expensive to conduct detailed ecological monitoring programmes, that measure entire biotic communities in a particular study area, it is necessary to select indicator species that could be used as proxies for evaluating ecosystem health. In studies throughout the world and also in South Africa, macroinvertebrate communities have been used successfully in assessing ecosystem health as part of monitoring programmes for wastewater disposal to the marine environment (ANZECC, 2000b; CSIR, 2001a; CSIR, 2003 a & b).

Meiofauna distribution patterns, in conjunction with related biogeochemical parameters, have also been used successfully in this regard, for example, in intertidal areas along sandy beaches (Skibbe, 1991; Skibbe, 1992). These are, therefore, useful measurement parameters to include in the monitoring programmes associated with wastewater discharges to the surf zone, in which the wastewater plume is typically trapped in the surf zone with significant risk of negative impact on inter-tidal zones.

In South African estuaries, macrophytes have also been used successfully as long-term indicators of ecosystem health (CSIR, 2003b).

Other biotic parameters that have also been used are fish (ANZECC, 2000b), particularly in areas that support resident populations, such as estuaries, shoals, reefs and settlements on moored substrates.

Where the study boundaries of a monitoring programme include areas that support biotic species of economic importance (e.g. the prawn populations on the Thukela banks off the Kwazulu-Natal coast), these species could also be considered as measurement parameters in ongoing monitoring programmes.

However, it is important that scientifically sound reasons are provided for the selection of specific biotic indicator species in a particular project. Before choosing a particular taxonomic group(s) as a measurement parameter of ecosystem health, it is important to also test it against the following criteria (ANZECC, 2000a &b):

- Sensitive to potential impacts for waste inputs
- Measured response will reflect the ecological condition or integrity of the study area
- Approaches to sampling and data analysis can be highly standardised
- Response can be measured rapidly, cheaply and reliably
- Response has some diagnostic value.

A useful checklist that can be used to assist in the selection of suitable measurement parameters, in general, is provided in Table 7.12 (ANZECC, 2000b).

TABLE 7.12: Checklist for selection of measurement parameters (from ANZECC, 2000b)

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 that need to be resolved?*

ii. Spatial scales

The selection of sampling locations largely depends on the transport and fate of the wastewater plume, both in the near and far field.

Traditionally, monitoring programmes to evaluate ecosystem health included intensive sampling grids to overcome the inherent uncertainties of the spatial (and temporal) variability of the system. However, with the use of numerical modelling, many of the inherent problems of the traditional approach can be overcome in that numerical modelling assists in defining the most critical space- and time-scales at which monitoring will need to be done in order to obtain the desired output.

Although long-term monitoring programmes may, initially, still require relatively intensive spatial scales to address uncertainties in a system's response, over a number of years, these can be reduced to only a few selected points, through an iterative process, as the predicted responses of the system are verified. These high sensitivity points, however, need to be justified on the basis of specific criteria, such as high concentrations of mud and silt that indicate a long-term depositional area.

Specific beneficial use sites within the study area can also indicate sampling locations. For example, recreational beaches and mariculture areas will be logical sampling locations in instances in which the transport and fate of a wastewater plume poses a potential risk to human health.

iii. Temporal scales (sampling interval or frequency)

The temporal scale of a monitoring programme (i.e. sampling frequency) largely depends on the:

- variability in the load of contaminants from waste inputs
- variability in processes driving transport and fate in the receiving environment
- temporal sensitivity of the ecosystem to contaminant loading, i.e. exposure time versus negative impact.

The sampling interval or frequency should at least resolve the main source of natural variability of the constituent under investigation. Scales of change over time differ widely in the water column (minutes – days) compared, for example, with sediments (days – seasons – decades). Non-periodic events, such as storms, can also have a dramatic influence that needs to be taken into account where appropriate.

In the water column, high frequency physical processes, such as tides, currents, wind and waves, mainly control variability. A sampling frequency that is too low relative to the underlying natural variability, therefore, will result in biased data which will make it difficult to separate a human-derived impact from a natural anomaly. In the same way, sampling at a frequency that is too low relative to the variability in waste inputs may result in marked negative impacts being missed. In order to resolve the variability in the water column, sampling frequencies generally have to be high (e.g. hourly-daily-weekly). As a result, the use of water column measurement parameters as part of monitoring programmes is usually not cost-effective.

An instance where water column sampling is considered to be the most appropriate means of addressing the issue of concern is the monitoring of the microbiological status (e.g. *E. coli*) of marine waters, e.g. at recreational or mariculture sites at which a waste input poses a potential risk to human health.

If the cost implications of a monitoring programme based on highly variable parameters, e.g. those in the water column are taken into account, it is usually more appropriate to focus effort on those environmental components that tend to <u>integrate</u> or <u>accumulate</u> impacts or change over time, such as sediments and organisms (e.g. filter feeders).

Sediment sampling frequency is strongly linked to the time-scale which the sampled sediments act as 'particle traps'. As with sampling of the water column, sediment sampling at a frequency that is lower than the periodic re-suspension events will make trends difficult to interpret and could lead to spurious conclusions. Therefore, where cost constraints necessitate limitations on sampling frequencies, it will be inappropriate to select sampling locations that are situated in areas reflecting short-term variability. In such instances, longer-term depositional areas should rather be targeted. For example, because sediment processes often show strong seasonal trends, sampling is often confined to a particular season.

NOTE:

Depositional sites can be designated both long- and short-term. For example, an open coast site may be a depositional site for a period of days to weeks whereas an estuary may be such for a period of months to years. The ecological impact of both does not have to be linearly related to the persistence. Both provide important insights into the sediment and pollutant dynamics of the coastal and estuarine environments. This information provided by both is key to the design of optimal monitoring programmes, particularly in terms of sampling frequency.

Another commonly used technique to partially overcome the problem of high frequency variability is to measure seasonal variability of constituents in filter feeders. Where filter feeders are used to measure variability of constituents, it is important to realise that the body mass of these organisms also has a strong seasonal variability related to spawning cycles. Thus, a variability factor for contaminants inside the organism's tissue is introduced that needs to be separated from potential long-term signals caused by human interference. To achieve distinction, two procedures have to be implemented:

- Samples need to be taken at 6-monthly intervals
- Long-term sampling needs to be undertaken within a narrow time window each year to reduce seasonal uncertainty.

iv. Sampling and Analytical Techniques

The choice of an analytical technique to apply in a monitoring programme is largely dependent on the selection of measurement parameters and the output that is required to evaluate properly whether monitoring objectives are complied with. Key requirements that need to be stipulated in a sampling programme include:

- Sampling technique
- Number of replicates (determined by the statistical technique used in analysis)
- Sample handling and storage.

It is strongly recommended that an appropriately accredited analytical laboratory conduct chemical analyses of marine biogeochemical parameters.

NOTE:

The WHO's suggestion for the design of a monitoring programme for recreational use (e.g. at bathing beaches) is provided below (WHO, 1999). The design is based on an intensive study to test the scientific validity of the long-term sampling approach. It is particularly important to determine the effect and influence of temporal, spatial, and environmental factors, and the intensive study must be adequate to capture all of these factors to ensure that the results of design exercise portray the actual situation at the specific beach.

Measurement parameters. Two microbiological indicators of faecal contamination are recommended by the WHO, namely faecal streptococci/enterococci and Clostridium/Clostridium perfringens. In addition to the microbiological indicators, data on salinity, pH, turbidity, water and air temperature are also collected to assist in the interpretation of results. Other important observations that also need to be recorded include:

- ***Rainfall***
- ***Cloud cover***
- ***Wave height, current direction and speed***
- ***Tidal state and magnitude***
- ***Wind direction and speed relative to beach***
- ***Bather and animal population***
- ***Beach debris and sanitation***
- ***Location and status of stormwater drains, wastewater discharges and beach facilities.***

Spatial scale. To determine the spatial scale on which samples need to be collected to be representative of conditions at a specific beach, a 20-m grid along the width of the bathing area and stretching from the high-water mark to the end of the swimming zone, should be sampled during the intensive study. The sampling grid should cover water depths from ankle deep (0.15 m), through knee deep (0.5 m) to chest deep (1.3 m).

Temporal scales. To determine the sampling frequency, intensive studies should be performed at least over the period of a typical bathing season, e.g. 3 months. A minimum of 50 days, during which sampling is done hourly over a 12-hour period (during the day), is recommended. Because conditions vary in response to the density of users and discharge patterns from WWTW (e.g. peak uses often occur on weekends and holidays), sampling days need to be representative of all situations. Because rainfall events can also affect conditions, typical rain events also need to be included in the sampling period. Samples should be collected from all sites along the grid.

Statistical analysis of the above data set and an evaluation thereof are then used to determine the temporal and spatial scales for a representative and cost effective long-term monitoring programme.

7.4 DATA EVALUATION

Computers and statistical software are valuable tools for the evaluation of environmental data. However, they are only tools: the ultimate assessment depends on scientific expertise, as well as a proper understanding of statistical procedures and their applicability to environmental data. Where statistical expertise is limited, commercially available software packages (or the techniques described in the following section) must be used cautiously. Statistical techniques that are applied inappropriately can result in erroneous results or interpretations.

7.4.1 Data qualification

To ensure compatibility with the wide range of statistical and analytical techniques that are available for the analysis of environmental data, it is crucial that different types of data be logged in pre-defined formats that have been specified internationally (UNESCO/WHO/UNEP, 1992). This will ensure that the transfer of data to targeted users (e.g. researchers, authorities and other interested and affected parties) occurs in the shortest, most effective manner.

Data types can be in any of the following formats:

- Date and time
- Numerical data, including measured, recorded, counted or estimated data
- Alpha-numeric or text format, including observations, description, comments or reference to stored visual material such as digital photos or videos
- Boolean, e.g. yes/no, true/false, present/absent, visible/not visible or occurrence/non-occurrence.

In the case of measured numerical data, they can be:

- Continuous data, for which theoretically an infinite number of values can be assumed between two given values, such as, temperature, salinity or constituent concentrations. For a given sample set, frequency distribution can be either normal or non-normal.

- Discrete data, which are typically counts that results in a discrete, discontinuous data point, e.g. biological data sets. In the case of discrete data, the frequency distributions will generally be non-normal and data manipulation will be required before distribution related statistics can be applied. Without manipulation, non-parametric procedures (e.g. median and percentiles) can be applied.

The required accuracy and precision of data need to be clearly defined before embarking on data acquisition exercises. Rounding-off and the number of significant figures must be defined for each type of data. A high level of confidence with regard to data accuracy is essential for any further analysis.

Where data sets contain values that deviate from the normal cluster of values (referred to as outliers), the validity of these needs to be determined. Where outliers cannot be rejected, approaches need to be applied to analyse such data, including (UNESCO/WHO/UNEP, 1992):

- Application of non-parametric techniques to the entire data set
- Comparison of different statistical analyses on the data set, with and without the outliers, to decide whether the outliers should be part of the data set or not.

Where the measured data are either below or above the detection limit of the analytical technique, they are typically recorded as ‘not detectable’ or ‘below limit of detection’ or ‘maximum detectable level’. Further analysis of such data, as part of a larger data set should be undertaken with caution. More than one approach should be followed to determine the extent of the influence of such data on results. Ways in which to deal with such data are (UNESCO/WHO/UNEP, 1992):

- Use all values
- Ignore all ‘not detectable’ or ‘below limit of detection’ or ‘maximum detectable level’ values
- Use all values, but replace ‘not detectable’ or ‘below limit of detection’ values by 0
- Use all values, but replace ‘not detectable’ or ‘below limit of detection’ values with the detection limit (e.g. 0.5)
- Determine median values (a parameter which is insensitive to extreme data).
- Truncation of data set, by trimming equal amounts of data on both ends of the distribution
- Replace ‘not detectable’ or ‘below limit of detection’ values by the same number of next greater values and replace the same number of greatest values by the same number of next smallest values.

7.4.2 Analysis of data

NOTE:

This section provides a basic overview of statistical techniques that are typically applied in the evaluation of environmental data, where and when appropriate. It highlights important factors that need to be taken into account when applying statistical analyses to data sets and is by no means exhaustive. It is strongly recommended that the reader apply the information provided in this section in conjunction with the detailed statistical sources or with the assistance of a qualified statistician. The information was primarily extracted from UNESCO/WHO/UNEP (1992), Devore and Farnum (1999), Spiegel (1972) and US-EPA (2002c).

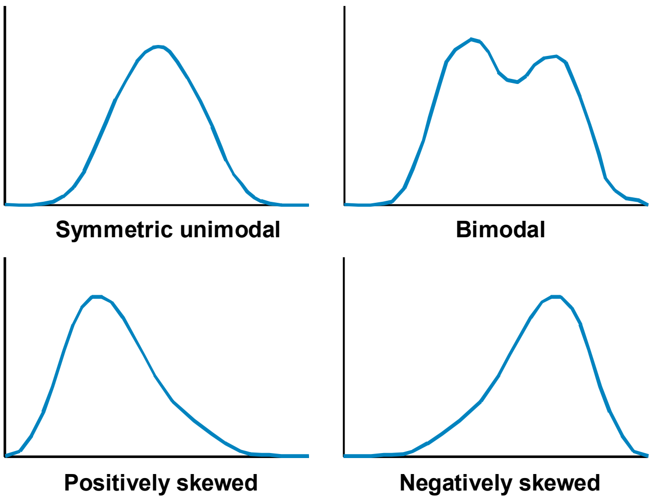
The selection of statistically techniques depends on the type of monitoring programme. Clarke and Green (1988), for example, demonstrate the complexity in the statistical design and analysis for biological studies.

i. Frequency distribution

The bulk of environmental data is in numerical format. These numerical data sets typically have either:

- No specific frequency distribution or
- Normal (or Gaussian) frequency distribution or
- Non-normal (skewed) frequency distribution.

The frequency distribution is determined by counting the number of times each value, or class of values, occurs in a data set. Typical frequency distributions are as follows:



These distribution patterns determine the type of statistical methods that can be applied to a data set with confidence. If the distribution of data does not conform to a specific distribution type, statistical output can be erroneous as a result of the application of invalid techniques.

The frequency distribution of data sets can be determined by:

- Visual methods
- Graphical methods
- Statistical methods.

Visual tests are the simplest procedures for reviewing the distribution of data sets.

This can be done by visual comparison of a histogram of a data set (values divided in a limited number of equal intervals) for the typical bell-shaped characteristics of a normal distribution. As a rule of thumb, the number of intervals can be taken as the square root of the number of values of the data set.

Graphical methods provide a more rigorous determination of the distribution frequency. For a normal distribution, the cumulative frequency distribution will plot as a straight line (at least between 20% and 80%) on a log/normal probability graph.

More rigorous statistical methods can be used, e.g. the Chi-Square Goodness of Fit Test, where the number of samples in a data set is more than 20. Many other statistical tests are available and a qualified statistician should be consulted in the selection of an appropriate method for a specific data set. Examples of such methods are provided in Devore and Farnum (1999).

Other well-known statistical methods that are applied to determine the extent of distortion from a normal distribution frequency are the:

- Skewness check, in which data are not symmetrically distributed about the mean (e.g. the median is not equal to the mean)
- Kurtosis check, in which a data set has a more or less “peaked” distribution curve compared to a normal distribution frequency.

Depending on the frequency distribution, either parametric or non-parametric descriptive statistical techniques can be applied. Non-parametric techniques can typically be applied to any data set. However, parametric techniques should only be applied to data sets showing a normal frequency distribution. Where a data set is non-normal, mathematical functions (e.g. logarithms and square roots) or more advanced transformation (e.g. Poisson) can be applied to transform the raw data into transformed values that may resemble normal frequency distribution. The parametric statistical methods can then be applied to the transformed values.

Where probability statistics are applied, it is important to provide an indication of reliability as well as the degree of confidence in estimating the mean. Standard statistical handbooks can be consulted for methods to determine of the standard error and confidence limits, e.g. Devore and Farnum (1999).

ii. Basic non-parametric descriptive statistical techniques

These techniques are usually suitable for application to any data set because:

- They are non-dependent on the frequency distribution
- Computations are simple and results are more easily understandable
- Missing data or outliers have a lesser effect on computations and results.

Basic non-parametric parameters that can be calculated for a data set include:

- **Sample (or arithmetic) mean** (X)**,** which is a measure of the central tendency:

  $X = \Sigma x_i/n$ (for data set x_i, i = 1 to n)

  (The sample mean should be applied with caution as the effects of outliers or missing data can produce errors in results)

- **Median**, which is a measure of the central tendency that resists the effects of outliers or missing data. The median (M) for n data values is obtained by first ordering the data from the smallest to the largest value, then:

  M = $(n=1)/2^{th}$ value on the ordered list if n = odd
  M = average of the $(n/2)^{th}$ and the $(n/2+1)^{th}$ value if n = even

- **Range**, which is the difference between the largest and smallest value in a data set.

  $R = x_{max} – x_{min}$ or
  Middle-range (R/2) which is also a crude estimate of the sample mean

- **Percentile,** which is a value on a scale of one hundred that indicates the percent of a distribution that is equal to or below it.

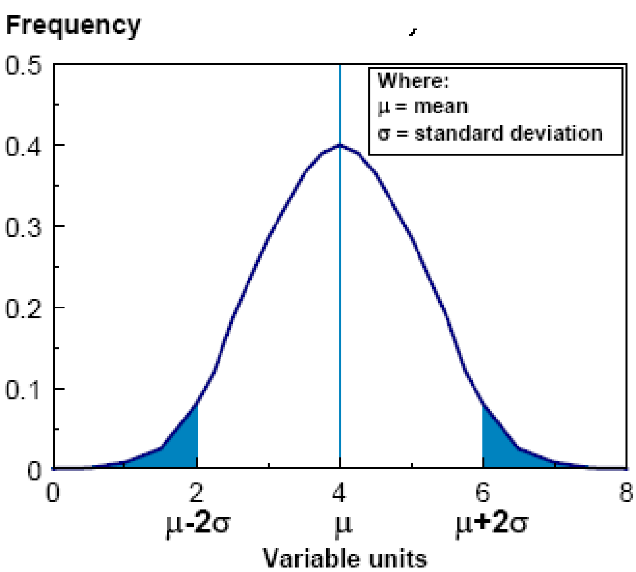
  P = p x 100/(n+1)

  For a given percentile the p value is:
  p = P (n+1)/100

- **Mode,** which is the value occurring most frequently.

iii. Parametric descriptive statistical techniques

Where data sets resemble a normal distribution (or where a transformation of the original data set conforms to a normal distribution), parametric statistical methods can be applied, in addition to non-parametric statistical techniques. A graphic illustration of a normal distribution is as follows:



Typical properties of a data set with a normal frequency distribution are:

- Values are symmetrical about the mean
- Mean (μ) = median = mode
- The curve is bell-shaped
- Extreme values will only constitute a small portion.

Basic parametric statistical parameters include:

- **Sample variance** (s^2), which is a primary measure of variability and which measures deviations from the mean ($x_1 - X$, $x_2 - X$, ..):

  $s^2 = \Sigma(x_i - X)^2/(n-1)$ (for data set x_i, i = 1 to n)

  (A simple way to obtain one quantity for the deviations is to sum the averages and divide by the number of samples. Because of the return of a 0 for this summation, the deviations are squared to obtain a measure of variability)

- **Sample standard deviation** (s), which is the square root of the sample variance (s^2):

  $s = \sqrt{s^2}$

- **Variance** (σ^2), which is the discrete distribution with a mean (μ) of a discrete data set (x_i) with probabilities p(x) is:

  $\sigma^2 = \Sigma(x_i-\mu)^2.p(x))$

  And for a continuous data set (x_i) with a distribution density function f(x) the **variance** (σ^2) is:

  $\sigma^2 = \int(x-\mu)^2.f(x)dx$

- **Standard deviation** (σ) is the square root of the σ^2:

  $\sigma = \sqrt{\sigma^2}$

  (The standard deviation is a measure of the spread of the data around the mean and is expressed in the same units as the mean.)

  An empirical rule for a data set with a normal frequency distribution is as follows:

  - Approximately 68 % of the values are within the standard deviation (σ) of the mean
  - Approximately 95 % of the values are within the 2σ of the mean
  - Approximately 99.7 % of the values are within 3σ of the mean

- **Sample co-efficient of variation**, which is the ratio of the standard deviation (σ) to the mean (μ) and is a measure of the extent of variability relative to the mean.

- **Probability density function.** Most environmental data are considered to constitute a stochastic process, which implicates an uncertainty in the results and that the process and the subsequent data from the process are random. Randomness is addressed by normal distributions and predictions can be made and hypotheses can be tested. The probability (between 0 and 1) provides an indication of the likeliness that an event will occur. The probability of an impossible event is 0 and a probability of 1 indicates that it is certain that the event will occur.

  By integration, the probability density function, f(x) for a normal distribution, can be expressed in terms of the mean (μ) and the standard deviation (σ) as:

  $f(x) = (1/\sigma\sqrt{(2\pi)}).\ e^{-0.5(x-\mu)/\sigma)2}$

vi. Time series analysis

Many mathematical techniques are available for the analysis of a time series (a stochastic process with a fixed sampling interval) for preliminary data exploration, detection of trends and forecasting. The most common method is using moving averages over appropriate intervals, which will eliminate irregular patterns and, depending on the width of the 'smoothing window', also cyclical variations. The method of least squares can be used to find the equation for a trend line or curve.

Other methods include the determination of the auto-covariance and the cross covariance functions.

The auto-covariance function provides an indication of any cyclic trends present in a data set, i.e. the dependency of a variable on itself in time. Testing for auto-correlation is important because, if it is present, conventional statistical analyses are not valid. The auto-covariance function is measure of the similarity between a variable and a time-shifted replica (lagged version) of itself and can be determined as follows:

For a time series x(t) $(t = t_1 \ldots\ldots t_n)$

$$\gamma(k) = \frac{1}{n-k} \Sigma \left[(x(t) - \bar{x})(x(t+k) - \bar{x})\right]$$

Where: **k = lag**
m = maximum lag
$\bar{x}$ = average of time series x(t)
n = number of data points

Characteristics of $\gamma(k)$ are:

- If x(t) and x(t+k) increase together then $\gamma(k) > 0$
- If x(t) decreases while x(t+k) decreases then $\gamma(k) < 0$
- If x(t) and x(t+k) independent then $\gamma(k) = 0$ for large number of samples
- If k = 1 then $\gamma(k)$ = sample variance ($s^2 = \Sigma(x_i - X)^2/(n-1)$)

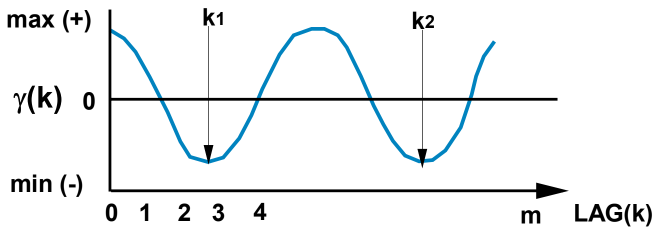
For a sampling interval Δt the periodicity can be determined as follows:

Period (units of Δt) = $\Delta t \times (k_2 - k_1)$

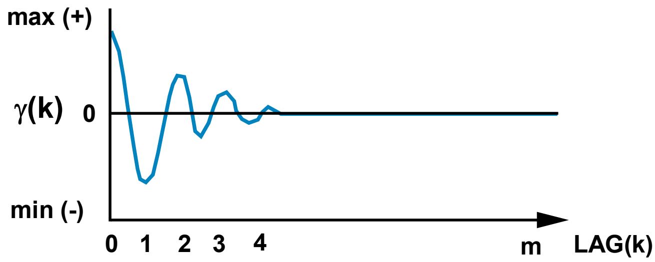
All cyclic events in the time series will be identified, if ($m \times \Delta t$) > period of the cyclic event.

The transformation of the auto-covariance to a discreet finite co-sinus series will result in the spectral density function (frequency domain).

A high correlation is likely to indicate a periodicity in the signal of the corresponding time duration. A schematic illustration of a time series that contains a periodicity and a random time series, is illustrated below:



Time series with a periodic component



Random time series

The cross-covariance function describes (or indicates) the inter-dependency of two variables in time (two time series with a fixed sampling interval). The cross-covariance function can be expressed as follow:

For two time series $x_1(t)$ and $x_2(t)$ $(t = t_1 \ldots\ldots t_n)$

$$\gamma_{x1x2}(k) = \frac{1}{n-k} \sum_{t=1}^{t=n-k} [(x_1(t) - \bar{x}_1)(x_2(t+k) - \bar{x}_2)]$$

$$\gamma_{x1x2}(-k) = \frac{1}{n+k} \sum_{t=1}^{t=n+k} [(x_1(t-k) - \bar{x}_1)(x_2(t) - \bar{x}_2)]$$

Where:
- **k = lag (k = 0, 1, 2, 3 ...... m)**
- **m = maximum lag**
- **$\bar{x}_1$ = average of time series $x_1(t)$**
- **$\bar{x}_2$ = average of time series $x_2(t)$**
- **n = number of data points**

Before a trend calculation is executed, the type of trend should be determined, i.e. if there was only a gradual change or where data before a specific time or event are distinctly different from the data collected after that time or event. It is beyond the scope of this guideline at this stage to discuss details of the numerous methods available, which are described in many statistical handbooks. A practical numerical solution (together with computer routines) for the calculation of the auto-covariance as well as cross covariance functions is described in Botes (1980), Botes (1986) and Liu (1974). Included in the program listings are practical methods to 'clean' time series data by applying digital band, low- and high-pass filters.

v. Hypothesis testing

Monitoring objectives can be stated as hypotheses that can be statistically answered, for example:

> There is no difference between sediment trace metal concentrations measured at two locations (e.g. one being a control site)

In proving the hypothesis it can either be the:

- Null hypothesis, showing statistically that the hypothesis is true, i.e. there is no difference, or
- Alternate hypothesis, statistically showing the hypothesis is false, i.e. there is a difference.

Statistical methods such as the *t*-test*, F*-test and the Analysis of Variance (ANOVA) can be used to test a hypothesis (UNESCO/WHO/UNEP, 1992).

A practical procedure of hypothesis testing from means of two data sets, with normal frequency distributions, is by using the *t*-test as provided below (derived from Spiegel, 1972):

For two data sets x_i and y_i (i = 1,2,3,.........,n_1 for x and n_2 for y)

Suppose: The means ($\mu 1$ and $\mu 2$) are unknown.

To test the null hypothesis: H_0: $\mu 1 - \mu 2 = G$

Where G is a given value.

$$\overline{x} = \frac{1}{n_1}\sum_{i=1}^{n_1} x_i \qquad \overline{y} = \frac{1}{n_2}\sum_{i=1}^{n_2} y_i$$

$$t = \frac{\overline{x} - \overline{y} - G}{\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}\sqrt{\frac{\Sigma x_i^2 - n_1\overline{x}^2 + \Sigma y_i^2 - n_2\overline{y}^2}{n_1 + n_2 - 2}}}$$

Using the *t*-parameter with a degree of freedom of (n1 + n2 -2) the null hypothesis (H0) can be tested (using standard *t*-distribution tables).

vi. Data correlations

An important consideration in the analysis of environmental data is to understand the relationships or interactions between different variables, as well as to determine temporal and spatial trends. The correlation between selected variables, as well as the behaviour over time, is important, for example for determining cause-and-effect relations or effectively predicting future change. The correlation between two parameters can be complex and not easy to establish. For example, temporal variations can bias the correlation of large data sets collected over extended periods.

The relationship (*r*) between two parameters can take many forms, for example:

- Linear positive
- Linear negative
- Logarithmic
- Exponential

- Power
- Trigonometric.

There could also be no relationship at all.

Typical standard methods of fitting a line to a set of data to test the relationship between two variables, which can be mathematically expressed, are:

Linear	$y = a + bx$
Non-linear	$y = a + bx + cx^2 + \ldots$
Exponential	$y = a.e^{(bx)}$
Logarithmic	$y = a + b.\ln(x)$
Power	$y = ax^b$
Trigonometric	$y = a + b.cos(x)$

A correlation coefficient is a quantitative measure of the strength of a relationship between two variables. The correlation coefficient (*r*) between two random variables ranges from -1 to 1. A value close to -1 indicates a strong negative correlation, i.e. the value of variable *y* decreases as variable *x* increases. When *r* is close to 1, a strong positive correlation exists between *x* and *y*; both variables increase or decrease together. The closer the value of *r* is to zero the poorer the correlation. For a normal distribution, a frequently used correlation coefficient is the Pearson's Sample Correlation Coefficient (Devore & Farnum, 1999).

Using the method of least squares, practical calculation or computation procedures of these typical relationships (derived from Spiegel, 1972) are as follows:

Straight line (Linear)

For variables x_i and y_i (i = 1,2,3,.........,n)

Equation: $y = a_1.x + a_0$

$$a_1 = \frac{\Sigma xy - \frac{\Sigma x \ \Sigma y}{n}}{\Sigma x^2 - \frac{(\Sigma x)^2}{n}}$$

$$a_0 = \overline{y} - a_1.\overline{x}$$

Where: $\overline{y} = \Sigma y/n$, $\overline{x} = \Sigma x/n$

$$r^2 = \frac{\left[\Sigma xy - \frac{\Sigma x \ \Sigma y}{n}\right]^2}{\left[\Sigma x^2 - \frac{(\Sigma x)^2}{n}\right]\left[\Sigma y^2 - \frac{(\Sigma y)^2}{n}\right]}$$

Logarithmic fit

For variables x_i and y_i (i = 1,2,3,.........,n)

Equation: $y = a + b \ln x$ ($x_i > 0$)

$$b = \frac{\Sigma y \ln x - \frac{(\Sigma \ln x)(\Sigma y)}{n}}{\Sigma(\ln x)^2 - \frac{(\Sigma \ln x)^2}{n}}$$

$$a = \frac{(\Sigma y - b\Sigma \ln x)}{n}$$

$$r^2 = \frac{\left[\Sigma y \ln x - \frac{\Sigma \ln x \ \Sigma y}{n}\right]^2}{\left[\Sigma (\ln x)^2 - \frac{(\Sigma \ln x)^2}{n}\right]\left[\Sigma y^2 - \frac{(\Sigma y)^2}{n}\right]}$$

continued…

Power curve fit

For variables x_i and y_i (i = 1,2,3,.........,n)

Equation: $y = a.x^b$ (x > 0, y > 0)

or: ln y = b ln x + ln a ... solve linearly

$$b = \frac{\Sigma \ln x \ln y - \frac{(\Sigma \ln x)(\Sigma \ln y)}{n}}{\Sigma (\ln x)^2 - \frac{(\Sigma \ln x)^2}{n}}$$

$$a = \exp\left[\frac{\Sigma(\ln y)}{n} - b\frac{\Sigma \ln x}{n}\right]$$

$$r^2 = \frac{\left[\Sigma(\ln x)(\ln y) - \frac{(\Sigma \ln x)(\Sigma \ln y)}{n}\right]^2}{\left[\Sigma(\ln x)^2 - \frac{(\Sigma \ln x)^2}{n}\right]\left[\Sigma (\ln y)^2 - \frac{(\Sigma \ln y)^2}{n}\right]}$$

Exponential fit

For variables x_i and y_i (i = 1,2,3,.........,n)

Equation: $y = a.e^{bx}$ (a > 0)

or: ln y = ln a + bx

$$b = \frac{\Sigma x \ln y - \frac{(\Sigma x)(\Sigma \ln y)}{n}}{\Sigma x^2 - \frac{(\Sigma x)^2}{n}}$$

$$a = \exp\left[\frac{\Sigma(\ln y)}{n} - b\frac{\Sigma x}{n}\right]$$

$$r^2 = \frac{\left[\Sigma x \ln y - \frac{\Sigma x \ \Sigma \ln y}{n}\right]^2}{\left[\Sigma x^2 - \frac{(\Sigma x)^2}{n}\right]\left[\Sigma \ln y - \frac{(\Sigma \ln y)^2}{n}\right]}$$

vii. Multivariate methods

Multivariate statistical methods are designed to evaluate more than one variable at a time. Examples of multivariate methods include (UNESCO/WHO/UNEP, 1922; US-EPA, 2002c):

- **Discriminant function analysis (DFA),** in which multiple variables are used to predict groups of cases
- **Multivariate analysis of variance (Manova**), testing if the groups are significantly different in terms of the independent variables
- **Principal components analysis (PCA)**, which takes multiple variables and defines a smaller number of new variables by constructing linear combinations of the original variables. The new variables separate the cases as much as possible.
- **Canonical correlation**, which uses two sets of variables and, like PCA, creates linear combinations of the initial variables in each set so that the number of variables is reduced. The new linear combinations are selected so that they maximize the correlation between the pairs of variables, one from each set.
- **Cluster analysis**, which defines groups of cases based on the similarity of multiple variables measured for each case

viii. Indices

Another method to assess environmental data is the use of such indices that are obtained by aggregating several variables into one number. This simplified expression of a complex set of variables can be used for a variety of purposes, including (UNESCO/WHO/UNEP, 1992):

- Communicating environmental information in a user-friendly format
- Planning tool for managing water resources
- Assessing changes in the quality of water resources
- Evaluating the performance of wastewater disposal systems.

Different indices are currently in use or under development. These can be grouped as follows:

- General water quality indices, in which physical, chemical and microbiological variables are aggregated to produce an overall index of water quality
- Specific use-quality indices, in which variables are aggregated on the basis of their importance to a specific beneficial use, e.g. a human health index
- Planning indices, which also include social and economic factors
- Chemical indices, which are calculated by grouping water quality variables according to their chemical nature, e.g. toxic metals.
- Biological indices, which are based on selected biological characteristics, including community diversity.

An example of a chemical and health water quality index system, based on a theory by House (1989), has been developed into a statistical computer model for the rivers of the Western Cape (Van Driel & Botes, 2003) to facilitate the transfer of monitoring assessments to Water Quality Forum meetings.

7.4.3 Presentation of data

To be useful from a management perspective, data must be presented in a clear format to provide the appropriate scientific and engineering knowledge for informed and effective decision-making. The most effective manner in which to communicate environmental data and information is through graphical presentation of numerical and statistical data. The advantages of the graphical presentation of data are:

- Large data sets can be illustrated effectively
- Qualitative aspects, such as correlations and trends, as well as quantitative aspects such as outliers, are illustrated effectively
- Provides a user-friendly means of communicating complex numerical and statistical outputs.

There are numerous ways to represent data graphically. Typical presentation formats include:

- Time series plots, which constitute a simple means to illustrate trends, cyclical variations and outliers.
- Plots to illustrate spatial and temporal variability (e.g. contour plots, scatter plots, and bar graphs), for example, to show the spatial or temporal effects of pollution sources in an area
- Statistical summary of variable, for example, using box and whisker plots (e.g. ranges, mean, percentiles)
- Seasonal or periodical variations, illustrated through plots of statistical parameters (e.g. ranges, percentiles, means) of discrete monthly or seasonal data sets collected over a long period
- Correlation plots between two or more variables, illustrating relationships between these different variables.

7.5 REPORTING

7.5.1 Compliance Monitoring

Evaluation of data on the wastewater volume and composition needs to occur in 'real-time' (i.e. as close as possible to the time of sampling). This is to ensure that management actions can be implemented, as soon as possible, to mitigate negative impacts on the receiving environment.

Evaluation of the data comprises a comparison of the measured data against the pre-determined critical limit specified for a particular constituent in the licence agreements issued by the DWAF under section 21 of the National Water Act.

7.5.2 System Performance Monitoring

Monitoring of a wastewater disposal system includes:

- Physical inspections of the outfall system (for offshore marine outfalls)
- Hydraulic performance (for offshore marine outfalls).

A short description of the reporting format is provided below.

Physical inspection. Underwater video recording supplemented by a short written report on the observations needs to be submitted to the DWAF.

Hydraulic performance. A report on the evaluation of the hydraulic performance of the outfall prepared by a qualified environmental consultant, needs to be submitted to the DWAF. In the event of non-compliance with design criteria, management strategies and actions to improve the situation need to be specified as well.

7.5.3 Environmental Monitoring

Depending on the measurement parameters selected, the monitoring data on the receiving marine environment that are typically collected as part of a monitoring programme for wastewater disposal to the marine environment, include:

- Biogeochemical parameters in the water column
- Biogeochemical parameters in sediments (including interstitial waters)
- Biogeochemical parameters (e.g. toxins) in biological tissue
- Data on relative abundance, species diversity and community structure and composition of biological indicators.

A detailed Monitoring Report must be supplied to the DWAF and other relevant authorities (where and if required). The report should contain:

- A list of monitoring objectives (or hypotheses) and how these relate to the overall Environmental Quality Objectives specified for the study area
- Details of the design and implementation of the monitoring programme (also indicating the relationship between selected measurement parameters and monitoring objectives)

- An evaluation of the monitoring data in relation to the monitoring objectives (or hypotheses). This evaluation should make use of data summaries and graphical presentations in order to enhance readability
- A statement on whether the monitoring objectives have been met
- In the event of non-compliance, possible reasons for the non-compliance
- Management strategies and actions required to address non-compliance
- Recommendations on refinements to the monitoring programme
- Appendices containing cruise and laboratory reports, raw data tables and other relevant background information.

A reporting schedule needs to be negotiated with the DWAF, but it is realistic to submit a Monitoring Report after each set of monitoring surveys. Where sampling components of the monitoring programme are conducted at frequent intervals, e.g. microbiological monitoring at recreation areas, the results of these samples need to be submitted in the form of monthly interim reports, clearly highlighting preliminary indications of non-compliance with monitoring objectives. In addition the results of these samples interim reports should contain management strategies and actions that will be taken to address any problem timeously.

Various communication routes can be utilised to communicate findings to wider stakeholder groups, which may not have the relevant scientific or engineering background, including:

- *Pamphlets*. The key findings of the main report can be summarised in layperson's language in the form of colourful pamphlets. These information-sharing documents are then distributed to the wider stakeholder groups either living in a particular study area or who are potentially affected by activities in the area.

- *Stakeholder meetings*. Stakeholders can also be informed through regular meetings at which the results of monitoring programmes are presented in a user-friendly manner.

- *Internet websites*. The Internet is a very powerful tool for disseminating information to a wide spectrum of interested and affected parties. Communication can either be undertaken on a case-by-case basis or it may be more appropriate to link these individual cases to a master website such as that maintained under the auspices of the DWAF.

- *Media reporting.* Media reporting, e.g. in a local newspaper, is another means of communicating information to the wider community that may not have access to the Internet.

7.6 DATA STORAGE FACILITIES AND QUALITY CONTROL

Monitoring data are expensive to collect and require substantial investments of both human and financial resources. As a result, such data must be made as usable, useful and retrievable as possible. The sheer volume of data generated as part of ongoing monitoring programmes dictates that computer-based data management systems must provide the basis for data storage and management (ANZECC, 2000b).

A good data management system should have (ANZECC, 2000b):

- Reliable procedures for the recording of analytical and field observations
- Procedures for systematic screening and validation of data (quality control)
- Secure storage of information
- Simple retrieval system
- Simple means of analysing data
- Flexibility to accommodate additional information.

Although detailed specifications are site-specific, the following are important factors to consider in the selection of the hardware framework and software platform (UNESCO/WHO/UNEP, 1992):

- Physical capability of handling the data load and the implications of expansion
- Ability to transfer data to other data storage systems
- Compatibility of the data storage with facility available data sources
- Choice of software (commercial or developed) with regard to:
  - Expansion and support
  - Delivery of suitable outputs
  - Robustness with regard to changing data formats
  - Data transfer capabilities between different platforms (e.g. Desktop PC, Server, Webserver, Palm devices, mobile phones and sampling/analysis equipment).

The needs of the end-users probably constitute the most critical consideration in the design of a data management system. Although the existing data management system of the DWAF (Water Management System [WMS]) may be suitable in some respects (e.g. storage of source data), the system caters mainly for rivers and groundwater. Consequently, WMS may not necessarily be suitable for capturing the type of data collected in the marine environment (including estuaries).

To ensure conformity across South Africa, it is, therefore, strongly recommended that investigations be undertaken regarding a national framework for setting up management systems to store data collected as part of ongoing monitoring programmes in the marine environment. As there are existing data management systems for the marine environment (e.g. South Africa Data Centre for Oceanography [SADCO] - http://fred.csir.co.za/ematek/sadco/) and the Department of Environmental Affairs and Tourism that are operational in this domain, it is crucial that such a framework be developed in consultation with other relevant departments, taking into account linkages to existing management systems.