Chapter 4

Water resources model inputs and pre-processing

4.1 Data and information in water resources modelling

4.1.1 Data capture and accumulation

Data constitute the foundation on which the knowledge of our environment is built. Unfortunately, southern Africa, like many other developing regions, has limited water resources, hydrological and climatological data, with data recording stations being sparsely distributed and in many cases having been closed (Lynch, 2003). The poor quality and distribution of data resources affect the type of modelling tools and the detail that can be accurately accounted for in water resources management and planning models. A number of models that require detailed data inputs such as the MIKE-SHE model (DHI, 2000), a distributed physical model, have little application in South Africa given the state of our data resources. The tendency is therefore to develop and utilise less detailed models and work towards building the data resources to ensure that the country will have adequate data in future. Decisions on the data characteristics have to be considered as part of many interrelated variables. These interrelated variables include: the nature of the water problem to be solved through modelling, the models available, expertise of modellers, data quality and availability, as well as the stakeholder expectations on the overall solutions to the water problem.

Water practitioners in water resources management and planning should ensure that the database developed for modelling processes, covers data in the following areas:

- Hydrological variables
- Soils
- Catchment demarcations and spatial information
- Legal instruments, policies and water infrastructure operation rules

- Water quality
- Groundwater variables
- Water storage systems and their characteristics
- Water demand and land uses
- Biota and abiotic aspects of the watercourses
- Results from previous water management studies
- Identified scenarios and stakeholder expectations
- Available resources for solving the water resources problem
- Water systems and linkages
- Watercourse and flow generation hydraulics
- Water transfer systems and their characteristics

Often, large amounts of data are gathered for specific projects without making provisions for how these data should be archived for later use. Ideally, data capturing and accumulation should have a long-term objective and should seek to store such data where they can be accessed by other users. The information contained in a robust data-management system is then available, not only for the use for which the data were collected originally, but also for a multitude of uses that may never have been anticipated. The NWA (Republic of South Africa, 1998) makes specific provision for the establishment of national information systems on water resources. DWAF must ensure that all water resources project clients and contractors integrate their data capturing and storage efforts with the national initiatives at all times. The need for national information management systems and the guidelines developed by the national body (DWAF) which was tasked, through the Water Act, to develop these data and information management systems, will in future provide the national guidance to all stakeholders who are involved in data generation and archiving.

Water resources models require larger databases when GIS is incorporated. The use of object-oriented databases with relational tables is one approach that is widely recommended (McKinney *et al.,* 1999). Polhill *et al.* (2002) also noted that the use of a relational modelling structure in the database allows for the possibility to vary the simulation scales in a single model, thus allowing detailed high resolution simulations as well as large-scale assessments within one water resources model. Database developers in water resources should clearly consider the implications of the various database architecture selection options to ensure that the final database is technically sound and appropriate to the project needs. Ideally the water resources database should:

- Be compatible with national initiatives in the water sector;
- Incorporate the preferred local and international data inventory trends;
- Allow for transparent linkages between data and the model; and
- Have an ability to statistically integrate and display uncertainties in spatial and temporal representations.

With large water resources databases, water resources modelling is faced with the problem of selecting the most appropriate data components. Winiwarter and Schimak (2002) pointed out that in the past model data inputs were limited by computing power but recent technological improvements have addressed the computer limitations. Rather, the problem centres on complexities of identifying and limiting data inputs to those data with the highest quality. However, these input data limits should not compromise the quality of model outputs, for example if they are applied beyond reasonable limits.

4.1.2 Data quality and formats

The use of a single set of high quality data by all users is a very important aspect of water resources management and planning which can directly eliminate unnecessary inconsistencies. Garry (2002) pointed out that efficiency and effectiveness regarding availability of information to business is enhanced by the existence of a “single version of the truth” across functions and disciplines as a one stop service to water stakeholders in different fields and at a variety of levels in the data and information hierarchy. Newman (2004), however, advised that, after establishing a “single version of the truth” in data management, the challenge remains that stakeholders have to be empowered, so that they share a common view of this truth, thus insuring acceptability and use of the recommended “truth”.

For each water resources problem, as much data as possible, especially the latest data should be identified and utilised in water resources modelling to give comprehensive and up to date solutions. However, exceptions, do exist. As an example, modellers may seek to compare outputs generated using the same data after changing model variables, in which case they will use the same old data inputs as those which were used in the previous study.

Local streamflow and water quality records are intermittent, irregular, and frequently contain periods of low accuracy observations. Modellers should note that streamflows and pollution loads are some of the only empirical clues to the true responsiveness of the catchment to rainfall and human impacts. Ideally, systematic efforts have to be made to re-process incomplete or low-accuracy flow records, to patch missing parts of flow records in plausible ways and to infill irregular water quality samples in a systematic manner. There is also a need to ensure that water quality constituents expressed as a monthly time series are calculated in a representative flow-weighted manner.

The presentation of data and their formatting affects what can be done with them, the choice of model, and compatibility of tools that will be required to access the data. The use of uniform standards to format and present data is an important component of accessible data. Samadi, Beukes and Remmelzwaal (2002) pointed out that data communication on a national and international platform based on a rigorous set of standards provides a common basis for all modelling and improves the applicability of performance measurement in model outputs. Any initiative to develop local data resources should account for present and possible future requirements to include such data in a common platform for the country and, ultimately, internationally.

Water resources data capturing and archiving should incorporate the associated meta-data, commonly referred to as "data-about-data". The meta-data should at least include the following where applicable:

- Recording institution,
- Methods of data collection used,
- Details of data monitoring programme/project in which data were collected,
- Relevant prevailing characteristics when data were collected,
- Processing done and tools used,

- Calibration details,
- Accuracy and range of measurement device and
- Details of modifications made including procedures used.

The absence of meta-data introduces low confidence in data users as they have to deal with the unknown risk of using the data. Additional challenges are also introduced in the modelling process as more attention must be paid to data and model verification, which increases project costs.

4.1.3 Water resources data transmission

The maintenance of agreed standards is important in water resources data collection and transmission. The World Meteorological Organization (WMO) recommended a set of standards to be adhered to when transmitting climate and hydrological data (WMO, 1994). These standards relate mostly to the data transmission formats, where codes are suggested to ensure uniformity and reliability of data transmission. The WMO codes were developed to meet the requirements for the exchange of meteorological data at basin, national and international level, and also to allow the data to be routed over the World Weather Watch telecommunications channels (WMO, 1994). Water resources practitioners should ensure consistent use of standards in data transmission to reduce data errors and reduce modelling risk levels.

4.1.4 Record length and scales

The length of time series hydrological data should adequately cover a wide range of hydro-meteorological events. DWAF (2001c) pointed out that 15 years represents an adequate minimum period for monthly streamflow simulation and, for daily streamflows, 10 years is recommended as a minimum. WMO (1994) also recommend that the basic recording gauges should be operated for relatively long periods of at least ten years. The length of period selected in water resources time series data should ideally cover all possible hydrological or climatic events within a water system, which are usually longer than the 10 to 15 year periods stated above as a minimum.

4.1.5 When to update a set of time series data

Updating of time series data is ideally done as soon as new data become available. However, this is time consuming and may not add value to the modelling process if the new records do not include new events that will improve the decisions made using the data. The internet has drastically improved the processes involved in updating time series data making the process time efficient. In cases where monthly data stretching over at least fifteen years are used, the user can update the inputs once every year. In the case where the available monthly data series are still less than the required fifteen years, data updates should be done more frequently to incorporate changes in each month. In the case of daily data, monthly updates should be preferred for any data sequences that are longer than ten years. For shorter records, daily updates will be very useful. Smaller time steps such as hourly data usually stretch over short periods of not longer than a few months or a couple of years. Users of such data are advised to use all the available data in their water resources modelling processes.

4.1.6 Model data input

The development of input data should ideally start with the development of suitable databases and data formats. Data, information collection and archiving should be structured in a way that enables easy utilization in the targeted model. Resources are wasted in cases where each data user has to format data before use. The conceptualisation of data input routines in water resources models should account for all of the readily available data formats to reduce the burden of converting data from one format to another. While the reduction of non-core activities in modelling such as data formatting have to be minimised, water practitioners are reminded to evaluate the benefits of developing data collection and capturing methods to suit specific tools.

4.1.7 Data review and analysis

Data quality controls should ideally begin at the data collection point such that primary data comply with high quality standards before they are given to users. A system of quality monitoring through inspections should be developed for all data recording stations.

Preliminary data review for data that are collected manually should ideally ensure completeness and correctness of the information supplied with the data. This information includes dates of collection, station name, station identification code as well as checking the completeness of the data and any calculations made by the observer. Checking the observed data against existing records is also recommended to highlight possible hidden inconsistencies.

Preliminary data review is usually followed by the use of specialist software and other simple computer based data assessments. The WMO guide to climatological practices (WMO, 2000) recommends ideal methods to be followed in assessing data. One important computer-based method for checking correctness of recorded elements involves the use of various mathematical relationships such as double mass plots, variability, variance, correlation, consistency. The use of more sophisticated software, especially in the key water resources inputs such as rainfall data, to plot station positions, together with their records to later interpolate data in space and then plot isolines, is recommended in cases where at least three data recording stations are available.

Even after taking stringent measures to ensure data completeness, missing and incomplete records as well as suspicious records are often found within the data-set. Missing records may be interpolated, estimated and patched in. It is important that all the estimated and interpolated values are indicated as such to ensure that users are adequately informed about the data prior to use. A description of the method used to estimate or interpolate missing data will also provide a reliable basis for the modelling process. Other errors and observations noted and resolved should also be recorded using meta-data codes or descriptions.

Preliminary data review should preferably be followed by a data validation process. In validation, standard checks are carried out on the data to detect errors in time and magnitude. Sequential readings or records are validated in the light of expected patterns or the simulated behaviour of related variables that have also been recorded. This assessment should result in the observer applying quality codes to the records indicating if the records are good quality, or faulty, and the level of confidence attached to them in terms of data accuracy. WMO (1994) pointed out that validation processes should never be completely fully automated but should be guided by experienced human judgement to avoid systematic errors in the automated process.

4.1.8 Data coding

Water resources management data requires coding to aid data processing as well as make the files more compact and less ambiguous. Consideration of existing national and international coding systems is important when deciding on a methodology to code data. Coding instructions and training should be made available to observers. The coding method selected should be suitable for further data use, including its use in models. The following is a listing of possible codes to be incorporated with the data:

- Location code (to indicate the place of recording)
- Variable/parameter code (The range of variable codes is enormous and includes: text definition of variable, and letters to represent other information about the data such as measurement units)
- Data-qualification codes (to qualify unusual or uncertain data. This code should also address current and background status of the data.)
- Missing data codes (to indicate data that were not recorded)
- Transmission codes (to ensure data are transmitted quickly and reliably)

4.2 Data processing

Poch (2002) pointed out that original raw data are often defective, requiring a number of pre-processing procedures before they can be registered in an understandable and interpretable way. WMO (1994) explained that data processing entailed transforming the raw data into forms that enable ready manipulation and efficient storage for prospective users.

Data processing involves several processes. Water resources managers have to ensure that the data utilised in their work are at least subject to the following data processes:

- Data preparation (data entry and coding)
- Data entry (input of data into electronic formats for immediate use or archiving)
- Validation (range checks, sum checks, consistency checks)
- Primary processing (standardisation of units, further data coding and data formatting)
- Database updating (adding data to existing databases by extending time series records)
- Secondary processing (data statistical summaries, routine reports, missing data infilling and interpolation)
- Retrieval processes (data and output device selection based on parameter type, location and period of record)
- Output processes (computer storage media, telemetry and plotting)

4.3 Data storage and dissemination

The present trend in water resources management is to build water resources databases that are linked to geographical information systems and modelling tools, which utilise the data. The water manager tasked with the development of the water management database should identify and comply with the initiatives on the national information systems on water resources provided for in the NWA (Republic of South Africa, 1998). DWAF is currently developing the basis for the national water resources information management systems which are discussed in DWAF (2003a).

Other important considerations in data storage and dissemination include the following:

- Identification of the data to be stored,
- Identification of data accumulation and dissemination methods,
- Development of data standards, formats and a data management plan in collaboration with the national processes,
- Definition of meta-data standards and development of a meta-data catalogue,
- Definition of intellectual property rights, confidentiality and other legal boundaries such as data exploitation exclusivity rights,

- Selection of the correct storage platforms with adequate storage facilities and a suitable storage and retrieval engine,
- Ensuring that the data storage and dissemination process allows easy data updates with software to track changes made during updates,
- Development of efficient and adequate dissemination processes which are made available to all users, with ease and at low cost,
- Compatibility with data use requirements,
- Ease of access by targeted users at all times and provision of appropriate restrictions for non-users,
- Development of a strategy for long-term data management,
- Development of quality assurance procedures,
- Installation of functional query handling processes and
- Installation of well maintained data and information computer backup systems.

4.4 Implications of data sources in WRM models

Available data and the data sources have major implications on the selection of modelling tools. The main data component of water resources models, rainfall, is usually available as point measurements from rainfall gauges but its application usually requires spatial representation. As a result rainfall representation in models remains one of the major challenges contributing to poorer model outputs. Alternative sources of data have to be considered depending on the intended accuracy levels in the models, the availability of other data sources and other constraints related to the sourcing of data from such sources. In the case of sourcing rainfall data, the following additional sources need consideration:

- *RADAR:* This has the advantage of giving spatial measurements, though availability is still very limited in South Africa. RADAR data are also prone to RADAR measurement errors which require more resources to resolve than the errors in rain gauge data.
- *Remote Sensing*: In South Africa, data from polar-orbiting satellites are often used in water resources. These data are not continuous as they are measured for a limited time when the satellite being used passes over a particular area. As an example

Landsat-7, provides data for each point on earth every 16 days. Polar-orbiting satellites, which are located closer to the earth (500 km-1500 km) than geostationary satellites (35,000 km) give better spatial resolution (30 m X 30 m resolution) and wider coverage (Dozier, 2003). The procurement of satellite data and its processing are expensive for most local water resources studies. Water practitioners are recommended to take advantage of recent initiatives in the Department of Agriculture which have made the data from the Landsat-7 Satellite freely available. Landsat images, including the archives that date back to 1986 are now freely available to all governments, research organisations and non-governmental organisations in SADC as a result of the South African Department of Agriculture's commitment to provide twelve million Rands towards procuring these data images (NDA, 2004).

The processing of data to provide a better representation of the study area, such as converting point data to give data series that can be applied over a defined catchment area, involves a number of possible techniques. Some techniques which water practitioners should take into consideration when dealing with rainfall data are as follows:

- *Inverse distance weighting*: weights are calculated depending on the distance between the location where an estimate is required and the locations where the rainfall is measured. However, this method is limited in that it never produces a rainfall value which is higher than the maximum value in the observed data-sets. This is not realistic.

- *Interpolation:* a mathematical relationship of surrounding point values with the area where aerial or point values are required is established and used to generate the required time series data. One such method, Kriging, utilises the covariance structure of the area where an estimate is required.

- *Multiple regression:* This method requires a sufficient spatial density of the locations at which the values are observed to explain the variation in the measured amounts. The method is most useful in mean annual precipitation estimations.

- *Stochastic generation*: The method generates synthetic data after establishing important statistics pertaining to the observed data for the different sites and preserving these statistics in the generated data estimates. The method can generate data series that are longer than available records; this is useful when long hydrological sequences are required in a water resources modelling scenario.

4.5 Summary of recommendations on model inputs and pre-processing

Water resources models are driven by data and , unfortunately, South Africa has limited hydrological and climatological data. The available records are characterised by sparse aerial coverages, poor quality, discontinuities in time, and unstructured data collection and archiving initiatives. Continuity of collection of time series data has been affected by the country's ongoing political transition where national priorities which targeted improvements in livelihoods and poverty alleviation have affected investments in data collection projects (Gill, 2004). The type, resolution and accuracy of modelling tools that can be successfully applied to any area are almost entirely dependent on the nature of available data. In the water resources simulation case study used in this research, a monthly hydrological modelling approach was selected rather than a daily approach that could have provided more detailed assessments, as this was the best resolution that could be used with the available data. In this case study, it was observed that the case study area, which exceeded 3 000 km^2 had only four reliable rainfall gauges and two runoff gauges. This sparse distribution of rainfall gauging points means that detailed physical models are not generally applicable in most catchments.

Apart from land-based recording gauges, other available sources of data have to be considered especially for the purposes of providing more reliable data. Here, data from remote sensing and radar technology are of particular importance. These forms of data are expected to become more widely applied and accepted in the future of Africa's water resources management (NDA, 2004). The development of tools for water resources management should seek to utilise remote sensing and radar data where appropriate.

The absence of a nationally supported and coordinated Earth Sciences data collection and archiving programme has meant that water resources data are not readily available. Water resources management projects need additional inputs to collate, process and format data in most water resources projects. Data secured from one project, must still be checked for quality and processing as there are no guarantees or applicable national standards on water resources data quality and data processing. The choice of tools used in providing information to decision making is limited by the state of the available data and the additional resources required to process these data. The process of data collection, updating, formatting, archiving, review, analysis, coding and dissemination should ideally be done according to nationally and internationally recognised standards. Important international guidance is provided in WMO (1994). On the local scale, direction will soon become available through the NWA-based initiatives (DWAF, 2002a).

Chapter 5

Model software selection and development

5.1 Policies and a framework in development and use of models

The NWA (Republic of South Africa, 1998) is often described as an "enabling" piece of legislation. While this Act provides little in the way of specific regulatory procedures, it does stipulate the types of standards and tools that should be used for integrated water management approaches, as emphasised in the National Water Policy (DWAF, 1997d), and this provides a platform for flexibility. The strength of this approach is that it enables the flexibility that is required in regulating the dynamic water sector. In spite of the National Water Act being enabling, the framework for the integrated management of water resources is adequately provided for via water resources strategies. At national level, the Act provides for the Minister to progressively develop a NWRS. This strategy must set out the objectives, plans, guidelines and procedures relating to the protection, use, development, conservation, management and control of water resources (DWAF, 2002a). Further provisions have been made in the NWA to develop IWRM on a catchment basis using the WMA units. This provision requires the formulation of water resources management strategies at WMA. The CMAs are expected to develop Catchment Management Strategies (CMSs) within each declared WMA. The establishment of common water resources approaches and techniques that will be guided by DWAF's national tools is ideally the key to the success of the 19 CMAs. Some of the main goals used by DWAF to provide providing common guidelines to CMAs include the following:

- To assist CMAs to adopt a consistent, technically sound and dependable approach to the evaluation of water resources for compulsory licensing,
- To assist in developing understanding of the scientific and technical information requirements of water resources management and,

- To provide a common basis for the assessment of modelling outputs from different stakeholders working in different WMAs as well as different water management institutions.

Modelling results from one area are often compared or related to findings from other areas. This immediately calls for the use of common definitions to variables and parameters. Typical cases include modelling scenarios which utilise vegetation characteristics. In South Africa, the Acocks vegetation definitions (Acocks, 1988) have been used as the baseline in model inputs. The Acocks vegetation water use data are currently considered very coarse for representing water use by different types of vegetation at spatial resolutions below the level of quaternary catchment. Data on water use values by different plants are currently being improved using field-based techniques such as those using large aperture scintillometer (LAS) methods (Dye and Le Maitre, 2004). Vegetation water use should therefore be derived on the basis of all the available field data including the LAS measurements as well as other historical definitions such as the Acocks system.

The definition of low flows in WRM models, where water flow regimes are model inputs or impact on other inputs, is often a source of discrepancies in deriving a common base in the simulations. An important assessment of methods for defining Low Flows in South Africa was presented by Smakhtin, Watkins, Hughes and Sami (1998) who identified the following software-based methods for low flow analysis:

- Flow duration curve construction (The method should be used in conjunction with the interactive facility to determine the flow rate and the percentage of time that this rate is equalled or exceeded).
- Analysis of continuous low-flow intervals and their deficient-flow volumes (The method looks at events/spell or continuous time series analysis).
- Extreme low flow events frequency analysis.
- Procedure to separate base flow from the total continuous daily stream flow hydrograph, and to estimate related base flow characteristics.
- Procedure to calculate recession characteristics of a stream (recession constant, half-flow period, distribution of recession rates).

DWAF is working towards the development of standards to guide model developers in their software coding. Model developers working on DWAF projects are expected to follow these coding standards in their modelling processes (DWAF, 2002b). Most other water institutions, water resources researchers, and consultants do not have any software standards and usually recommend that their contractors comply with DWAF's standards.

DWAF, through the Planning Directorate, has set up a system to define some of the most important boundaries in model selection, through a process of model accreditation. In this initiative, water resources management models are evaluated and recommendations are made as to which models should be used for national water management and planning projects (DWAF, 2003c). A national advisory committee to "police" the process of model selection for DWAF water resources projects is still to be set up.

International trends also have a major bearing on local modelling frameworks. McKinney *et al.* (1999) reported that general water quality simulation capability is now a standard feature of river basin models. River basin water assessments in South Africa are now expected to integrate groundwater, surface water and water quality issues into the main modelling components, without the sorts of bias in one or more areas as was usually the case in previous studies. Most previous studies on surface water resources failed to adequately account for groundwater, and in some cases groundwater was treated simply as a percentage loss to the water system; this is clearly not a correct representation of the complex groundwater processes or the hydrological cycle (WRC, 2003a).

Water management policies in arid and semi-arid countries, seldom are able to use catchment or river basin based water resources management approaches. These approaches are not suitable in arid areas which seek to address groundwater with its own demarcation boundaries such as aquifer boundaries (Moriarty, Batchelor and van Wijk, 2001). Moriarty *et al.* (2001) also noted that the optimisation models used in South Africa usually exclude groundwater. This is unfortunate since groundwater provides a significant contribution to water supply. In fact, groundwater currently contributes more than 15 % of all water supplies that are used in South Africa which is an important component of the water balances used in water resources modelling. The NWA promotes a holistic and integrated approach to water resources management and it is therefore essential that groundwater should be adequately addressed in all modelling efforts. If groundwater is excluded in modelling, data on groundwater are never collected and modelling routines remain undeveloped. The exclusion of groundwater in basin models provides a poor basis for further work on the total water system or on groundwater specifically.

The water resources management conceptual framework is the foundation of model development and should be updated continuously as water management trends change. This includes changes in the types of technology and software that can be used. In cases where relative locations of areas being modelled and other referral spatial points are involved, the modeller should ensure that a GIS is incorporated in the water management model. The use of Digital Elevation Models (DEMs) is also becoming increasingly important in all water resources projects where more accurate and accessible terrain data are required, and which naturally give a platform for better model outputs. The changing water environment, such as reduced water availability, increased demands for water, and the ever growing concerns on deteriorating environmental qualities including general natural resources quality, has seen a significant demand for policy reform. McKinney *et al.* (1999) pointed out that in spite of all the changes in policy, most water management tools still fail to answer questions on feasibility, costs and the likely implications of alternative water management policies in developing countries. Improved understanding of the different variables in developing countries will result in the development of more applicable modelling tools.

Davis and Hirji (2003) pointed out that the development of models should take place within an environment that accounts for stakeholder needs, especially those who are targeted to use the tools or gain from their use. Without such interaction with stakeholders, modelling tools are likely to fail to gain acceptance, which is crucial for their usefulness to be fully realised.

5.2 Topography, watercourses and climatic factors in South Africa

Topography: Water resources modelling is sensitive to the topographical characteristics of the study area. The topography directly influences water flow regimes, runoff distribution, temperatures, wind speeds and directions, rainfall types and intensities as well as many other water resources variables. The mostly mountainous coastal areas of South Africa have very sharp difference in rainfall patterns over short distances. Hilly terrain with relatively high rainfall patterns reaching annual levels above 1000 mm are common in the south eastern coastal areas of South Africa while a generally flatter terrain with drier conditions where annual rainfall is below 500 mm, cover 65 % of the country; in other words: most of the inland area. A large portion of South Africa, 21 % of the land area, is very arid with annual rainfall below 200 mm. Model settings and parameter values will need to be modified for each river basin under consideration as most river basins have dissimilar climatic and topographic characteristics. The use of high resolution and accurate DEMs is recommended to ensure that the topography is accurately and completely represented in water resources models.

River systems: South Africa's water demand is concentrated in relatively few locations and rising water demands in these areas has already or will soon exceed the water available in surrounding catchments. A system of water transfers from distant catchments is utilised to meet the excess demand in areas of high water demand. McKinney *et al.* (1999) observed that the complexity of the water transfers in South Africa requires more resources to be invested in the modelling of integrated water systems rather than the separate component approach that has characterised local water studies in the past. Water practitioners should appreciate the linkages of water system components across the different boundaries to be able to provide dependable and reliable models.

The classification of river systems according to resource quality objectives followed by a process to determine and provide for the basic human needs and ecological reserves (Republic of South Africa, 1998), are key variables that should be accounted for in the development and use of water resources models.

Climate: Water resources modelling in South Africa has to account accurately for the country's unique climatic characteristics. In South Africa, the main driver of water resources model, rainfall, is highly unpredictable, and accurate estimations of rainfall are complicated by the sparse distribution of rainfall gauging points. Large fluctuations in the average annual precipitation are very common in most areas of the country. Some 21 % of the country receives a total annual rainfall of less than 200 mm, 48 % of the country receives between 200 mm and 600 mm, while only about 30 % of the country records more than 600 mm. In total, 65 % of the country has an annual rainfall of less than 500 mm - usually regarded as the absolute minimum for successful dry-land farming. With such differences in rainfall the use of average hydrological data inputs is usually a major source of errors in the water resources planning and management models. A maximum limit of one rainfall gauge per 25 km^2 is provided for daily models such as the ACRU model (Seed *et al.,* 1995).

Most modelling data, especially rainfall, are available from several organisations and more resources are usually needed to acquire and format these data before any modelling can start. Important data sources that have to be consulted for water resources simulation assignments include: the Department of Water Affairs, SAWS, the Agricultural Research Council (ARC), Hydrology and soils research groups - especially the CSIR and University of Natal, Agricultural land users - especially the South African Sugar Association (SASA) and a large number of town councils and municipalities.

Continuous, smaller time step data such as hourly data and, to a large extent daily data, are seldom available. Modellers are advised to thoroughly investigate model data requirements and the available data before they select or develop a specific model for use. While the tendency in the water sector has been to use models that do not require detailed inputs, the introduction of water management and planning at WMA level (Republic of South Africa, 1998) has meant that the preferred monthly models will be too coarse for the temporal and spatial scales required. The NWA, with its CMA-based planning and operational management, requires modelling and decision making to be made at smaller scales in both time and space to accommodate the daily operations at field or plot scales.

Landuse and vegetation: The vegetation mapping and research in South Africa's water resources management is usually based on the research and work conducted by the botanist J. P. H. Acocks (Acocks, 1988). The Acocks vegetation database has been used as the basis for simulating land cover in South Africa (Lumsden, Jewitt and Schulze, 2003). Modellers should take note of the results of recent field-based research to determine water use by vegetation, using techniques such as the Bowen Ratio Energy Balance Systems and the Large Aperture Scintiliometer (LAS) (Dye and Le Maitre, 2004). These techniques are expected to improve the accuracy of vegetation water use model coefficients as well as other model inputs related to energy fluxes over different land uses.

Soils: The targeted level of accuracy and resolution as well as available resources are some of the key factors to be considered when deciding on the types of soils data to use in the modelling process. An important local source of soils data for water resources modelling is the soils classification developed for the South African version of the SCS model (Schmidt and Schulze, 1987). Schmidt and Schulze (1987) classified South African soils according to their hydrological responses to suit the requirements of the SCS model. The SCS model gives a scale of 1 to 100 for the classification of soils according to curve numbers that describe the soil on the basis of antecedent soil moisture conditions and its ability to absorb water.

South African soils information inventories include the BNHSZ inventory where the country's soils were mapped into 84 Broad Natural Homogenous Soils Zones (BNHSZ) (Schulze, 1996). This soils inventory is coarse and useful for regional parameterisation. Schulze (1996) also presented drainage rates indices, plant available water; soils texture classes and soils depths for different areas in South Africa. This data is however too coarse when one is using higher resolution models such as those simulating field scales.

An internationally accepted method of defining soils that should also be considered is the Global Soil and Terrain (SOTER) database. The SOTER database incorporates soils data from all corners of the world presented in a single database that is characterised by a single set of rules using reference keys based on the *'soil characteristics', 'soil properties' and 'soil horizons'*(FAO, 1995). In South Africa, water resources modellers should aim to define soil characteristics in their models in a way that follows local and international standards. Van Huyssteen and du Preez (2004) have recently provided some detailed insights into some local soils characteristics in relation to the SOTER database. The use of internationally accepted standards including those in soils is crucial for model reliability and acceptance as it provides a platform for current and future reference with other local, regional and international research.

Groundwater: With more than 15 % of all water use in South Africa being supplied from groundwater, the groundwater–surface water interactions must be considered as important components of the water balance in water resources models. Water practitioners should note that they introduce weaknesses and inaccuracies in water resources models if they inadequately represent the complex nature of any constituent components of the water resources system, including both surface water and groundwater processes.

5.3 Water resources institutional frameworks in South Africa

DWAF is the primary water resources institution in South Africa. This department has made provisions for new approaches in the water sector through the application of the NWA (Republic of South Africa, 1998). The approaches provided for in the NWA require water management tools that are suitable for higher resolution water assessments and planning to be implemented through CMAs, as well as the existing larger-scale water assessment and planning methods at river basin level. Apart from the challenges encountered in developing a framework for delegating water resources management and planning activities to CMAs, DWAF is also faced with the challenge of integrating all stakeholders and involving them in decision making and solution developments within their WMAs (Merrey, 2000; Schreiner and van Koppen, 2000; Moriarty *et al.,* 2001;). Merrey (2000) also noted that the majority of the people within catchments are mostly poor urban or rural communities who are unaware of the provisions of the new water law and the CMA process, and are often left out of the initiatives to develop the CMAs. Water practitioners must appreciate the risks involved in developing solutions for water resources problems in an environment of transformation where the long-term goals remain largely undefined or poorly understood. It is important to ensure that all affected parties are involved and take ownership of the decisions and solutions.

While the WMA approach is based on river catchments, water resources managers must realise that the catchment approach is not always logical in arid areas such as the Karoo in South Africa. Arid areas usually depend on groundwater resources whose boundaries do not always coincide with surface catchment boundaries. Moriarty *et al.* (2001) noted that in such arid areas, the focus should be on managing water at the lowest appropriate levels using IWRM principles such as aiming to maximise the economic value of water, rather than simply applying catchment-based approaches.

5.4 Socio-economic, political and trans-boundary issues

5.4.1 Socio-economic issues in modelling

A holistic approach in water resources management requires the incorporation of socio-economic factors in the management tools. Freebairn (2004) explained that water has a variety of values, for example, either directly in household drinking, bathing and gardens, or it is a valuable input used in helping to grow irrigated vegetables, rice and cotton or as environmental flows to sustain native flora and fauna. Because of its scarcity, allocating more water to one use, say industry, means that less water is available for other uses such as the environment. This calls for the application of all inclusive decision making processes which involve the evaluation of all the advantages and disadvantages in water resources planning and management decision making.

In water resources management in an arid country, the inclusion of economic considerations should account for the allocation and reallocation of limited water volumes among the competing uses so as to increase economic efficiency and national well being. Economic efficiency, or national productivity, is maximised by allocating water among the different uses so that the marginal social value of the last litre used in each different use is equalised (Freebairn, 2004). In South Africa, in addition to efficiency, economic management of limited water must also consider equity and poverty alleviation issues. Schreiner and van Koppen (2000) pointed out that the concept of water as an economic good should never be considered in isolation from other social costs as this will cause considerable cost to society through societal issues such as water deprivation that is aggravated by inequities. Good water resources planning should be robust in the sense of being able to understand present scenarios and address future changes in the availability of water, market conditions, technology, incomes, equity and resource quality objectives, particularly where many of these changes cannot easily be forecasted. Modelling tools must be able to accommodate the inevitable changes in the water resources variables which take place over time.

5.4.2 Political issues in WRM

Water management takes place within the framework of a political environment that is also guided by the water legislation and other statutory instruments such as water policies and regulations. Bate and Tren (2003) reported that the allocation of water in South Africa has been used as a political weapon in pre-independence South Africa. The attachment of land rights to water licenses was one measure that resulted in major discrepancies in water allocations that were meant to support the politics of the time. The water allocation discrepancies are further aggravated by the skewed spatial and temporal distribution of water when expressed against the spatial distribution of human population and water needs. High population densities in South Africa are located far from adequate water resources. These population density patterns have tend to follow the distribution of economic resources such as minerals and the forced settlement patterns of previous political regimes. The challenge to water practitioners is to ensure that they can account for water outside of the boundaries set by the political framework. This will involve cases such as the integrated consideration of water systems for previous homelands and other connected systems, and applying equitable water allocation rules while addressing the differences in data availability and quality as well as major knowledge gaps.

Since independence in 1994, South Africa has shifted from the more Euro-centric water legislation based on riparian rights and replaced it with the NWA, which is more suitable for the new democracy and also suits the semi-arid conditions over most of South Africa. While every effort has been made by the legislators to define new approaches to water management, the water resources manager is left with the challenge of translating the legal instruments into real life practices. The tendency to take existing tools and apply them without adequately addressing the new legal requirements in the water sector hampers compliance with the spirit and content of the legislation. Ideally, water managers should seek to develop permanent solutions; this includes the development of new water resources planning and management tools that are suitable for the new water environment.

5.4.3 Trans-boundary issues

Integrated water resources management requires a holistic approach that covers all stakeholders and all components of the water system being investigated. Since most watersheds cross administrative and political boundaries, modellers have to ensure that the tools they develop or use are not constrained by unnatural boundaries such as country or other administrative borders. The NWA makes provision for water to meet international obligations as one of the two most highly prioritised allocations of water; the other important allocation being water for the basic human needs and ecological reserve. South Africa is also signatory to a number of international legal instruments on shared water courses. These legal instruments include water specific tools such as the Revised SADC Protocol on Shared Watercourse Systems (SADC, 2001) and the Helsinki rules (ILA, 1996) and instruments addressing a wide range of sustainability issues including water such as Agenda 21 (UNCED, 1992). South Africa also has other bi- and multi-lateral treaties and other targeted international agreements with its neighbouring countries to cater for the detailed requirements of specific water systems, such as the Lesotho Highlands, the Limpopo Basin and the Orange River System.

In water resources systems that affect different stakeholders who are separated by administrative boundaries, models must reflect the affected parties' perspective of their water resources system. The water resources models should allow different stakeholders to understand model assumptions, content, capabilities and output, have confidence in the model's validity, and view it as a useful decision support tool. A model developed within these characteristics presents a "shared vision" which is an important characteristic of successful water resources models involving transboundary decisions, and indeed any other decision involving different stakeholders.

The management of water resources within the SADC region under the Revised SADC Protocol on Shared Watercourse Systems should observe a number of principles (SADC, 2001). Ideally, the principles that should be accounted for in the case of model-based water resources management should include the following:

- Respect for the sovereignty of member states in the utilisation of a shared watercourse.
- Application of rules of general or customary international law and equitable utilisation.
- Maintaining a proper balance between development and environment protection and conservation.
- Co-operation on joint projects and studies.
- Information and data sharing.
- Equitable and reasonable utilisation of shared watercourse systems.
- Use of discharge and abstraction permits or licences.
- Obligation to notify neighbouring countries about emergency situations, protection against pollution and use of installations for peaceful purposes.

The principle on information and data sharing is very important in water resources modelling. It is aimed at levelling the playing field and creating an enabling environment for negotiations for equitable utilisation of shared watercourses. The SADC Hydrological Cycle Observing Systems (SADC-HYCOS) was developed to address this principle (Mokuoane, 2000). Information sharing is central to the co-operation and economic integration envisaged by the SADC Treaty. The development of water resources models for the SADC region is expected to strengthen co-operation and information sharing within the SADC countries (SADC, 2003a).

More detailed water management requirements within SADC are expected to be handled by a water sector coordination unit which was established under the original SADC Protocol (SADC, 2001). Its vision is: To attain the sustainable, integrated planning, development, utilisation and management of water resources that contribute to the attainment of SADC's overall objective of an integrated regional economy on the basis of balance, equity and mutual benefit for all member states.

The water sector unit's overall objective is to promote co-operation on all water matters in the SADC region for the sustainable and equitable development, utilisation and management of water resources, and contribute towards the uplifting of the quality of life of the people of SADC region. Water resources management initiatives that cross SADC country boundaries, including modelling have to be formulated within the framework of the revised SADC Protocol on Shared Watercourse Systems.

5.5 Recommendations on models and software

Due to the complexity and variability of factors affecting water resources planning and management, data observations alone are often insufficient for decision making; hence the need to use models. Models are recommended for use as tools for a wide range of tasks which include: to compensate for the lack of measured data, to simplify the complexity of the mostly unpredictable interaction of water resources variables, and to assess the implications of possible water resource management scenarios.

The decision to use a model must be based on a sound understanding of the problems to be solved. Schulze (1998) reported that a number of problems occur where models are used to solve problems that do not warrant the use of such tools. Simple discussions and consultations, supported by good data, can easily provide many solutions. Many water resources managers are confronted with the need to decide on the use and selection of suitable modelling tools. A decision to use a model should be guided by the following:

- Complexity of the problem and the number of dependent variables and fixed parameters,
- Presence of uncertainties or situations involving approximate knowledge where predictions are difficult,
- Presence of conflicting goals and the need to incorporate several viewpoints or options,
- Cases where multiple scales need to be evaluated,

- The need to evaluate many scenarios,
- Problems where distinct boundaries of prevailing phenomena cannot be established,
- Decision-making situations that require the use of long sequences of data,
- The need to provide multi-objective outputs,
- Problems involving complex relationships, including those where such relationships are continuously changing such as weather and climate.

The decision to use a model has to be made with a clear understanding of the available resources, including data and tools or the expertise required for solution development. As a minimum requirement, the decision on which model to develop or select should fall within the following guidelines:

a) *What decisions are to be made using the model outputs?* Models can be developed to answer a specific question pertaining to a unique problem, for example flow regime problems on a river supplying water to a specific hydropower station, will require a customised model. Other scientists believe that the idea of developing specific models to answer specific questions is expensive and time consuming. Parkinson (2003) suggests that developers must focus more on developing general modelling solutions that can be used to answer many questions. The advantages of having one proven model ready to answer a variety of questions is that time can be spent more efficiently on simulating and analysing outputs rather than devoting this time to endless model developments. McKinney *et al.* (1999) identified the differences between "holistic" and "specific" models. Holistic models are data-intensive and may involve many other processes that a model user may not need to simulate but ends up simulating because he cannot run the model without those processes. "Specific" models were identified as less comprehensive and less demanding in terms of setting up and running. Ideally, modellers should weigh the benefits of each choice of modelling approach to determine whether or not a specific or a general model are most appropriate to the problems to be solved.

b) *What is the best model resolution (spatial and temporal scales) to address these problems?* Model choices are passively or actively made on the basis of the concept of control volume using the relationship between the complexity of the mathematical equations and the spatio-temporal resolutions involved. The very detailed scales, for example a two metre wide water course require the use of complex equations in hydrodynamic models, whilst larger spatial scales, for example, areas which are as large as the physical system with sparse data would be adequately addressed using rules of thumb. Khatibi, Moore, Booij, Cadman and Boyce. (2002) categorised modelling techniques using the concept of control volume. Figure 5.1 below shows the model categorisation based on the control volume technique.

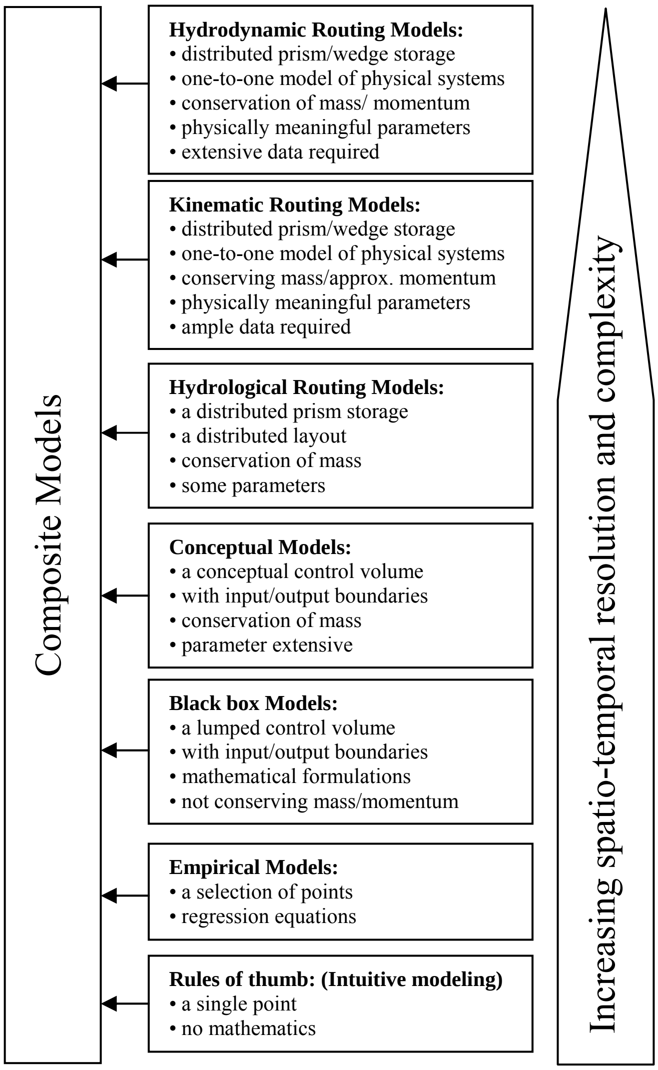

Figure 5.1 Categorising modelling techniques (Adapted from Khatibi *et al.*, 2002).

National water resources planning in South Africa has tended to use conceptual models run at monthly time steps with the quaternaries as the smallest area units. The limitations in available data have also meant that stochastic hydrological time series data are used instead of the mostly short and patchy records. The NWA however makes provision for more detailed water resources studies at WMA level. The WMA approach has generated a need to evaluate water use at field level and plan daily water operations as well as implementing non-permanent water use licenses that seek to enhance water use efficiency. Ideally these higher resolution objectives will best be handled using physical models that give more accurate and detailed catchment characterisation.

c) *What objectives are to be met?* The most appropriate model to meet the project objectives must have an optimal balance between uncertainties resulting from model assumptions (or fundamental uncertainties), and uncertainties resulting from the data (or operational uncertainties) (Willems, 2003).

d) *What are the model costs? Do these costs cover model support?* Availability of model user support, model user-friendliness, the costs of procuring the model as well as model version control requirements are other considerations to make when evaluating model costs. Many internationally developed models require high levels of expenditure on licences and continuous model updates. Model users are required to renew their licenses annually or more frequently than that to be able to continue using such tools. In most cases users are not supplied with the model source code so that they never have a chance to customise the tools or connect these to other modelling tools. Backward compatibility is a major problem in some commercially developed software where users are forced to purchase frequent releases of new model versions and updates.

e) *How accurate and reliable are the inputs?* When errors dominate in say the distributed rainfall inputs, a simpler (lumped or conceptual) model can provide more robust water resources simulations. (Khatibi *et al.*, 2002) pointed out that process descriptions alone do not compensate for shortfalls in the data and vice-versa. When rainfall data are made available through a dense gauging network generating high quality data, the more detailed distributed grid type or physical model may be used to provide improved output performance.

f) *What are the quality and confidence levels expected in outputs?* Models generating high margins of errors are usually unsuitable in cases involving economics and financial investments as well as other cases where very low levels of risk in water availability are required. Low risk levels for water availability are required in industries of a national strategic nature such as hydropower stations or other highly prioritised users such as transboundary commitments. "Holistic" models that combine different objectives such as economics and hydrology face challenges in sub-model output quality differences caused by complexities in handling information exchange between the sub-models. While hydrologic models often use simulation techniques, economic models usually use optimization techniques. The two sub-models often have different spatial development horizons, which refer to the area over which the impacts and developments extend, as well as the area over which the model can be validated. Time horizons are also different, with economic models using large time horizons spanning years, while hydrologic processes use small time intervals that reflect physical processes. Water practitioners should aim to utilise object-oriented programming combined with relational databases to capture the scale hierarchies in economic-hydrologic models.

g) *How complex should the model be?* In rainfall-runoff models, simple models that involve fewer parameters or weights to be evaluated, and which rely on simple mathematical procedures (e.g. least squares solution), are often better able to forecast discharge than those models which involve a significantly higher number of parameters or weights to be evaluated and which rely on complex mathematical computations (Goswami, O'Connor and Shamseldin, 2002). However, simple models pose the risk of errors of exclusion or over-simplification and they may overlook important factors. A process of value management is required which would reveal that a further reduction in the number of variables or level of detail in model components would create an unacceptable difference between the model and the real system where model output is distinctly unrelated to the physical system characteristics.

h) *Which models are accredited?* National model accreditation processes such as those being implemented by DWAF have major implications on the models available for use in DWAF projects. In DWAF's model accreditation system, models to be used on specified types of studies were identified and evaluated for use in these areas. Recommendations were then provided to modellers on which models to use, when, and how to use them (DWAF, 2003c). These recommendations tended to prefer models that the selecting teams were familiar with, which is a source of undue limitations to the promotion of innovation and allowing rapid developments that are important for the sustainability and continued acceptance of modelling tools and their outputs.

In cases where a new model has to be developed, the model developer is expected to design the model development process. The model development process should be guided by the following:

- The level of expertise available to develop or run the model. In the case of developing models for CMAs, one must appreciate the level of skills of the CMA personnel who will be tasked to use the model. Many models developed and used in South Africa have little to no user support and, in some cases, the models are too complex for a user community that is inadequately trained and supported (Hughes *et al.*, 2004).
- Ideally, model development processes need to be standardized on best practice. DWAF is currently establishing a system of guidance to modellers. This will involve the development of guiding documentation which will be recommended by an advisory committee.
- Models should be sufficiently detailed to capture the dominant processes and natural variability, but should not be unnecessarily refined as this compromises the computation time and wastes resources (Booij, 2003). The model developer must first identify the dominant processes and associated key variables. Second, the appropriate spatial and temporal scales for each key variable are determined. Furthermore, relationships between key variable scales and the output variable are used to combine the appropriate variable scales to one appropriate model scale. In the third step, mathematical process descriptions consistent with these model scales are selected or formulated.
- The model developer and users must have a common or “shared vision” in the model. The ownership of a model will need to be transferred to the stakeholders and the model users through carefully planned engagement during the model development process.

5.6 Routines, objects, tool integration and interfacing in modelling

Effective integration of data sources, numerical tools, application of intelligence analysis and knowledge are the key to good water resources modelling. At the level of data and background-information collation, numerous and often incompatible bits of information from disparate sources have to be brought together. Ideally, model development should be structured to incorporate integration early in the development stages.

McKinney *et al.* (1999) recommended that modelling at basin level should ideally be based on a GIS decision support system that integrates economic, agronomic, institutional, and hydrologic components. To achieve this, comprehensive modelling frameworks that integrate agronomic, institutional, and hydrologic components need to be developed at basin level. This will facilitate the national provision of policy instruments, national economic assessments of water use as well as hydrological assessments. Figure 5.2 below presents a suggested framework for river basin management modelling, including relationships, and decision items at various levels.

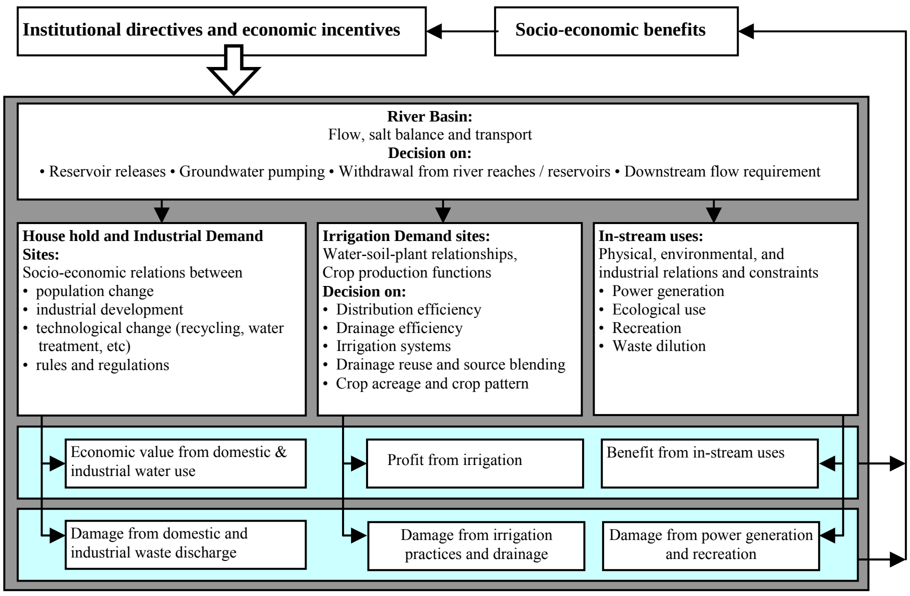

Figure 5.2 Framework for river basin management modelling (Adapted from Mckinney *et al.*, 1999).

Model developers and water managers should appreciate that water resources models tend to become components of water resources decision support systems through the integration of data components, simulation routines, multi-criteria decision aids, as well as GIS and DEMs capabilities. Ideally, the approaches followed in model development should allow parallel integration rather than sequential connectivity/integration. Sequential integration involves a unidirectional linkage of routines and sub-models such that outputs from each component are fed as input into a subsequent component. This process was common in early modelling practices and does not allow the more realistic bidirectional interactions between the various system components (McKinney *et al.*, 1999). Integrating the model components in a parallel way allows status information to be exchanged between the sub models continuously during simulation, such that feedback loops and external linkages are updated as the simulation progresses. Sequential integration will usually give flawed results in water resources modelling where backward linkages and external influences occur with time (McKinney *et al.,* 1999).

Standardization of the databases has to be addressed as a key aspect of integrated models. The goal is to provide an environment in which all computations made from different modelling components from a variety of institutions with similar data inputs but simulating different aspects of water resources modelling, converge and share the same data resources. The overall goal should be to establish a consistent and always realistic representation and comprehensive database for the river network, its critical reaches, water transfer routes, water sources, pollution points, water storages and operation rules, climate, landuse, topography, water provision scenarios and other water resources variables (Pistocchi and Mazzoli, 2002).

The representation of multi-spatial scales in water resources models is best handled using agent-based models where object-oriented programming is combined with a relational database (Polhill, Gotts and Law, 2002). This approach allows user-configurable scale hierarchies to be enabled using a relational model, where relational tables are used to link the groups of entities (spaces) at different levels of detail. While this method is a flexible approach to the representation of the various scales at which processes influence land use changes and how land uses at a variety of scales influence model parameters, the method results in reduced flexibility on how the objects can be redefined at run time. Any changes to the modelling objects will require the changes to be reflected at the different scales.

5.7 User platforms and model packaging

Model users tend to prefer models that present model platforms and user interfaces which blend well with their existing software functionalities and working environment (Eric, 1999). It is also easier to integrate models and other applications if they were developed using the same source code or have a similar architecture such as an object-oriented approach. A model should at least emulate (have the same "feel and touch") the commonly used platforms for existing tools to improve user acceptance and reduce training needs. Another important approach in the development of user platforms or model working environments is to create user profiles with different privileges and responsibilities in the interaction with the water resources management tools. This will lead to the definition of different levels of interaction between the user and the models, thus reducing user support requirements and improving model security and integrity.

The use of **h**yper **t**ext **m**ark-up **l**anguage (HTML) and e**x**tended **m**ark-up **l**anguage (XML) files in user platforms to support model users will ensure that the documentation of a user defined model exactly matches its implementation. Discrepancies are commonly encountered between the description of a model and its actual implementation, leading to inconsistencies in the model building processes. South African model users frequently encounter such discrepancies when they attempt to use the WRPM, WRSM90 and Shell models with updated input data. The documentation is usually inadequate for one to be able to determine how the different model components were connected and the correct source data including information on how they were manipulated to give the revised data used in the model. The recommended approach is to incorporate software into the model for automatic documentation of the model development process, inputs and simulation information. In this automated document, changes are automatically tracked, recorded and stored each time a user modifies or adds anything to the existing model.

The use of an open data model is important in improving user access. The user will be able to access such data using other external software and can easily integrate these data into their reports. A web browser can also be used to access input and output data stored in an open model.

5.8 Guiding thoughts on model software selection and development

The NWA (Republic of South Africa, 1998) as well as the National Water Resources Strategy (DWAF, 2002a) make provision for monitoring and information systems for water resources and set responsibilities for providing water related information. In these provisions, general boundaries to areas and issues that can be handled by water resources models are provided. Chapter 5.1 of this study unpacks the national legal and policy provisions to make them useful to a model developer.

The efficiency and effectiveness of water resources model applications are influenced by the topography, watercourse systems and climatic variables. On the other hand, how the water resources model handles socio-economic, political and trans-boundary issues influence the softer issues in modelling such as model acceptance by stakeholders and even the technical correctness of the solutions derived from the model. It follows, then that, the prescription of models or their development should ideally account for all the variables which affect the output, acceptance of outputs and applicability of derived solutions. The application of model results is usually left to water institutions. Chapter 5.3 also discusses the constraints within these institutions, their capacities and preferences which have to be accounted for in the model development or selection.

As part of the model development or selection a series of questions should be asked to determine the most appropriate solution. These questions include the following:

- What are the problems to be solved using the model?
- What is the best model resolution (spatial and temporal scales) to address these problems?
- What objectives are to be met?
- What are the model costs in terms of licences, copyrights and user support

- What are the input requirements?
- How complete, accurate and reliable are the available input data?
- What is the level of expertise available to run the model and to receive as well as to evaluate model outputs?
- How complex should the model be?
- Are there any models or model components already prescribed or preferred to handle the problem?

These questions should be addressed in such a way that the model development process generates a surplus of benefits and meets the set objectives as well as complying with the requirements of relevant water resources projects. One of the key modelling requirements is to generate reasonably accurate results that can be defended inline with the existing legislation, such as the NWA in South Africa and transboundary legal instruments in cases where international waters are involved.

The development of water resources models that use stochastic data or remotely sensed data is an important approach in most of South Africa's drier catchments where data availability is very limited. Detailed physical models that require high-resolution data should be restricted to small data-rich catchments. In the absence of other specific recommendations, the 25km^2 aerial unit size recommended for the ACRU model's driver rainfall approach (Seed *et al*, 1995) should be used as the maximum size of each sub-catchment unit to be simulated separately.

Water resources model development and use in recent years has focused on tools that are useful in national planning. The WMA approach required in the NWA looks at the WMA as the largest spatial unit such that new tools should now seek to simulate catchments at higher spatial and temporal resolutions.

The globalisation of sustainability and development issues as provided for in initiatives such as the global Millennium Development Goals, as well as the enactment of enabling national legislation such as the NWA in South Africa, have brought about new challenges to water managers, thus increasing the complexity of existing water resources problems. Political boundaries are no longer expected to be the limit of water resources management and planning programmes. Water resources managers are now expected to ensure that the water resources management and planning tools in use are holistic and incorporate issues beyond the political, regional, socio-economic or other boundaries which do not coincide with hydrological or water management boundaries.

Model development, like any other part of information technology, is very dynamic. The use of MS-DOS text commands that dominated earlier models is no longer a preferred option. Many new users coming out of colleges and universities have no training on older IT tools such as MS-DOS. On the other hand, the component model approach or OOP has meant that modellers do not have to reinvent many model components as they can now use existing components developed by other specialists in their own space and time for specialised functions in their own modelling tools. A lot of these components are freely available and allow for full third party interfacing and use. The Internet has also revolutionised model user support and development. Modelling tools developed today should ideally take advantage of the Internet functionalities, to provide remote user support, online feedback and further model development, WWW-based model dissemination and user support through online forums, as well as WWW based output presentation and publishing of results. Presently, there are no strict restrictions on the material that is posted on the internet. Model developers are expected to go through the process of peer-reviewing and ensuring high quality levels in their internet postings. On the other hand users of internet based modelling material should ideally exercise caution.

Chapter 6

Verification, Calibration and Validation

6.1 Model Verification

Once a model is developed the numerical techniques in the computer code will need to be examined to ascertain that they are accurate representations of what is being modelled, the concepts involved, the optimisation methods applied, as well as the relations between variables. These model examinations constitute model verification. Most model coding software comes with debugging applications to identify and suggest solutions for some model code inaccuracies. Modellers should ideally make full use of debugging software to handle shortcomings in the coding and model equations, as well as numerical discrepancies.

Model verification should ideally utilise actual input data. DWAF (2001c) recommends that the data used in model verification should at least display a representative range of events to be simulated. Further recommendations in DWAF (2001c) are that verification should attempt to capture all the model efficiency parameters handled in calibration but may utilise shorter time series data than those used in calibration. Verification inadequacies should be handled early in the modelling process, ideally as part of the model development, in which case model equations can be improved. In model verification, adequate examination of the mass balances and flow routing can resolve most of the model coding shortcomings.

In some cases, additional model verification is done as part of the modelling process, preferably after model calibration. In such cases, if the verification is not acceptable according to a defined set of criteria, then a second round of calibration is required which should account for the lessons learnt from the verification. This should be followed by a second round of verification. If the model was developed to allow model improvements through repetitive verification and calibration, the cyclic repetitions should continue until the model parameters are "acceptable". Acceptability will mean that the simulated values "mimic" as closely as possible corresponding observed values either in a time series or for individual discrete events/output.

6.2 Model Calibration / Validation

Model calibration and validation are essential steps in any water resources model application. During calibration, model parameters, for which data may not be available, are estimated and adjusted until the model outputs are equal to or relate closely to recorded or observed measurements. Calibration simply involves adjusting parameters in the model so that the model reproduces the measurements. Calibration should involve iterative procedures of parameter evaluation and refinement, as a result of comparing simulated and observed values of interest. However the values of the calibration parameters must be within a range that makes sense to the physics, chemistry or other scientific principles involved. Model validation is in reality an extension of the calibration process (Donigian, 2001). Its purpose is to assure that the calibrated model properly assesses all the variables and conditions which can affect model results, and demonstrate the ability to predict field observations accurately for periods separate from the calibration effort. By definition, model validation involves the comparison of model results with numerical data that have been derived independently, either from experiments or observations of the environment.

6.2.1 Observed and field data in model calibration

Model efficiency and the accuracy of the calibration process are highly dependent on the available observed data. Identifying abnormalities in the observed data should be the first step of the calibration and validation processes. Calibration in water resources models should ideally include comparisons of daily, monthly and annual values as well as individual events, whenever sufficient data are available for these comparisons. All of these comparisons should be performed for a proper calibration of hydrology and water quality parameters. In addition, when a continuous observed record is available, such as for streamflow, simulated and observed values should be analyzed on a frequency basis and their resulting cumulative distributions (e.g. flow duration curves) compared to assess the model behaviour and agreement over the full range of observations.

Ideally, modellers should always utilise the latest available data in model calibration. In cases where current sources differ from those used in other to-be-compared studies, the modeller should illustrate and examine the differences by using a data-set that is common to both sets of analysis. Seeking such common ground between comparative analysis facilitates the understanding of differences and similarities in model outcomes and findings, which improves the interpretation of model outputs (DWAF, 2001c).

In recognition of the inherent variability in natural systems and often unavoidable errors in field data observations, water practitioners should ideally establish and use data accuracy characterisations that are consistent in water resources management. In cases where specific local guidance on characterisation of accuracy in data is not available, international documented characteristics may be used. Important documented characteristics include those from the United States Geological Survey (USGS), which provided the following characterization of the accuracy of its streamflow records in all its surface water data reports (Donigian, 2001):

Excellent Rating	95 % of daily discharges are within 5 % of the true value
Good Rating	95 % of daily discharges are within 10 % of the true value
Fair Rating	95 % of daily discharges are within 15 % of the true value

The WMO Commission for Hydrology also provides recommendations for accuracy levels in hydrological simulations as presented in Table 6.1 below.

Records that do not meet these accuracy criteria are rated as ‘poor’ and water practitioners are cautioned against their use. According to Donigian (2001), model results for flow simulations that are within the accuracy tolerances indicated above can be considered acceptable calibration and validation results, since these levels of uncertainty are inherent in the observed data.

Table 6.1 Recommended accuracy (uncertainty levels) expressed at the 95 % confidence interval (Adopted from WMO, 1994).

Measured parameter	Accuracy
Precipitation (amount and form)	3-7 %
Rainfall intensity	1 mm per hour
Evaporation (point)	2-5 % or 0.5 mm
Wind speed	0.5 m/s
Surface water level	10-20 mm
Wave height	10 %
Water depth	0.1 m or 2 %
Width of water surface	0.5 %
Velocity of flow	2 – 5 %
Discharge	5 %
Suspended sediment concentration	10 %
Suspended sediment transport	10 %
Bed-load transport	25 %
Water Temperature	0.1 - 0.5 ^{o}C
Dissolved Oxygen	3 %
Turbidity	5 – 10 %
Colour	5 %
pH	0.05-0.1 pH unit
Electrical conductivity	5%
Soil moisture	1 kg/m^3 - 20 kg/m^3

In the calibration of time series routines, the observed data and simulated data should attempt to cover all the possible scenarios of climatic conditions, hydrological systems, anthropogenic conditions and other important variables depending on the objectives of the modelling. DWAF (2001c) recommends that calibration should be based on at least 10 years of data for daily time steps, and at least 15 years for monthly steps, in order to evaluate parameters under a variety of climatic, soil moisture, and water quality conditions. Generally, the longer the record, the better the quality of simulation outputs. The selection of a suitable record length should be considered in relation to rainfall variability. Semi-arid areas and mountainous areas, which experience high rainfall variability, are better represented by longer rainfall record lengths (Seed *et al.*, 1995).

6.2.2 Model parameters in calibration

Calibration focuses on determining the most suitable values and ranges for model parameters. A number of methods are available for determining the parameter values.

Automatic methods or optimisation techniques such as the use of genetic algorithms and the shuffled complex methods are some of the many techniques that modellers can use. In the absence of better techniques, modellers are expected to use their own judgement based on experience and existing literature to estimate model parameters. In water resources models at basin scale, modellers should ideally follow a three step process in selecting parameters. In the first step, key hydrological parameters are determined; the next step should involve sensitivity analysis of the parameters to come up with an optimal parameter set. In the third step the parameters should be regionalised to derive specific parameters for each sub-basin.

Model parameters that cannot be deterministically, and uniquely, evaluated from topographic, climatic, physical, chemical or other scientific characteristics of the watershed and compounds of interest may require other parameter estimation techniques such as automatic calibration. A number of software tools are available for automated parameter estimation. Modellers must however be aware of the limitations of the automatic method selected. Most automated calibration methods do not capture all the dependent variables, they often fail to appreciate the parameter sensitivities and users cannot easily comprehend the processes involved in these methods. It is therefore advisable to use either manual calibration alone or automatic calibration together with manual methods. Some automatic calibration methods based on the genetic algorithm can adequately replace manual calibration (Ndiritu and Daniell, 1999). Ultimately, calibration should result in parameter values that give the best overall agreement between simulated and observed values throughout the calibration period.

6.2.3 Model objective functions

The objective functions or criteria selected to evaluate model performance must be relevant to the project objectives or decisions to be made using the model. The performance measurement criteria used in water resources modelling are often referred to as objective functions.

In South Africa, DWAF (2001c) recommends that the minimum criteria for goodness-of-fit of the simulated values when compared to recorded values should at least address the limits in the following:

- Annual mean, standard deviation, dry season mean (streamflows, constituent concentrations and loads),
- Percentile curve (streamflows, constituent concentrations and loads),
- Mean monthly distribution (streamflows, constituent concentrations and loads),
- Cumulative mass curves,
- Gross yield-storage curve for streamflows, and
- Deficient flow-duration frequency curves.

The NWA (Republic of South Africa, 1998) makes provision for a number of objective functions in water resources modelling. In the water ecosystem area, the NWA provisions include the classification of water resources and resource quality objectives, water reserve provisions and pollution prevention targets. In Chapter 4, of the NWA (Republic of South Africa, 1998) provisions for water use are presented to cover issues such as permissible water use, equitable and beneficial water allocation, water use authorisations and licences, lawful water use, controlled water use activities and water use allocation schedules. These legal requirements should be accounted for in setting out the model boundaries as well as in the definition of objective functions where appropriate.

Model calibration/validation usually involves statistical and graphical comparisons of model outputs with actual observed or measured data. Schulze (1998) pointed out that visual comparisons of model results are very subjective. Donigian (2001) pointed out that in water resources management models which involve water quality routines and biotic constituents, model performance should preferably be measured using, primarily, visual and graphical presentations rather than the frequency of observed data which is often inadequate for accurate statistical measures. The use of statistical measures is usually recommended where adequate data are available. In this case, predetermined criteria of goodness of fit in the objective functions are established. Some of the main statistical measures or objective functions in the rainfall runoff component of water resources modelling include the following:

Conservation of the mean (Ob$\bar{x}$): This is expressed as the percentage difference between means of observed and of simulated values. For a good simulation this objective function has to be minimised, zero being the ideal level.

$$Ob\bar{x} = 100.\frac{1}{N}\left(\sum\nolimits_{i=1}^{N} Q_{obs,i} - Q_{sim,i}\right) \qquad \text{Equation} \qquad 6.1$$

Overall volume error (F): For best results *F* has to be minimised.

$$F = \sum_{i=1}^{N}(Q_{obs,i}\Delta t - Q_{sim,i}\Delta t) \qquad \text{Equation} \qquad 6.2$$

Root mean square error (RMSE): The closer the RMSE is to zero the better the simulation.

$$RMSE = \left(\frac{1}{N}\sum_{i=1}^{N}(Q_{obs,i} - Q_{sim,i})^2 \right)^{\frac{1}{2}} \qquad \text{Equation} \qquad 6.3$$

Coefficient of regression (R^2):

$$R^2 = \frac{\frac{1}{N}\sum_{i=1}^{N}\left(Q_{obs,i} - \overline{Q_{obs}}\right)^2 - \frac{1}{N}\sum_{i=1}^{N}\left(Q_{obs,i} - Q_{sim,i}\right)^2}{\frac{1}{N}\sum_{i=1}^{N}\left(Q_{obs,i} - \overline{Q_{obs}}\right)^2} \qquad \text{Equation} \qquad 6.4$$

The coefficient of regression, measures the degree of association between the simulated values and the values estimated by the regression model. This objective has to be maximised to 1 for a good simulation.

As a guide, Table 6.2 below presents an example of the criteria recommended when using the coefficients of regression values (R and R^2) for water flow comparisons in the United States' Environmental Protection Agency (EPA) projects (EPA, 2003).

Table 6.2 Comparisons of coefficients of regression values (R and R^2) for water flow (adopted from EPA, 2003)

R	← 0.75	0.80	0.85	0.90	0.95 →
R^2	← 0.6	0.7	0.8	0.9 →	
Daily Flows	Poor	Fair	Good	Very Good	
Monthly Flows	Poor	Fair	Good	Very Good	

Nash and Sutcliffe (1970) recommended another important objective function for the measurement of efficiency in Rainfall-Runoff models. Equation 6.5 presents the "Nash and Sutcliffe" objective function equation; the aim is to generate a simulation where the values of *E* is very close to one. Values of *E* above 0.7 are classified as acceptable.

$$E = 1 - \frac{\sum_i (Q_{sim_i} - Q_{obs_i})^2}{\sum_i (Q_{obs_i} - \overline{Q_{obs}})^2} \qquad \text{Equation 6.5}$$

Where Q_{sim} is the simulated streamflow, Q_{obs} is the measured streamflow and $\overline{Q_{obs}}$ the average streamflow value in the measured period.

In monthly and annual simulations the EPA also recommends some rough guidance for calibration/ validation targets as shown in Table 6.3 below:

Table 6.3 Calibration/ Validation targets in environmental modelling (Redrawn from EPA, 2003)

Criteria	% difference between simulated and recorded values		
	Very Good	**Good**	**Fair**
Hydrology / Monthly Flow	<10	10-15	15-25
Sediment	<20	20-30	30-45
Water temperature	<7	8-12	13-18
Water Quality/ Nutrients	<15	15-25	25-35
Pesticides/ Toxics	<20	20-30	30-40

6.2.4 Guidelines for effective model calibration

The United States Geological Survey provides additional recommendations on guidelines for the calibration of water resources models (USGS, 1998). These guidelines have been adapted and presented in Table 6.4 below, where guidelines pertaining to specific U. S models such as the model MODFLOWP have been excluded. Model calibration guidelines, such as these (Table 6.4) are not intended to be followed sequentially, but may be repeated many times during model calibration. Ideally, modellers should use their own judgement of their modelling objectives to select those guides which provide the best relationship with their specific modelling scenario.

Table 6.4 Guidelines for effective model calibration. (Adapted from USGS, 1998).

Guideline	Description
1. Apply the principle of parsimony	Start simple and add complexity as warranted by the hydrogeology and the inability of the model to reproduce observations.
2. Use a broad range of information to constrain the problem	For example, in ground-water model calibration, use hydrology and hydrogeology to identify likely spatial and temporal structure in, for example, aerial recharge and hydraulic conductivity, and use this structure to limit the number of parameters needed to represent the system. Do not add features to the model to attain model fit if they contradict other information about the system.
3. Maintain a well-posed, comprehensive regression problem	a) Define parameters based upon their need to represent the system, within the constraint that the regression remains well-posed. Accomplish this using composite scaled sensitivities and parameter correlation coefficients. b) Maintain a comprehensive model in which as many aspects of the system as possible are represented by parameters, and as many parameters as possible are estimated simultaneously by regression.
4. Include many kinds of data as observations in the regression	Adding different kinds of data generally provides more information about the system. In ground-water flow model calibration, it is especially important to provide information about flows. Hydraulic heads simply do not contain enough information in many circumstances, as indicated by the frequency with which extreme values of parameter correlation coefficients occur when using only hydraulic heads.
5. Use prior information carefully	a) Begin with no prior information to determine the information content of the observations. b) Insensitive parameters (parameters with small composite scaled sensitivities) can be included in regression using prior information to maintain a well-posed problem, but during calibration it often is advantageous to exclude them from the regression to reduce execution time. c) For sensitive parameters, do not use prior information to make unrealistic optimized parameter values realistic.
6. Encourage convergence by making the model more accurate	Even when composite scaled sensitivities and correlation coefficients indicate that the data provide sufficient information to estimate the defined parameters, nonlinear regression may not converge. Working to make the model represent the system more accurately obviously is beneficial to model development, and generally results in convergence of the nonlinear regression. Use model fit and the sensitivities to determine what to change.
7. Evaluate optimized parameter values	a) Unreasonable estimated parameter values could indicate model error. b) Identify parameter values that are mostly determined based on one or a few observations using dimensionless scaled sensitivities and influence statistics. c) Identify highly correlated parameters.

8. Test alternative models	Better models have three attributes: better fit, weighted residuals that are more randomly distributed, and more realistic optimal parameter values.
9. Evaluate potential new data	Use dimensionless scaled sensitivities, composite scaled sensitivities, parameter correlation coefficients, and one-percent scaled sensitivities. These statistics do not depend on model fit or, therefore, the possible new observed values.
10. Evaluate the potential for additional estimated parameters	Use composite scaled sensitivities and parameter correlation coefficients to identify system characteristics for which the observations contain substantial information. These system characteristics probably can be represented in more detail using additional estimated parameters.
11. Use confidence and prediction intervals to indicate parameter and prediction uncertainty	a) Calculated intervals generally indicate the minimum likely uncertainty. b) Include insensitive and correlated parameters, perhaps using prior information, or test the effect of excluding them. c) Start by using the linear confidence intervals, which can be calculated easily. d) Test model linearity to determine how accurate these intervals are likely to be. e) If needed and as possible, calculate nonlinear intervals f) Calculate prediction intervals to compare measured values to simulated results. g) Calculate simultaneous intervals if multiple values are considered or the value is not completely specified before simulation.
12. Formally reconsider the model calibration from the perspective of the desired predictions	Evaluate all parameters and alternative models relative to the desired predictions using prediction scaled sensitivities, confidence intervals, composite scaled sensitivities, and parameter correlation coefficients.

6.2.5 Selection of model objective functions

Different objective functions are applicable to different models, and to the same models used for different goals. The selection of objective functions should at least be guided by the following:

- Errors between simulated and observed values must be minimal.
- The selected objective functions must be related to the specific aim and relevance of the modelling application.
- The characteristics of the simulations that require the most accurate representation should be accounted for in the objective functions.

- Objective function selection should have a bias towards those objectives that were identified as important by the stakeholders/users as well as those that are important to the problem owners.
- Selected objectives must not conflict with other objectives including those in the sub-models.
- Objective functions should be practical to model and must relate as closely as possible to the physical system.

In the case of the ACRU model, and other similar water resources models in South Africa, Jewitt and Schulze (1999) recommended a selection of objective functions for application. The objective functions in this list and other ACRU related functions are explained in more detail in Schmidt, Smithers, Lynch, Schulze, and Pike. (1995). These authors advised practitioners to, use the ACRU-model based objective function selection, listed below, for local water resources models, where applicable:

- Total observed/simulated flows (mm)
- Mean observed/simulated flows (mm)
- Correlation coefficient
- Students “t” value
- Linear regression coefficient
- Base constant for regression equation
- Standard error of simulated flow
- Variance of observed flow
- Variance of simulated flow
- Standard deviation of observed flow
- Standard deviation of simulated flow
- % difference in standard deviation
- Coefficient of determination
- Coefficient of efficiency

There are different concepts used in objective functions for river basin or water resources models based on the differences in types of models, especially the simulation and optimisation models. In simulation models, the objectives are centred on the evaluation of the water system performance relative to set criteria which include ecological sustainability, climate change implications, changing water demands and water supply prioritisation. In optimisation models, the objective functions are interrelated, with constraints driving the model internal processes. A typical example of constraints used in South Africa is the penalty system in the WRYM and WRPM. Modellers should however note that optimisation models in water resources should also contain a simulation component to characterize the hydrologic regime, unless another method of incorporating the hydrological process is utilised.

An insight in the processes involved in optimisation models such as the WRYM and the WRPM will assist water practitioners in their modelling. These models use the dynamic programming algorithm where the main problem to be solved by the model is decomposed into a sequence of smaller problems ("sub-problems"). A system of interrelated objective functions, utilising a penalty system for solution optimisation are set for the "sub-problems". The model user defined penalty system forces the internal model solutions to follow the route with the least penalties when the model is run. Solutions derived for the "sub-problems" are linked to the overall objective which may be as an example to maximise the yield of a river basin or ensuring consistent water flows in the river throughout the year. In the case of yield maximisations, examples of objectives set in the sub-problems include increasing the water storage potential of individual dams just before the rain season, thus maximising water storage in the upper reaches of the catchment. To maintain a consistent flow regime in the river you may however need to keep higher levels of storage, or operate larger dams with capacities of say four times the mean annual rainfall (MAR) which seldom spill but have enough water to maintain consistent flows. The optimisation of objectives in the WRPM is done using two sets of algorithms:

i) Network Algorithm
ii) Tree Algorithm

The network algorithm which utilises an out-of-kitler routine solves the flows in each channel after the tree algorithm has resolved the outflows at each system node (DWAF, 1987).

In cases where variables are constrained by several issues, the best model choices are usually optimisation models. In a river basin, these models can allow objectives to be set on a variety of water resources issues, which include:

- Hydrologic specifications (e.g. minimum flow levels)
- Social value systems (e.g. river basin stakeholders' choice)
- Economics (e.g. maximize economic return per unit of salinity)
- Equity (e.g. justified free water supply limit)
- Environmental quality (e.g. reserve allocation)

The scale at which problems are intended to be solved in the model influences the type of models to be developed and applied, as well as the objective functions to be used. In most detailed studies such as the simulation of evaporation processes of a small field (e.g field covering one hectare), mechanistic models which simulate the physical processes in detail are preferred. Objective functions for such detailed models should aim at high levels of detail and accuracy. Variables such as water table with accuracies in terms of centimetres, moisture content to millimetre detail, number of plants and their individual species, leaf cover per square metre of area, temperatures and rain water interception per square metre are of critical importance in detailed mechanistic models. On the other hand, modellers working at larger scales such as research projects involved in simulating water availability in southern Africa over the past 100 years will mostly utilise conceptual models where the level of detail and model objectives are coarser. As an example the model inputs may be such that the representations of large spatial areas, for example, thousands of square kilometres of land cover, may be entered in the model in a generalised format as Savannah grassland without giving further details. Objectives set will also have to depict a generalised format that can capture average parameters over thousands of square kilometres.

6.2.6 Stages in water resources model calibration

In water resources management and planning models, the calibration process should be handled as a hierarchical process which begins with the hydrology calibration of both runoff and streamflow. In cases involving water quality and sedimentation, the next stage should be to calibrate the sediment erosion and sediment transport, and finally calibration of non-point source loading rates and water quality constituents. Other hydrologists (Donigian, 2001; EPA, 2003) suggest that when modelling land surface processes, hydrologic calibration must precede sediment and water quality calibration since runoff is the transport mechanism by which non-point pollution occurs. Likewise, adjustments to the in-stream hydraulics simulation must be completed before in-stream sediment and water quality transport and processes are calibrated.

In the hydrologic calibration stage at least five characteristics should be calibrated in successive examinations of the river basin, in the following order: (1) annual water balance, (2) seasonal and monthly flow volumes, (3) water quality, (4) base-flow, and (5) storm events. Simulated and observed values for each characteristic are examined and critical parameters are adjusted to improve or attain acceptable levels. Table 6.2 and Table 6.3 show examples of calibration target levels recommended by the United States of America's Environmental Protection Agency.

Sediment calibration should preferably follow the hydrologic calibration. In sediment calibration, sediment parameters should be modified to increase agreement between simulated and recorded monthly sediment loss, deposition and storm event sediment removal. Estimated loading rates based on measured rates of sediment depositions are also used to calibrate the sedimentation parameters in cases where continuous loading rates have not been measured.

The calibration of water quality and non-point source loading should aim to obtain acceptable levels of agreement for observed and simulated concentrations as well as meeting the agreed criteria. Table 6.2 and Table 6.3 gives some recommendations on water quality calibration targets. The calibration process should ensure that parameters remain within physically realistic bounds, and in the case of non-point source loading, the expected parameter ranges as presented in literature should be used as guidance.

6.3 A summary of guidance on model verification, calibration and validation

In this chapter, model verification, calibration and validation is discussed, with the aim of improving the quality of the modelling process. A key aspect of ensuring high quality outputs in modelling is to verify the numerical techniques in the computer code, thus ascertaining that they are accurate representations of what is being modelled, the concepts involved, the optimisation methods applied, as well as the relations between variables.

A process of calibration and or validation should ideally follow the verification process before the model is finally used. In calibration, the modeller is expected to use the available input records and then enter and adjust the estimated model parameters, for which data may not be available until the model outputs are equal to or relate closely to physical observations. Based on this research, the following are some of the considerations to be made in evaluating the efficiency and accuracy of the calibration process:

- The observations used in calibration and validation must be of appropriate quality. A good rating is recommended for daily flows where 95 % of the daily flows are within 10 % of the true value. Flow records of this level of accuracy are seldom available in South African catchments. As an example, an area exceeding 3 000 km^2 that was used in the case study of this thesis in 2001 to 2003 had only two flow gauges. One of these gauges had incomplete flow records ending in 1988 (Figure 9.6).
- The use of the WMO (1994) guidelines is recommended for the accuracy of variables covered in this document; these are presented in Table 6.1.
- The use of conservation of the mean, overall volume error, root mean square error and coefficient of regression are recommended for use in comparing simulated and recorded flows.
- In calibrating sediment loads and deposition, an accuracy of 20-30 % difference between the simulated and recorded values is recommended. This figure is also recommended as a good accuracy value by EPA (2003).

- On the basis of the investigations in this study, the calibration process should ideally follow the USGS (1998) recommendations (Table 6.4) where project specific guidance are not provided.
- In surface water resources models involving water quality the calibration process should ideally involve the following three calibration stages, carried out in the same order.
  1. Hydrological calibration
  2. Sediment calibration
  3. Water quality calibration

The above calibration sequence allows for sedimentation processes to utilise calibrated flow and other hydrological data. The water quality calibration is likely to utilise the hydrological data and calibration parameters. Sedimentation which is a sub-component of water quality processes is ideally handled before the overall water quality calibration.

Chapter 7

Spatial data and stakeholder inputs in water resources modelling

7.1 Geographical Information Systems (GIS)

Water resources management and planning problems are always characterised by a spatial dimension. A system to handle the spatial dimension, GIS, is an important component of effective solutions in water resources problems. GIS is a general-purpose technology for handling geographic data in digital form, with the ability to pre-process data into a form that is suitable for problem analysis, to support analysis and modelling directly, and to post-process and present results in a useable and user-friendly format (McKinney *et al,*1999). GIS offers a spatial representation of water resources systems, and it also offers predictive and analytical capacities for solving complex water resources planning and management problems. Bivand and Lucas (1997) observed that GIS is generally classified as a technological tool, while modelling is seen more as a scientific activity and that these different perspectives of GIS and models affect their integration. GIS provides the platform for integrating water resources variables of various modelling aspects, which include hydrological, social, demographic, economic and environmental.

GIS can be used in various ways to support water resources modelling. Some of the uses where GIS can be applied include the following:

- *Store and manage data* – GIS performs geospatial data-management tasks (data storage, manipulation, preparation, and extraction) and spatial data processing (overlays and buffering) (Maidment, 2001).

- *Extract parameters* – GIS provides characteristics and properties of watersheds and river reaches for hydrologic modelling (Maidment, 2001).

- *Provide visualisation* – GIS displays can be used in three modelling stages
    - Pre-processing:- to verify the basic data and information for the model.
    - During modelling:- to visualise physical processes against a time line e.g. flood propagation, pollution flume propagation and sediment loading.
    - Post processing:- for evaluation of the results of the modelling. For example, floodplain mapping in GIS shows the extent of areas damaged by floods (Collins and Campbell, 2003).

- *Documentation support* – GIS provides documentation support for geographic images, mapping, drainage files and meta-data.

- *Model surfaces* – A GIS can be used as a mapping or terrain analysis tool and for delineation of catchment areas as well as representing channel shapes based on elevation models (Doan, 2003).

- *Develop interfaces* – Map-based interfaces to hydrologic models can be developed using GIS tools (Doan, 2003).

GIS and modelling routines are interfaced using a number of techniques. The lowest level of interfacing is achieved through "loose-coupling", followed by "tight-coupling" and then "full integration" or "embedded coupling". Modellers and stakeholders should have adequate information on the level of interfacing that will give the best results in the different water resources modelling projects. Jun (2000) and McKinney *et al.* (1999) provided some basic characteristics of the different levels of interfacing which can be used as guidance in water resources modelling projects. Figure 7.1 illustrates the integration classifications presented by Jun (2000).

"Loose-coupling" is the simplest approach where two systems exchange files such that a GIS uses data from the other modelling system as its input data, and vice-versa. At this level of integration, the two systems run independently and no system modification or programming takes place except that the data or outputs of one system need to be formatted for use as inputs to the other system. The GIS processes in "loose-coupling" are separated from the main model such that they use a separate database with information being transferred between the GIS and the main water resources model (McKinney *et al.*,1999). “Loose-coupling” does not involve coding or other complicated handling of source code of the two systems being linked.

In “tight-coupling” the two systems share the communication files as well as a common user-interface. The development of the common user interface is achieved by using macro languages such as Arc Macro Language (AML) which is provided by the Arc/Info GIS package. This approach can also support minimal numerical manipulations since AML is not suited to perform complex numerical manipulations. The two systems however still remain separate.

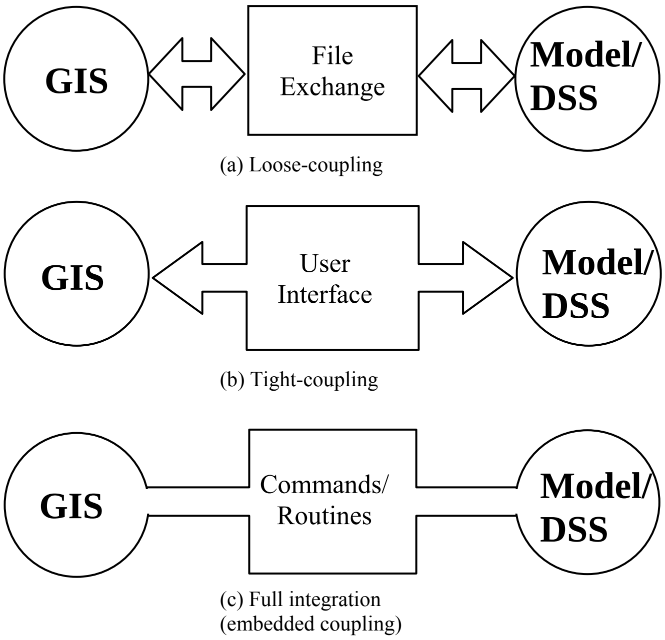

Figure 7.1 Schematic classification of GIS-integration methods in terms of the extent of integration (Adapted from Jun, 2000).

Techniques that result in “tight-coupling”, thus producing an integrated system where the GIS routines and other model routines share the same database, do not require the availability of the source code or an in-depth understanding of the GIS code. These techniques are less demanding in programming and are accessible to most modellers. Djokic and Maidment (1993) recommended that the best method to achieve “tight-coupling” is to use an application programming interface (API). In an API for GIS, a library of routines allows the user to access and integrate most of the functional capacities of the GIS in a standard programming language; this allows the user to write analytical programs, which, through the API functions, directly handle spatial data management, graphic display, and user interaction. Another important technique for "tight-" or "deep-" coupling is the use of object-oriented programming which was recommended by Raper and Livingstone (1996).

The idea behind the use of object-oriented programming is that a river basin is perceived as consisting of objects that interact in specific ways. In the river basin the coupling will involve spatial objects and thematic objects. Spatial objects represent real world entities, and thematic objects include attributes, methods and topics. The attributes include spatial, external physical, environmental and socio-economic data related to the spatial objects, the methods are the rules or functions describing the relationships between the objects, while the topics represent the objectives or tasks to be reached or completed.

In the most complex form of coupling, embedded coupling or full integration, a more complete integration can be achieved by creating user-specified routines through generic programming languages such as FORTRAN or C and adding them into the existing set of commands or routines of the GIS package. This requires such resources as source codes or command libraries and relatively complicated programming, which is not available to most GIS users.

Jun (2000) recommends that if resources permit, users should apply embedded coupling which gives the best results through ease of operation, facilitating quick achievement of modelling objectives. He further pointed out that embedded coupling is less error prone as data transfers are not done and has the added ability of looking at both the spatial system and the environmental system as one. Bivand and Lucas (1997) also explained that full integration gives better, faster, easier systems and enhanced water resources analysis. However modules should not be so tightly integrated that the potential universality of the design is seriously reduced.

The tendency to build the best model to represent the physical processes in reality, and then worry about finding data that would fit into that model should be avoided. Modellers should first identify the data and then build a spatial hydrological model that uses the data that are actually available or can be obtained within the project boundaries.

GIS integration is therefore expected to address input data requirements and the specific roles of GIS in the modelling.

The selection of a GIS coupling technique in South Africa should be guided by a number of issues which include the following:

- The GIS procurement costs including the licences of the GIS components and source code to the model developers and users. As an example a complete GIS package (Arc-GIS) costs more than R193,000 (GIMS, 2003) which most water resources practitioners in South Africa find unaffordable.
- Availability of adequate skills in the developer team to address short-term needs during development and the long-term demands of maintaining the water resources model and its GIS components.
- Data availability and the formats of existing GIS data. Most projects in South Africa tend to utilise the digitised, relatively coarse 1:50,000 maps in GIS. These input data are major constraints in GIS integration. Efforts to develop complete integration of the water resources model and GIS may not add value in cases with poor input data, such that "loose-coupling" will be the most appropriate approach. In cases where continuous DTMs are available covering the whole study area, it will be important to consider full integration.
- Availability of associated software for use with the GIS applications. As an example, most GIS users in South Africa are using ESRI GIS software, however these users usually lack the Arc-Objects and other modules required for integrating this type of GIS software with models. In many cases they have licences for limited functionality of the GIS package thus limiting the possibilities of how the GIS integration can be handled.
- The final modelling output presentation requirements, targeted audience expectations and the preferred model packaging and distribution.

Water resources practitioners should ideally seek to identify ways to address the challenges of integration when developing solutions to water resources management and planning problems. Some of the ideal requirements of an integrated system of water resources modelling and GIS include the following:

- All GIS data must be held in a common geospatial coordinate system,
- The primary structure used for spatial representation of large regions must be vector data (points, lines, and areas), supported by raster and Triangular Irregular Network (TIN) surface data where necessary,
- Correct and complete datasets in suitable formats must be available,
- Relationships among geographic features in different data layers are needed to trace water movement from feature to feature through the landscape, and
- Geospatial information describing the water environment should be linked with time series information about water measurements to form a complete information system for water resources.

7.2 Digital Elevation Models (DEMs)

Digital Elevation Models (DEMs), also called Digital Terrain Models (DTMs), are used in water resources management and planning for hydrological applications to simulate overland flow routing and provide a three dimensional topography model. DEMs are used to delineate watersheds, analyze channel networks, predict soil water content, predict erosion potential, model non-point pollution and carry out flood and hydrograph analysis (Duke, Kienzle, Johnson and Byrne, 2003).

The most common DEM data structure is the raster or grid structure. This normally consists of a matrix of square grid cells with the mean cell elevation stored in a two-dimensional array. Other DEM structures, such as the triangulated irregular network and contour-based structure have very limited use in water resources and are not discussed further in this study. Water resources practitioners should be aware of the basic processes involved in DEMs, especially how the water flows are simulated in the models to appreciate some of the implications of using DEMs. Because DEM cells at their best spatial resolution may be as coarse as 30 m X 30 m (e.g. the United States Geographical Survey's commonly used 7.5 minute DEMs), they often do not represent artificial linear features. As a result of the coarse nature of DEMs, flow direction matrices that are derived from DEMs alone are often inaccurate (Duke *et al.*, 2003). DEMs with a grid or raster data structure use the deterministic eight neighbour (D8) flow direction algorithm to simulate flows (Figure 7.2). Other flow determination methods are also used but the dominant flow direction determination technique is through the comparison of the elevation of each cell with its eight neighbours and allocating a single flow direction from each cell to the neighbouring cell presenting the steepest gradient.

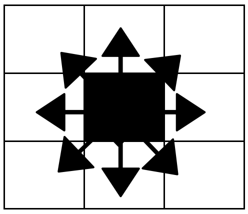

Figure 7.2 Deterministic eight neighbour (D8) flow direction in DEMs.

Each DEM cell is allocated one uniform ground level irrespective of other possible height variations within the cell. Depending on the sizes of each cell many linear features such as tillage furrows, culvert drains and road embankments are not accounted for in the DEMs. Representation of artificial or man-made linear functions on DEMs is possible through the use of specialised and separate algorithms. In the case of linear depressions, the modeller is advised to use ancillary stream data to impose or “burn” the stream vector data on to the DEM. “Burning” stream flow data involves the use of software to lower DEM cells accordingly to produce a manipulated flow direction matrix which is more accurate. However, there are terrain pits where, say, a cell-X will be lower than all the surrounding cells such that flow from this cell-X cannot be determined. Oliveria and Maidment (2000) recommended that such depressions should be filled first before imposing any vector data for linear features. They advised on a method of converting the wholly raster data into vector based polyline features to define the river reaches and other linear features as well as maintaining a raster domain for polygon features to define the sub-basins. In the case of embankments or other linear high points, a road enforcement algorithm (REA) or other software should be used to produce a secondary single flow direction matrix that accounts for roads and elevated linear features. The matrix is then imposed on the topographically derived (DEM) flow direction matrix.

The accuracy of DEMs cannot be better than the source data. The data used in DEMs, are usually derived from aerial photographic interpretation, topographic maps, field surveys, remote sensing and geographical positioning systems (GPS). In many cases there is no quantitative assessment of DEM accuracy, and error propagation to secondary parameters such as slope and aspect is not addressed (Monckton, 1994). The result is that poor decisions are made on the basis of poor input data.

To determine DEM accuracy, independent knowledge of the topography is required to determine the difference between the digital surface and the real elevations of the same locations on the ground. This requires both a suitable sample of ground truth points, and suitable statistics from which to derive error terms (Barringer and Lilburne, 1997). Modellers are cautioned against taking such ground truth points from the same topographic database as the contours, in the form of local spot heights recorded at trig stations and local peaks. Trig beacons and spot heights do not provide a good sample of the landscape since they over-represent peaks, under-represent low areas, and may be non-randomly distributed with a bias towards hilly areas. Acquisition of ground truth points should preferably be derived by independent survey, either photogrammetric, traditional field survey or by using Global Positioning Systems (GPS) (Barringer and Lilburne, 1997).

The RMSE between DEM and ground truth elevations can be used to measure DEM accuracy:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} d_i^2}{n}} \qquad \text{Equation 7.1}$$

where n = number of points

$d_i = Z_{ground.i} - Z_{dtm.i}$

$Z_{ground.i}$ = ground elevation recorded at point i

$Z_{dtm.i}$ = DEM elevation at point i

Alternatively Li (1988) advocates the use of the standard error (s) and mean error ($\bar{d}$) to determine DEM accuracy:

$$s = \sqrt{\frac{\sum_{i=1}^{n} (d_i - \bar{d})^2}{n}} \qquad \text{where} \qquad \bar{d} = \frac{\sum_{i=1}^{n} d_i}{n} \qquad \text{Equation 7.2}$$

Where n and d_i are as explained for equation 7.1.

RMSE is the more widely used statistic but assumes a zero mean error ($\bar{d}$), and therefore assumes no systematic bias in the DEM. This is not a justified assumption, according to Monckton (1994), who pointed out that systematic errors occur frequently. A good example of the use of RMSE is the classification of DEMs in the USA. In the United States, DEMs are available in three levels. The classification of these levels is mainly based on accuracy (Garbrecht and Martz, 2000). In level 1, a vertical RMSE of 7 metres is the targeted accuracy standard and a RMSE of 15 metres is the maximum permitted. Level 2 DEMs have identifiable systematic errors removed and the maximum permitted RMSE is one-half of the original map contour intervals. The level 3 DEMs have a maximum permitted error of one-third of the contour interval. Water resources modellers are advised to adhere to specified DEM accuracy levels and to use the recommendations provided above in cases where standards have not been specified.

In selecting DEMs, water resources or hydrological modellers should consider both quality and resolution (Garbrecht and Martz, 2000). Quality refers to the accuracy of the elevation data, while resolution refers to the precision of the data, specifically to the horizontal grid spacing and vertical elevation increase. Quality and resolution must be consistent with the scale and model of physical processes under consideration and with the study objectives.

7.3 Databases and data models in water resources modelling

Traditionally, water resources data usually consisted of time series data on point observations of water resources phenomena, including rainfall, streamflow, water quality, and climate. Integration and the use of GIS among many other recent developments in water resources management and planning have changed the concept of water resources data. Data for water resources management are now expected to cover the management of river basin erosion, water resource quality objectives, floods, water quality, sedimentation and land use. The trend in integrated water resources management is such that data, database systems, modelling tools and water system rules must now be incorporated into a single system that can be modelled and run in unison. The design of the databases used in such integrated water systems should take into consideration the following:

- **Data formats:** The data should comply with prescribed standards in terms of formats and must be in those formats that allow integrated approaches. As an example, Maidment (2001) expressed the need for geospatial data and temporal water resources data to be captured and archived in the same formats and within the same environments. He explained that, rather than just applying GIS in water resources, spatial data and temporal data can now be viewed as just similar information sources that the water resources practitioner can access and use.

- **Data accuracy:** The maintenance of data accuracy should ideally start early, that is at data collection stages. At the water resources modelling stage, the modeller should ensure that the data used are within acceptable accuracy levels and that meta-data are also entered into the database at the same time as the data. Some data are useless when information about the data (meta-data) is not available. Meta-data in water resources should ideally include:

    - Source of data (Name of the institution supplying the data and the location of the gauge used for data observations giving the name and again using the geographic coordinates).
    - Data accuracy information (use of standard flags for each data element is preferred especially in time series data. As an example, flags for rainfall and stream flow data may include the symbol "M" for missing data, "999" for a data element that was beyond measurement capacity and "P" for patched data).
    - Confidence levels associated with observed data discrepancies. As an example the USGS gives an "excellent" rating if 95 % of daily discharges are within 5 % of the true value (Donigian, 2001).
    - The dates when the data were collected and reasons for the collection process if it influenced the data
    - Details of anything that was done to the data, for example corrections to some data elements, patching of missing values, data extensions or other forms of data manipulation
    - An accurate and up-to-date directory of station characteristics and changes that occurred during the period of data recording. The introduction of a new recording

technique or the employment of new personnel to read and record data usually introduce new trends in the data.

- **Database structure:** The structure must be simple and flexible, allowing secure storage, easy user access to data elements, and generic to accommodate linkages to other related uses that may not directly relate to water resources. As an example a person working on an agriculture study may wish to utilise the data on soils in the water resources database. If the soils data components in the water resources project are not accessible such that they have to be developed again for the agriculture project, then this database is considered to be poorly structured and formatted. The use of object-oriented programming and a relational database structure for database and model interfaces was noted in this study to be the best approach in water resources management and planning database architecture. This approach accommodates the different spatial and temporal scales as well as providing a suitable platform for direct integration with different models that will access and process the data. Another approach that may be pursued in cases where models have to link to existing databases is the use of e**x**tensible **m**arkup **l**anguage (XML) to act as the link between the existing database and the models. In this case the user interacts with the XML database management system which uses XML generated data files to transfer information between the models and the database. A methodology of developing an open modelling framework using the XML approach is presented and recommended by Kokkonen, Jolma and Koivusalo (2003) for hydrological and climatic models. The advantage of using XML or other intermediate platform is that data and model components developed independent of each other by different developers can be linked. However this method lacks the full integration that is possible when object oriented programming and a relational database structure are used.

- **Level of detail, scales and resolution:** McKinney *et al.* (1999) recommended the use of object-oriented programming coupled with a relational database to allow the possibility to vary the simulation scales in one model. This allows detailed high resolution simulations (field or plot scale) as well as large scale (regional) assessments within a single water resources model.

- **Targeted uses and output formats**: Database developers should clearly consider the implications of the various database architecture selections to ensure that the final database is technically sound and appropriate to the project needs. A database must at least satisfy the core objectives behind its development, before attempting to satisfy other peripheral needs.

- **Organisational and industrial compliance:** Databases should comply to organisational requirements as well as the industrial requirements to allow other external users to be able to access the data and also to ensure that the database can be updated from external sources. The provisions in the NWA (Republic of South Africa, 1998) on the development of national information systems should guide water resources database developments to ensure industry wide compliance.

- **Database updating methods and data entry quality management**: The methods used to update the database are crucial in database design. As an example, databases that are updated by inputs from different users in different locations will require additional and more stringent data quality management systems than those that use a single automatic entry such as a cellphone connection relaying data from an automatic water level gauge in a dam. The different requirements in database updating affect the nature of the database to be used, including: software, data structure and data storage and interfacing hardware.

- **Background of original data collection:** Data are usually collected for a specific use in a particular project such as a water resources model with unique data needs. As an example, stations for measuring hydrological and climatic variables have been established for once-off projects such as dam developments and closed soon after the project is finished. The archiving of these data should be accompanied by an explanation indicating that the data were collected for a once-off project and that no further data can be obtained from that station. Another example is that of the water quality database developed for use in the model WQ2000 (Herold, 2003) for the Vaal River. Salinity was the focus of water quality assessments in the WQ2000 study while other water quality variables received insignificant attention. Users of such a database may end up developing or selecting a model that is strong on salinity thus forcing their modelling exercise to focus only on salinity, which is just one of the many water quality problems in the Vaal River. Other major water quality problems such as the high levels of phenols and sulphur oxides from the electricity power stations and mining activities in the upper Vaal River catchment area will need to be accounted for to ensure completeness in the water resources modelling process.

7.4 Stakeholder factors in water resources modelling

In the past decade, water resources management has faced a paradigm shift, from a top-down approach to participatory management (Flügel and Staudenrausch, 1999). Stakeholder participation now plays a significant role in the development and use of water resources management tools. The range of human related issues influencing the water resources tools include: experiences and preferences by developers, owners of the problems and stakeholders, as well as knowledge and abilities of developers and users. The model conceiver's perceptions, which are influenced by his/her value system, are also crucial as they are the basis for most of the choices made during model development. Kloprogge and van der Sluijs (2002) pointed out that a model and its outcomes may be difficult for users and stakeholders to accept which will lead to conflict if the model does not adequately reflect their knowledge and perspectives.

The process of solution development in water resources management, including the development of models, should take into account the following stakeholder factors:

- Information made available for decision making must be supported by adequate knowledge, data and a good appreciation of the problem. Poch (2002) recommended that environmental models should not only seek to process numerical aspects but should include reliable knowledge and experiences from experts and the wider public participants.
- Project proponents should ensure that they are prepared to negotiate with stakeholders and provide alternatives and offer compromises. A case study involving poor consultation is presented in the Orange River Basin Development study by the Secretariat of the World Commission on Dams. This study revealed that the little consultation that took place before the implementation of the Orange River

Development Project (ORDP), and provided no alternatives to stakeholders. Stakeholders in the Orange River Basin are still disgruntled by the fact that they were given no alternatives during the consultations (WCD, 1998).

- Suitable human expertise with skills in all the important disciplines are important in water resources model development. Winiwarter and Schmak (2002) advised that models that simulate natural systems such as water resources models need to be assembled by at least two types of experts: those that understand the natural system, and those that are able to transform the concepts into computing algorithms.
- Methods should not be developed simply to please influential groups of stakeholders, but rather the groups should be empowered to understand all the variables involved so that they can contribute competently to the development of the best methods.
- For models to be used effectively, stakeholders should understand model assumptions, content, capabilities and output. They also need to have confidence in the model validity and view it as a useful decision support tool that is transparent, flexible and accessible (Palmer and Woods, 1999).
- Stakeholders should be involved early in the solution development process to accommodate their participation in the solution development and hence provide a basis for long-term ownership of solutions. In South Africa, platforms for stakeholder involvement include workshops, conference sessions, questionnaires, user forums and targeted publications. The allocation of resources to involve stakeholders should be carefully planned and managed as it can delay or derail a project if poorly planned and implemented. Modellers should avoid over commitment of resources to stakeholder involvement. It has been observed and documented that stakeholder involvement cannot result in a perfect decision. d'Aquino (2002) advised that a perfect decision can never be found in catchment modelling which aims to fulfil all stakeholder perceptions. Giraud *et al*. (2002) recommended the use of socio-economic behavioural models by stakeholders as platforms for group contributions to the design of solutions in catchment based water resources management. Giraud *et al.* (2002) also argued that the stakeholders have a strong influence on catchment based water management but usually lack the tools to make well-informed and well-structured group contributions.
- Long-term commitment by problem owners, project funders and other involved stakeholders is crucial for the sustainability of water resources solutions.

- Authorities or those responsible for financing solutions in water resources tend to target the reduction of costs in project implementation by using all possible actions including limiting stakeholder involvement and even scaling down the Environmental Impact Assessments. The WCD (1998) pointed out that the South African parliament quickly authorised the implementation of the first phase of the ORDP after inadequate consultation of stakeholder and environmental impact study to reduce costs and avoid possible delays. However, water resources practitioners should note that the acceptance and ownership of water resources solutions by stakeholders is in fact one key parameter used to measure the success of water resources management and planning. In evaluating the risks associated with IWRM projects and stakeholder needs, Rees (2002) also pointed out that one of the key objectives in integrated water resources management is the maximisation of total social welfare.

7.5 Summary of recommendations on spatial data and stakeholder inputs

All water resources management problems inherently have a spatial dimension. The development of models for water resources management tools should always seek to include a GIS approach for handling spatial data. The use of GIS should not be limited to mapping but should ideally involve the following:

- Data analysis and model output predictions
- Storage of data
- Model interfacing
- Post processing
- Results presentation
- Model documentation support
- Data and study area displays

Model user and developers should note that the type of GIS coupling used in water resources management tools will affect the kind of use that will be possible, as well as data management. In this study, the model developers and users are encouraged to develop full integration of models and the GIS tools, where spatial data are available and are required in the modelling process. This allows a single database to be developed for spatial data and other types of input data. Seamless integration of GIS and water resources models has the added advantage of reducing sequential batch processing of model modules, thus eliminating the potential for input errors, and allowing multiple module simulations at each time step.

In cases where GIS software costs restrict its use, modellers and developers in water resources projects should ideally seek suitable open source GIS software and utilise it at no added costs in their projects. Listings, information and download linkages of hundreds of such free open source GIS software are available on the world wide web. Ramsey (2004) also provides technical assessments of several GIS tools that can be utilised in various GIS projects.

When selecting an Open Source GIS or Freeware GIS tool, the user or model developer should consider a number of factors. Some of the most important issues for consideration deal with the specifications of the tool. These include:

- What the GIS software can be used for.
- The implementation language used in the tool.
- The users or developer operating system requirements.
- The spatial data input file sizes and formats.
- Specifications of the computer requirements for running the GIS tool.
- Availability, format and adequacy of GIS user or developer support.

Digital Elevation Models (DEMs)

Ideally, DEMs should be used in models where the water resources planning and management questions to be addressed involve variables that are dependent on land elevation. Modelling variables for which DEMs will be required include: water flows, orographic rainfall, air current flows, groundwater levels and river recharge.

In selecting a suitable DEM, water resources or hydrological modellers should consider both quality and resolution, where quality refers to the accuracy of the elevation data, while resolution refers to the precision of the data. Since the accuracy of DEMs cannot be better than the source maps, photos and other documents, the selection of GIS data sources should receive special attention.

Databases

In the development of water resources databases, modellers should aim to develop a single database for all input data as well as input parameters. Ideally, a single database that can be applied in a number of related models is required. The NWA already promotes the development of such national databases.

Another important aspect of good data management is to provide meta-data for all the model input data.

Stakeholders

The development of solutions that are appropriate and acceptable to all stakeholders is also discussed in this chapter as an important characteristic of good water resources management and planning. In most cases, stakeholders have different levels of understanding and appreciation of the issues surrounding the water resource modelling processes. Water resources projects should therefore include initiatives to improve the levels of stakeholder understanding, thus empowering the different stakeholders to ensure that they share a common understanding and are able to provide meaningful support and feedback in water resources solution developments.