Guidelines for
water management modelling
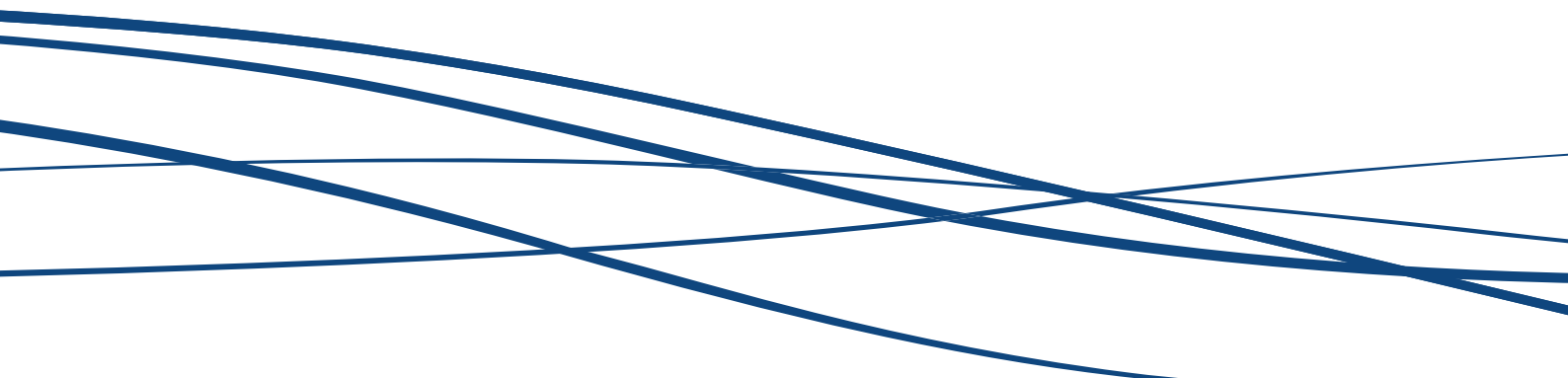
Towards best-practice model application

Australian Government

Department of Innovation, Industry, Science and Research

Department of Sustainability, Environment, Water, Population and Communities
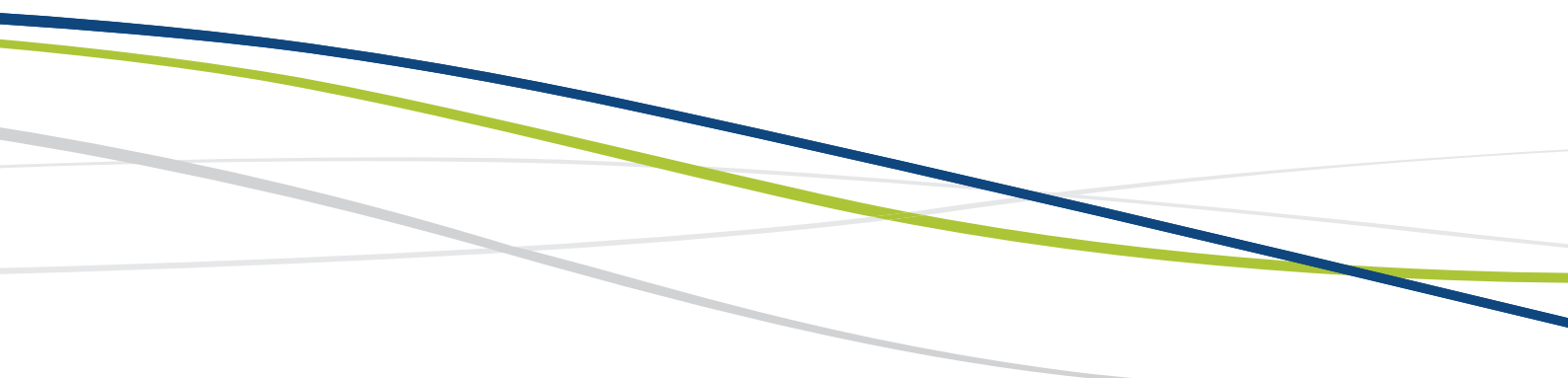
National Water Commission

Contents

1

Introduction

Background

As a major provider of modelling products eWater has adopted a policy (eWater, 2010) which recognises its responsibility to foster a best practice approach to the use of its products.

There is also now a growing expectation in Australia (and elsewhere) that there should be a consistent approach to applying models used to support water management decisions across the nation, with the aims (amongst others) of:

- Improving modelling practice;

- Removing inconsistencies between model applications, including in adjoining catchments where the same model code is used in both, and in situations such as managed river systems that interact with one another;

- Providing quality assurance, including evaluation of uncertainty;

- Improving decision making, including improving the use of science to improve the quality and robustness of decisions made and outcomes;

- Improving communication with end-users of model results: water managers, decision makers and the wider community; and

- Providing a process that is transparent, robust and repeatable.

A common thread through other water-related modelling guidelines (discussed further in) is an emphasis on Quality Assurance as a way to achieving Best Practice Modelling. Considering this point, the user needs outlined above and the potentially wide range of modelling domains relevant to water management for which guidelines could be developed, an appropriate approach is to develop a hierarchy, or family, of guidelines. At the highest level are overarching guidelines that provide a generic procedure to underpin delivery of quality assured, best modelling practice outcomes. This procedure would provide a framework for a consistent set of supporting domain-specific guidelines which, in turn, would be supported by a set of model specific guidelines. The level of detail increases from the highest level to the lowest. The guidelines in this document represent the highest level in that architecture.

This document is not meant to be prescriptive, nor to re-invent material and concepts that are better explained or already available elsewhere. Hence, these high level guidelines draw as appropriate on other material already available from the international literature (eg Bay-Delta Modeling Forum (2000), CREM (2008), Jakeman et al, (2006), Middlemis et al, (2000), Packman and Old (2005), Scholten et al (2007), USEPA (2002), Van Waveren, et al (2000), Water New Zealand (2009)) and on the Internet.

Definition of Best Practice Modelling

Best Practice Modelling can be defined as a series of quality assurance principles and actions to ensure that model development, implementation and application are the best achievable, commensurate with the intended purpose (eWater, 2010).

What is in practice "best achievable, commensurate with the intended purpose" may be subject to data availability, time, budget and other resourcing constraints. Hence, what is meant by the term "Best Practice Modelling" can vary. Not only does it depend on the circumstances of the project, particularly fitness for purpose, but it also depends to a great degree on interpretation in peer review. This, in turn, will be influenced by the general state of knowledge and technology in the modelling field, which can be expected to progressively develop over time (such as new remote sensing data sources coming on line, and new computing hardware), as well as data, time, budget and resourcing constraints. Best Practice Modelling provides for a strategic approach to modelling which enables the trade-offs that may be imposed by these constraints to be better managed, and assists in identifying priorities for addressing model and data limitations.

Scope   This document – hereafter referred to as "this guidance" – has been prepared by eWater to provide guidance on the application of modelling tools (software or other) to solve problems and on providing decision support to end-users of model results. It proposes a high level generic procedure that is intended to result in quality assured model applications. The overall decision framework is illustrated in *Figure 1*.
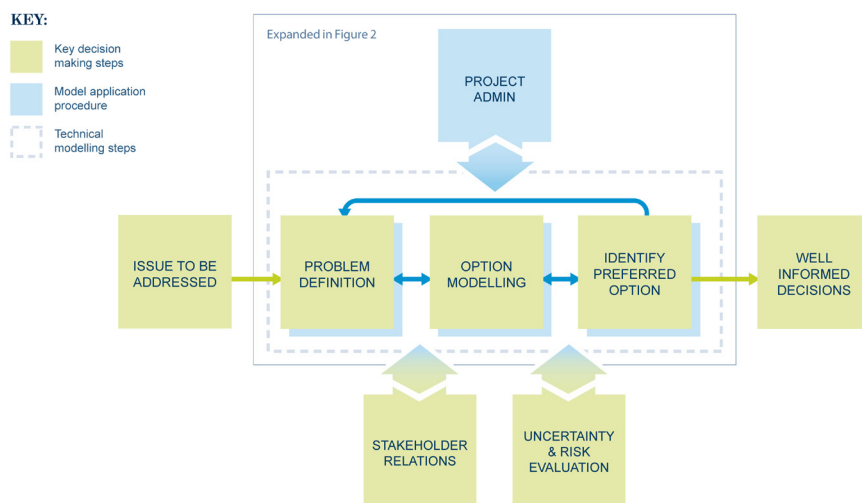
Figure 1
Decision framework for model application

This guidance is intended to provide a framework and common structure for consistent sets of guidelines for various modelling domains and eWater products relevant to water management (in the broadest sense of the term), although in principle it could be relevant to other modelling domains and other modelling products. It covers risk assessment, decision support, and communication and interaction between modellers and end-users of model results as well as technical aspects of modelling. However, this guidance does not address quality-assured development of modelling software as other eWater procedures cover this aspect.

The target audience for this guidance is chiefly practising modellers with appropriate background; ie it is not a textbook. It is also intended for managers, reviewers and

decision makers, and stakeholders more generally, to indicate how model application should be approached, to assist in managing expectations of what quantity and quality of information is achievable and meaningful for a given modelling project, and to assist in assessing whether the modelling is fit for purpose.

Consistent with eWater's identified responsibilities to its partners, and the expectations of water managers in Australia and elsewhere, the aim is to support a best practice approach to modelling and achieving consistent practice, especially when implementing a suite of models (such as in the context of the Murray-Darling Basin). Based on Refsgaard et al (2005a), this guidance may be classified as public interaction guidelines.

This guidance provides for an integrated approach that enables interactions and feedbacks between all domains relevant to water management (eg hydrological, ecological, engineering, social, economic and environmental) to be considered. It provides for outputs of analyses in these domains to feed into multi-objective decision analysis, which links outputs with views and preferences of multiple stakeholders and enables decisions to be made with the aim of satisfying the objectives of all sectors. In addition, the procedure in this guidance is intended to be flexible enough to accommodate variations in the meaning of the term "Best Practice Modelling" and also allow for continuous improvement as the state of knowledge and technology in the modelling field develops and improves.

Note   Stakeholders are individuals, organisations or groups with an interest
         in a project and its outcomes; these can include the organisation
         commissioning the project (the client), water managers, decision makers,
         community groups and individual members of the public.

The rest of this document is organised as follows: *Chapter 2 (page 9)* provides guidance on quality assured model application, and is the core of this document. *Chapter 3 (page 49)* provides supplementary material on model choice, *Chapter 4 (page 57)* provides information on further reading and a case study is described in *Chapter 5 (page 61)*. A glossary of some key terms, which draws on a number of sources including the eWater Glossary on the Internet (*www.ewater. com.au/glossary/*), is provided in *Chapter 6 (page 73)*. *Chapter 7 (page 81)* contains a list of references.

2

Procedure for Quality Assured Model Application

The generic procedure proposed for quality assured model application is adapted from Blackmore et al (2009) and is summarised in *Figure 2*. The project administration phase shown has requirements relevant to all steps in the other three phases. The figure shows feedback occurring between the phases and between steps within each phase but it has been simplified in that there will also be feedback between certain steps in different phases.

While there is a logic to the sequencing of this procedure, circumstances may arise where the needs of a given project are best met by considering certain steps out of sequence and this guidance does not seek to constrain the flexibility to do this where it is needed. In particular it is quite likely that not all steps in the project administration phase will need to be considered until many of the steps in the problem definition phase, at least, have been considered. Likewise, feedbacks between every step within each phase will not always be needed, though it is likely that some review of the outcomes of each phase in terms of the other phases, will greatly enhance the benefits of this process.

It is also recognised that it will not be necessary or appropriate to go through all steps in the procedure in detail in every project; this may be particularly true when the work is part of an ongoing long term modelling program. However, every step should at least be thought about and conclusions documented. Indeed, this guidance advocates a "horses for courses" approach, in that the extent to which the proposed quality assurance procedure needs to be implemented will vary depending on the scope and intended purpose of the project, risks associated with project outcomes and, to some extent, on the magnitude of the project as well.

Project Administration

It is essential that some elements of this phase are undertaken at the very start of any project. Critical among these is the appointment of a project manager who is assigned responsibility for the delivery of the project. Decisions on budgets and timeframes are also likely to be made early in the life of the project, although if these can be deferred until later in the problem definition phase better outcomes may be able to be obtained from the project.

Project Governance

Governance arrangements that are appropriate will vary depending on the size and significance of the project. As a general rule, any project requires a project manager and a project director; for small projects, the project manager may also be the person who does the work. The project director has an overseeing and review function, and this will be particularly important in small projects where there may be no other formal review mechanism in place.

For major projects, or projects that are important or potentially sensitive, a Steering Committee with stakeholder representation should be set up. The role of this committee is to provide overall direction and it should also help in gaining stakeholder acceptance of the project. The project director may delegate technical review of some or all elements of the project to a technical reference panel or its equivalent. It is the

Procedure for Quality Assured Model Application

responsibility of the project director and/or steering committee to appoint technical reviewers who have an appropriate level of technical expertise to review the work undertaken. Project directors should review technical work only in areas where they themselves have an adequate level of technical competence.
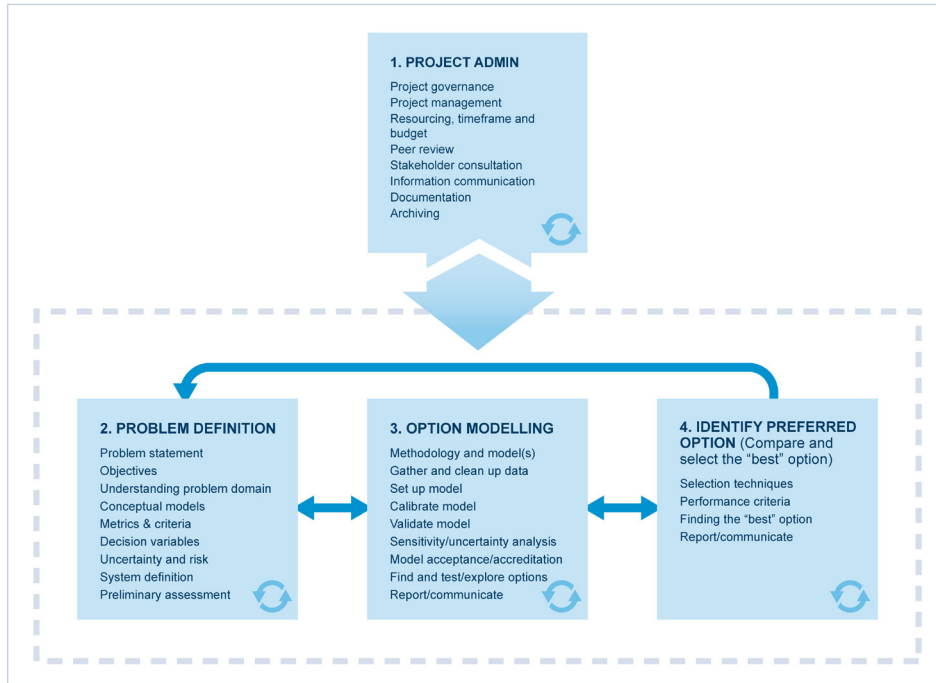
*Figure 2
Flow chart of procedure for quality assured model application (adapted from Blackmore et al, 2009)*

Project Management

Project management can be described as a process based on use of management and project domain relevant skills and knowledge to organise, plan, control, monitor, evaluate and deliver a project to achieve the project objectives on time, on budget and to agreed quality and performance levels. Detailed requirements for project management are commonly contained in corporate project management procedures and these should be followed where available. Where these are not available there are numerous publications and other sources of information on project management procedures; prominent among these are the PRINCE2 methodology (Office of Government Commerce, 2009) and a Guide to the Project Management Body of Knowledge (Project Management Institute, 2008). Hence, discussion of project management in this guidance is confined to aspects particularly relevant to model application for water management projects.

Elements of project management include developing a project control plan, achieving agreement on milestones, defining performance metrics and criteria and monitoring progress against all these. Two important aspects are to manage any changes in the scope of work as the project progresses, and to regularly update estimates of the cost of completion. Relevant stakeholders should be alerted to changes in the scope or costs, particularly increases, and their implications for the project budget, schedule and quality as early as possible. Allied to this is the need to manage stakeholder expectations to ensure they do not increase as the project progresses, and also

to identify and consult about any trade-offs in scope and methodology that may be appropriate.

Another need, which also relates to project governance, is to ensure that decisions relating to the project and its outcomes, and the reasoning underlying decisions, are well documented and made available to all participants. This ensures that all participants are working from the same understanding, and that someone coming along later can see what decisions were made, what was done to underpin the decisions and why. Sufficient detail should be provided to enable the project to be reproduced or extended by an independent project team.

Resourcing, Timeframe and Budget

Financial and physical resources can limit the options available to stakeholders which in turn can influence the problem space. Similarly, the timeframe may also bound the problems that can be addressed (ie "what can happen over this time?") and also influence the solution space (ie "what can be done in a specified time?"). This is discussed further in *Problem Statement (page 17)*.

Peer Review

Peer review is important for establishing the credibility, reliability and robustness of results and the methodology used to obtain the results. It is undertaken by people with specialist understanding in fields relevant to the project. It thus differs from stakeholder consultation, where people without specialist understanding will also be involved. The following levels of peer review can be considered:

• Internal peer review, where the reviewers are from the organisation undertaking the project;

• External peer review where the reviewers are appointed by the organisation undertaking the project; and

• External peer review where the reviewers are appointed by a third party, such as stakeholders or an external regulator.

Different levels of peer review may be appropriate at different stages of the project and for many projects the third level listed above will not be needed at all. However, peer review at key points in the project will help ensure it is on track, and remains that way, and will minimise if not eliminate the need to repeat work. Evidence of peer review of deliverables from the project should always be documented.

Internal peer review should be undertaken by a suitably qualified individual or group in the organisation for each step or deliverable. For important steps, such as model calibration and scenario analyses, review should be undertaken at intermediate stages to ensure results make sense and that any unexpected or counter intuitive results can be explained, even if only via informal discussions.

The need for external peer review will depend on a number of factors, which may not have any relationship to project size (however measured), such as:

- The confidence or trust that stakeholders or the client have in the internal peer review processes;

- Whether there is a regulatory requirement for peer review;

- Whether decisions made based on project outcomes are likely to be legally challenged by stakeholders or other parties;

- Whether decisions made based on project outcomes are likely to become part of a political agenda, or could be otherwise sensitive; and

- How concerned stakeholders or the client are with the scientific robustness and credibility of outcomes versus a more pragmatic attitude to "just get things done".

Where external peer review is required it should be undertaken at completion of key steps and at other appropriate times. If a project is likely to be highly contentious or sensitive then it is often prudent for the external peer reviewer to be actively engaged early in the process to avoid a situation where peer reviewers need to be appointed later by a third party.

However, there is a danger with external peer review, especially where the project involves multiple objectives and many different disciplines. In these cases the project management team will have developed a unique understanding of the issues involved through working together, which allows them to make judgements on appropriate analysis and methods. If external reviewers do not have this understanding, they may have difficulty in suggesting or advising on appropriate methods and levels of detail. This is particularly true if the reviewers come from specific disciplines. The ambit of the review should be defined to match the skills and knowledge of the reviewer.

Peer review is further discussed in relation to *Project Governance (page 10)*, with reference to *Model Acceptance/Accreditation (page 42)*, and at various other points in this guidance. More detailed guidance on external peer review is available from other sources such as CREM (2008, Box C1).

Stakeholder Consultation

Stakeholder consultation is important for building acceptance of, and ownership in, the project. It should start early in the project, ideally at the problem definition stage, with further consultation occurring when the project methodology is being developed and then during the process of identifying and analysing solution options, and presenting results. Early involvement is likely to lead to better outcomes, and active participation by stakeholders will give greater benefits for ultimate acceptance than if stakeholders just take a passive role.

In addition to building acceptance and ownership, involving stakeholders in problem definition and in developing the project methodology will have the added benefit of enabling their background knowledge to be captured in the process (this includes

corporate knowledge and local experience). It should also ensure that the approach is relevant to the question being asked, assist in clarifying exactly what that question is, and ensure expectations match time frames and budgets.

Management of expectations will be needed throughout the project and ongoing consultation will facilitate this. Ongoing consultation will also help ensure that if changes in methodology or scope are needed, with or without changes in budget and timelines, these needs are understood and accepted.

Information Communication Clear communication is necessary at all levels and throughout the process. Amongst other things, it supports stakeholder consultation. The aim is to ensure that:

- The objectives of the project are agreed and understood by all stakeholders;

- All project participants have the same understanding of what they are doing, have the information they need to do their work and are using the same data etc;

- The decision makers have all the information they need in a form that is easy to ;understand;

- The project is well documented so that if questions arise later there is a clear record of the reasons for decisions;

- Everyone who is affected by or influences the outcome of the project understands what is going on and what they need to do to make it work.

To ensure adequate communication, a communication strategy should be developed early in the project in consultation with appropriate stakeholders. The communication strategy should be commensurate with the sensitivity and risks associated with the project and its outcomes. It should at least indicate how, when, where and to whom information about the project, particularly results, will be presented. There are a number of modes by which this can occur and, from the beginning, it is important to consider who will be affected by the decision making, potential audiences more widely, which stakeholders to involve in the process and how to best communicate to the range of different audiences. Where appropriate, communications specialists can advise on these aspects. If successful delivery of the project depends on stakeholder participation (such as the actions of individual farmers), then early involvement and transparency are likely to lead to better outcomes.

It is also important to consider privacy and security issues. It may be inappropriate to report or otherwise communicate some information for these reasons.

Where presentation of project results is concerned, there are many options available. For example, presentation could be passive – such as a final report for decision makers to read – or it could be more interactive as in workshops and meetings. Input from stakeholders should be sought on how to communicate results during communication strategy development, as they will know how best for project outputs to be presented to them.

Communication can be an iterative process, with an agreement to review the process at intervals. The important thing is to enable rigorous, transparent, and defensible decision making of the best quality given the tools and information available at the time. In addition, it is important to establish the scope and timing of the reporting task early in the project.

Documentation Documentation plays a number of important roles, which are:

- To keep a record of what was done so that it can be reviewed and reproduced;

- To provide source or background material for further work and research;

- To effectively communicate the results from models; and

- To effectively communicate the decision making process, including decisions made, and the reasoning underlying decisions.

Good documentation supports the exchange of information with stakeholders, thereby supporting transparency of process and contributing to gaining acceptance of project findings. It will also enable someone coming along later to see what decisions were made, what was done to underpin the decisions and why, particularly if aspects of the project need to be revisited.

Archiving Model input data, model results and the version of software (either the version number, if a software versioning and archiving system is maintained centrally, or otherwise the software itself) used to create the results for adopted scenarios and any other scenarios of potentially enduring interest should be archived in a scenario management system. All reports from the project, including documentation of decisions relating to the project and its outcomes, and the reasoning underlying decisions, also need to be archived. Archiving of model output data, the software and input data used to create it, and documentation on decisions made is essential for making information from models available to stakeholders and also for ensuring repeatability if model runs need to be redone or updated.

As an example, the high level architecture of the main components of a system implemented in 2010 and the interactions between them are shown in *Figure 3*. This system links models, in this case denoted an Integrated River System Modelling Framework (IRSMF), with systems to save model scenarios and results to databases that can subsequently be interrogated by reporting tools to produce reports (Podger et al, 2010a).

The following conventions apply in *Figure 3* (Podger et al, 2010a):

- rounded rectangles are computer applications/programs;

- ovals are information stores (eg files on disk, a Subversion repository, or a database);

- solid lines depict the flow of information – direction shows where the data originates and its destination;

- dashed lines show data retrieval – the Reporting Tools obtain data from the Summary DB and the Model Run Results;

- dotted lines show data linkages, an association from one dataset to another (eg the information contained in the summary database is linked to the provenance details describing the IRSMF configuration used to produce it);

- reports are Excel spreadsheets, combining the "standard reports" predominantly used by modellers and the "reporting spreadsheets" typically used by the policy planners and the reporting group the system was developed for.
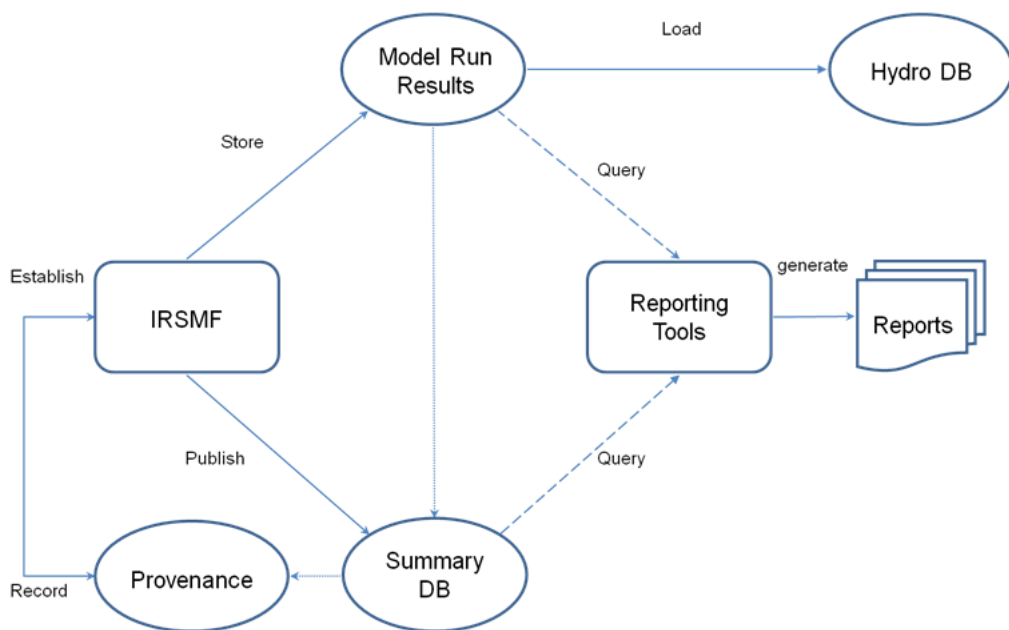
Figure 3
Overview of
architecture of
example archiving
system (after
Podger et al, 2010a)

Problem Definition

This phase involves defining the purpose and scope of the work, setting performance criteria and gaining early stakeholder buy-in and support. It is needed whether the work is a new project or an extension or update of previous work. It is also likely that many of the steps will be part of an iterative process that could go across the other two main phases in this procedure.

Often the project objectives, scope, timeframe and budget are defined up front, via project Terms of Reference (or Project Brief). Material in this section and in *Project Administration (page 10)* is relevant to determining the Terms of Reference in the first place, but Terms of Reference are rarely sufficiently detailed to allow a project to be implemented without further consideration and definition. Consultants typically use their Proposal (Bid) for undertaking the project to assist this and many projects include an inception phase as well to confirm and further clarify the objectives, scope and methodology (and sometimes timeframe and budget as well). The material in this section is also relevant to preparation of project proposals and the inception phase of projects, as are the parts of the material in *Option Modelling (page 24)* and *Identify*

relating to methodology and the material in .

Problem Statement

The problem to be addressed needs to be clearly articulated. This should involve consultation between water managers/policy officers and modellers, at a minimum, and, where appropriate, the wider community, especially as in some cases the problem will be defined differently by different stakeholders. Getting this step right will minimise the risk that the wrong tool will be used for the job.

Note	Defining the problem is the most important step in any solution finding strategy.

The problem and possible solutions should be initially considered free of time, resource and budget constraints, so that the true nature of the problem and wide-reaching or imaginative solutions are not overlooked; then the effect of these constraints on project scope and methodology should be clearly explained. This also provides an opportunity to demonstrate that more detailed approaches to solving the problem have been considered and explain why they were rejected based on scientific, technological, time, resource or budget constraints. Where multiple options are available for constraining the scope, it is also good practice to explain how priorities were assigned to addressing particular issues in the modelling process. Amongst other things, this will enable cost/expediency issues to be better managed. From these considerations a preliminary understanding of solution options and scenarios to be analysed should be obtained.

An extremely valuable question at this early stage is "Is there a role for modelling in the project and what is this role?" Firstly, it should not be assumed that there is inevitably a role for modelling. Secondly, the answer to this question provides part of the context for many subsidiary questions relating to model choice and implementation. It is very important not to approach this question with the view that a particular model or modelling tool is the end point.

Objectives

The objectives of the current project and the goals should be identified in a consultative process involving water managers/policy officers and modellers, and also the wider community where appropriate. As many water management decisions will often have more than one goal it will be important to ensure these are all identified.

Sometimes it can be useful to express objectives in a hierarchy that shows primary objectives, secondary objectives and so on. In this regard, consideration should also be given to possible additional future objectives and goals that could be met based on this project or on future projects that build upon the model established in this project. The decision on which option offers the best solution will be based upon whether, or how well, each option meets the agreed objectives.

Understanding the Problem Domain

The range of disciplines that needs to be brought to bear to address the problem at hand should be identified and agreed between water managers/policy officers and modellers and, where appropriate, with the wider community. Social acceptance, adaptation, environmental, and economic considerations are also part of the problem domain, in addition to issues of water management per se, that impact on the effectiveness of solutions. An example of the importance of the way that social and economic considerations interact with water system performance modelling would be in the consideration of the effectiveness of implementing a rebate scheme for householder installation of rain water tanks to reduce potable water consumption: simply providing tanks at subsidised cost may not mean that all of those tanks are used effectively by all of the householders that receive them to provide the optimum reduction in potable water usage that might otherwise be projected by a purely "physical" model of the water supply system.

System Definition

System definition requires identification of: system components, their behaviour and appropriate levels of abstraction; the interactions between components, including any feedbacks; and system boundaries, forcings, states and outputs. System definition is critical to model applicability and should be done in consultation with stakeholders. Explicit decisions about what is inside and what is outside the system boundaries is critical for all components of the system, including economic and social components as well as institutional arrangements and biophysical aspects, and should be guided by the problem at hand, state of knowledge and available information and resources. System definition also includes consideration of temporal and spatial scales.

The system definition would usually be documented in the form of a conceptual model of the system (see *Conceptual Models (page 19)*).

Selection of the appropriate temporal and spatial scales will very often be informed (if not determined) by the scale and extents in time and space of available input data. It is therefore common at the system definition stage to explore and then list the data that is or might be available for the project to aid in defining the spatial and temporal extent of the analysis.

Analysis of available data for the purposes of problem definition could be undertaken to varying levels of rigour, from a cursory statement of the likely available data based upon the experience of the modelling team and stakeholders to a detailed investigation of data, literature review into additional data sources and analysis of the gaps and quality of available data. It is particularly important that this activity should bring prior knowledge (such as corporate knowledge and knowledge from sources such as former staff of an organisation who have retired, and long term residents of the project area, as necessary) to bear at this point as doing this has the potential to avoid repeating previous work (including past mistakes) thereby saving time and effort, and ultimately lead to a better solution. There will also be intangible benefits in that making use of prior knowledge, particularly local knowledge, will enhance the credibility of the work

and make gaining stakeholder acceptance easier. The information could be either quantitative or qualitative. However, care needs to be exercised when using it as its reliability may be hard to assess. Agreement should be reached with stakeholders on how this information should be used; the process could be supported by formal protocols where appropriate. Use of prior knowledge should also flow through to methodology development, discussed in *Methodology Development (page 25)*. Workshops, meetings and background research are among the possible mechanisms for accessing this knowledge. A site inspection should also be done at this stage unless there are good grounds for not doing so.

Conceptual Models

Understanding of how the system to be analysed works, the relative importance of various components of the system and the functional relationships between these components, should be developed into a conceptual model. Assumptions should be clearly stated, particularly any key simplifications, and any noteworthy exclusions should be identified.

As different stakeholders may well have different perspectives on how the system works and which aspects are important, it is important that all stakeholders are consulted and agreement reached on the most appropriate conceptual model applicable to the project at hand. Modellers and stakeholders should also consider dealing with multiple, alternative conceptual models as this is in most cases the best way to address model structural uncertainty (Refsgaard et al., 2006). Developing and agreeing on conceptual models is a key aspect of problem definition, and may entail an iterative procedure involving at least the other steps in *Problem Definition (page 16)*.

Clarifying the conceptual model, or alternative conceptual models, is important even where there is an existing computer based numerical model available. This may show up limitations in adopting the existing model and possibly indicate means of mitigating these (Jakeman et al, 2006).

Representations of the conceptual model could include schematics, diagrams, maps, plans, drawings, flow charts, graphs, written and verbal descriptions, and equations. As stated by Jakeman et al (2006):

> *"Initially the conceptualisation may be rudimentary, with details postponed until the results of knowledge elicitation and data analysis can be exploited. A tentative initial conceptualisation and a visualisation such as a block diagram may be a great help in showing what else must be found out about the system."*

Relevant points for consideration in the process include the following (Packman and Old, 2005):

- Is the conceptual model (domains, boundary conditions, space/time scales, etc.) adequately defined, are all relevant processes/interdependencies addressed and assumptions clearly stated?

- Has the need for alternative conceptual models been assessed?

- Has the soundness of the conceptual model been assessed and does it make sense?

- Has this or a similar conceptual model been successfully applied in previous studies?

- Is the conceptual model consistent with the project objectives and required model complexity?

The corollary to these points is that numerical models adopted need to be consistent with the conceptual model.

Metrics and Criteria

Performance criteria and indicators that demonstrate compliance with the agreed objectives should be identified and agreed between stakeholders and project officers (eg modellers). When setting system performance criteria, due consideration should be given to socio-political, economic and environmental aspects. Model performance criteria and metrics are also important; needs for these are discussed in *Methodology Development (page 25)*.

In some cases, in addition to assessing the performance of the system in terms of the agreed metrics, certain levels of performance must be met. Some of these criteria are set by legislation (such as minimum acceptable water quality standards), others are set and agreed on by stakeholders. For example, an objective might be to reduce drinking water usage, where a target of a 20% reduction in mean annual per-capita drinking water usage by the year 2020 might be agreed by the stakeholders. In these cases, the baseline period for establishing the criteria also needs to be defined, for example the required reduction might be relative to the mean annual per-capita drinking water usage over the period from 2005 to 2009. This then becomes the criterion against which the success of different options in achieving the objective in question is assessed. Criteria might not be set for other metrics, and it might be sufficient to simply compare the performance of different options. For example, one option might be more socially acceptable than another, although no absolute level of social acceptance has been set.

Performance criteria and indicators are also needed from a project management perspective. These could take many forms and could be fairly broad, such as meeting delivery milestones, or could be very specific in terms of accuracy of modelling results.

As well as an interest in the ability of the system to deliver, stakeholders will also be concerned about the risk of the system failing to perform. Hence, it is advisable to include risks as metrics or criteria as well. Examples of these include urban water supply reliability criteria and environmental watering frequencies for wetlands. Other examples from environmental management (Maier et al, 2008) include likelihood of failure (the complement of reliability), vulnerability (degree of failure) and resilience (inverse of the expected duration of failure). Including risk based metrics or criteria will

enable an understanding of the risks and implications of failure to be taken into account in the option selection process.

In general, performance criteria and indicators will embody the expectations of the stakeholders in terms of what information they will get from the project and the quality of this information, to at least some extent. It will be important, therefore, to ensure expectations are realistic. A common understanding of expectations of what information can be delivered by the project should be established during stakeholder consultation early in the project, preferably at the problem definition stage, and should be confirmed and agreed in writing. Ongoing communication and consultation with stakeholders will be needed, as discussed in *Stakeholder Consultation (page 13)*, to reinforce this understanding and avoid the often encountered problem that stakeholder expectations tend to rise during the life of a project.

Another issue to consider in selecting metrics is HOW they will be evaluated. If there are alternative metrics which could suit the purpose, the one with an evaluation method that is familiar to the project team - even data from previous evaluations – might be more expedient. While models should not dictate metrics, there needs to be a reasonable means of evaluation (and data available).

Decision Variables

Decision variables include anything that the stakeholders can adjust to influence the performance of the system. For this reason it is important to identify and agree decision variables during stakeholder consultation early in the project, preferably at the problem definition stage and in conjunction with establishing stakeholder's expectations and performance metrics and criteria.

Decision variables might include social and economic instruments and incentives, as well as institutional arrangements and the biophysical components of the system. Different solutions are generated by considering different states of decision variables.

Note	*"The notion of uncertainty includes both subjective and objective aspects. Becoming confident or establishing lack of confidence is an act of subjective judgement about the validity of some information. However, the judgement might be supported and informed by the evaluation of 'objective' facts and other forms of evidence."* (Refsgaard et al, 2005b.) As a general principle, transparency and good reporting are essential for satisfactory uncertainty assessment.

Uncertainty and Risk

Uncertainty

Uncertainty needs to be considered in the context that the models are being applied to support a decision making process that involves selecting a "preferred" course of action by weighing performance against competing objectives (Blackmore et al, 2009). "While in some circumstances it might be sufficient to make decisions based on fixed (often mean) values, very different choices might be made if the extent of uncertainty

on inputs to the decision process and its impact on outcomes were better understood." (Blackmore et al, 2009).

There are many sources of uncertainty relevant to decision making processes. For models themselves, and application of models, relevant sources of uncertainty include:

- the science underlying the model;

- model assumptions and simplifications of what the model is representing;

- model input data including parameters, constants and driving data sets;

- code uncertainty such as numerical approximations and undetected software bugs;

- stochastic uncertainty (this is addressed under "variability" below);

- variance in model ensemble results (where an ensemble is used) and multiple parameter realisations (where obtained); and

- other unknown sources.

Additional study and collecting more information allows error that stems from the types of uncertainty other than stochastic uncertainty – these are referred to collectively as epistemic uncertainty - to be minimised/reduced (or eliminated). In contrast, stochastic uncertainty – more commonly referred to as variability (see glossary in *Chapter 6 (page 73)*) - is a natural phenomenon and is irreducible but can be better characterised or represented with further study (CREM, 2008).

The existence of variability ensures the future cannot be predicted exactly. Predictions can be expressed in terms of probabilities of exceedance or non- exceedance of certain outcomes (events), such as floods, and can be expressed either quantitatively or qualitatively (eg see Refsgaard et al, 2005b, p18).

With no epistemic uncertainty, the probability of a defined event (and the converse, which is the magnitude of an event of a defined probability) has a single value. The effect of epistemic uncertainty is that there will be a range of estimates of the probability of exceedance or non- exceedance of a given event, and the actual probability will lie within this range with some level of confidence (or that there will be a range of estimates of the magnitude of an event of a defined probability of exceedance or non- exceedance, and the actual magnitude will lie within this range with some level of confidence). This is analogous to the classic flood frequency analysis problem in hydrology where the expected magnitude of a flood of a given probability of exceedance is calculated, such as the 100-year flood, and confidence limits describing the possible range of magnitudes are also calculated.

Expressing and quantifying uncertainty arising from application of models has a number of benefits, including that this:

- Provides input to socio-economic evaluations

- Supports prioritising data collection

- Enables articulation of limitations of modelling

Uncertainty needs to be considered at a number of places in the modelling process. Sources of uncertainty should be identified and assessed during the problem definition phase, and more detailed consideration and analysis of uncertainty should be undertaken during the option modelling phase. In the comparison of options phase (see *Identify Preferred Option (page 43)*), where the aim is to identify the preferred option, uncertainty considerations should be factored into the risk assessment and other analyses undertaken.

In the problem definition phase, it is particularly important to take uncertainty into account when deciding metrics and performance criteria, although it may only be possible to provide a qualitative assessment at this stage. Uncertainty considerations are also relevant to identification of decision variables and to system definition, where uncertainty will have an effect on decisions about aspects such as spatial and temporal scales.

In the option modelling phase, uncertainty analyses should be undertaken as an adjunct to model calibration and validation. The calibration and validation performance measures are an important component of this, as are the results of any sensitivity analyses. Uncertainty is an important input to determining whether a model is fit for purpose and therefore to getting the model accepted by stakeholders; and to gaining accreditation where this is required. It is also an important consideration when exploring solution options, as uncertainties might alter choices; and always results should be reported to no more significant figures than can be justified given the uncertainties that apply. Uncertainty analysis in the modelling phase is discussed in *Sensitivity/ Uncertainty Analysis (page 39)*.

A practical example of a comprehensive uncertainty analysis is available in a report by Van Dijk, et al (2008). The report describes the analysis of uncertainty in hydrologic river system models based on a multiple lines of evidence approach. It also describes the analysis of uncertainty in scenarios modelled, relative to the uncertainty in the models, and the amplification of change and uncertainty in the river systems modelled.

Risk

Good management decisions do not only focus on the way a system performs when all is going well. They also take into account the risk of the system failing, and its inherent resilience. Risk has two components, frequency (or likelihood) and consequence (or impact), which together inform our expectation of undesirable outcomes, and how to manage for these. Models are useful tools for supporting the evaluation of risk scenarios and testing the performance of risk management strategies (Blackmore, et al, 2009).

Risk can be defined as anything that may have a negative impact on the ability to achieve objectives. The process of Risk Assessment involves consideration of both likelihood and consequence, and is described in the Australian/ISO Standard for Risk Management, ISO31000:2009 (AS/NZS, 2009). Communication and consultation are critical to the success of the risk assessment process.

Understanding uncertainty is an essential part of risk assessment, as without uncertainty (including stochastic uncertainty) there is no risk (Blackmore, et al, 2009). Uncertainty can be expressed in terms of a probability distribution function provided sufficient information can be obtained from model results and other relevant sources to define the distribution. Otherwise likelihoods can be expressed qualitatively (eg high, medium and low) and are often derived from expert knowledge or risk workshops.

A potential consequence of uncertainty is that control measures could be developed which are partially or completely ineffectual. For example, a conceptual model of salinity processes may have been developed that is partially or completely incorrect due to limitations of knowledge about the system being studied. This could lead to partially or completely incorrect estimates of likelihood and/or consequences, and hence risk, and in turn lead to identification of partially or completely ineffective control measures.

However, modelling uncertainty considerations and model results are only part of the risk assessment process, as relevant social, economic and environmental factors need to be taken into account as well. These add to the dimensionality of the risk assessment process, but they can be accommodated satisfactorily in risk matrices or risk curves. Information on these factors should be sought from all relevant sources, including the opinions of experts and other stakeholders, and published data and knowledge, in addition to results from models and any other analyses. Bringing in additional information can mitigate potentially misleading effects of uncertainty as well.

Preliminary Assessments

This step is a valuable preliminary to deciding on a project methodology; in some cases it might even lead to a review of the project objectives and metrics. It should consist of a simplified preliminary appraisal of the likely results from the planned modelling, with sufficient rigour that the preliminary results could be used as an independent check on the results from the modelling. Where the modelling results differ from the preliminary assessment then further investigation would be required to determine if the model was behaving correctly. If a project is an extension or update of previous work, or there is a precedent of similar work elsewhere and consistency is required, this step may not always be needed.

Option Modelling

This phase is essentially about developing a modelling methodology, building and calibrating models then running scenarios covering solution options of interest. It also includes presenting results from the modelling.

In some situations there may be a package of actions that could be undertaken. These may be made up of a number of individual options and each potential combination of individual options would in itself be an option. In this situation, development of the required package from the more attractive of these options would occur in the next phase of the project. Hence, the role of this phase is to provide results for each option that can be used in evaluations in the next phase to find the "best" (optimum) individual option or package of actions.

| Note | "Decision making is a dynamic, anthropogenic process that is, at most, informed by scientific analysis. While simulation models can provide valuable assistance, the accuracy of their outputs and the way in which the outputs are presented and used can substantially alter the decision being made and the value of the outcome" (Blackmore et al, 2009). |

Methodology Development

General principles

Determination of project methodology needs to be based on consideration of the scope of the project as a whole, and its objectives and metrics. That is, the needs of the next phase, where the aim is to identify the preferred option, need to be considered as well as the needs for analysis in this phase (see *Identify Preferred Option (page 43)*). *System Definition (page 18)*, the findings of the preliminary assessment (*Preliminary Assessments (page 24)*) and conclusions drawn at the problem definition stage (*Problem Statement (page 17)* will be relevant in this regard. The output from this step should at least be a methodology statement all stakeholders can see, if not a methodology report.

In this step it will be important to have resolved any differences of views on the conceptual model of the system that various project participants may have, otherwise gaining agreement on, and acceptance of, the methodology will be difficult. This may lead to changes in conceptualisations of the problem or how the system to be modelled works (either by modellers or by stakeholders), and it could lead to revision of the scope of work as well. Hence, one or more feedbacks to, or iterations with, steps in the Problem Definition phase (*Problem Definition (page 16)*) may be needed before the methodology can be finalised. The approach adopted will depend to some extent on whether the project is an extension of previous work or it is a new project but, in either case, stakeholder input and agreement, and peer review as discussed in *Peer Review (page 12)*, should be obtained.

The behaviour and performance of the system can be analysed in many different ways, ranging from complex, geographically explicit computer models, through simple lumped models to surveys, consultation and expert opinion. It is the appropriateness of the approach to provide a relevant level of understanding to address the original questions posed that matters, not necessarily its ability to mimic reality. It is easy to waste time and resources modelling to a higher resolution or level of accuracy than is necessary.

For example, sometimes decision makers do not want a numerical result, they might just want qualitative information such as on a scale from low to high, in which case a computer-based model may not be needed (although, taking the term "model" in a broader sense, some other sort of model is likely to be needed in any case).

Uncertainty in all its forms, methodology for investigating it (see *Sensitivity/Uncertainty Analysis (page 39)*), and the likely implications of uncertainty for interpreting analytical results, should also be considered during this step. Where it is apparent that uncertainty levels could be unacceptably high and suitable alternative models are available, particularly in the context of catchment rainfall-runoff modelling, use of a multiple model ensemble to explore uncertainty could be considered. While this approach can have benefits it also has costs in that it increases the workload as every model in the ensemble has to be calibrated, validated and applied to each scenario to be analysed. Current indications are that for catchment rainfall-runoff modelling, in Australia at least, ensemble modelling provides improvements that are useful in a small scale or research context but the benefits are not sufficient to warrant the extra effort involved for large scale applications (eg basin to national scales) due to capacity constraints, especially where the models must be run frequently (Vaze et al, 2011). Uncertainty is discussed further in *Uncertainty and Risk (page 21)*.

Where project participants have access to an existing model that is suitable for the problem at hand, or they are familiar with the use of a particular model, or models, for addressing similar problems, then it is generally advisable to use these, adapting the models and methodology as necessary. However, it needs to be ensured that this course of action is appropriate for the problem at hand, with support from stakeholder consultation and peer review.

The subject of model choice, with respect to numerical, computer-based models, is discussed in *Chapter 3 (page 49)*. The critical point in this regard is that whatever the choice, there must be sufficient data available to support model calibration, validation and application to option modelling.

Assuming the adopted methodology entails use of one or more numerical, computer-based models, there are other aspects that need to be considered. These are discussed below.

Model calibration considerations

Model calibration measures and statistics should be decided at this stage and they should be relevant to project (and model) objectives. More than one measure/statistic should be used and placing too much reliance on one should be avoided as this may distort or bias results. Measures and statistics should be chosen that best reflect the intended use of the model; this will be particularly important for defining the objective function if optimisation is to be used in calibration. For example, if it is a flood hydrograph model then high flows are the main interest; if it is a yield model,

flow patterns are the main interest, with more emphasis on low flows; in both cases the mass balance is a key consideration.

An acceptable level of calibration, expressed in terms of values of these adopted measures/statistics, should also be decided at this stage and this will guide the calibration process; this will also be guided by considerations of acceptable levels of modelling uncertainty (discussed further in *Sensitivity/Uncertainty Analysis (page 39)*). Leaving these decisions to the calibration step is too late as it will potentially lead to the calibration process dictating what the model can be used for, which could be inconsistent with the project objectives. If an acceptable calibration cannot be achieved then it may be that the chosen model is not useful, and it may be necessary to make another choice or make some other adjustment to the methodology.

An important, and difficult, question is how to translate soft objectives and stakeholder wishes (eg a desire for a "healthy river") into model performance metrics and criteria. Workshops, meetings and other discussions may be needed to attempt to resolve this question. Even then it may not always be possible to achieve a resolution initially; subsequent *Sensitivity/Uncertainty Analysis (page 39)* and scenario analysis (*Find and Test/Explore Options (page 40)*) may assist in reaching a solution. The temptation to avoid considering this aspect and use generic criteria, say from literature, in a "one size fits all" approach should be avoided as far as possible. Apart from any other problems this may cause for gaining acceptability of model results, this would ignore the fact that all model applications are unique with respect to data availability, hydrological regime and modelling purpose and hence the criteria should vary from one application to another.

Equally importantly, the period of record of historical data to be used for model calibration should be chosen at this point. Ideally, the period of record chosen should be representative of the range of variability of input data that could occur; also, preferably, some of the data should be reserved for use in validation testing of the calibration. However, in reality there is not always enough data to match the ideal in which case compromise will be necessary, and care will need to be exercised in interpreting model results. There could be several reasons for this insufficiency, such as the total period of record being short and changes in physical or management characteristics being great enough to render early data inappropriate for use in model calibration.

A related issue is that it is also important to ensure that data sets used for calibration and validation are consistent with data sets available for option modelling, especially when hindcasting, otherwise bias will be introduced into results. For example, in rainfall-runoff modelling there may be data from several rainfall stations with a good spread of elevations available for the calibration period, but data from only one station at low elevation available for long term hindcasting. In cases such as this only the data from the long term station should be used for calibration, although the long term data could be conditioned using the recent data to derive improved estimates of

representative catchment rainfall patterns over the long term (eg conditioning could involve adjusting the average of the long term data or some other regression based technique). The resultant calibrated parameter set should then be compatible with the data available for long term hindcasting.

The key is to ensure that the calibration provides as sound a foundation for extrapolation, when modelling solution options, as possible. Otherwise the range of validity, where the model can be trusted, will be limited.

Option modelling considerations

Initial agreement on likely solution options and scenarios to be modelled should be obtained, firming up the preliminary understanding from the problem statement stage, recognising these may get changed later in the light of results obtained (discussed further in *Find and Test/Explore Options (page 40)*). This will minimise the chances of building models that cannot analyse the desired scenarios, or worse, having to force scenarios to fit the modelling tool.

Where relevant, the baseline period for establishing performance criteria needs to be defined. For example, the requirement could be to achieve a 20% reduction in observed mean annual per capita drinking water usage for the period 2005-2009, by 2020.

Requirements such as this, and whether solution options are going to be analysed based on hindcasting or forecasting, will have a bearing on the modelling period adopted. For example, when hindcasting it may be necessary to use over 100 years of input data to provide a sufficiently representative sample of conditions that could occur and influence performance of solution options. When forecasting it may be sufficient to use shorter periods (as little as ten years may be enough for the above example performance criterion) and also use multiple replicates to obtain multiple realisations of the forecast outcome from which statistics such as exceedance probabilities of levels of performance can be derived for input into risk analyses; alternatively a long single data sequence (say, 1,000 years or more) could be used depending on requirements. When using a long single sequence like this, it needs to be kept in mind that making the data set longer may not necessarily provide much, or any, additional information from the point of view of evaluation of options against performance criteria and statistics (eg estimates of the mean, standard deviation, exceedance probabilities of events, and reliabilities of supply).

Note	The main point is that models are being used to examine alternatives or options so that a problem can be understood and solved which therefore requires a clear specification of the problem and that the models used are fit for purpose.

Overview

Data clean up can be by far the most time consuming step in a project. The importance of this step cannot be overemphasised.

Data should be obtained from all relevant data bases and other sources, including prior knowledge (as discussed in *System Definition (page 18)*) and original records where necessary. Data needs to be critically reviewed, irrespective of its source, and suspect data items need to be adjusted or removed. Statistical tests are available to check for outliers and other idiosyncrasies such as non-stationarity (discussed further in *Stationarity and change detection (page 30)*), and these should be utilised where needed.

The overriding issue is that of data uncertainty, and all the following discussion in this section relates to it; data uncertainty is inextricably linked with data quality. It must be considered in any analysis of uncertainty or sensitivity analyses, as discussed in *Sensitivity/Uncertainty Analysis (page 39)*, as it potentially has the largest impact on model results of all the sources of uncertainty.

Infilling data gaps and data generation

If there are data gaps, especially in time series data, then these should be filled using recognised techniques and expertise. Data from a variety of other sources could be used where appropriate, such as via Multiple Lines of Evidence, data assimilation and data fusion approaches. Other data sources include: regionalised data, which could be transposed from outside the project area or from another part of the project area using an appropriate technique; results from other models (eg infill or replace streamflow data with rainfall-runoff model results); remote sensing and other spatial data; published literature and local knowledge. They also include palæo-hydrological data which can be valuable for extending or generating time series data as it can provide insights into long term behavioural cycles and other characteristics that cannot be obtained from other sources, thereby providing a more representative data set and statistics than can be obtained from instrumental records alone.

When filling gaps in time series data, it should be ensured that the statistical characteristics of the data after infilling are not changed inappropriately. This includes statistics such as the mean, standard deviation and skewness, and also event frequencies and high and low spell lengths (eg wet and dry days); for some data types, such as streamflow, serial correlation should also be checked. A number of techniques are available to test for this. For example, starting with a complete data set, by arbitrarily removing progressively greater parts of it and recalculating statistics it is possible to test to see when the new statistics are significantly different from those for the complete data set. Regression analysis should be used to compare the data set to be filled with the source data set, or sets, for filling; if correlation is sufficiently poor then other sources of data for infilling should be sought or the mean of the data to be

filled should be used instead. If a data set under consideration to be filled has a high proportion missing, such that its statistics are not reliable, then it may be preferable to discard it and use another data set; an indication of what proportion is "high" can also be found from testing with a suitable complete data set and removing progressively greater parts of it.

Generating time series data based on the statistical and other behavioural characteristics of historical data is essential where modelling is to be applied in forecasting mode, and there are many techniques available for doing this. Techniques such as analysing the statistics of observed data and data inferred from other sources, such as palæo-hydrological data, for the same period can help show whether mixing data from various sources for the purpose of data generation is appropriate or not. In circumstances where it is desired to start a model run at a date in the past and then model through the present and into the future, it may be appropriate to add generated data onto the end of observed or otherwise inferred data to make mixed data sets for input to models. Otherwise combining generated and other time series data is not recommended.

Infilled, extended and transposed data needs to be re-reviewed to ensure it is reasonable, bearing in mind that there could be a trade-off between completeness of data set and data quality involved. Data that is inferred from other sources (eg remote sensing or other models) and observed data will usually be of different quality. Mixing data from different sources into one data set (eg mixed modelled and "observed" streamflow data) is not preferred if data being used in model calibration or validation, although there will sometimes be little choice but to use mixed data. However, mixed data sets may be suitable where they being used as input for analysing "historical" scenarios; perhaps with results based on observations weighted differently to results based on modelled data.

If key data is missing and no suitable substitute is available then it may be necessary to modify the project methodology. In particular, it may influence the model time step adopted or the length of the period of historical data that can be used as input to model runs for calibration, validation or hindcasting, or even the choice of model.

Stationarity and change detection

It also needs to be remembered that data sets may not be stationary for a variety of reasons, even where the location or area where data is collected does not change (eg changes in measurement techniques or instrumentation; land use and management - especially vegetation; stream channel geometry; long term trending in groundwater; long term climate variability and/or climate change). Non-stationarity may limit the length of the period of historical data that can be used as input to model runs, particularly for model calibration and validation where results may be sensitive to this. However, if the purpose of the modelling exercise is to model effects of changes such

as changes in vegetation or trends in groundwater levels then it is valuable to have access to data that is non-stationary due to it reflecting the changes of interest.

As is pointed out by WMO (2009; Chapter 6): "Detection of changes in long time series of hydrological data is an issue of considerable scientific and practical importance. It is fundamental for planning of future water resources and flood protection. If changes are occurring within hydrological systems, existing procedures for designing structures such as reservoirs, dams and dykes will have to be revised; otherwise systems will be over or under-designed and will either not serve their purpose adequately or will be more costly than necessary. … Change in time series can occur in numerous ways: gradually (a trend), abruptly (a step change) or in a more complex form." The case for taking non-stationarity into account in water management is also made by Milly et al (2008). Climatic drivers of non-stationarity and implications for the assessment of flood risk in Australia are discussed by Westra et al (2010); the discussion is highly relevant to water management as a whole.

A number of parametric and non-parametric statistical tests for change detection are available. General guidance on the methodology of change detection in time series of hydrological records is given in sources such as WMO (2009; Chapters 5 and 6) and Kundzewicz and Robson (2004). Yue and Pilon (2004) offer guidance on the selection of a test for non-normally distributed data by comparison of test power. Other facets of detectability of trends are dealt with by Radziejewski and Kundzewicz (2004), who examine how strong a change (gradual trend or abrupt jump) must be and how long it must persist in order to be detected by different tests.

Setting up and Building a Model

This step entails defining the layout of the model for a given application in detail: its boundaries and its internal layout in a spatial sense (such as the layout of links and nodes for a link-node model; catchment boundaries and internal subdivision (if any) for a rainfall-runoff model), determination of parameter values that are fixed and initial estimates of parameter values for calibration. The layout developed should not only be suited to the intended purpose for the current project, including being able to accommodate scenarios planned to be run, but should also take into account possible future applications; this should minimise need to redo the work and more importantly, it will avoid issues of inconsistencies between results from different versions of the same model. For example, in setting up a model of a river system for modelling water quantities, it could be borne in mind that modelling of water quality may be needed in the future which may necessitate more detail and use of more rigorous modelling techniques and/or a shorter modelling time step than needed for modelling water quantity alone.

All relevant constraints that apply should also be taken into account (in part, this relates to the model parsimony issue discussed in *Model Parsimony (page 50)*). Data limitations are a common constraint but there are others, and some of these may have conflicting effects. For example, in some cases model detail may need to

be greater than can be justified technically, such as in cases where there are socio-political sensitivities or issues of model credibility with stakeholders that can only be resolved in this manner; however, even in cases such as this the additional detail must be supported by suitable data. For example, in river system modelling, where a system with existing water resource and irrigation developments are to be modelled, irrigation enterprises are usually grouped together for representation in the model. This is done for a variety of reasons, including model and data limitations, and privacy issues. However, on occasions, gaining stakeholder acceptance of the model and the results it produces may require representing each irrigation enterprise individually. In this situation the onus may be on the stakeholders to provide the data needed; obviously, it will also be important to ensure the modelling software has the necessary functionality.

Calibrate Model

Overview

Model calibration entails adjusting model parameter values until satisfactory agreement is obtained between model results and observed data, as expressed in terms of the calibration measures and statistics chosen when developing the project methodology. Adjustments can be made manually, via an automated optimisation technique, or by a combination of these.

Irrespective of the approach adopted, a more robust and reliable result will be obtained if the modeller has, or develops, a sound understanding (ie a good feel) for the model and the behavioural characteristics of the system being modelled. This includes thinking critically about all results obtained and reviewing these (especially where automated optimisation is used), checking that results are being obtained for the right reasons, and ensuring parameter values are realistic. The modeller should have a rough idea of the expected results and from this should do order of magnitude checking and/or sanity checking of results – the preliminary assessment will assist with this.

Sometimes the model may produce some results that appear to be counter-intuitive. It will be particularly important to check these results and understand the interactions that may have brought them about, and then determine whether the results are valid or not. At the very least it will be necessary to fully explain such results to stakeholders.

General issues

While the aim is obviously to produce a calibration that is sufficient for the purpose of the model, it is possible to over calibrate, and this should be avoided as it could severely impair the usefulness of the model for analysing solution options. In general, over-fitting is more likely to be an issue when there is limited calibration data available than when representative data sets are available (or when the model has more parameters requiring calibration – degrees of freedom – than the available data can support; this relates to the issue of model parsimony discussed in *Model Parsimony (page 50)*). Having a good feel for the model and the system being modelled, critical

review of results, and model validation testing are all useful for avoiding or overcoming this problem.

Where calibration data sets are used that are non-stationary, for whatever reason (see *Gather and Clean–up Data (page 29)*), care needs to be exercised in interpreting results. For example, there may be apparent trends in model performance (eg ability to reproduce low flows) that are an artefact of systematic changes in the data and it would be inappropriate to try and redress these. In this situation, greatest weight should be given to the most reliable data and to the data which represents system conditions most relevant to the needs of the project. On the other hand, some apparent trends in model performance may be due to artefacts of the model and its parameter values, such as drawing down a modelled storage during calibration runs (a situation that may not be valid or applicable when analysing solution options), which could distort the calibration statistics.

Modellers should also beware of the non-uniqueness (or "equifinality") problem – see *Model Parsimony (page 50)* for more discussion on this topic. Usually there are more parameters (unknowns) than there are known data sets to calibrate to (in algebraic terms, there are usually fewer equations than there are unknowns), therefore more than one combination of parameter values is likely to give good results for the calibration measures and statistics. However, not all combinations will necessarily be realistic, or valid, and it is important to check on this, particularly to avoid "getting the right result for the wrong reason": in which case any scenario information derived by the model is likely to be in error. The validation step will assist with this, but this does not negate the need for, or the importance of, critical review by the modeller. This is especially true when optimisation is used to assist with model calibration.

As indicated in *Methodology Development (page 25)*, calibration is more than just minimising a goodness-of-fit statistic and other metrics and checks should be used to ensure the model provides reliable results. An additional metric could be as simple as also minimising differences between observed and modelled mass balances. As discussed in relation to optimisation below, if there is additional data available to constrain the model then advantage should be taken of this as well. It might also influence the choice of model (see *Chapter 3 (page 49)*). Using this additional data could be expected to reduce problems of equifinality and lead to a more robust calibration.

Role of optimisation

There are varying views as to which stage of model calibration optimisation should be applied. These range from automatic calibration – letting the optimisation procedure do it all – through to optimise first and then manually fine tune, calibrate manually first and then use optimisation to fine tune, and not use optimisation at all.

The most important step in the use of optimisation is defining the objective function, and this is based on the calibration measures and statistics. Hence, deciding the calibration measures and statistics is particularly important when optimisation is being used.

Of similar importance is defining the bounds of the allowable ranges of parameter values. Model responses are more sensitive to the values of some parameters than others and it is particularly important that the ranges of the more sensitive parameters are bounded realistically; for the others it may not matter very much. Here, there could be considerable benefits from the use of data assimilation and data fusion techniques. In catchment rainfall-runoff modelling, these have provided the ability to close the water balance at catchment scale by use of spatially integrated data sets. For example, the use of independent estimates of actual areal evapo-transpiration when calibrating one of these models provides the ability to constrain modelled processes by use of multiple objective functions; this has the benefit of assisting the identification of suitable parameter sets and reducing the influence of uncertainty in observed time series data (R. Nathan, pers. comm., 2010).

Automated optimisation can be a valuable tool for assisting and expediting model calibration but it should not be used blindly. It should be avoided unless the modeller has a very good feel for the model and the behavioural characteristics of the system being modelled; placing realistic bounds on parameter value ranges is essential if this approach is to be employed. Without a good feel for the model or the system being modelled, manual adjustment of parameter values towards achieving a calibration should be undertaken first. This is also true if the approach of optimising first and fine tuning manually is being considered; this approach is also not advisable unless the modeller has a very good feel for the model and the behavioural characteristics of the system being modelled, as the same problems of defining parameter value ranges apply.

The logistics of manual calibration versus optimisation may influence the choice of approach and these logistics will be affected by the number, sensitivity and interdependence of parameters. The end result may be multiple sets of parameter values, due to the "equifinality" problem discussed above.

Single and multi-stage approaches

For single disciplinary models, particularly where only one modelled variable is of interest (eg streamflow), then single stage calibration will be appropriate. Single stage calibration may also be appropriate in integrated modelling where more than one modelled variable is of interest; such as single stage calibration of an integrated surface and groundwater model for streamflow and groundwater level, by optimisation. A pre-requisite for this is that the coupling between the domains must be fully integrated and not sequential. A challenge in this situation is how to weight the multiple performance criteria in the objective function used in optimisation.

However, for integrated models such as river system models or where more than one modelled variable is of interest, and some variables are dependent on others (eg streamflow and salinity), a staged approach to calibration may be necessary which would be based on considerations of independence (eg calibration for streamflow would precede calibration for salinity).

This may entail forcing some dependent variables, such as by reading them in rather than allowing them to be modelled, holding their value constant or ignoring them, while calibration is undertaken for other variables, or temporarily breaking the model into segments and calibrating the variable of interest for each segment separately. Once calibration for a given variable is completed then the model could be allowed to run with these while other variables are calibrated. Examples include calibrating the routing procedure separately for each reach of a multiple reach river model, and reading a sequence of historical water demands in as data while calibrating water supply storage behaviour.

A potential danger with this approach is that errors will compound and parameter values obtained from calibrations undertaken in later stages in the process could be very distorted. This highlights the importance of critical review at every stage to ensure that results are being achieved for the right reasons and that parameter values make sense. If necessary, earlier stages in the calibration process may have to be revisited and some trade-offs in quality made in order to achieve a satisfactory outcome overall. This is another reason for avoiding over-calibration: a point which was discussed earlier (mentioned under "General issues" above).

Review

While the need for the modeller to critically review model results after every model run has already been discussed, more formal peer review should also be undertaken and it should be undertaken more often than just at the end of the model calibration process. Needs for peer review are discussed more fully in .

Note	The key is not only to "think before you compute" but also to think while you compute (… and think after you compute).

Validate Model   The term "validation", as applied to models, typically means confirmation to some degree that the calibration of the model is acceptable for the intended purpose. Sometimes this is termed "verification" but this is erroneous and the term should never be used as it is not logically possible. Terms such as "model evaluation", of which part is testing against data, expectations, etc, and "corroboration" may be better alternatives – what is being tested in this step is usefulness, not truth (A. Western pers. comm., 2010; CREM, 2008, p41; Oreskes et al, 1994; Silberstein, 2006). Refsgaard and Henriksen (2004) define model validation as: "Substantiation that a model within its

domain of applicability possesses a satisfactory range of accuracy consistent with the intended application of the model".

Refsgaard (2001) proposes four types of validation tests; these are proposed in the context of catchment rainfall-runoff modelling but in principle they are relevant to other hydrological modelling domains as well. The tests are relevant to different situations with data availability for model calibration and validation, and whether impacts of some changes in catchment conditions have to be modelled or not. Caveats about data stationarity and consistency, discussed in relation to model calibration (*Calibrate Model (page 32)*) also apply here. The four types of validation tests are:

- *Split sample test:* this is the most commonly used test and it is applicable where there is sufficient data for model calibration and validation, and where catchment conditions to be modelled are unchanging. The available record is split into two parts; the model is calibrated using one part and validated using the other, and both calibration and validation should give acceptable results.

- *Proxy basin test:* this test is applicable when there is insufficient data for calibration and validation on the catchment of interest. If, for example, streamflow has to be predicted for an ungauged catchment and there are gauged catchments within the region then two gauged catchments should be selected. The model should be calibrated on one catchment and validated on the other, and vice versa. Only if the two validation results are acceptable and similar can even a basic level of credibility be attached to the ability of the model to simulate streamflow from the ungauged catchment adequately. This test should be supported by other checks (some of which are mentioned at the end of this list) which should help improve credibility, including ensuring that the transposition of parameter values does not occur outside the region of applicability (such as by checking the catchments are at least generally similar hydrologically and their rainfall regimes are comparable); if it is possible to derive an approximate estimate of average depth of runoff for the ungauged catchment based on data from surrounding catchments and downstream gauging stations, then this is also a useful check. Several approaches are available for choosing catchments for use in regionalisation of parameter values, of which the nearest neighbour technique is the most common; if necessary independent expert advice should be sought on the most appropriate.

- *Differential split sample test*: this test should be applied when a model is to be used to simulate streamflows or other variables (such as soil moisture patterns) in a given gauged catchment under conditions different to those corresponding to the available data. The testing is structured on the basis that the model should have a demonstrable ability to perform through the transition from current to proposed future conditions. The test may have several variants depending on the specific nature of the modelling study. If, for example, the requirement is to model the effects of a change in climate regime, two periods with different values of the climate variables of interest should be identified in the historical record,

such as one with a high average rainfall and another with a low average rainfall. If the requirement is to model streamflow in a wet climate scenario then the model should be calibrated on a dry period and validated on a wet period (If in this case the ability to model the transition is not needed then the model could be calibrated and validated on wet periods). Other test variants can be defined for the prediction of changes in land use, effects of groundwater abstractions and other such changes.

- *Proxy basin differential split sample test:* this is the most difficult test for a hydrological model as it is applied to cases where there is no data available for calibration and where the purpose is modelling effects of conditions that are subject to change. The test is a combination of the two previous tests.

Refsgaard (2001) also points out that if the accuracy of the model for the validation period is significantly worse than for the calibration period, this is an indication of over-fitting; there could be a problem with model structure causing the parameters to be specific to the conditions used for calibration (or it could be simply that parameter values are over calibrated). The ratio of accuracy during the calibration period to accuracy during the validation period is sometimes used as a measure of the degree of over-fitting (Refsgaard, 2001). Keeping the number of parameters as low as possible will help minimise the potential for problems such as this when calibrating and validating models.

For integrated water management models, the type of validation testing that will be appropriate is likely to vary from one component to another, depending on whether that component is modelling an aspect of the system which is subject to change or not, and availability of data for calibration.

Irrespective of the validation tests used, other approaches such as uncertainty analysis, sensitivity testing (discussed in *Sensitivity/Uncertainty Analysis (page 39)*), checks based on independent data (such as checks on the average water balance based on remote sensing and other spatial data), cross-validation with other models and techniques such as "leave-one-out" calibration and validation using multiple gauged catchments should be applied as well. It may also be appropriate to seek understanding from domains beyond modelling to provide independent checks on model calibration, such as remote sensing or application of tracer methodologies.

Irrespective of which approaches to model validation are adopted, peer review is an important part of the process, as discussed in *Peer Review (page 12)*.

Part of best practice is to explain how and to what level of detail model evaluations that have been undertaken can be interpreted. This involves careful assessments of: the level of agreement between model and measurement to the degree possible; the level of uncertainty in the processed measurements (eg if comparing against water quality constituent loads there will be significant uncertainty in the observation);

and how comparable the model predictions and observations are; as well as acknowledgement that accurate aggregated performance does not guarantee accurate disaggregated performance.

In other words, it should be possible to explain how data is informing the modelling and to be explicit about what particular model-data comparisons actually show. For example, the fact that a distributed, catchment rainfall-runoff and water quality constituent export model gets the catchment scale runoff about right does not mean the runoff from each land use is right and similarly for water quality constituent loads etc. Such an explanation would provide insight on the relationship between model and data to the stakeholders, and insight on what can and cannot be concluded on the basis of model-data comparison; it also forces the modeller to critically examine the relationships between the model and the data. This point should be initially considered when the project methodology is being developed, as indicated in *Methodology Development (page 25)*, but explanations and conclusions will not be able to be finalised until this step and uncertainty or sensitivity analysis (discussed in *Sensitivity/ Uncertainty Analysis (page 39)*) are undertaken.

As above, the issue here is that the model predictions are correct for the right reasons, at the appropriate scale. This clearly matters where the decision making depends on those reasons being right. In the example above, catchment scale data is unlikely to provide useful information on land use change effects at a local scale given confounding influences of riparian, instream and other factors, although useful information may possibly be obtained for the catchment outlet. Such evaluations need to consider both hard and soft data sources.

A related issue is that of models being used beyond their originally intended purpose (field of competence, or domain of applicability) and this needs to be managed. This applies particularly to situations where an existing model is to be applied in a new project but it is also important when a new model is planned to be applied. For this reason, the possibility of the model being misused should be anticipated and uses for which the model is suited and those for which it is not suited should be clearly stated when reporting on model calibration and validation. The reporting could be based on a categorisation system, such as that used in the Danish guidelines (Refsgaard et al, 2010) which includes three categories (Refsgaard, pers. comm., 2010): (a) fields of documented applicability; (b) fields of potential applicability; and (c) fields where model is not likely to be applicable; validation tests are a pre-requisite for establishing fields of documented applicability. In some cases the bounds of what a model can and cannot do may be dependent on the skills of the individual modeller, and this needs to be borne in mind when evaluating fields of applicability.

More detailed guidance on model validation is available from other sources; such as CREM (2008, Appendix C, where this is called "corroboration").

Sensitivity/ Uncertainty Analysis

Uncertainty and sensitivity analysis are closely related; however, while uncertainty is parameter specific, sensitivity is algorithm-specific with respect to model "variables." By investigating the "relative sensitivity" of model parameters, a user can become knowledgeable of the relative importance of various parameters in the model. By knowing the "uncertainty" associated with parameter values and the "sensitivity" of the model to specific parameters, a user will be more informed regarding the confidence that can be placed in model results. (CREM, 2008.) Background information on uncertainty is available in *Uncertainty and Risk (page 21)*.

Uncertainty analysis

There has been increasing acceptance of the need to include uncertainty analysis in hydrologic and hydraulic modelling applications, although the approach is not new (eg Stedinger and Taylor, 1982; Stedinger et al, 1985). Many methodologies and tools suitable for supporting uncertainty assessment have been developed and some useful reviews of these are available (eg Matott et al, 2009; Pappenberger and Beven, 2006; Refsgaard et al, 2005b; Refsgaard et al, 2006; and Refsgaard et al, 2007). These note that no single methodology is suitable for addressing all the different aspects of uncertainty assessment. A number of relevant methods are discussed in downloadable files available on the website of the Uncertainty Analysis in Environmental Modelling Workshop, 2004 (*www.es.lancs.ac.uk/hfdg/uncertainty_workshop/uncert_methods. htm*), and by Matott et al (2009), Refsgaard et al (2005b, Chapter 4), and Refsgaard et al (2007). Core information from Matott et al (2009), with links to more information and a number of tools, is available at *www.epa.gov/athens/research/modeling/ modelevaluation/index.html*. Potentially relevant approaches include ensemble modelling (discussed further in *Methodology Development (page 25)*). Van Dijk et al (2008) present results of practical application of a number of uncertainty analysis techniques in the context of river system modelling.

The selection of an adequate methodology depends on (Refsgaard et al, 2005b):

• Where in the modelling process the analysis is to be carried out.

• The type, nature and source of uncertainty.

• The priority that addressing each of the identified sources of uncertainty has, according to their importance for the decision-making process ("policy relevance").

• The available resources and level of ambition with respect to completeness of the analysis.

Unfortunately, choice of methodology is not necessarily straightforward, guidance on selection of methods and applications is limited (Pappenberger and Beven, 2006) and there is a "lack of a coherent terminology and a systematic approach …" Montanari (2007). However, some guidance regarding the selection of appropriate methodologies/tools for different purposes, as well as more detailed guidance on uncertainty analysis for model applications overall, is provided in Matott et al

(2009), Refsgaard et al (2005b) and in Refsgaard et al (2007). In addition, Brugnach et al (2008) provide guidance on strategies for addressing uncertainty when the model purpose is, respectively: exploratory analysis, communication and learning. Their discussion includes a number of examples from the broad environmental modelling domain.

It is also noteworthy that the USEPA (CREM, 2008) recommend sensitivity analysis as the principal evaluation tool for characterising the sources of uncertainty, ranging from the most important to the least important, in environmental models. Sensitivity analysis is discussed further below.

Sensitivity analysis

Sensitivity analysis is a popular technique for use in evaluating uncertainty in models. It is the study of how the response of a model is affected by changes in a model's input data or parameter values. Sensitivity analysis will enable uncertainties in model output to be systematically apportioned to different sources of uncertainty in the model inputs (particularly parameter values), and modellers and stakeholders will gain insight into the relative importance of these in the model. It is usually undertaken as an adjunct to model calibration and validation.

Most commonly, sensitivity analysis entails varying the value of one or more model parameters in a systematic way and then reporting and evaluating the changes in key model outputs. However, it can also entail changing input data sets, such as making systematic adjustments to input rainfall data, and perhaps varying certain constants as well where these are not well defined. As an example, in catchment rainfall-runoff modelling it is not unusual for the greatest sensitivity to be shown to uncertainties in the input rainfall data, and for sensitivities to certain parameter values to be relatively less, although perhaps still significant as well. More detailed guidance on sensitivity analysis for model applications is available from other sources; such as CREM (2008,Appendix C).

Find and Test/ Explore Options

Options, which could be a single action or a package of actions, for addressing the problem at issue need to be identified and agreed between stakeholders and project officers, and expressed in terms of scenarios to be modelled. This may entail one or more workshops involving all interested parties in an iterative process involving identifying new options or modifying options previously identified, analysis and reporting back on results. Several iterations may be necessary to ensure each option is thoroughly explored to the satisfaction of all stakeholders. In some cases a new option may be identified that requires a change in model set up, in which case it may be necessary to revisit the model calibration and validation steps as well.

Initial identification of the options likely to be of interest, and how they are to be analysed, should be undertaken at the problem definition stage (see _Problem Statement (page 17)_) with firming up occurring when the project methodology is

being developed (*Methodology Development (page 25)*). By the time the model is calibrated and validated, and ready to analyse scenarios, it is necessary to be able to define scenarios in sufficient detail for them to be modelled.

Irrespective of whether modelling to predict scenario performance is based on hindcasting (using historical data) or forecasting (using generated or otherwise derived data representing the future), model results should only be interpreted in a statistical sense. In particular, when hindcasting, too much weight should not be placed on comparing modelled behaviour between scenarios or with observed data on given historical dates. On the other hand, close inspection and comparison of individual results between scenarios or with observed data is valuable for finding apparent anomalies and explaining or rectifying these. In general, more and better statistical information can be extracted from model results when forecasting is used than when hindcasting is used, particularly about extreme events, especially when a number of replicates of input data sets are used. The issue of hindcasting versus forecasting is further discussed in *Factors Influencing Model Selection (page 53)*.

Great care is needed when interpreting results for scenarios when the period of record available for model calibration is not representative of the range of hydrological conditions that could occur (usually because the record is short). In this situation, scenario modelling is likely to involve taking the model outside the range of conditions for which it was calibrated. While undesirable, this is sometimes unavoidable but it needs to be remembered results may not always be reliable and uncertainty considerations will be paramount. This has been highlighted by the difficulties encountered in simulating the exceptionally low streamflows that occurred during the ten years of drought in south eastern Australia, ending in 2010, using models calibrated with data from periods which were not as dry. Similar difficulties arise when attempting to model flood conditions which are outside the range covered by the data.

If risk-based measures are being used in the decision process, then scenarios should include those needed for the risk assessment process, including extreme events. Where Monte Carlo or other methods are being used that involve sampling from a specified distribution, scenario selection may not be needed.

Results for each scenario should be reviewed by the modeller, if no-one else, to ensure they make sense. If results are counter-intuitive or have some unexpected features, but are found to be valid, these can often be the most informative and the reasons for these need to be explained; this requires that the modeller has a good feel for expected model responses and the characteristics of the system being modelled. If the conclusion is that such results are not valid then it may be necessary to revisit certain steps in the modelling procedure, potentially including the model set up, calibration and validation, but most likely the parameterisation for the relevant scenario. Where results for certain individual scenarios are seen to be critical to achieving a solution, formal

peer review may be needed, as discussed in *Peer Review (page 12)*, to improve confidence in them by all interested parties.

Report/ Communicate Scenario Analyses

Results of scenario analyses should be reported to stakeholders using approaches that facilitate their understanding and acceptance, as discussed in *Information Communication (page 14)*. The reporting could provide some comments on the performance of each scenario, including some preliminary comparisons between scenarios, from a technical point of view but detailed comparisons should wait until the next phase of the project (*Identify Preferred Option (page 43)*). However, scenario analyses may involve some trial model runs by the modeller to adjust or fine tune a given scenario before a useful result is obtained and, if only for reasons of clarity, it may not be necessary or desirable to report the results of all the trial runs.

Suitable approaches for reporting results may include workshops and bringing in external expert opinion to facilitate performance assessment of options. Results may be further processed, as would be needed where the output from scenario analysis informs a risk assessment process.

When presenting and communicating interim results it may be adequate to do this in an informal manner, such as via presentations, or in a form that facilitates the "next step". For final results formal documentation should be prepared, commensurate with need and potential audiences. Care needs to be taken in reporting to ensure results are commensurate with the capabilities of the model. For example, when hindcasting using historical hydrological data as input, a model will produce results for given historical dates but model limitations may mean that reporting results in a statistical sense is all that is appropriate. In addition, care needs to be taken to avoid reporting results in a manner that inappropriately influences decision making or goes beyond the terms of reference (agreed scope) of the project.

Model Acceptance/ Accreditation

The aim of model acceptance is to gain agreement that the model is fit for purpose. Model accreditation is seen as a regulatory issue, with its own process needs, and as the province of governmental agencies. The quality assurance procedure in this guidance should support the accreditation process but does not purport to be a model accreditation procedure in its own right.

The fitness for purpose of a model will often be subject to caveats due to constraints, such as limitations on data availability, which could in turn affect the accuracy of results. These caveats, and implications for reliability of results, should be clearly stated in all reporting (for a practical example, see Podger et al, 2010b).

Model acceptance entails peer review which could be external or "in-house", or both, as discussed in *Peer Review (page 12)*. Where stakeholder acceptance is being sought, it involves review by stakeholders as well. This latter could be achieved via the project governance process where a project steering committee with stakeholder representation has been set up. In the normal course of events, stakeholder review

would then consist of considering the findings of the peer review process including recommendations from the technical reference panel, where one exists. In some circumstances, stakeholders may prefer to have an independent peer review conducted, either by a panel of experts of their own choosing or by a panel agreed with the project principals.

In the event that acceptance is not achieved then, depending on the seriousness of deficiencies identified, it may be necessary to revisit the model calibration, model set up and the project methodology, and even go back and revisit the whole problem definition step.

Identify Preferred Option

Increasingly, water management models are being used in decision making contexts that involve selecting a "best" course of action (ie the preferred option) by weighing performance against competing objectives. Typically these include socio-political, economic and a variety of environmental considerations.

For the process of comparing options and selecting the preferred option to be robust and reliable, and to minimise the chances of making a poor selection, model results and other sources of knowledge need to be placed into a decision making context. The process should explicitly take into account the issue of uncertainty (eg Maier et al, 2008; Myšiak et al, 2008). An example process is illustrated in *Figure 4* (feedback loops are left out to avoid overly cluttering the figure).

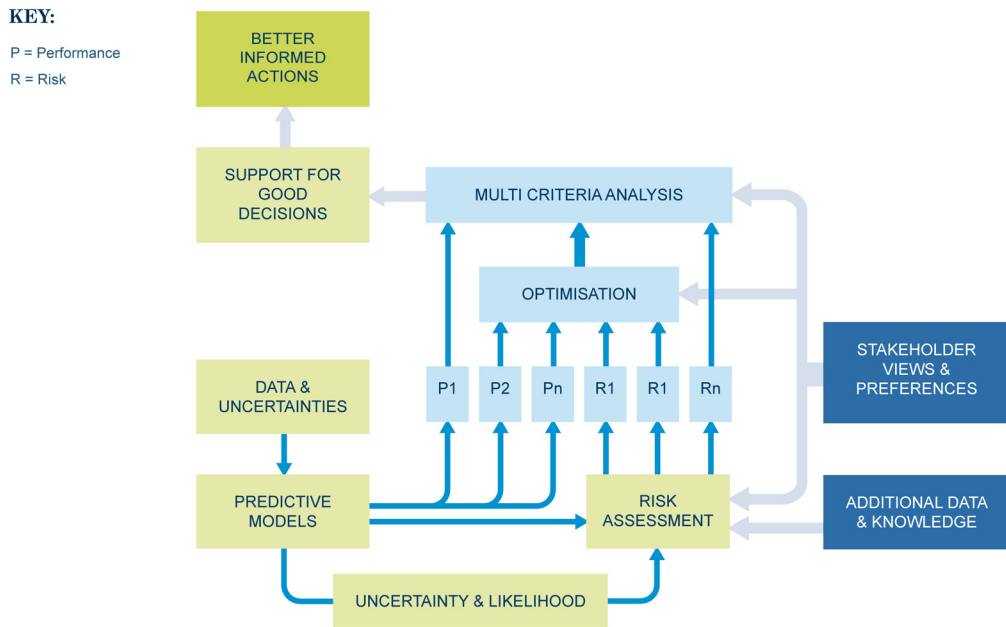*Figure 4 Process for selecting the best option; feedback loops not shown (after Blackmore et al, 2009)*
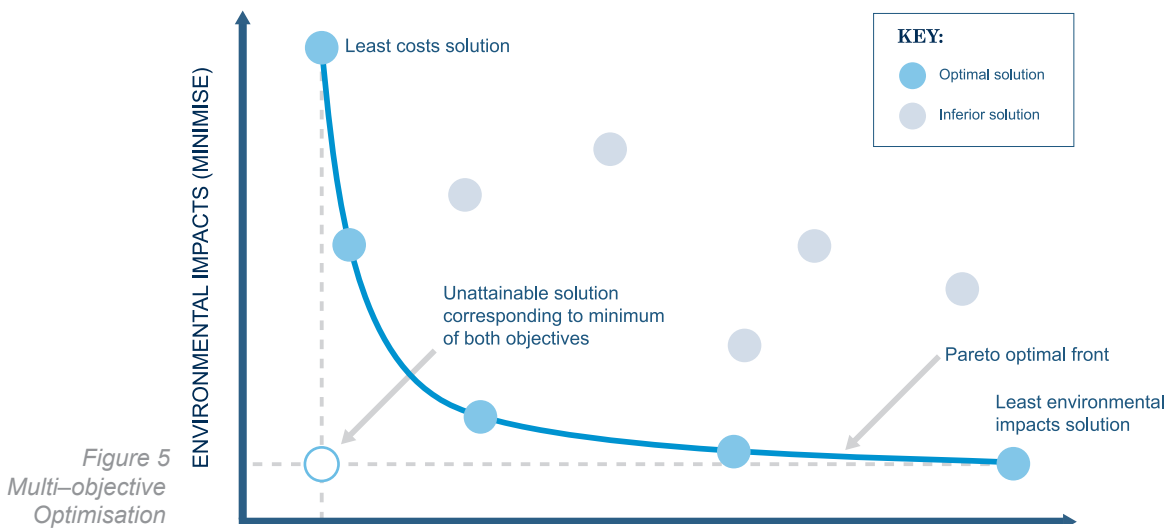
Techniques for Selecting the "Best" Option

There are many techniques that can be used, either individually or in combination, to help select the best option. Optimisation and multiple criteria analysis (MCA) are two such techniques which are being used increasingly. Others include expert opinion,

uncertainty analysis and risk assessment. Optimisation and MCA are discussed further below while uncertainty and risk are discussed in _Uncertainty and Risk (page 21)_.

Optimisation

Selecting the "best" option from a multitude of possible solutions, in the face of many possibly conflicting objectives, is a challenge faced by many water managers. For example, in the urban planning area the question might be asked "how many houses should be connected to a stormwater reuse system to give the most effective performance in terms of water supply reliability, energy consumption and capital cost?", or, in river management, "how can we operate our dams so that the months per year over which water levels are less than 20% of total system storage, the months per year over which water restrictions apply, and the operating cost, are optimized?" Without guidance and a rigorous process in place, viable options can be overlooked and "guesses" can result in sub-optimal solutions. Computational approaches to optimisation can provide comprehensive consideration to the whole decision space (thus avoiding missed opportunities), rejecting solutions that are obviously less satisfactory than (ie inferior to) others and generating a Pareto front of optimal solutions that satisfy multiple objectives. _Figure 5_ shows an example of a Pareto front where the requirement is to minimise both costs and adverse environmental impacts. Solutions selected from the Pareto front can then be taken to the decision makers to consider for final selection (Blackmore et al, 2009).

KEY:
 ● Optimal solution
 ● Inferior solution

Least costs solution

Unattainable solution corresponding to minimum of both objectives

Pareto optimal front

Least environmental impacts solution

ENVIRONMENTAL IMPACTS (MINIMISE)

Figure 5
Multi–objective
Optimisation

A variety of multi-objective optimisation techniques are available, such as genetic algorithms and simulated annealing. More information on these techniques and applications can be found in numerous sources including Collette and Siarry (2003), Loucks and Van Beek (2005), Simonovic (2009), Soncini-Sessa et al (2007) and Vàzquez and Rosato (2006). In some sources the discussion also includes some guidance on choice of method (eg Collette and Siarry, 2003). However, given the wide range of techniques available and the range of contexts where they could be applied, it

is recommended that expert advice should be sought as to the appropriate technique to use for a given application.

Multiple Criteria Assessment (MCA)

Multiple criteria analysis (MCA) is a well established methodology for ranking or scoring the overall performance of decision options against multiple objectives. The approach has widespread and growing application in the field of water management (Hajkowicz and Collins, 2007) and is potentially capable of improving the transparency, auditability and analytical rigour of water management decisions. The MCA framework ranks or scores the performance of alternative decision options against multiple criteria based on a set of performance measures, which are the raw scores for each decision option against each criterion. It can be represented by an evaluation matrix of the form shown in *Figure 6*.
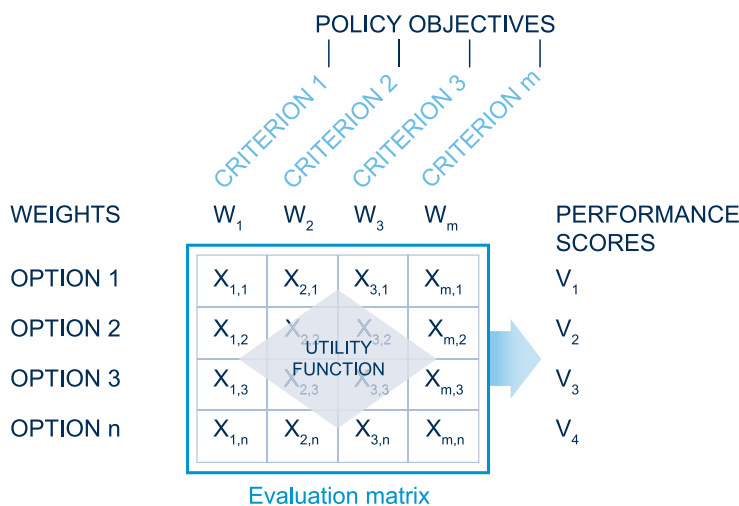

Figure 6
Evaluation matrix for multiple criteria analysis (after Hajkowicz, 2005)

A variety of MCA algorithms can be used to either rank or score the decision options. The application of a large number of techniques in 113 water related MCA studies for a variety of purposes in 34 countries is reviewed by Hajkowicz and Collins (2007). Fuzzy set analysis, paired comparison and outranking methods were found to be the most commonly applied techniques. However, it is clear that the choice of appropriate MCA technique to use is potentially very wide. Therefore, expert advice should be sought as to the appropriate technique to use for a given application.

More information on MCA techniques and applications can be found in numerous sources including Collette and Siarry (2003), Loucks and Van Beek (2005), Simonovic (2009), Soncini-Sessa et al (2007) and Vàzquez and Rosato (2006). In some sources the discussion also includes some guidance on choice of method (eg Collette and Siarry, 2003, and Soncini-Sessa et al, 2007).

Performance Criteria

Performance criteria for use in the context of a decision making process should be decided in consultation with stakeholders at the start of the project, and preferably during the problem definition phase of the project, as discussed in *System Definition*

. Where techniques such as optimisation and multiple criteria analysis (MCA) are used, setting performance criteria will include defining metrics such as objective functions for optimisation and weightings for factors included in MCA.

Where MCA is concerned, performance criteria should be chosen that are relevant to current needs. Hajkowicz and Collins (2007) list examples from water management related studies they reviewed. These include various combinations of: cost (including net present value); economic considerations (including employment, income, productivity); technical feasibility; biodiversity and wildlife protection; water quality enhancement; water supply reliability; fairness and equity; political and legal feasibility; energy supply; and human health. Hajkowicz and Collins (2007) also note that there are few methods to help with selection of performance criteria and decision options; this reinforces the guidance in *Techniques for Selecting the "Best" Option (page 43)* that expert advice should be sought when using MCA. However, irrespective of method adopted, stakeholder involvement should be obtained, such as via a workshop.

For optimisation, criteria expressed in quantitative terms are needed. For example, some of the listed criteria including cost, some economic factors, water quality enhancements and reliability of water supply could be expressed in quantitative terms, where the concept of reliability of supply could be expanded to include considerations such as delivery of environmental flows.

A critical aspect of the iterative process that should be followed to identify the preferred option will be to review these criteria in the light of results from models and other lines of evidence, and amend them as necessary. Factors that should be considered when setting performance criteria are further discussed in *System Definition (page 18)*.

Finding the "Best" Option

The methods used to analyse system performance should be closely linked to the agreed performance measures. Often, available models and methods determine which measures are evaluated, but whatever methods are used, they should be only as complicated as necessary to answer the question in hand.

Using a process such as the one illustrated in *Figure 4*, water-management stakeholders can identify and address problems by iterating around a cycle that defines objectives based on the initial problem statement and determines what metrics will be used to ascertain that the objective has been achieved within the context of a well-defined system. Different proposed solutions are then evaluated in terms of the agreed metrics and criteria (with uncertainty explicitly taken into account) and the outcomes are compared to select the "best" solution (Blackmore et al, 2009).

At the core of the process illustrated in *Figure 4* are simulation models. The models use input data sets (that include uncertainty) to predict system performance (P1 – Pn in *Figure 4*). With uncertainty protocols included in the models, multiple runs (for example, Monte Carlo simulation) can provide evaluation of the consequences and probabilities

of events. Model outputs can also be used to inform a wider risk assessment, including providing a better understanding of "worst case" scenarios, for example.

Different aspects of performance are then optimised, and a Pareto surface of optimal solutions is generated. From this, a shortlist of possible solutions is selected, any additional analysis is undertaken if required, and knowledge of how each option will perform is considered in the MCA process for the selection of a preferred option. If probability distribution functions (or any understanding of likelihood) are available for the model input data, risks associated with each of the shortlisted options can be calculated (R1 – Rn in *Figure 4*), and considered as performance measures in the MCA. Additional data and knowledge might be needed to assess all the risks that are of interest to the stakeholders, and stakeholders' views and personal experience provide additional input to the multi-criteria assessment. An understanding of the quality of the evidence used to support each step of the process enhances the value of the decision support (Blackmore et al, 2009).

It should be noted that the level of rigour with which each element of the process shown in *Figure 4* is applied, and therefore choice of approach to use, can vary with circumstances. Considerations such as cost-effectiveness of more rigorous analyses, and consequences if a less than optimal selection is made, are among relevant factors influencing choice of approaches.

In some cases it may be sufficient to work through the process "intuitively"/interactively such as in a workshop with stakeholders and relying only on expert opinion (eg via a Citizens Jury – Jefferson Center, 2004), but even in this situation a simple multiple criteria assessment, where weightings have been considered and agreed by stakeholders, will improve the quality of the final outcome. Formal mathematical techniques such as optimisation, uncertainty analysis and MCA will greatly improve outcomes, but should be carefully explained so that they do not alienate stakeholders, and might be inappropriate for evaluating sociological aspects. Whatever method is used, the important thing is to document outcomes and reasons for them.

Report/ Communicate Comparison of Options

Results of the comparison of options should be reported to, and discussed with, stakeholders using techniques that facilitates their understanding and acceptance (see *Information Communication (page 14)*). Information should be provided in a form that makes the task of integrating or combining it with additional information from other sources (ie additional to all the data used in this project and the project results) which may have a bearing on the final decision as to the preferred option to adopt.
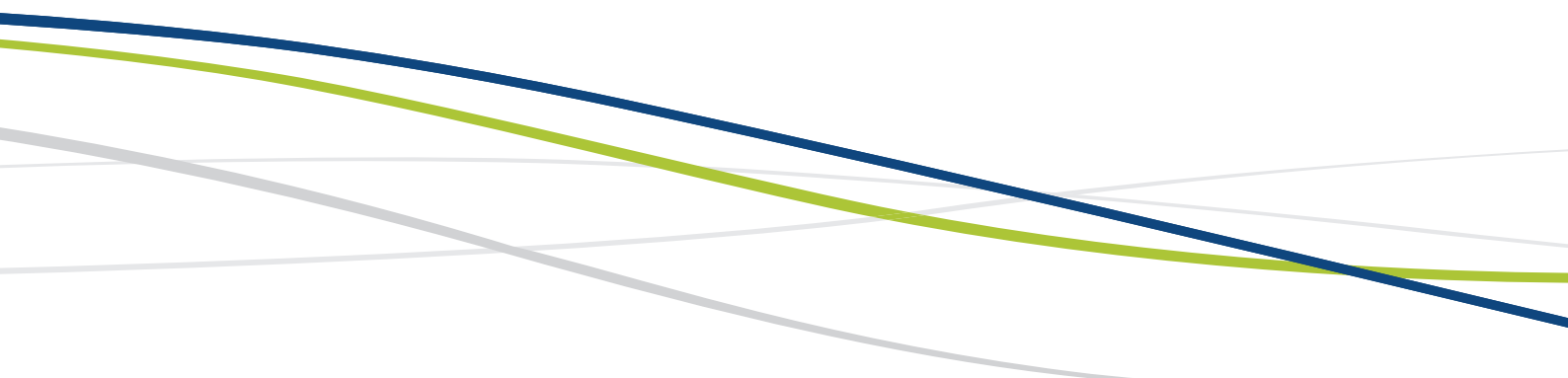
Model Choice

Overview   Models may be used to support activities in a number of contexts and the types of models and the requirements for model application may vary greatly. Generic guidance on model choice is provided in the first of two CRC for Catchment Hydrology reports on this subject (CRC for Catchment Hydrology, undated) and guidance relevant to two water quality modelling domains is provided in CRC for Catchment Hydrology (2005).

When selecting a model, the objectives of applying the model and all relevant constraints that apply should be taken into account. It is often difficult to identify and assess the relative advantages and disadvantages of models that are potential candidates for application for a given purpose. In addition, the decision to adopt a certain model for a given purpose may have wider implications; for example, adoption of a certain model for water resources planning and management may have implications for short and medium term hydrological forecast modelling, and also for future directions of research in modelling.

Hence, it is likely that pragmatic choices will have to be made when choosing which model or models, and the appropriate level of model complexity, to adopt. As models gain complexity, or expand the processes represented, the demand for data to calibrate and validate them increases (Silberstein, 2006), and this data is often not available or inadequate. Hence, a balance has to be struck between model complexity, availability of data for model calibration and validation, and model predictive performance (CRC for Catchment Hydrology, undated). The conceptual relationship between these three factors is illustrated in *Figure 7* (after Grayson and Blöschl, 2001), where the optimum combination lies along the "ridge" that runs upwards from about the intersection of the data availability and model complexity axes (highlighted by the dot on the red line running across the figure); this relates to the well known issue of model parsimony discussed below.

Model Parsimony   The issue of model parsimony, and the related problems of model equifinality ("non-uniqueness" in groundwater modelling terminology; Grayson and Blöschl, 2001), parameter identifiability and scale, are extensively discussed in the international literature. These apply in principle to all modelling domains relevant to water management (eg Blöschl, 2006; Oreskes et al, 1994), including groundwater modelling as well as surface water modelling, notwithstanding that groundwater modelling presents some intrinsically different challenges to surface water modelling. In the field of surface water modelling, much of the discussion is with reference to catchment modelling (eg Beven, 1989; Beven, 1993; Beven, 1995; Croke and Jakeman, 2001; Grayson and Blöschl, 2001 (Section 3.3.5, Chapter 3); Hairsine and Sander, 2009; Jakeman et al, 2006; Kirchner, 2006; Perrin et al, 2001; Silberstein, 2006; Son and Sivapalan, 2007; Young et al, 2006). Review and practical guidance in the context of groundwater modelling is provided by Hill and Tiedeman (2007), for example.

Unfortunately this principle is often overlooked. As stated by Jakeman et al (2006):

> "Model structures with too many parameters are still endemic. Models with too many degrees of freedom incur serious risks. Among them are: fitting to inconsistent or irrelevant "noise" components of records; severely diminished predictive power; ill defined, near-redundant parameter combinations; and obscuring of significant behaviour by the spurious variation allowed by too much freedom. Even so, model testing for redundancies and possible model reduction are seldom reported. Data paucity should limit the model complexity. For example, in modelling of flow and transport for prediction, spatial data on landscape attributes may be useful to structure and discretise a model in fine detail, but detail is unwarranted if the flux measurements available for model calibration cannot support it".

(Noting that data for calibration could include other data such as storage, soil moisture, vegetation patterns and other data derived from fluxes, as well as fluxes). *Figure 7* illustrates the point, where the shaded area shows the zone where models are too complex and with insufficient data, such that an optimum set of parameter values cannot be defined with confidence (CRC for Catchment Hydrology, undated).

Figure 7 Conceptual relationship between model complexity, availability of data for model calibration and validation, and model predictive performance (after Grayson and Blöschl, 2001)
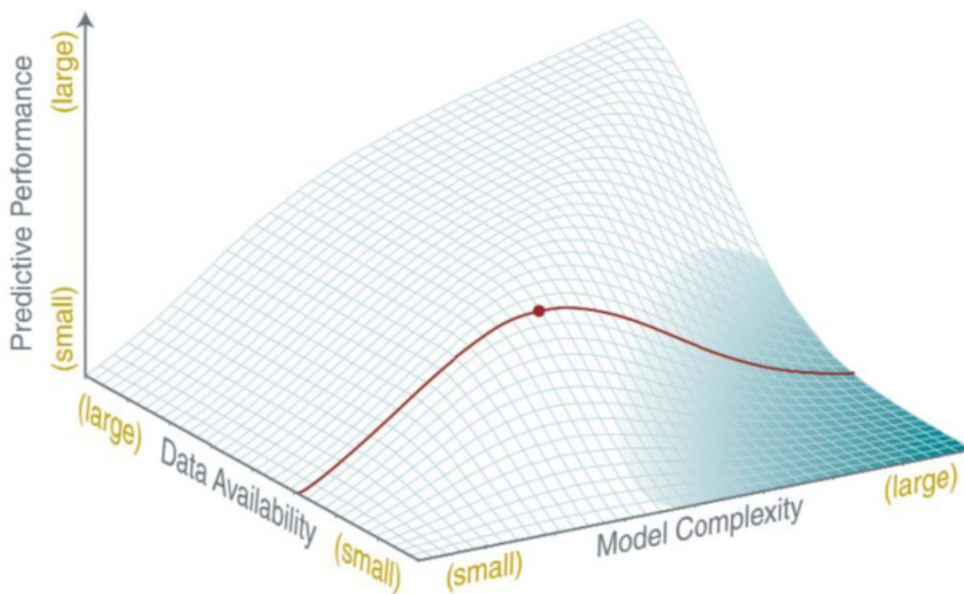
Uncertainty and the trade-off between uncertainty and model complexity is the nub of the issue. This is particularly true when the requirement is to minimise predictive uncertainty, for example, of a time series of streamflows, but may be less so when the requirement is to gain a qualitative understanding of an aspect of system behaviour. With reference to predictive uncertainty, Silberstein (2006) shows that the complexity of the model adopted should be just enough to minimise uncertainty, but any more or less complexity will increase uncertainty. Ultimately, with increasing complexity, a point will be reached where model complexity becomes so much greater than can be justified,

given the uncertainty constraints which apply, that calibration and interpretation of results of model application become intractable problems (ie due to the equifinality problem). Hence, often a simple model will be more appropriate than a complex model and the results from it, even if superficially less detailed, will be more meaningful.

Conversely, there is a limit to model simplicity and this limit is reached when the model fails to adequately explain the observations (Perrin et al, 2001). If the model fails to explain observations this may be due to incorrect process specification, incomplete process specification (lack of feedbacks), or data limitations. The model should be just complex enough to capture important forcings and feedbacks that can dominate behaviour, but no more. The appropriate level of complexity in a given system varies depending on the modelling time step and this needs to be commensurate with the response characteristics of key factors of interest; eg if environmental flows are of interest and these are sensitive to daily flow patterns then modelling of these should be at a daily time step. A balance therefore needs to be struck between model identifiability and simplicity to avoid over-simplification, even though use of a simplified model may be attractive in order to minimise data requirements. This also highlights the importance of conceptual model development, discussed in , for identifying the important feedbacks and controls so that (at least) the limitations of a modelling approach can be evaluated, or the need for greater investment in data or alternative model approaches can be considered. In cases where sufficient data is not available ,or feasible to collect, to avoid using an over-simplified model then the limitations of this need to be clearly communicated (L.E. Band, pers. comm., 2010).

In relation to catchment models, much of the discussion concerns "physically-based" models and distributed models. With reference to "physically-based" models, Beven (1989) argued:

> *"… that there are fundamental problems in the application of physically-based models for practical prediction in hydrology. These problems result from limitations of the model equations relative to a heterogeneous reality; the lack of a theory of subgrid scale integration; practical constraints on solution methodologies; and problems of dimensionality in parameter calibration."*

Beven (1993) also stated:

> *"Difficulties in defining truly mechanistic model structures and difficulties of model calibration and validation suggest that the application of distributed hydrological models is more an exercise in prophecy than prediction."*

That is, problems arise when oversimplified process representations are used which fix important feedbacks as constant, calibrated parameters; examples include a lack of phenology, ignoring the role of stomatal controls and inadequate terrain resolution (L.E. Band, pers. comm., 2010). The lines of argument put forward by Beven (1989 and

1993) are supported by Grayson and Blöschl (2001), Hairsine and Sander (2009), and Kirchner (2006), amongst others.

In the context of groundwater modelling, Hill and Tiedeman (2007) discuss the trade-off between model fit and prediction accuracy with respect to the number of model parameters requiring calibration and also show there is an optimum, in a similar fashion to Silberstein (2006). To assist with finding this optimum all model fit statistics proposed by Hill and Tiedeman (2007: Section 6.3.2) include a penalty as the number of parameters requiring calibration increases.

Siberstein (2006) also discusses the issue of using models as a substitute for data collection and argues:

> *"that improvement in the management of our environment and water resources will not come with improved models in the absence of improved data collection because we cannot manage what we do not measure."*

Put another way: lack of data may justify using a simple model, but doing this does not make up for the lack of data.

Factors Influencing Model Selection

WMO (2009) include the following factors and criteria as being relevant when selecting a model:

a)   The general modelling objective; eg hydrological forecasting, assessing human influences on the natural hydrological regime or climate change impact assessment, or a combination of these.

b)   The type of system to be modelled; eg small catchment, aquifer, river reach, reservoir or large river basin.

c)   The hydrological element(s) to be modelled; eg floods, daily average discharges, monthly average discharges, groundwater levels, water quality, and emergent areas such as environmental watering and aquatic ecosystem health, amongst others.

d)   The climatic and physiographic characteristics of the system to be modelled.

e)   Data availability with regard to type, length and quality of data versus data requirements for model calibration and operation.

f)   Model simplicity, as far as hydrological complexity and ease of application are concerned.

g)   The possible need for transposing model parameter values from smaller catchments/hydrological units/systems to larger ones (or other ungauged ones).

h)   The ability of the model to be updated conveniently on the basis of current
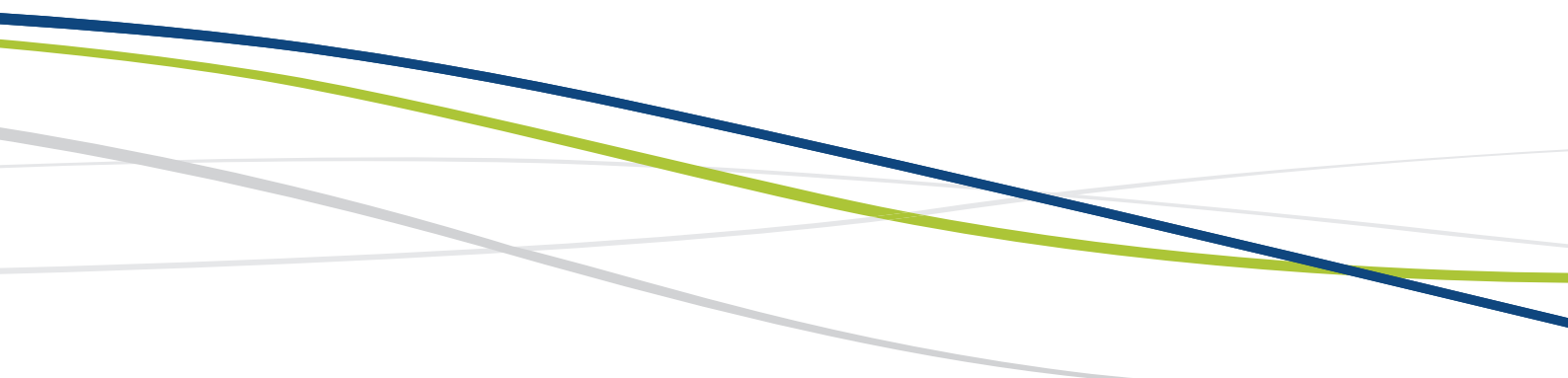
hydrometeorological conditions.

Based on these, the following factors are relevant:

- The level of modelling expertise available. If necessary, additional skills may have to be brought in (eg through hiring consultants with appropriate expertise).

- Uncertainty issues, discussed in *Sensitivity/Uncertainty Analysis (page 39)*, the broad purpose of modelling (eg obtain quantitative predictive results with uncertainty minimised or provide qualitative understanding) and the balance that needs to be struck between model parsimony and adequately representing key responses, as discussed in *Model Parsimony (page 50)*.

- The time interval for which results are wanted (eg daily, monthly or seasonal), also discussed in *Model Parsimony (page 50)*. A rule of thumb, which applies particularly to catchment rainfall-runoff modelling, is that to obtain reliable monthly time series results the modelling should be undertaken at a daily time step; for reliable daily time series results, modelling should be undertaken at a sub-daily time step (preferably no more than hourly) although it is recognised this is not often practicable. For reliable annual or seasonal time series results, modelling at a monthly time step may suffice. However, if only statistical results are required (eg flow duration curve and event frequencies) then modelling at the time step for which the results are required may be sufficient (eg modelling at a daily time step should be sufficient for obtaining daily statistics).

- Whether the model is going to be used on its own, or if it is going to be used in conjunction with other models. For example, if one or more models need to be linked together to obtain the results required then there may be incompatibilities that will have to be reconciled. Even where two implementations of the same modelling tool are involved (for example, connecting a model of an upstream catchment to a model of a downstream catchment), the need to avoid or reconcile incompatibilities may become a constraint on project methodology. Where models for different domains are involved (such as for modelling surface water and groundwater), and the way they operate is fundamentally different, then some innovative approaches may have to be adopted to avoid problems of incompatibility. In some cases, such as with surface water and groundwater models, there is generally more flexibility in the ways surface water models can be applied and the solution may lie in adapting the project methodology to suit the more constrained groundwater model.

- Freedom of choice may be limited by a desire to minimise problems of different models for much the same purpose in the same project area, or to avoid problems of different models in adjoining project areas, particularly where the models may need to be linked in some way in the future or results compared in some way.

- Whether a deterministic or a stochastic modelling approach is required. While most model applications are deterministic, the use of an approach where model

inputs or parameter values, or both, are varied stochastically (ie involving use of random numbers) can sometimes be appropriate. An issue with using truly random numbers is that results are not repeatable which makes interpreting results and communications with stakeholders more difficult. To some extent these problems can be overcome by using pseudo-random numbers which will give a repeatable sequence if a fixed seed number is used. This partly negates the rationale for using a stochastic approach but the approach still enables provision to be made for modelling the components of system behaviour that cannot be explained via the deterministic approach.

- Whether simulation or optimisation, or a combination of both, is needed.

- Whether the model is to be used for hindcasting (see glossary, *Chapter 6 (page 73)*) or forecasting (see glossary) when being applied in predictive mode; this is a decision that may be affected by other considerations such as whether a surface water model is to be linked to a groundwater model. Hindcasting involves the use of historical data as input (and must be used in model calibration and validation). When forecasting, the appropriate data source(s) will depend on whether the application is for short term, medium range or long range hydrological forecasting. Short term and medium range hydrological forecasting typically entail use of current, real time, data together with some predicted data such as a flood recession, while long range hydrological forecasting typically involves the use of stochastically generated input data based on a prescribed set of statistics or data synthesised from another model to represent possible future conditions. In water planning, the choice of hindcasting or long range hydrological forecasting has implications for modelling scenarios with trends or step changes in them, including scenarios considering trends in groundwater conditions, climate change, growth in urban water demands, land use/cover changes such as fires and forestry impacts, and new structural features coming on-line at various times. Modelling these is easier and more transparent when long range forecasting is used but interpreting and communicating results may be more difficult (use of long term forecasting may also overcome some of the compatibility issues discussed above). Irrespective of whether hindcasting or forecasting is used, model results should only be interpreted and reported in a statistical context.

4
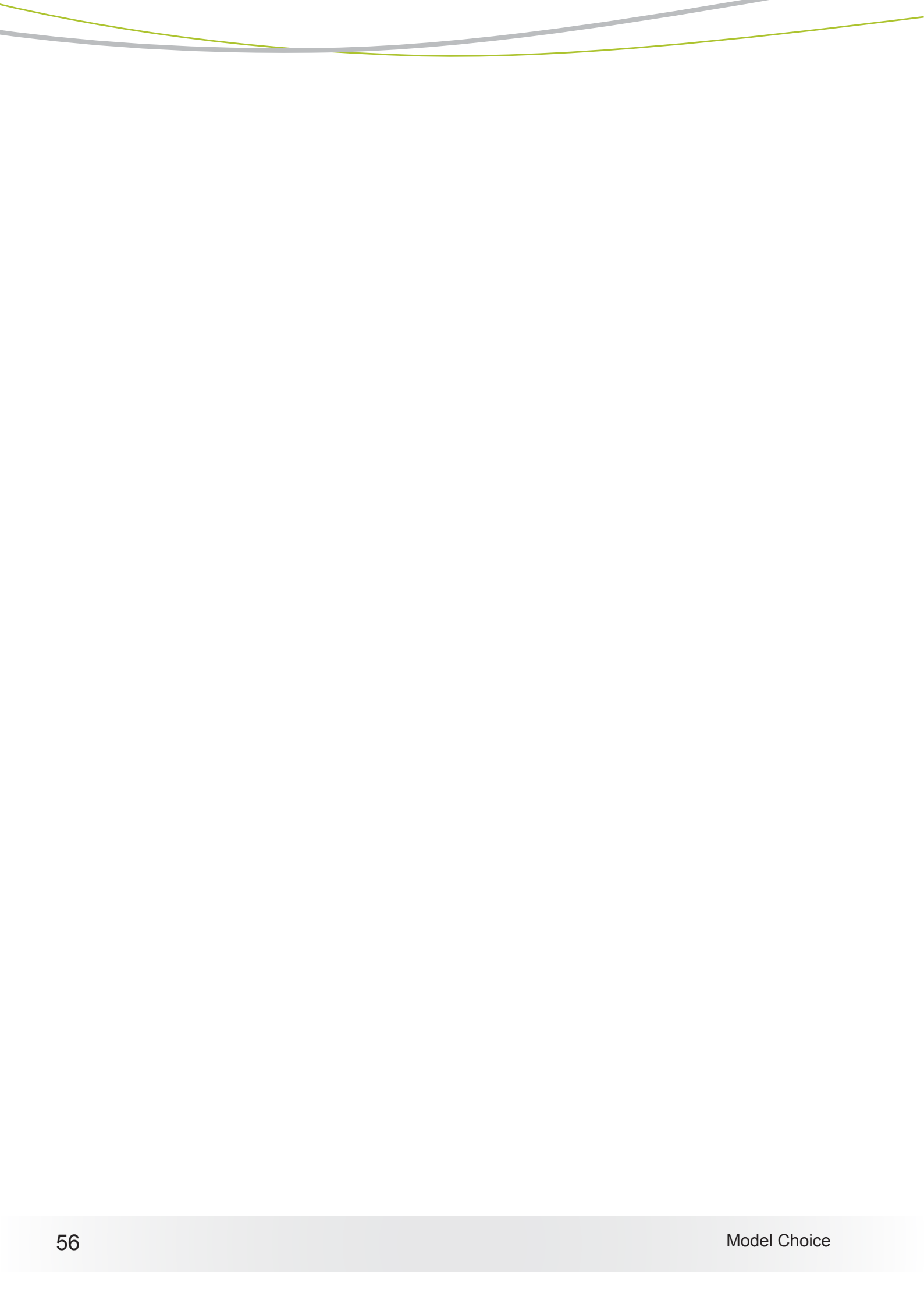
Further Reading

Guidelines for modelling have been developed by a number of organisations. To varying degrees, these cover technical issues of development, implementation and use of models (mainly domain specific), and also issues relating to interaction between modellers and end users of model information, where the content may be more general. However, all are basically quality assurance procedures.

Guidelines relevant to surface water modelling found include USEPA Guidelines (USEPA, 2002 and CREM, 2008), Californian Guidelines (Bay-Delta Modeling Forum, 2000), Dutch guidelines (Van Waveren, et al, 2000) and European Union Guidelines (Scholten et al, 2007). The USEPA guidelines are completely generic, and are relevant to domains such as air quality and public health modelling as well as to water related domains. The CREM document does not describe a quality assurance process as such, but it complements the USEPA guidelines with information, scientific background and best practice guidance on model development, evaluation and application (Packman and Old, 2005). The Californian guidelines are water-specific, but are sufficiently generic to be relevant to just about any surface water or groundwater modelling domain. The Dutch guidelines are also generic but cover a number of specific modelling domains as well, although they do not address river system modelling for planning purposes. Their development involved all the main players in the Dutch water management field (Refsgaard and Henriksen, 2004). It is notable that all these guidelines are generally seen as promoting "good practice" rather than "best practice".

The European Union Guidelines were developed under the HarmoniQuA project and are based on input from pre-existing guidelines such as the Dutch Guidelines and MDBC Groundwater Flow Modelling Guideline (Middlemis et al, 2000), amongst others. The HarmoniQuA project has delivered a range of other products including a range of software to support users (see *harmoniqua.wau.nl/public/Products/software.htm)*, particularly MoST (Scholten et al, 2007). The guidelines are in the form of a Knowledge Base that can be downloaded as an integrated part of the MoST software or as a standalone text file (see *harmoniqua.wau.nl/* and *www.harmoni-ca.info/toolbox/index.php*). In common with the Dutch Guidelines, these guidelines are generic but have elements that are specific to a number of domains as well. The main elements of the guidelines are also described in a paper which discusses the possibilities for establishing a European quality assurance standard for modelling (Packman and Old, 2005).

Of the groundwater guidelines available, the most widely cited internationally is the MDBC Groundwater Flow Modelling Guideline (Middlemis et al, 2000). The drivers for these were a perception among end-users that model capabilities may have been over-sold, and that there is a lack of consistency in approaches, communication and understanding among and between modellers and water resources managers, often resulting in considerable uncertainty for decision making.

Other groundwater modelling guidelines available include those produced by Hill and Tiedeman (2007); these have developed from short courses conducted since 1991. In particular, guidance is provided on:

- Sensitivity analysis to evaluate the information content of data;

- Data assessment to identify (a) existing measurements that dominate model development and predictions and (b) potential measurements likely to improve the reliability of predictions;

- Calibration to develop models that are consistent with the data in an optimal manner; and

- Uncertainty evaluation to quantify and communicate errors in simulated results that are often used to make important societal decisions.

A good summary of the steps in the procedure advocated (Hill and Tiedeman, 2007: Table 10.1), a review of previous work, worked examples and exercises are provided as well.

Refsgaard et al (2005a) identify three types of quality assurance guidelines for modelling:

1. Internal technical guidelines – examples cited include user manuals for particular models;

2. Public technical guidelines – often containing the same substance as internal technical guidelines but they differ in the sense that they have been prepared through a consultative and consensus building process involving many persons and organisations;

3. Public interactive (sic) guidelines - established through a public consultative and consensus building process, like the public technical guidelines but differing in that they have an additional focus on regulating the interaction between the modeller and the water manager, who often have the roles of consultant and client, respectively.

Examples of public interactive guidelines cited in Refsgaard et al (2005a) include the MDBC Groundwater Flow Modelling Guideline (Middlemis et al, 2000), the Dutch guidelines (Van Waveren, et al, 2000) and the Californian Guidelines (Bay-Delta Modeling Forum, 2000).

Other relevant documents available include:

- a HarmoniQuA project review of the state of the art in quality assurance in modelling related to river basin management (Refsgaard, 2002);

- a position paper on steps needed for quality assurance in the development and evaluation of environmental models (Jakeman et al, 2006), and an evaluation
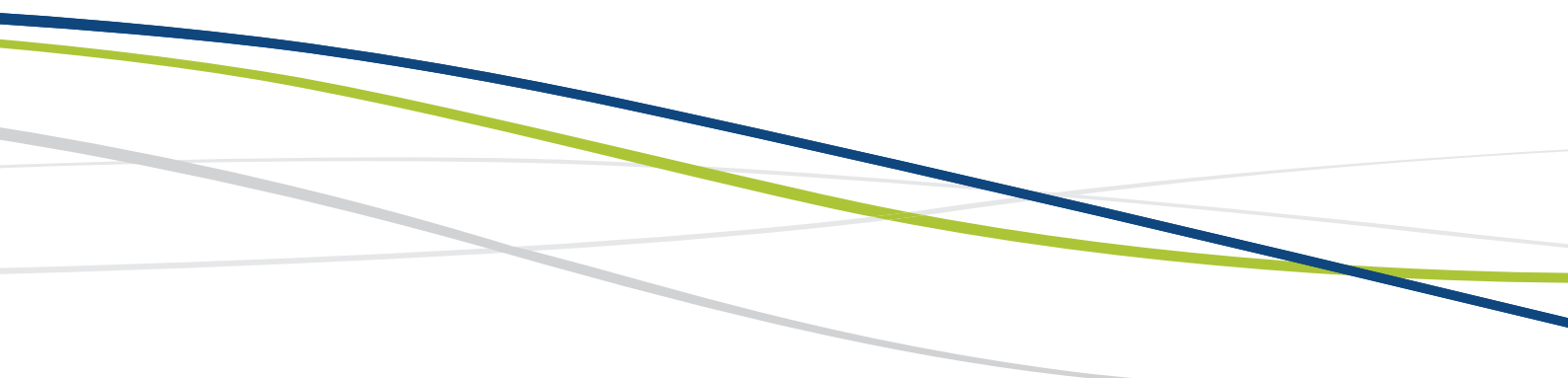
by Robson et al (2008) of the usefulness of these steps in the context of the development and application of process-based biogeochemical models of estuaries;

- a book chapter on best practice modelling (Crout et al, 2008);

- book chapters on environmental decision making (Maier et al, 2008; Soncini-Sessa et al, 2007) and policy implementation under uncertainty (Myšiak et al, 2008);

- text books on modelling for water management (eg Loucks and Van Beek, 2005; Simonovic, 2009); and

- National Modelling Guidelines for Water Distribution Network Modelling from New Zealand (Water New Zealand, 2009).

The focus of the position paper by Jakeman et al (2006) is mainly on catchment modelling although the ten steps only go as far as the model evaluation or testing phase. Robson et al (2008), in their evaluation of the ten steps, discuss the need to include consideration of fitness of the model for supporting the answering of questions posed in project requirements. The focus of the New Zealand guidelines (Water New Zealand, 2009) is on water distribution networks, but the principles articulated are relevant to river system modelling as well.

There is a fair degree of consistency between all the guidelines and the position paper of Jakeman et al (2006), in terms of the QA process they promote. However, it is notable that only the New Zealand and USEPA guidelines give much prominence to the need for data review and clean-up, and the effort that can be needed for this step is generally grossly underestimated where it is addressed at all. Apart from the New Zealand and the California guidelines, all the QA processes are based on the assumption of discrete modelling projects rather than ongoing modelling activities.

5

Case Study: Great Barrier Reef Catchment

This chapter was prepared by Dave Waters and Chris Carroll. It describes a case study applying the eWater Best Practice Modelling decision framework (*Figure 1*) to an on-going real-world modelling project in the Great Barrier Reef (GBR) in Queensland. The case study provides an example of best practice, given the understanding of the project requirements, and the time, science and methodology limitations which applied at the time the project was started. The project team acknowledge issues with the modelling have been identified as the project has been rolled out. In accordance with the best practice approach, it is intended to address these in the light of lessons from the first round of modelling and data improvements, through a process of continuing improvement in later rounds of modelling.

The project itself is an important component of the Queensland Government's Paddock to Reef Integrated Monitoring Modelling (Reef M&M) Program. This program has been established to measure and report on progress towards the targets set in the Reef Plan (Queensland Department of the Premier and Cabinet, 2009). It combines monitoring and modelling at paddock through to catchment and reef scales. The program area comprises six Natural Resource Management (NRM) regions (*Figure 8*). These regions include all 35 catchments that drain to the GBR lagoon. Summary statistics for these six regions are provided in *Table 1 (page 66)*. More contextual information is provided in *Problem Definition (page 65)*.
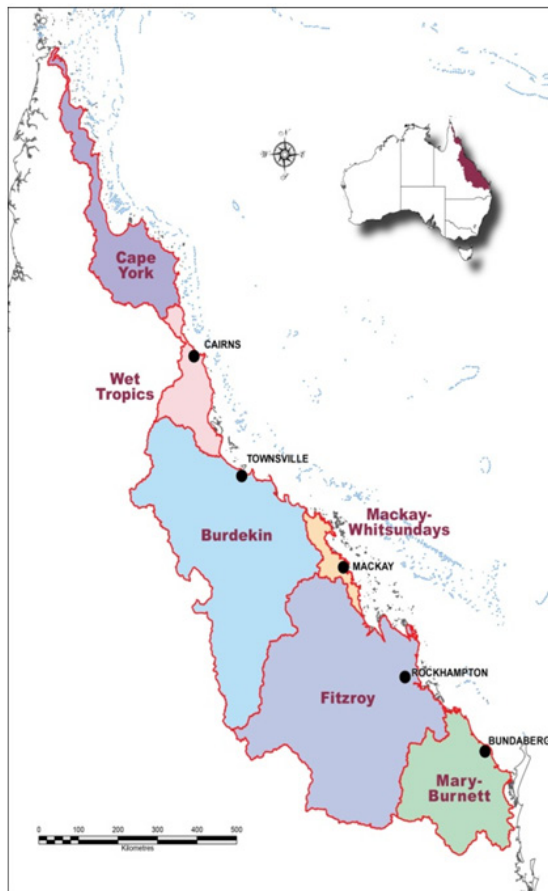

*Figure 8
Location and extent of the six NRM regions (35 catchments) being modelled with Source Catchments*

Project Administration

Project Governance

The Queensland and Commonwealth government administer the Reef M&M Program through an Intergovernmental Organisational Committee (IOC), supported by a Reef IOC Monitoring & Evaluation sub-committee, and Reef Coordination Advisory Group (CAG), as shown in *Figure 9*. This Committee is chaired by the Reef Secretariat in the Queensland Department of the Premier and Cabinet and consists of state and commonwealth scientists, Great Barrier Reef Marine Parks, and Natural Resource Management Group representatives. An Independent Science Panel (ISP) provides scientific oversight of the overall M&M program. Delivery of the modelling project is monitored by the CAG, and the CAG must authorise any variations in this project.
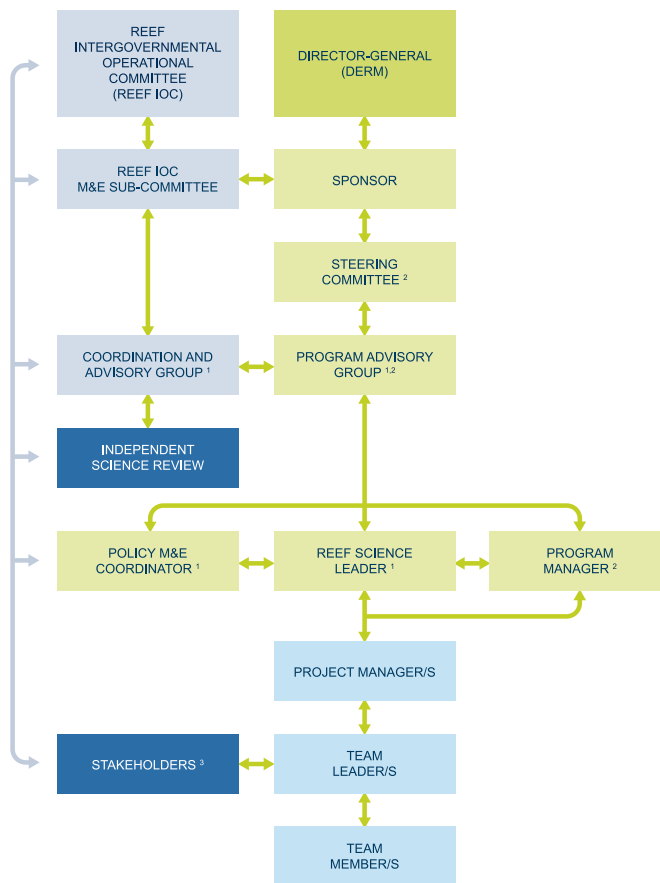
Figure 9
Project Governance arrangements which include Government, Scientists and local stakeholders

Notes on *Figure 9*

1. The Reef Science Leader and the Policy M&E Coordinator are representatives on the Coordination and Advisory Group and the Program Advisory Group.

2. The Project Manager provides secretariat support to the Steering Committee and the Program Advisory Group.

3. Various stakeholders are engaged as appropriate by each level in the hierarchy.

Project Management

The Queensland Department of Environment and Natural Resource Management (DERM) are responsible for undertaking and delivering the catchment modelling. A Reef Science Team Leader and Catchment Modelling Leader provide the project management and report to the Reef CAG and the ISP.

Resources, Timeframe, Budget

Six catchment modellers have been allocated across the six Reef NRM regions (shown in *Figure 8)*. Experienced modellers were appointed due to the complexity of the work and were appointed for three years as at early 2011. The timeframe for initial model development and reporting was extremely short and this has limited the extent to which new science could be incorporated during model development. Updated modelling is required every year for reporting on Reef Plan water quality targets.

Peer Review

There are three levels of peer review of the reef catchment modelling: internal (within the State Government), external (Industry and Research partners) and through the Independent Science Panel (ISP), with the ISP appointed through the IOC M&E sub-committee.

Stakeholder Consultation & Communication

A series of reef-wide regional workshops were conducted to develop the overall Reef M&M Program. More than 100 scientific and technical personnel from 18 organisations were involved in the program design. A Stakeholder reference group is also part of the overall reef governance arrangement, as shown in *Figure 9*.

For the modelling project, each of the six catchment modellers was allocated a Reef NRM region to model, with four based in the actual region. The modellers have a role in consulting and communicating with the Natural Resource Management Group for the region and capturing local corporate and industry knowledge of regional land uses, hydrology and water quality. Outputs from the catchment modelling are reported in an Annual Reef Report Card and through technical reports and regional workshops.

Documentation and Archiving

For day-to-day model development and documentation of modelling methodology a centralised wiki has been established. The wiki was established for the modelling team and software developers to ensure a single point of truth for documentation and to provide an efficient mechanism for project staff to keep abreast of updates and progress.

Technical reports will be developed for each regional model which will include the detailed modelling methodology, assumptions and results. In addition, a summary report will be developed for the entire GBR. The aggregated model outputs from the technical reports will be included in the Annual Reef Report Card.

For long-term documentation and archiving, a Spatial and Scientific Information Management for Reef (SSIMR) project has been established for all paddock and catchment modelling and monitoring data. This project was funded as the department saw the importance of ensuring all data and model runs were easily accessible into the future. All catchment modelling input point data and spatial layers, modelled outputs for

all scenario runs, assumptions and relevant versions of the software and associated documentation will be captured and archived through the SSIMR project. All associated data and model runs are available to reef collaborators online.

Problem Definition

Problem Statement

Over the past 150 years the Great Barrier Reef (GBR) catchments have been extensively modified for agricultural production and urban settlement, leading to a decline in water quality entering the Great Barrier Reef lagoon. A scientific consensus statement concluded 'water discharged from rivers to the GBR continues to be of poor water quality in many locations'; and 'land derived contaminants, including suspended sediments, nutrients and pesticides, are present at concentrations to cause environmental harm' (Brodie et al., 2008). In response to these water quality concerns the Reef Water Quality Protection Plan (the Reef Plan) was initiated in 2003 and updated in 2009 through a joint Queensland and Australian government initiative (Queensland Department of the Premier and Cabinet, 2009). A clear set of water quality and management practice targets are outlined for catchments draining to the Great Barrier Reef, with the immediate goal to halt and reverse the decline in water quality entering the reef by 2013; and the long-term goal to ensure that by 2020 the quality of water entering the Reef from adjacent catchments has no detrimental impact on the health and resilience of the Reef.

Objectives

Catchment modelling is being used to report catchment pollutant loads for each catchment draining to the Great Barrier Reef, for:

1. A baseline of 2008/2009 catchment land use and management conditions; and

2. Changes relative to the baseline for each subsequent year from 2010 to 2013 post investment in improved management practices.

Problem Domain

The Reef M&M Program is a collaborative arrangement between state and commonwealth government research organisations, Natural Resource Management regional groups, universities and agricultural industry (cane, grains, grazing and horticulture). It was identified from the start that all relevant stakeholders must be included in the process to address the water quality issues in the GBR.

One significant feature is that the program is a continuum which seeks to both influence and change on-farm management practices, through incentives and policy, with the aim that improvements in management practice will result in improvements in water quality at the end of catchments discharging to the reef lagoon. Another important aspect of the program is that the government are investing a significant amount of money in water quality monitoring at a range of scales and modelling to ensure the best available data is used to parameterise and validate models.

System Definition

The GBR modelling project covers a large geographical extent from Cape York in the north to the Mary Catchment in the south, as shown in Figure 8, an area of approximately 405,000 km^2. Large climatic variation occurs across the study area with average annual rainfall ranging from 500 – 3100 mm. The modelled pollutants of concern to the GBR ecosystem are sediments, speciated nutrients and pesticides. For the high rainfall areas such as the Wet Tropics and Mackay-Whitsundays regions, nutrients and pesticides from canelands are the major pollutants of interest. For the Cape, Burdekin, Fitzroy and Mary-Burnett regions, which are predominantly grazing and nature conservation areas (> 80%) with small areas of cropping, sediment and nutrients from hillslope and gully erosion are a major source of pollutants in comparison to the wet tropics catchments. The Burdekin, Fitzroy and Mary-Burnett have major storages with significant irrigation extraction. The models are being developed to demonstrate the long–term improvement in water quality resulting from investment in improved farm management practices such as conservation tillage, riparian management and increasing ground cover. The models are constructed to model generation, delivery and transport of the major pollutants of interest. The modelling takes into consideration the complex interactions of climate, soils, and land use and land management. The modelling exercise is building on many years of literature, research, and monitoring, modelling and expert knowledge available across the GBR.

Table 1
Summary statistics for the six regions being modelled with Source Catchments

Region	Catchment Area (km^2)	Climate Zone	Rainfall (mm/yr)	Dominant Land Uses
Cape York	42,793	Tropical	800 – 2400	Grazing 52 %, Nature Con. 46 %
Wet Tropics	22,000	Tropical	716 – 3015	Grazing 46 %, Nature Con. 33 %, Sugar cane 9 %
Burdekin	138,245	Dry tropics	649 – 1800	Grazing 94 %, Forest & Nature Con. 3 %, Sugar cane 1 %
Mackay/ Whitsundays	9021	Dry tropics	600 – 3000	Grazing 47 % , Forest & Nature Con. 26 %, Sugar 19 %
Fitzroy	142,000	Dry tropics/ sub-tropics	526 – 2065	Grazing 82 %, Forest & Nature Con. 10 %, Cropping 7 %
Burnett/Mary	51,722	Sub-tropics	600 – 2000	Grazing 78%, Forest 16 %, Cropping 4 %

Conceptual Models

The Reef M&M framework requires the ability to link paddock and catchment monitoring and modelling outputs from a local, to sub catchment, catchment and through to the marine scale. The key components of the conceptual model from paddock to reef are captured in *Figure 10*. These include plot and paddock scale research examining the effectiveness of alternative management practices, water quality monitoring from farm to end of catchment, utilising the latest satellite imagery to track spatial and temporal changes in cover to paddock and catchment scale modelling of the response to the significant investment in improved on farm management practices.

Case Study: Great Barrier Reef Catchment

The updated Reef Plan (Queensland Department of the Premier and Cabinet, 2009) clearly outlines water quality and management practice targets. The specific water quality targets are:

By 2013 there will be a:

- minimum 50% reduction in nitrogen and phosphorus loads at the end of catchments;

- minimum 50% reduction in pesticide loads at end of catchments;

- minimum of 50% late dry season groundcover on dry tropical grazing land.

By 2020 there will be a:

- Minimum 20% reduction in sediment load at the end of catchment.

The 2008/2009 year is being used as the baseline year. The changes in water quality in subsequent years will be modelled and assessed against the baseline year.
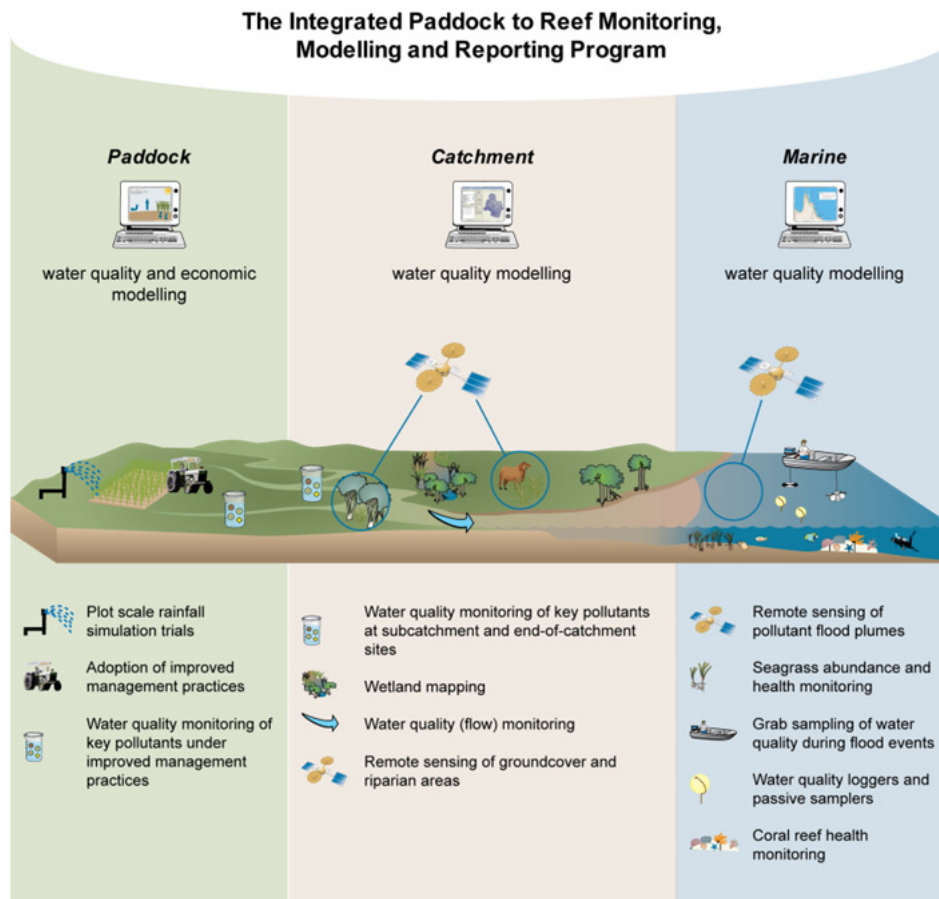

Figure 10 Paddock to Reef Monitoring & Modelling Program

The reef targets will be assessed using four lines of evidence (Carroll et al, 2011):

1. Effectiveness of land management practices;

2. Prevalence of improved land management practices;

3.   Improvements in water quality (long-term monitoring); and

4.   Improvements in water quality (modelling).

Decision Variables

Identifying farm management practices that reduce sediment, nutrient and pesticide loads at a plot/paddock scale is the first step towards improving water quality at the larger catchment scale and subsequently in the GBR marine lagoon. It is a "no regrets approach" in that even though changes in water quality might not be detected at the marine scale in the short-term the changes at the paddock scale will be evident sooner and in themselves provide local environmental and economic benefits.

Ground cover presence and persistence on dry tropical grazing lands, the extent and connectivity of intact riparian areas, and the location, persistence and inundation frequency of wetlands are important catchment attributes that play a role in the water quality leaving paddock and properties, and entering streams and the reef lagoon. Each of these attributes has specific management practice targets associated with them under the Reef Plan.

Uncertainty and Risk

In relation to the modelling program, it is acknowledged that there is a high degree of uncertainty around the modelling including input data, model structure, measured data. The project objectives require the reporting of the "relative" change in pollutant loads as opposed to absolute loads entering the reef as a result of implementing improved management practices. All the various sources of uncertainty associated with model inputs and outputs are not specifically reported, due to tight timelines for year one. However, Ellis et al (2009) demonstrated how the PEST (Parameter ESTimation Tool) optimisation software could be used in conjunction with Source Catchments to examine parameter sensitivity and prediction uncertainty related to flow and pollutant loads. As a result PEST is used in the calibration process for the hydrological modelling, and this enables further work to be undertaken in subsequent years to provide an estimate of the associated uncertainty in pollutant loads. In addition, a Reef M&M funded project examining all aspects of uncertainty in model development will commence in 2011 looking at all aspects of model uncertainty.

Preliminary Assessment

Previous modelling used the SedNet/ANNEX model (Cogle et al, 2006; Brodie et al, 2003; Prosser et al, 2001) to generate long term average annual sediment and nutrient loads leaving GBR catchments, and run management practice scenarios to predict sediment and nutrient loads. In contrast, Ellis et al (2009) used the Source Catchments model (formerly known as WaterCAST) in the Fitzroy catchment to predict the flow and load of constituents at any location in the catchment over time, with the ability produce reports at varying temporal scales (from daily to annual) and spatial scales (from a single sub-catchment to a whole-of-catchment). The Source Catchments model has also been applied in the Barron, Burdekin, and Pioneer reef catchments (Carroll et al., 2010). Hence, there is a history of modelling and monitoring within the GBR region which has formed the basis for preliminary assessment and model validation. As part of a process of continuous improvement it is intended to further refine the first

round of Reef M&M Source Catchment modelling through access to new and improved data sets.
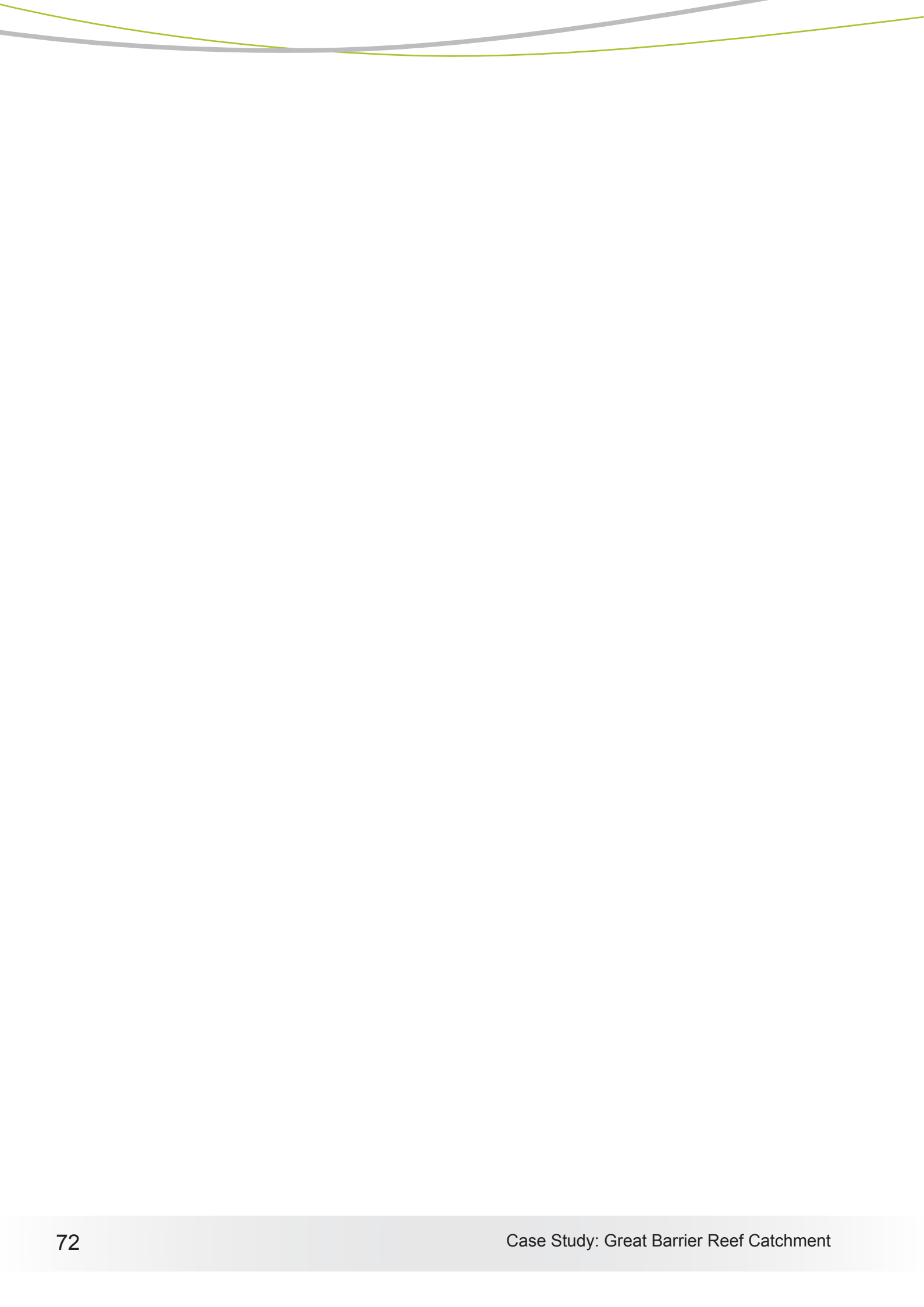
Modelling Methodology

This section outlines the model methodology used; major features of the model configuration are summarised in *Table 2*.

Model Feature	Approach	Rationale	Limitations
Model output timestep	Average annual	Whilst Source Catchments generates runoff and constituent loads at a daily timestep, for reporting purposes, modelled pollutant loads are only required as average annual. Secondly a number of the sediment and nutrient generation algorithms incorporated into the model are designed to be used as average annual outputs, with disaggregation of these processes still rudimentary.	Construction of the model to estimate average annual loads, limits the use of the model to examine modelled outputs at a finer temporal resolution.
Modelling extent	The GBR area (420,000 km2) was split into 6 modelling areas	The 6 regions were based on the regional NRM group boundaries. Consolidating to 6 models made model construction manageable (albeit lengthy), ensured modelled areas aligned with regional NRM body reporting areas and enabled direct comparison of modelled loads to previous modelling in the region.	Aligning to NRM regions resulted in models being constructed at quite different spatial extents (9,000 km2 – 140,000 km2). This created some run time issues for the larger catchments with many subcatchment – land use combinations.
Modelling period	1986 – 2009	A fixed climate period was chosen from 1986 – 2009. This 24 year period included both wet and dry periods which are important for hydrology calibration. A fixed climate period was chosen for all modelling scenarios to normalise for climatic influences. In addition the bare ground index satellite imagery (used to derive yearly cover layers used in the SedNet erosion model) was available from 1986 onwards.	Running the model over a much longer period may have given a better representation of the long term variability in loads
Functional units or landuse category selection	9 – 11 landuse categories were represented in the model	Nine common landuse categories were used across all regions with two additional local categories where required. The land use categories were chosen based on two criteria (a) They were required for reporting purposes and or (b) they were the dominant land use by area.	A number of the landuse categories included for reporting purposes may not have been included normally if the criteria were based purely on their relative contribution to end of catchment loads. The inclusion of 9 – 11 land use categories had a significant impact on run time for the bigger catchments.
Constituents modelled	Fine and coarse sediments, nitrogen and phosphorus species, TN DIN, DON, TP, DIP, DOP, 8 frequently detected herbicides	For reporting purposes, suspended sediment, 'coarse' sediment, particulate and dissolved nutrients and 8 pesticides were required to be modelled. Fine and coarse sediments and speciated nutrients were required for hillslope, gully, stream bank and floodplain erosion/deposition/ entrainment processes to be represented in the model.	The inclusion of 17 constituents impacted on model run time. Model validation was also extremely time consuming given this number of constituents.

Model Feature	Approach	Rationale	Limitations
Minimum stream threshold	30 – 50 km2	This stream threshold (which drives automated sub catchment creation) was chosen primarily to allow use of the SedNet stream bank erosion algorithm and to correspond to the scale at which a number of the erosion algorithms are applicable. The fine resolution enables reporting of modelled outputs at numerous scales. It is also easy to aggregate up from a finer scale.	A number of the validation data sets were not available at this finer scale. The small subcatchment threshold significantly increased model run time in the larger catchments.
Rainfall runoff model selection	Simhyd	Previously applied in numerous reef catchments successfully. Only requires calibration of 6 parameters which rationalise the number of rainfall-runoff parameters that needed to be optimised in calibration. Future work will include investigation of a number of alternative rainfall runoff models available in Source Catchments to look at improving the hydrology calibration.	Simhyd uses only 6 parameters to represent the runoff generation process. The small number of parameters used in the model has highlighted its limitations in representing ground water losses.
Hydrology calibration	PEST parameter estimation software	The inbuilt Source Catchments calibration tool was still under development at the time of model building. PEST had been previously used for model calibration in the Fitzroy catchment. PEST enabled calibration of all identified hydrology parameters in relation to model outputs. Thus, the parameter sets derived were immediately ready for use in the Source Catchments project.	Simultaneous calibration requiring many model runs was time consuming. Run times were in the order of days for the largest projects. These will improve as computer speed improves.
Input climate data	SILO 5km2 gridded daily data	Enabled a consistent approach to climate data collation and model input, easily updated and is repeatable into the future. Pluviometer data only available in selected locations and highly variable temporal spread.	Interpolated grid data will not always capture the spatial and temporal rainfall variability due to network coverage limitations, eg where there are significant rainfall gradients over small areas or orographic effects.
Hydrology data	DERM daily gauging station data	Consistent and quality assured flow data for model calibration collected over 10-40 years at each site. Generally 20-70 gauges located in each region. Data quality assessed based on duration of data, quality of rating, percentage of missing data.	Can be errors in rating curves used to derive runoff rates, particular at highly flows need to identify these prior to calibration.
Routing and decay modelling In-stream and storages	Laurenson routing model used for flow routing, exponential decay models used for nutrients, sediment trapping model for storages	Laurenson routing model was chosen as it is widely used, previous modelling in the region had derived routing parameters for this model and it is a relatively simple model to apply. SedNet defined in stream processing of sediments and nutrients have been inbuilt where possible.	It was identified that the process for determining where routing parameters should be applied and the approach for simultaneous calibration with the rainfall runoff parameters is an area requiring further work.

Case Study: Great Barrier Reef Catchment

Model Feature	Approach	Rationale	Limitations
Sediment, nutrient and pesticide generation models	Dynamic SedNet, Howleaky, APSIM	Existing SedNet approaches to constituent generation have been adapted to operate within the Source Catchments framework for sediment and nutrient generation in grazing lands. In other situations simplified Source Catchments models have been employed to provide constituent loads (eg concentration x runoff). In the case of intensive agricultural land uses (such as 'cropping,' and sugarcane), specialist models such as Howleaky (Rattray et al. 2004a) and APSIM (Keating et al. 2003) have been used to provide daily timeseries pollutant loads for combinations of soil, climate and management practice, with these timeseries applied to the Source Catchments projects through purpose built import tools.	Combining constituent generation (and in stream processing methods) from a vast array of sources and techniques is not easy to implement or manage, but the flexible nature of Source Catchments and its modelling framework, TIME makes this possible.
Modelling of weirs/ storages	Modelling of all storages > 10,000 ML capacity	A minimum capacity of 10,000 ML was chosen as the cut off for storages to be included in the models and this was applied across all regions. This rationalised the number of storages included in the models while allowing major storage impacts to be accounted for.	This assumption may result in an overestimation of modelled runoff in dry periods with low flow. This will be investigated further in year 2.
Modelling of extractions, losses and inflows	IQQM model estimates of extractions and losses were used	Previous DERM IQQM modelling for water resource planning purposes are the most accurate estimate of extractions and losses for major irrigation areas considered in this model.	Time series extraction data sets were derived from IQQM model runs under a full development scenario. In a number of the regulated catchments "full development" may not have been reached during the time period used for model calibration. This may result in an overestimation of runoff in low flow years.
Model size and run times	Automation of model runs	Model runs have been automated to speed up the model running process due to the size of the models.	Automation will speed up the run process but will reduce the time and effort spent to examine each component in detail.
Modelling scenarios	Pre-European, baseline (ie 2008/09), 2009/10, 2010/11 etc	Modelling requirements were to model the anthropogenic load (pre European - baseline) then the relative change from the baseline year each subsequent year.	No clear consensus on a pre- European land use.

Glossary

This Glossary contains definitions of a number of terms used in this guidance which may not be otherwise generally known, but it is not necessarily comprehensive. For a more comprehensive set of definitions, refer to the eWater Glossary, available at: *www.ewater.com.au/glossary/*

Term	Definition
Accuracy*	Closeness of a measured or computed value to its "true" value, where the "true" value is obtained with perfect information. Due to the natural heterogeneity and stochasticity of many environmental systems, this "true" value exists as a distribution rather than a discrete value. In these cases, the "true" value will be a function of spatial and temporal aggregation.
Automated calibration@	Process of calibration using an optimisation procedure with model parameter values constrained to physically defensible ranges to find a set of parameter values that minimises a pre-defined objective function. The set of parameter values found may or may not be unique (see "equifinality"), may represent the global optimum or merely a local optimum on the response surface of model parameter values, and may or may not be robust and fit for purpose (this is also termed "inverse modelling" in some domains, eg groundwater).
Boundaries*	Boundaries specify the area or volume (spatial boundary) and the time period (temporal boundary) to which a model application will apply.
Boundary conditions*	Sets of values for state variables and their rates along problem domain boundaries, sufficient to determine the state of the system within the problem domain.
Calibration	Process of adjusting the values of model parameters within physically defensible ranges until the model performance adequately matches observed historical data from one or more locations represented by the model (ie a match is obtained that is robust and fit for purpose).
Confidence interval	The interval which includes the true value [of a data item, whether observed or estimated] with a prescribed probability and is estimated as a function of the statistics of the sample (WMO, 2008).
Confidence level	The probability that the confidence interval includes the true value (WMO, 2008).

Term	Definition
Credibility@	The confidence that (potential) users have in a model and in the information derived from the model such that they are willing to use the model and the derived information. Specifically, credibility is a function of the performance record of a model and its conformance to best available, practicable science.
Data assimilation	The combining of diverse data, possibly sampled at different times and intervals and different locations, into a unified and consistent description of a physical system, such as the state of the atmosphere (AMS Glossary, 2010).
Data fusion	The use of techniques that combine data from multiple sources and gather that information in order to achieve inferences, which will be more efficient and potentially more accurate than if they were achieved by means of a single source (Source: Wikipedia, *en.wikipedia.org/wiki/Data_fusion*, accessed: 2 Nov. 2010).
Domain boundaries	The limits of space and time that bound a model's domain and are specified within the boundary conditions (see also definition of "boundary conditions").
Equifinality	The principle that in open systems a given end state can be reached by many potential means. In environmental modelling studies, and especially in hydrological modelling, two models are equifinal if they lead to an equally acceptable or behavioural representation of the observed natural processes (Source: Wikipedia, *en.wikipedia.org/wiki/Equifinality*, accessed: 30 June 2010).
Epistemic uncertainty$	Uncertainty due to imperfect knowledge; this form of uncertainty can be reduced by further studies such as research and data collection.
Forecasting	See "Hydrological Forecasting".
Hindcasting@	Modelling of scenarios representing the past, the present or possible future conditions using historical time series data as input. [This term is in common use in other fields such as coastal hydraulics modelling, where the context is exactly the same as here: eg modelling historical ocean wave patterns/regimes for different beach and/or coastal engineering scenarios].
Hydrological forecasting	WMO (2009) lists three categories: a) Short term: periods of up to two days b) Medium range: periods ranging from 2 to 10 days c) Long range: periods exceeding 10 days

Term	Definition
Internal technical guidelines#	Model application guidelines that usually focus on the technical aspects of modelling and are mainly intended for use by modellers. Examples are QA procedures in particular organisations, manuals for software packages which might include hints on how to best use the package, and some text books.
Inverse modelling	See "automated calibration".
Levels of evidence	A ranking system used [in evidence-based medicine] to describe the strength of the results measured in a [clinical trial or] research study (Wikipedia, *en.wikipedia.org/wiki/Levels_of_evidence*, accessed: 3 Nov. 2010).
Manual calibration@	The process of model calibration where parameter values are adjusted manually by trial and error within physically defensible ranges, based on the judgement of the modeller, to obtain a match between model results and observed data that is robust and fit for purpose. The resultant set of parameter values may or may not be unique (see "equifinality") and may or may not represent a local or global optimum.
Model*	A simplification of reality that is constructed to gain insights into selected attributes of a physical, biological, economic, or social system. A formal representation of the behaviour of system processes, often in mathematical or statistical terms. The basis can also be physical or conceptual.
Model application@	Application of a fit for purpose model to address a real world problem such as supporting natural resource management decision making.
Model code@	The mathematical representation of a conceptual model in the form of a functioning computer program.
Model coding*	The process of translating the mathematical equations that constitute the model framework into a functioning computer program.
Model development@	The conceptualisation, specification in mathematical terms, coding, testing and verification of a modelling tool, with the end product intended to be model code which is ready to be implemented and applied to addressing real world problems.
Model implementation@	The setting up, calibration and validation of model code for a particular purpose, with the aim of producing a model that is fit for purpose.
Model uncertainty	Uncertainty related to the hypotheses that underlie the model itself and the model structure (After Grayson and Blöschl, 2001; Section 3.4).

Term	Definition
Multiple lines of evidence	The use of several independent evaluation strategies to address the same evaluation issue, relying on different data sources, on different analytical methods, or on both (Centers for Disease Control and Prevention, 2006).
Non-uniqueness	Groundwater modelling term synonymous with "equifinality" (see earlier definition) (After Grayson and Blöschl, 2001; Chapter 3).
Over-fitting@	Calibrating a model to the point where goodness of fit statistics for the calibration period are maximised/optimised but the calibrated model is not robust when used with input data for other periods.
Parsimony	The principle of parsimony calls for keeping the model as simple as possible while accounting for the system processes and characteristics that are evident in the observations and are important to the predictions, and while respecting all system information – (Hill and Tiedeman, 2007; Chapter 11).
Phenology	The study of periodic plant and animal life cycle events and how these are influenced by seasonal and interannual variations in climate (Wikipedia, *en.wikipedia.org/wiki/Phenology*, accessed:11 Jan. 2011).
Precision*	The quality of being reproducible in amount or performance. With models and other forms of quantitative information, precision refers specifically to the number of decimal places to which a number is computed as a measure of the "preciseness" or "exactness" with which a number is computed.
Predictive modelling@	Applying a model to analyse scenarios representing the past, the present or possible future options.
Predictive uncertainty@	Uncertainty in prediction of hydrological responses associated with: uncertainty in the input data due to sampling or interpolation error; uncertainty in simulated responses due to errors in model parameter values; and uncertainty related to the hypotheses that underlie the model itself and the model structure (After Grayson and Blöschl, 2001; Section 3.4).

Term	Definition
Project management [alternatives from the web via Google, except the last which is made up from the others]	• The process by which projects are defined, planned, monitored, controlled and delivered such that the agreed benefits are realised, or • Planning, monitoring and control of all aspects of a project and the motivation of all those involved in it to achieve the project objectives on time and to the specified cost, quality and performance, or • The discipline of planning, organizing, and managing resources to bring about the successful completion of specific project, or • The application of modern management techniques and systems to the execution of a project from start to finish, to achieve predetermined objectives of scope, quality, time and cost, to the equal satisfaction of those involved, or • The use of skills and knowledge for co-ordinating the organisation, planning, scheduling, directing, controlling, monitoring and evaluating of prescribed activities to ensure that the stated objectives of a project are achieved, or • A process based on use of management and project domain relevant skills and knowledge to organise, plan, schedule, direct, control, monitor, evaluate and deliver a project to achieve the project objectives on time, on budget and to agreed quality and performance levels.
Public interaction guidelines#	Modelling guidelines established through a public consultative and consensus building process, like public technical guidelines (see definition) but differing in that they also give guidance on the interaction between the modeller and the water manager, who often have the roles of consultant and client, respectively.
Public technical guidelines#	Modelling guidelines often containing the same substance as internal technical guidelines (see definition) but differing in the sense that they have been prepared through a consultative and consensus building process involving many persons and organisations.
Reliability*	An expression of the degree to which, and consistency with which, a model meets quantitative performance criteria following calibration, validation testing and other checks, particularly sensitivity/uncertainty analysis. Statistically, reliability is inversely related to random error. Reliability is necessary but not sufficient for validity (Based loosely on: Wikipedia, *en.wikipedia.org/wiki/Reliability_(statistics)* accessed:16 Nov. 2010).

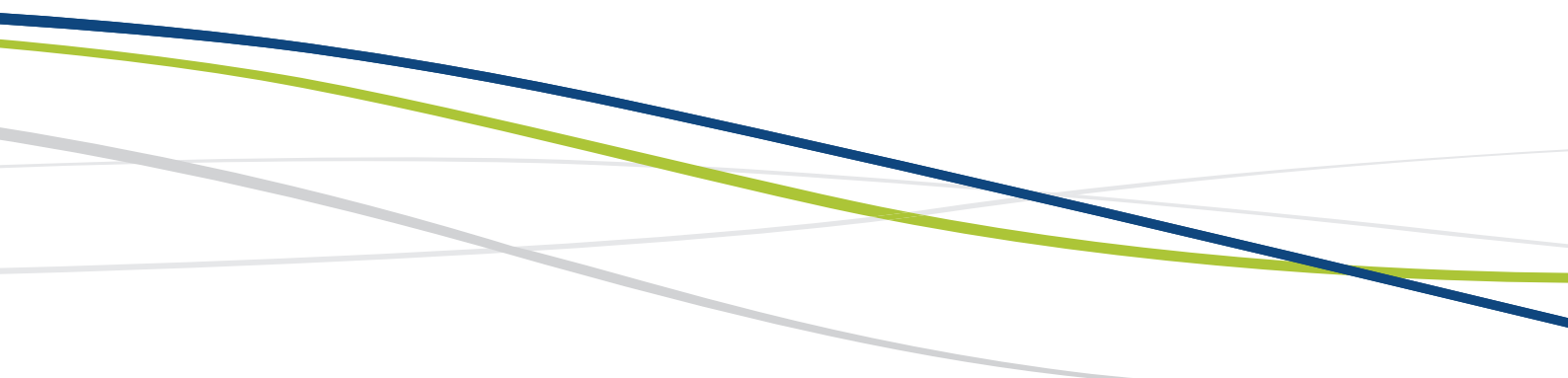| Term	Definition
Robustness*	The capacity of a model to perform well across the full range of environmental conditions for which it was designed.
Scenario@	A set of parameter values describing the characteristics of practices and entities represented by a model; these values may remain constant throughout a model run or they could change (eg when modelling effects of tree growth in a rainfall-runoff model, or a new water source comes on-line part way through a run of a water planning model). Examples of practices include water management rules and irrigator responses to these; examples of entities include catchments, dams and rivers.
Sensitivity*	The degree to which the model outputs are affected by changes in the value of selected input parameters.
Sensitivity analysis	Analysis of the changes in one or more outputs of a model (eg flow or constituent loads) with variations in the assumptions of the model. Most commonly this is achieved by varying the values of one or more model parameters in a systematic way and then reporting the changes in selected key model outputs (Modified from definition in eWater glossary).
Stakeholder@	Any individual, organisation or group with an interest in a project and its outcomes; these can include the organisation commissioning the project (the client), water managers, decision makers, community groups and individual members of the public. An individual or organisation with an interest in the success of a project (Modified from definition in eWater glossary).
State variables*	The dependent variables calculated within the model, which are also often the performance indicators of the models that change over the simulation.
Stochastic uncertainty$	Uncertainty due to natural variability; this form of uncertainty is not reducible (see also "Variability").
Transparency*	The clarity and completeness with which data, assumptions and methods of analysis are documented. Experimental replication is possible when information about modelling processes is properly and adequately communicated.

Term	Definition
Uncertainty* (1)	[A] term used … to describe lack of knowledge about models, parameters, constants, data, and beliefs. There are many sources of uncertainty, including: the science underlying a model, uncertainty in model parameters and input data, observation error, and code uncertainty. Additional study and collecting more information allows error that stems from uncertainty to be minimised/reduced (or eliminated). In contrast, variability (see definition) is irreducible but can be better characterised or represented with further study.
Uncertainty$ (2)	A person is uncertain if s/he lacks confidence about the specific outcomes of an event or action. Reasons for this lack of confidence might include a judgement of the information as incomplete, blurred, inaccurate or potentially false or might reflect intrinsic limits to the deterministic predictability of complex systems or of stochastic processes.
Uncertainty analysis*	Investigates the effects of lack of knowledge or potential errors on the model (eg the "uncertainty" associated with parameter values) and when conducted in combination with sensitivity analysis (see definition) allows a model user to be more informed about the confidence that can be placed in model results.
Validation	Where observations and simulation results are compared using data that were not part of the calibration. A model is validated for a particular application and a successful validation in one example does not imply that the model is validated for universal use. Validation is a test of usefulness and not truth (Modified from definition in eWater glossary).
Variability*	Variability refers to observed differences attributable to true heterogeneity or diversity. Variability is the result of natural random processes and is usually not reducible by further measurement or study (although it can be better characterised).

* USEPA-CREM glossary of frequently used modelling terms
  (available at: *www.epa.gov/crem/index.html*).

@ Drafted for this glossary.

After Refsgaard et al (2005a)

$ After Refsgaard et al (2005b)

7

References

This chapter lists all works referenced in these guidelines. Please note that all URLs are current as at the time of writing; however URLs are subject to change without notice.

AMS Glossary (2010) Definition of data assimilation. American Meteorological Society Glossary of Meteorology.
Available at: amsglossary.allenpress.com/glossary/browse?s=d&p=4 Accessed: 2 Nov. 2010.

AS/NZS (2009) Risk management principles and guidelines. ISO 31000:2009. ISBN 0-7337-9289-8.

Bay-Delta Modeling Forum (2000) Protocols for Water and Environmental Modeling. Ad hoc Modeling Protocols Committee. Available at: cwemf.org

Beven, K.J. (1989) Changing ideas in hydrology - The case of physically-based models. Journal of Hydrology, 105(1-2): 157-172.

Beven, K.J. (1993) Prophecy, reality and uncertainty in distributed hydrological modelling. Advances in Water Resources, 16(1), Research Perspectives in Hydrology: 41-51. ISSN 0309-1708. doi:10.1016/0309-1708(93)90028-E.

Beven, K.J. (1995) Linking parameters across scales: Subgrid parameterizations and scale dependent hydrological models. Hydrological Processes, 9(5-6): 507–525.

Blackmore, J.M., Dandy, G.C., Kuczera, G., Rahman, J. (2009) Making the most of modelling: A decision framework for the water industry. 18th World IMACS / MODSIM Congress, Cairns, Australia. 13-17 July 2009. Available at: mssanz.org.au/modsim09

Blöschl, G. (2006) Hydrologic synthesis: Across processes, places, and scales. Water Resources Research, 42, W03S02. doi:10.1029/2005WR004319.

Brodie, J., Binney, J., Fabricius, K., Gordon, I., Hoegh-Gulderg, O., Hunter, H., O'Reagain, P., Pearson, R., Quirk, M., Thorburn, P., Waterhouse, J., Webster, I. and Wilkinson, S. (2008) Synthesis of evidence to support the Scientific Consensus Statement on Water Quality in the Great Barrier Reef. Available at: www.reefplan.qld.gov.au

Brugnach, M., Pahl-Wostl, C., Lindenschmidt, K.E., Janssen, J.A.E.B., Filatova, T., Mouton, A., Holtz, G., van der Kuer, P. and Gaber, N. (2008) Complexity and uncertainty: rethinking the modelling activity. Chapter 4: 49-68. In: Jakeman, A.J., Voinov, A.A., Rizzoli, A.E. and Chen, S.H. "Environmental Modelling, Software and Decision Support: state of the art and new perspectives". Elsevier. ISBN-13: 978-0-08-056886-7.

Carroll, C., Ellis, R., Hateley, L., Fentie, B., Rohde, K. and Waters, D. (2010) Great Barrier Reef Watercast Application Project: Stage 1. eWater CRC Technical Report.
Available at: www.ewater.com.au

Carroll, C., Waters, D., Searle, R.D., Vardy, S., Turner, R., Silburn, D.M., Shaw, M, Robinson, B., Attard, S., Thorburn, P.J., Armour, J.D., Rohde, K.W., Milla, R., Halpin, N., Davis, A.M., Schmidt, M., Wilson, B. and Clark, A. (2011) A Paddock to Reef Monitoring and Modelling framework for the Great Barrier Reef: Paddock and Catchment. Agric. Eco. and Environ., Special Issue: Catchments and coral reef (in press).

Centers for Disease Control and Prevention (2006) Introduction to program evaluation for public health programs: Evaluating appropriate antibiotic use programs. Office of the Director, Office of Strategy and Innovation, & National Center for Infectious Diseases, Division of Bacterial and Mycotic Diseases. Atlanta, GA, USA. Glossary accessed on 3 Nov. 2010. Available at: www.cdc.gov/getsmart/program-planner/Glossary-Eval-Res.html

Cogle, A.L., Carroll, C., Sherman, B.S. (2006) The use of SedNet and ANNEX models to guide GBR catchment sediment and nutrient target setting. Department of Natural Resources, Mines and Water. Report no. QNRM06138.

Collette, Y. and Siarry, P (2003) Multi Objective Optimization – Principles and Case Studies. Springer-Verlag, Berlin. ISBN 3-540-40182-2.

Cotsell, P., Gale, K., Hajkowicz, S.J., Lesslie R., Marshall, N. and Randall, L. (2009) Use of a multiple criteria analysis (MCA) process to inform Reef Rescue regional allocations. Proc. 2009 Marine and Tropical Sciences Research Facility Annual Conference, 28-30 April, Townsville. Reef and Rainforest Research Centre Ltd, Cairns.

CRC for Catchment Hydrology (2005) Series on model choice. 2: Water quality models – sediment and nutrients. February. Available at: www.toolkit.net.au/modelchoice

CRC for Catchment Hydrology (undated) Series on model choice. 1: General approaches to modelling and practical issues of model choice. Available at: www.toolkit.net.au/modelchoice.

CREM (2008) Guidance on the Development, Evaluation and Application of Environmental Models. Council for Regulatory Environmental Modeling, Office of the Science Advisor, U.S. Environmental Protection Agency. Draft. August. Available at: www.epa.gov/crem

Croke, B.F.W. and Jakeman, A.J. (2001) Predictions in catchment hydrology: an Australian perspective. Marine and Freshwater Research, 52: 65–79.

Crout, N., Kokkonnen, T., Jakeman, A.J., Norton, J.P., Newham, L.T.H., Anderson, R., Assaf, H., Croke, B.F.H., Gaber, N., Gibbons, J., Holzworth, D., Mysiak, J., Reichl, J., Seppelt, R., Wagener, T. and Whitfield, P. (2008) Good modelling practice. Chapter 2: 15-31. In: Jakeman, A.J., Voinov, A.A., Rizzoli, A.E. and Chen, S.H. "Environmental Modelling, Software and Decision Support: state of the art and new perspectives". Elsevier. ISBN-13: 978-0-08-056886-7.

Doherty, J. (2009) PEST: Model-Independent Parameter Estimation. Watermark Numerical Computing, Brisbane, Australia. Available at: www.sspa.com/software/pest.html

Ellis, R.D., Doherty, J., Searle, R.D. and Moodie, K. (2009) Applying PEST (Parameter ESTimation) to improve parameter estimation and uncertainty analysis in Watercast model. 18th World IMACS/MODSIM Congress Cairns, Australia. 13-17 July.
Available at: mssanz.org.au/modsim09

eWater (2010) Best Practice Model Development and Use in eWater CRC – Overview and Policy. Version 1. May.

Grayson, R.B. and Blöschl, G. (2001) Spatial Patterns in Catchment Hydrology: Observations and Modelling. Cambridge University Press. 404pp. ISBN 0-521-63316-8.
Available at: www.catchment.crc.org.au/special_publications1.html

Hairsine, P.B. and Sander, G.C. (2009) Comment on 'A transport-distance based approach to scaling erosion rates': Parts 1, 2 and 3 by Wainwright et al. Earth Surface Processes and Landforms, 34: 882–885. doi:10.1002/esp.1782.

Hajkowicz, S.J. (2005) Techniques for supporting natural resource investment decisions. A Consultancy Report for the Swan River Trust. 33 pp. 22 November.

Hajkowicz, S.J. and Collins, K. (2007) A Review of Multiple Criteria Analysis for Water Resource Planning and Management. Water Resources Management, 21: 1553–1566. doi:10.1007/s11269-006-9112-5.

Hill, M.C. and Tiedeman, C.R. (2007) Effective groundwater model calibration: with analysis of data, sensitivities, predictions and uncertainty. John Wiley & Sons, Inc., Hoboken, NJ, USA; 455 pp. ISBN-13: 978-0-471-77636-9.

Jakeman, A.J., Letcher, R.A. and Norton, J.P. (2006) Ten iterative steps in development and evaluation of environmental models. Position Paper. Environmental Modelling & Software, 21: 602-614. doi:10.1016/j.envsoft.2006.01.004.

Jefferson Center (2004) Citizens Jury Process. Available at: www.jefferson-center.org/ Accessed 28 April 2010.

Keating, B.A., Carberry, P.S., Hammer, G.L., Probert, M.E., Robertson, M.J., Holzworth, D., Huth, N.I., Hargreaves, J.N.G., Meinke, H., Hochman, Z., McLean, G., Verburg, K., Snow, V., Dimes, J.P., Silburn, M., Wang, E., Brown, S., Bristow, K.L., Asseng, S., Chapman, S., McCown, R.L., Freebairn, D.M. and Smith, C.J., (2003) An overview of APSIM, a model designed for farming systems simulation. European Journal of Agronomy, 18: 267-288.

Kirchner, J.W. (2006) Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology. Water Resources Research, 42, W03S04. doi:10.1029/2005WR004362.

Kundzewicz, Z.W. & Robson, A.J. (2004) Change detection in river flow records— review of methodology. Hydrological Sciences Journal, 49(1): 7–19. Available at: www.iahs.info/hsj/hsjindex.htm

Loucks, D.P. and van Beek, E. (2005) Water resources systems planning and management: an introduction to methods, models and applications. UNESCO, Paris and WL | Delft Hydraulics, The Netherlands. 680 pp. ISBN 92-3-103998-9. Available at: ecommons.library.cornell.edu/handle/1813/2799

Maier, H.R., Ascough II, J.C., Wattenbach, M., Renschler, C.S., Labiosa, W.B., and Ravalico, J.K. (2008) Uncertainty in environmental decision making: issues, challenges and future directions. Chapter 5: 69-86. In: Jakeman, A.J., Voinov, A.A., Rizzoli, A.E. and Chen, S.H. "Environmental Modelling, Software and Decision Support: state of the art and new perspectives". Elsevier. ISBN-13: 978-0-08-056886-7.

Matott, L. S., Babendreier, J. E. and Purucker, S. T. (2009) Evaluating uncertainty in integrated environmental models: A review of concepts and tools. Water Resources Research, 45, W06421. doi:10.1029/2008WR007301. Available at: www.epa.gov/athens/research/modeling/modelevaluation/index.html

Middlemis, H., Merrick, N., and Ross, J.B. (2000) Groundwater Flow Modelling Guideline. Prepared for Murray-Darling Basin Commission by Aquaterra Consulting Pty Ltd. Project no. 125. Available at: www.mdbc.gov.au

Milly, P.C.D., Betancourt, J., Falkenmark, M., Hirsch, R.M., Kundzewicz, Z. W., Lettenmeier, D.P. and Stouffer, R.J. (2008) Stationarity is dead: whither water management? Science, 319: 573-574. 1 February.

Montanari, A. (2007) What do we mean by uncertainty? The need for a consistent wording about uncertainty assessment in hydrology. Hydrological Processes. 21(6): 841–845. doi:10.1002/hyp.6623.

Myšiak, J., Brown, J.D., Jansen, S.M.L., and Quinn, N.W.T. (2008) Environmental policy aid under uncertainty. Chapter 6: 87-100. In: Jakeman, A.J., Voinov, A.A., Rizzoli, A.E. and Chen, S.H. "Environmental Modelling, Software and Decision Support: state of the art and new perspectives". Elsevier. ISBN-13: 978-0-08-056886-7.

Office of Government Commerce (2009) Managing Successful Projects with PRINCE2. 5th ed, 327 pp. The Stationery Office on behalf of Office of Government Commerce, London, UK. ISBN 9780113310593.

Oreskes, N., Shrader-Frechette, K. and Belitz, K. (1994) Verification, validation and confirmation of numerical models in the earth sciences. Science, 263: 641–6. 4 February.

Packman, J. and Old, G. (2005) European Quality Assurance Standards in River Basin Modelling Practice. HarmoniQuA project. Contract no. EVK1-CT2001-00097. December. 18 pp. Available at: harmoniqua.wau.nl/public/Reports/Standards/EQS-final.pdf

Pappenberger, F. and Beven, K.J. (2006) Ignorance is bliss: Or seven reasons not to use uncertainty analysis. Water Resources Research, 42, W05302. doi:10.1029/2005WR004820.

Perrin, C., Michel, C. and Andréassian, V. (2001) Does a large number of parameters enhance model performance? Comparative assessment of common catchment model structures on 429 catchments. Journal of Hydrology, 242(3-4): 275-301.

Podger, G.M., Barma, D., Neal, B., Austin, K. and Murrihy, E. (2010b) River System Modelling for the Basin Plan - Assessment of fitness for purpose. CSIRO: Water for a Healthy Country National Research Flagship, Canberra, Australia. December. 64 pp.
Available at: www.mdba.gov.au/files/bp-kid/1534-Final-Fitness-for-purpose.pdf

Podger, G.M., Yang, A., Brown, A., Teng, J., Power, R. and Seaton, S. (2010a) Proposed River Modelling Methods and Integrated River System Modelling Framework Design for use in Basin Plan Modelling. CSIRO: Water for a Healthy Country National Research Flagship, Canberra, Australia. July. 100pp.
Available at: www.mdba.gov.au/files/bp-kid/1527-Proposed-river-modelling-methods-CSIRO.pdf

Project Management Institute (2008) Guide to the Project Management Body of Knowledge. 4th ed, 459 pp. Newtown Square, Pa, USA. ISBN 1933890517.

References

Queensland Department of the Premier and Cabinet (2009) Reef Water Quality Protection Plan 2009 – for the Great Barrier Reef World Heritage Area and adjacent catchments. Published by the Reef Water Quality Protection Plan Secretariat. September. 100 George Street, Brisbane Qld, 4000. Available at: www. reefplan.qld.gov.au

Radziejewski, M. & Kundzewicz, Z. W. (2004) Detectability of changes in hydrological records. Hydrological Sciences Journal. 49(1): 39–51. Available at: www.iahs.info/ hsj/hsjindex.htm

Rattray, D.J., Freebairn, D.M., McClymont, D., Silburn, D.M., Owens, J. and Robinson, B. (2004) HOWLEAKY? The journey to demystifying 'simple' technology. Paper 422. In: Conserving Soil and Water for Society: Sharing Solutions. Proc. 13th International Soil Conservation Organisation Conference, Brisbane, July (Eds S.R. Raine, A.J.W. Biggs, N.W. Menzies, D.M. Freebairn, P.E. Tolmie). ASSSI/IECA: Brisbane.

Refsgaard J.C. (2001) Towards a formal approach to calibration and validation of models using spatial data. In: Grayson R and Blöschl G (Eds) Spatial Patterns in Catchment Hydrology: Observations and Modelling. Cambridge University Press: 329-354.

Refsgaard, J.C. (ed) (2002) State-of-the-Art Report on Quality Assurance in modelling related to river basin management. HarmoniQuA project. Contract no. EVK2-CT2001-00097. October.

Refsgaard, J.C. and Henriksen, H.J. (2004) Modelling guidelines—terminology and guiding principles. Advances in Water Resources, 27(1): 71–82. doi:10.1016/j. advwatres.2003.08.006.

Refsgaard, J.C., Henriksen, H.J., Harrar, W.G., Scholten, H., Kassahun, A. (2005a) Quality assurance in model based water management – review of existing practice and outline of new approaches. Environmental Modelling & Software, 20: 1201–1215. doi:10.1016/j.envsoft.2004.07.006.

Refsgaard, J.C., Troldborg, L., Henriksen, H.J., Højberg, A.L., Møller, R.R. and Nielsen, A.M. (2010) God praksis i hydrologisk modellering. Geo-guide 7. Geological Survey of Denmark and Greenland. 83pp. ISBN print: 978-87-7871-280-6. ISBN web: 978-87-7871-281-3. In Danish.

Refsgaard, J.C., van der Sluijs, J.P., Brown, J., van der Keur, P. (2006) A framework for dealing with uncertainty due to model structure error. Advances in Water Resources, 29(11): 1586–1597. doi:10.1016/j.advwatres.2005.11.013
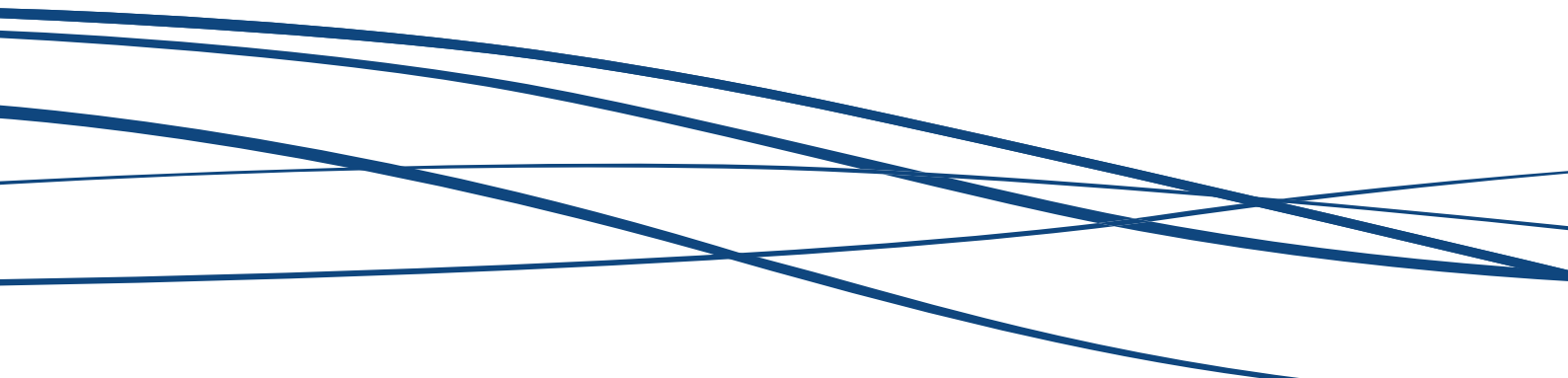
Refsgaard, J.C., van der Sluijs, J.P., Højberg, A.J., and Vanrolleghem, P.A. (2005b) Harmoni-CA Guidance: Uncertainty Analysis. Harmoni-CA Guidance 1. European Commission Contract no: EVK1-2001-00192. Available at: www.harmoni-ca.info/toolbox/index.php

Refsgaard, J.C., van der Sluijs, J.P., Højberg, A.J., and Vanrolleghem, P.A. (2007) Uncertainty in the environmental modelling process - A framework and guidance. Environmental Modelling & Software, 22(11): 1543-1556. November. doi:10.1016/j.envsoft.2007.02.004.

Robson, B.J., Hamilton, D.P., Webster, I.T. and Chan, T. (2008) Ten steps applied to development and evaluation of process-based biogeochemical models of estuaries. Environmental Modelling & Software, 23: 369-384.

Scholten, H., Kassahun, A., Refsgaard, J.C, Kargas, T., Gavardinas, C., Beulens, A.J.M. (2007) A methodology to support multidisciplinary model-based water management. Environmental Modelling & Software, 22: 743-759.

Silberstein, R.P. (2006) Hydrological models are so good, do we still need data? Environmental Modelling and Software, 21(9): 1340–1352. doi:10.1016/j.envsoft.2005.04.019.

Simonovic, S.P. (2009) Managing Water Resources: methods and tools for a systems approach. UNESCO, Paris and Earthscan, London. ISBN 978-92-3-104078-8.

Son, K., and Sivapalan, M. (2007) Improving model structure and reducing parameter uncertainty in conceptual water balance models through the use of auxiliary data. Water Resources Research, 43, W01415, doi:10.1029/2006WR005032.

Soncini-Sessa, R., Castelletti, A. and Weber, E. (2007) Integrated and Participatory Water Resources Management – Theory. Developments in Integrated Environmental Assessment Series (ed: Jakeman, A.J.), Elsevier, Amsterdam, The Netherlands. ISBN-13: 978-0-444-53013-4.

Stedinger, J.R. and Taylor, M.R. (1982) Synthetic Streamflow Generation: 2. Effect of Parameter Uncertainty. Water Resources Research, 18(4): 919-924, August. doi:10.1029/WR018i004p00919.

Stedinger, J.R., Pei, D. and Cohn, P.A. (1985) A Condensed Disaggregation Model for Incorporating Parameter Uncertainty into Monthly Reservoir Simulations. Water Resources Research, 21(5): 665-675, May. doi:10.1029/WR021i005p00665.

USEPA (2002) Guidance for Quality Assurance Project Plans for Modeling. EPA QA/G-5M. December. Available at: www.epa.gov/quality

Van Dijk, A.I.J.M., Kirby, J.M., Paydar, Z., Podger, G.M., Mainuddin, Md, Marvanek, S. and Peña Arancibia, J. (2008) Uncertainty in river modelling across the Murray-Darling Basin. A report to the Australian Government from the CSIRO Murray-Darling Basin Sustainable Yields Project, CSIRO Australia. October. 98 pp. Available at: www.csiro.au/resources/RiverModellingUncertaintyMDBSY.html

Van Waveren, R.H., Groot, S., Scholten, H., Van Geer, F.C., Wosten, J.H.M., Koeze, R.D., Noort, J.J., (2000) Good Modelling Practice Handbook, STOWA Report 99-05, Utrecht, RWS-RIZA, Lelystad, The Netherlands. Available at: harmoniqua.wau.nl/public/Reports/Existing%20Guidelines/GMP111. pdf (in English).

Vaze, J., Chiew, F.H.S., Perraud, J.M., Viney, N., Post, D.A., Teng, J., Wang, B., Lerat, J., Goswami, M. (2011) Rainfall-runoff modelling across southeast Australia: datasets, models and results. Australian Journal of Water Resources. February.

Vàzquez, J.F. and Rosato, P. (2006) Multi-Criteria Decision Making in Water Resources Management. In: Giupponi, C., Jakeman, A.J., Karssenberg, D., and Hare, M.P. (eds) (2006) Sustainable Management of Water Resources – an Integrated Approach. Fondazione Eni Enrico Mattei Series on Economics, the Environment and Sustainable Development. Edward Elgar Publishing Ltd, Cheltenham, UK. ISBN-10: 1-84542-745-9.

Water New Zealand (2009) National Modelling Guidelines for Water Distribution Network Modelling. Modelling Special Interest Group. Draft Version 01 Revision 04 – April. Available at: www.waternz.org.nz

Westra, S., Varley, I., Jordan, P., Nathan, R., Ladson, A., Sharma, A. and Hill, P. (2010) Addressing climatic non-stationarity in the assessment of flood risk. Australian Journal of Water Resources, 14(1): 1-16. Paper W09-796. Institution of Engineers, Australia.

WMO (2008) Guide to Hydrological Practices. Volume I: Hydrology – From Measurement to Hydrological Information. WMO No. 168. 6th ed. World Meteorological Organisation, Geneva, Switzerland. ISBN 978-92-63-10168-6.

WMO (2009) Guide to Hydrological Practices. Volume II: Management of Water Resources and Application of Hydrological Practices. WMO No. 168. 6th ed. World Meteorological Organisation, Geneva, Switzerland. ISBN 978-92-63-10168-6.

Young, W., Brandis, K. and Kingsford, R. (2006) Modelling monthly streamflows in two Australian dryland rivers: Matching model complexity to spatial scale and data availability. Journal of Hydrology, 331, 242– 256. doi:10.1016/j.jhydrol.2006.05.014.

Yue, S. & Pilon, P. (2004) A comparison of the power of the t-test, Mann-Kendall and bootstrap tests for trend-detection. Hydrological Sciences Journal. 49(1), 21–37. Available at: www.iahs.info/hsj/hsjindex.htm