<u>**The Winston Churchill Memorial Trust of Australia**</u>

Report by

Hugh Middlemis

2004 Churchill Fellow

BENCHMARKING BEST PRACTICE FOR GROUNDWATER FLOW MODELLING

INDEX Page

1 INTRODUCTION 1
1.1 PRECIS 1
1.2 ACKNOWLEDGEMENTS 1
2 FELLOWSHIP PROGRAMME SUMMARY 2
2.1 CZECH REPUBLIC - SEPTEMBER 13 TO 17, 2004 2
2.2 ENGLAND - SEPTEMBER 20 TO 24, 2004 2
2.3 NETHERLANDS - SEPTEMBER 27 TO OCTOBER 1, 2004 2
2.4 GERMANY - OCTOBER 4 TO 8, 2004 2
2.5 ENGLAND - OCTOBER 11 TO 20, 2004 2
3 BENCHMARKING BEST PRACTICE MODELLING GUIDELINES 3
3.1 BACKGROUND ON GROUNDWATER MODELLING 3
3.2 REVIEW OF NOTABLE GUIDELINES 4
3.2.1 European Water Policy Setting 4
3.2.2 HARMONIQUA and MoST 6
3.2.3 Dutch Good Modelling Practice 8
3.2.4 UK Guidance Notes 9
3.2.5 German Guidelines 11
3.3 MDBC GROUNDWATER FLOW MODELLING GUIDELINE 11
3.3.1 Background 11
3.3.2 Need for an Update? 12
3.3.3 Model Review 12
3.3.4 Parameter Optimisation and Uncertainty Assessment Methods 13
3.3.5 Leading Modelling Codes 13
4 INTEGRATED MODELLING 14
4.1 ISSUE 14
4.2 INTEGRATED MODEL SELECTION CRITERIA 14
4.3 IQQM-MODFLOW 17
4.4 ZOOMQ3D 17
4.4.1 Local Grid Refinement 18

4.4.2 Other ZoomQ3D Functionality Needs 19

5 CONCLUSIONS 20

6 RECOMMENDATIONS 20

7 BIBLIOGRAPHY 21

FIGURES

1 Ongoing Process of Model Refinement

2 Structured modelling process steps in the HarmoniQuA knowledge base

APPENDICES

A Brief Overview of Selected Groundwater Modeling Guidelines. FEM-Modflow Conference paper by Hill et al.

B Critical Review of Modelling Support Tool (MoST)

1 INTRODUCTION

1.1 PRECIS

This report outlines the programme of investigation and the findings from a 2004 Churchill Fellowship study tour to England, the Czech Republic, the Netherlands and Germany in September and October 2004. The aim was to benchmark the Murray-Darling Basin Commission (MDBC) Groundwater Flow Modelling Guidelines against international best practice, and to identify whether there are areas where improvements may be required for adaptation to Australian conditions. It is important to note that the Guidelines detail best practice methodologies for devising models as practical decision support tools for natural resources management, rather than detailing advanced analysis techniques that might be applied in highly complex projects.

Specific areas of study included:

- Invited speaker at an international conference in the Czech Republic, and member of the panel for a Special Focus Session to workshop modelling protocols, which also provided the opportunity to interact with leading modellers from the USA

- Attendance at two training courses on modelling software packages (Feflow and ZoomQ3D) Review of interim outputs (papers and software) from the HarmoniQuA project funded by the European Union (EU), which is developing a knowledgebase and software tool for multi-discipline model quality assurance, project management and communication

- Review of German groundwater modelling protocols

- Review of UK groundwater modelling protocols, and of their approaches to integrated surface and groundwater modelling, and use in catchment abstraction management and decision support systems.

This study has found that the existing MDBC Groundwater Flow Modelling Guidelines are fundamentally still relevant and fit for their intended purpose as a best practice guide. Further, the MDBC guide has influenced the development of best practice guidance internationally, including the next generation of combined quality assurance (QA) and guidance software, the Modelling Support Tool (MoST) from the European HarmoniQuA project.

It is considered likely that MoST will be widely adopted as it becomes recognised as a useful tool for improved QA, project management and communication, but it is currently not configured with adequate guidance content. As MoST is likely to become a widely adopted project management tool, it's guidance content should be improved by including more of the information from the MDBC, UK, German and other selected guidelines, which have been proven to be effective on practical projects.

1.2 ACKNOWLEDGEMENTS

I am very grateful for the financial and other support provided by the Winston Churchill Memorial Trust, which made this project possible. Many thanks also to specific support from Noel Merrick (University of Technology, Sydney), Rebecca Knol (Churchill Fellow, 1998), and Scott Keyworth (Murray-Darling Basin Commission) in helping to make my Churchill Fellowship a success.

I also acknowledge the efforts and support of colleagues and friends (old and new) who have taken an interest in this endeavour, particularly in the modelling community. My colleagues and management at Aquaterra Consulting have also been very supportive and indulgent.

By no means least, I feel humbled by the whole-hearted support from my family, Cas and Roger.

2 FELLOWSHIP PROGRAMME SUMMARY

2.1 CZECH REPUBLIC - SEPTEMBER 13 TO 17, 2004

- FEM-Modflow: International Conference "Finite-Element Models, MODFLOW and More 2004", Carlsbad (http://www.natur.cuni.cz/utf8/fem_modflow). Panel member on Special Focus Session "*Systematic approach to groundwater modelling: formalized frameworks, guidelines, and structural quality assurance*", convened to explore selected sets of guidelines for model development published in recent years.
- Attended a training course on FEFLOW (finite element groundwater modelling software).

2.2 ENGLAND - SEPTEMBER 20 TO 24, 2004

- Attended the inaugural training course on ZoomQ3D software (object-oriented, groundwater modelling software, with localised finite difference grid refinement) at the British Geological Survey. The ZOOM family of models represents one of the first applications of object-oriented groundwater modelling in the world, where the data is grid-independent, which is a key need to advance groundwater modelling software through better integration of models with GIS methods.

2.3 NETHERLANDS - SEPTEMBER 27 TO OCTOBER 1, 2004

- Wageningen University: Discussions with the HarmoniQuA project team (http://www.harmoniqua.org/, Huub Scholten, Co-ordinator of HarmoniQuA project, and Ayalew Kassahun, Knowledgebase specialist). Topics of discussion included the functionality of the Modelling Support Tool (MoST), comprising a knowledgebase and associated software for multi-discipline model quality assurance, project management and communication. Facilitated contact from Joost Herweijer (Reservoir Team, Adelaide) on similar knowledgebase systems developed in Australia.

2.4 GERMANY - OCTOBER 4 TO 8, 2004

- Stuttgart University: Discussions with Dr Johannes Riegger (Institute for Water Engineering), chair of the modelling guidelines working group for the Hydrogeology Section of the German Geological Society (FH-DGG www.fh-dgg/ak-hgm). Reviewed the German guide and provided some feedback on the English translation, and discussed the functionality of MoST. Facilitated interaction between the HarmoniQuA project and FH-DGG.

2.5 ENGLAND - OCTOBER 11 TO 20, 2004

- UK Environment Agency: Discussions with Paul Hulme (Senior Modeller, Science Group, and editor of the Groundwater Resources Modelling Guidance Notes), Stuart Kirk (Groundwater Advisor, EU Water Framework Directive), David Johnson (Principal Scientist) and Fenella Brown. Topics of discussion included modelling guides and integrated surface-groundwater (especially wetland) modelling approaches implemented for their Resource Assessment Management (RAM) Framework and Catchment Abstraction Management Strategies (CAMS: http://www.environment-agency.gov.uk/cams). Facilitated contact from Joost Herweijer (Reservoir Team, Adelaide) on knowledgebase and workflow process software systems developed in Australia.

- British Geological Survey: Discussions on the ZoomQ3D modelling package (www.bgs.ac.uk) with Andrew Hughes, Chris Jackson, Majdi Mansour, Ann Williams and Ilka Neumann, BGS staff on the Groundwater Systems and Water Quality Programme. Discussed the need for upgrades such as fully integrated (close-coupled) surface and groundwater simulation, with local-scale vertical grid refinement. Discussed model archiving and other protocols.

3 BENCHMARKING BEST PRACTICE MODELLING GUIDELINES

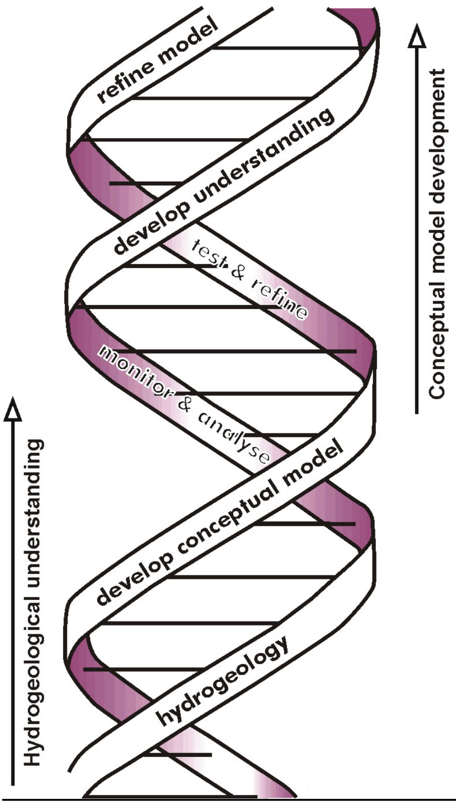
3.1 BACKGROUND ON GROUNDWATER MODELLING

The following introductory comments on groundwater modelling are taken from the MDBC Groundwater Flow Modelling Guideline (2001), of which I was the principal author:

> "*A groundwater model is a computer-based representation of the essential features of a natural hydrogeological system that uses the laws of science and mathematics. Its two key components are a conceptual model and a mathematical model. The conceptual model is an idealised representation (ie. a picture) of our hydrogeological understanding of the key flow processes of the system. A mathematical model is a set of equations, which, subject to certain assumptions, quantifies the physical processes active in the aquifer system(s) being modelled. While the model itself obviously lacks the detailed reality of the groundwater system, the behaviour of a valid model approximates that of the aquifer(s). A groundwater model provides a scientific means to draw together the available data into a numerical characterisation of a groundwater system. The model represents the groundwater system to an adequate level of detail, and provides a predictive scientific tool to quantify the impacts on the system of specified hydrological, pumping or irrigation stresses.*"

> "*Hydrogeological investigations and groundwater modelling are dynamic and inexact sciences. They are dynamic in the sense that the state of any hydrological system is changing with time, and in the sense that we are continually developing new scientific techniques to evaluate these systems. They are inexact in the sense that groundwater systems are complicated beyond our capability to evaluate them comprehensively in detail, and we invariably do not have sufficient data to do so (even if we had the capability). The ability of the data to provide an increasingly accurate representation (model) of the groundwater system increases with time, money, and the technical expertise applied. The study scope and objective needs to be balanced against the budget, time and data resources available, to develop an appropriate modelling study approach.*"

> "*The guide should be seen as a best practice reference point for framing modelling projects, assessing model performance, and providing clients with the ability to manage contracts and understand the strengths and limitations of models across a wide range of studies (scopes, objective, budgets) at various scales and in various hydrogeological settings. The intention is not to provide a prescriptive step-by-step guidance, as the site-specific nature of each modelling study renders this impossible, but to provide overall guidance and to help make the reader aware of the complexities of models, and how they may be managed.*"



The underpinning philosophy of best practice model methodologies can be summarised by the continuous improvement / quality assurance cycle of "***plan-do-check-act***". In generic modelling terms, this can be described as "***conceptualise-simulate-review-refine***". Figure 1 uses the double-helix concept (based on an original spiral concept by Hulme, 2003) to illustrate the ongoing process of model refinement. The double helix concept links hydrogeological knowledge with model development, monitoring and evaluation. Thus, developments in hydrogeological knowledge underpin model development, which provides a tool to further develop insights and understanding, and guide monitoring and analysis, which further improves understanding, and so on.

Figure 1 Ongoing Process of Model Refinement

The fundamental guiding principle for best practice modelling is that model development is an ongoing process of refinement from an initially simple representation of the aquifer system to one with an appropriate degree of complexity. Thus, the model realisation at any stage is neither the best nor the last, but simply the latest representation of our developing understanding of the aquifer system.

There are many sets of guidelines devised by various parties around the world for the development of groundwater models. Hill et al (2004), identified and summarised information about selected (key) sets of these guidelines. That paper (Appendix A), which I helped co-author, formed the basis for the guidelines workshop session at the recent FEM-Modflow international conference, which was the first stage of my Churchill Fellowship. The paper outlined differences in terminology, which can be trivial, and yet also have tended to “get in the way” of discussions over the last few decades about appropriate methodologies. However, there is widespread agreement about the basic modelling methodology, which is illustrated in Figure 2. Figure 2 is taken from the HarmoniQuA project (Scholten et al, 2004), but is quite consistent with the methodologies of all the best practice guidelines.

Many discussions of best practice methodologies have foundered on the issues of model calibration and validation, and related terminology. While all agree that calibration is fundamental (eg. Bredehoeft, 2003: “calibration involves fitting the model output to a set of data”), there has been less agreement on whether models can be (universally) validated. I believe that these issues have been clarified by Refsgaard and Henriksen (2004), the groundwater specialists on the HarmoniQuA project team. Their paper is notable in that it provides a scientifically sound and philosophically logical argument in justification of the position that models can be calibrated and validated for a specified “domain of applicability”. Hence, “conditional validation” should be accepted as a pragmatic and valid approach. Another recent paper (Hassan, 2004b) independently considers these same issues and comes to similar conclusions. As a modelling pragmatist, I concur with de Marsily (cited in Refsgaard and Henriksen, 2004) that *we do not seek universal truth from models, simply engineering confidence*. Thus, I believe that conditional validation of models should be accepted as valid in itself.

Bredehoeft (2003) agrees that “good modelling is an iterative process”, and that “our conceptual model changes with advances in science” (eg. see Figure 1), but also that “a model involving a wrong or incomplete conceptual model can be adequately calibrated”. This last observation highlights potential problems with the model calibration/validation process. However, there are indications from my discussions during the Churchill Fellowship that the philosophy of *conditional validation* that underpins the papers by Refsgaard and Henriksen (2004) and Hassan (2004) will be widely accepted (calibration being already universally accepted as a fundamental modelling approach).

The implication is that the fundamentals of the current best practice guides (including the MDBC guide), do not need substantial revision, as they are consistent with Refsgaard and Henriksen (2004), which I believe sets out conclusively the best practice framework.

3.2 REVIEW OF NOTABLE GUIDELINES

The next six sections provide brief background on the policy setting in Europe and provides review comments on notable guidelines that were considered as part of this Churchill Fellowship, from the viewpoint of trying to identify where the MDBC guide could be updated. The implications of this review for updates to the MDBC guide are discussed in the penultimate section, leading to the Conclusions and Recommendations. It is important to recognise the purpose for which the individual model guides were developed, and I have tried to present the review comments from that context.

3.2.1 European Water Policy Setting

Before discussing the European guidelines selected for study during this project, a brief description of the policy setting in terms of the Water Framework Directive and the Habitat Directive is in order. The Water Framework Directive was approved in December 2000, and is regarded as “the most important piece of European water legislation for over 20 years.” (Environment Agency, 2003). It promulgates a holistic or “integrated river basin management” approach to water resources management. The Water Framework Directive establishes a framework for the protection of inland surface waters (including transitions to coastal waters), plus groundwater.

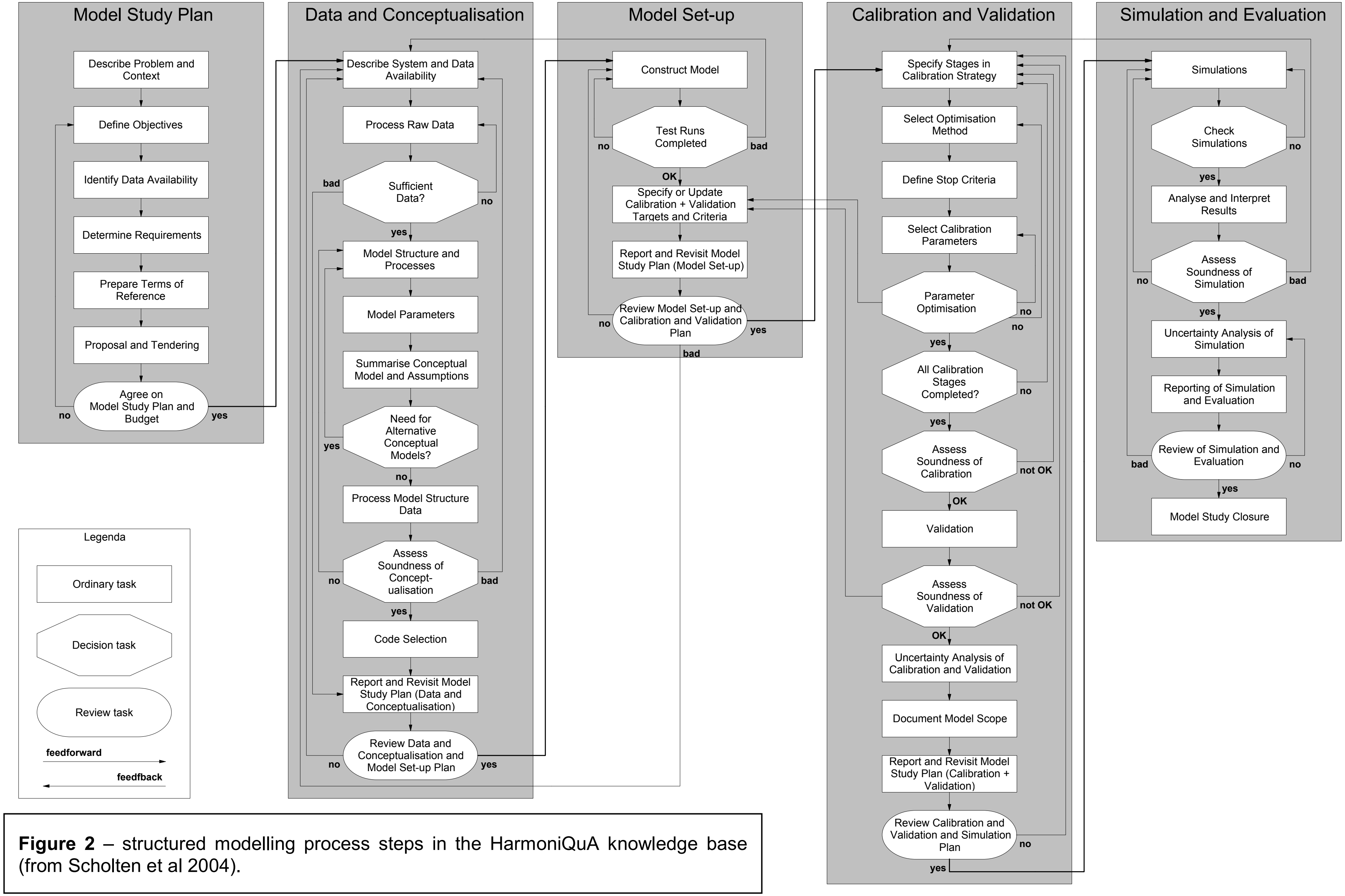


Figure 2 – structured modelling process steps in the HarmoniQuA knowledge base (from Scholten et al 2004).

The purposes of the Water Framework Directive include:

- To prevent further deterioration and protect and enhance the status of aquatic ecosystems and associated wetlands
- To promote sustainable water use, based on long term protection of available water resources
- To ensure the progressive reduction of pollution of groundwater and prevent further pollution.

Under the Water Framework Directive, EU Member States are required to achieve, within specified time frames, "good surface water quality status" and "good groundwater status", and prevent deterioration of existing "good status" systems, in terms of ecological quality and chemical quality. For groundwater, "good status" includes resource quantity, and involves consideration of "Groundwater Bodies" as follows:

- Groundwater that is in continuity with ecosystems and can place them at risk, either through the transmission of pollution or by unsustainable abstraction that reduces baseflow (to rivers and wetlands)
- Groundwater that can provide for the abstraction of significant quantities of water for human use (defined as greater than 10m^3/day).

The Habitats Directive was promulgated in 1992, requiring Member States to maintain in "a favourable condition" designated habitats and species, which form a European network called Natura 2000. For example, this requires Member States to review all licences and permits (eg. for water abstraction) to assess whether they are likely to have an "adverse effect" on such sites. The assessment includes the application of the "precautionary principle".

In summary, the Water Framework Directive and Habitats Directive together drive an integrated water resources management approach of focusing on sustainable support for water-dependant ecology, while protecting the quality and quantity of the water resources for human use.

The Australian policy setting is broadly consistent with the European, although the recently promulgated National Water Initiative (NWI) (http://www.pmc.gov.au/nwi/index.cfm and www.deh.gov.au/water/publications/case-studies/index.html) has the objectives of increasing the productivity and efficiency of water use to sustain rural and urban communities, as well as to ensure the health of river and groundwater systems. That is, the NWI aims acknowledge sustainable resource management and development goals and encourage water trading markets, as well as provision of environmental water needs. The four main aims of the NWI are to:

- improve the security of water access entitlements, including by clear assignment of risks of reductions in future water availability and by returning over-allocated systems to sustainable allocation levels;
- ensure ecosystem health by implementing regimes to protect environmental assets at a whole-of-basin, aquifer or catchment scale;
- ensure water is put to best use by encouraging the expansion of water markets and trading across and between districts and States (where water systems are physically shared), involving clear rules for trading, robust water accounting arrangements and pricing based on full cost recovery principles; and
- encourage water conservation in our cities, including better use of stormwater and recycled water.

3.2.2 HARMONIQUA and MoST

HarmoniQuA is the latest international development in best practice modelling guidelines, and is due for completion at the end of 2005. This report provides feedback, through the notes in Appendix B, on the functionality and guidance content on the HarmoniQuA tool.

HarmoniQuA (www.HarmoniQuA.org) is one of about 10 projects under the CATCHMOD cluster. These projects aim to support implementation of the EU Water Framework Directive (http://europa.eu.int/comm/environment/water/index.html), for integrated and sustainable water resources management and planning. While the Water Framework Directive encourages integrated river basin planning (ie. a clear surface water emphasis is apparent), groundwater is a recognised key component, and it is also explicitly recognised that successful implementation will involve integrated surface and groundwater modelling (a challenge for the modelling industry, but one where we are developing our capabilities).

HarmoniQuA aims to provide guidance for multi-disciplinary teams working in model-based integrated water management. A software tool (**Mo**delling **S**upport **T**ool or **MoST**) is being developed by HarmoniQuA to support the work of these teams, using XML for network delivery of the tool. A simplistic description of MoST is that it is a workflow management tool. It outlines the overall workflow process, with descriptions of the five main steps, 50 tasks and many more activities (sub-tasks), plus outlines of the best practice methods and associated guidance for specific activities. MoST is more than a simple workflow management tool, however, as it provides a means of recording the assumptions, data sources, model results and decisions throughout a study, with QA steps. The tool provides both a means of communicating progress throughout the modelling lifecycle, and an audit trail, which helps deliver "transparency" (a much sought after attribute of model-based studies). The version of the software (11 Sept, 2004) that was tested showed excellent functionality, which augurs well for the usefulness of the final tool.

Some recent papers provide excellent background on the HarmoniQuA approach. Scholten et al (2004, in press) provides a good overview of the HarmoniQuA methodology to support multi-disciplinary model-based water management. The HarmoniQuA papers outline the commitment within the HarmoniQuA process to the following universally important issues, related to the need for flexibility in modelling because it is an essentially creative process:

- ability to apply the methodology in a flexible manner, not as a rigid set of regulations, which seems to be a universal concern of the modelling community, expressed at the FEM-Modflow conference (Carlsbad, September, 2005), and also by all those that I have visited on this Churchill Fellowship

- ability to select an individual domain (eg. groundwater, which currently includes groundwater quality), and/or the other six domains in the overall framework, which includes rainfall-runoff, hydrodynamic (surface water), flood forecasting, (surface) water quality, biota-ecology and socio-economics (and acknowledgement that little is known about the last one); it also allows certain schemes to be closely linked/coupled as a "multi-domain" (eg. groundwater and hydrodynamic); thus, it is designed for stand-alone and/or integrated model studies

- application of a specific level of complexity to the study (basic, intermediate or comprehensive); although it is not yet implemented, there is the planned ability to also identify the application purpose (planning, design and operational management); the establishment at an early stage of the scope and context for the study in terms of purpose/complexity is very important, and consistent with a key part of the MDBC guideline

- various parties can review progress by logging on as modeller, manager, auditor, stakeholder or public, with access according to their authorisation (client-server and web-based tool not yet implemented), and filtering of the information to suit their profile (not yet implemented)

- the ability to proceed through the work process, with progress reports and reviews and communication between parties, and the ability to iterate back to a point to revisit matters if needed (eg. to revise the conceptual model or calibration performance assessment)

- semi-automatic documentation of the progress and decisions through the model journal (software), with a well-documented model archive at the end; although it is a pedantic point, this “model journal” is actually a model study journal, which differs from a “modeller’s journal” of simulation runs and data files etc, as discussed in more detail in Appendix B.

- comprehensive glossary, with a "synonym" feature to provide for cross-referencing of terms between jurisdictions, and listing of references for further information.

The fundamental scientific philosophy that underpins the methodologies of HarmoniQuA and MoST is presented in Refsgaard and Henriksen (2004). As stated earlier (Section 3.1), Refsgaard and Henriksen clarify many of the issues related to terminology that have tended to get in the way of discussions of best practice methodologies over the last few decades. They also provide a sound scientific and logical argument regarding the philosophy of model calibration and validation. It is interesting to note that another recent paper (Hassan, 2004a,b) independently considers the same issues as Refsgaard and Henriksen (2004) and comes to quite consistent conclusions. There are indications from my discussions that the philosophies underpinning these papers will be widely agreed and confirmed. The implication is that the fundamentals of the current best practice guides (including the MDBC guide), do not need substantial revision, as they are consistent with the above papers.

While it is noted that development and testing of MoST is ongoing, the current version of the guidance content of MoST is mainly based on the (Dutch) Good Modelling Practice Handbook (Van Waveren et al, 1999; addressed later this section). Additional content has reportedly been drawn from selected information from the Bay-Delta modelling protocol (BDMF, 2000), and also from the MDBC guide (MDBC, 2001), the latter particularly relating to model reviews and calibration performance assessment.

The major reliance on the Dutch Handbook for current guidance content means, in my view, that MoST is not (yet) practically as useful as it could be, mainly because the Dutch Handbook provides guidance at a high level rather than practical level. That is not to say that the MoST tool is not good (on the contrary, the tool is a major advance), but the guidance content of MoST should be improved. This can be done simply and easily by inclusion of appropriate content from the UK Guidance Notes (Hulme, 2002), which are very practical, and also including more from the BDMF and the MDBC guides. Although the German guide (Arbeitskreis, 2004) generally tends to address more philosophical than practical issues, there are also areas where the MoST tool could be enhanced by content from the German guide (notably the problem specification, commissioning of work, and the hydrogeologic framework). Detailed suggestions are provided in Appendix B, this report will be forwarded to the HarmoniQuA team, and improved consultation was encouraged during my Churchill Fellowship between HarmoniQuA and the German guideline team.

One of the main aims of both the MoST tool and any guidance content is to help educate non-specialists, and to support and inform interactions between modellers, the community and project managers. With improved content on specialist issues that will be filtered by MoST for the user, problem type, domain, and study stage, MoST should become a tool that is seen as practically useful, and not just as a set of generic or high level guidelines. Thus, the technology is likely to be taken up more rapidly and widely, and this should result in improved project outcomes for the community.

3.2.3 Dutch Good Modelling Practice

As pointed out in Refsgaard et al (2004):

> "*The Dutch guidelines (Van Waveren et al, 1999) are the most generic of the existing guidelines in that they cover all the domains relevant for the Water Framework Directive. The technical guidance for different modelling domains exists, but is not as detailed as some of the guidelines that only cover one domain (eg. ASTM guides or Australian guidelines on groundwater modelling). The Dutch guidelines emphasise the dialogue process between modeller and water resources manager, including the review procedures.*"

The "all domains" aspect of the Dutch guide refers to its aim of providing generic guidance in relation to simulation of the following general categories of physical problems ("domains"): groundwater (saturated and unsaturated), precipitation-runoff, water distribution, hydrodynamic (surface water), surface water quality, flood forecasting, calamity (eg. dam break), morphological, wastewater purification, ecological, water-related economics, and emissions. In the spatial dimension, the Dutch guide differentiates (again, in a generic manner) between one-, two-, three-dimensional and point models, as well as the local, regional, national and international scales. The Dutch guide also differentiates between stationary and dynamic temporal conditions, and analytical or numerical methods (and "everything in between") in terms of the mathematical solution technique. The two main sections of the Dutch guideline total just 120 pages, with another 40 pages of checklists and bibliography. Thus it is obvious that the guidance that is provided across all these domains is at a high level (ie. "guiding principles", rather than detailed or practical methods).

The core of the Dutch guide is the "structured ontology" (Scholten et al, 1998). In this case, the structured ontology could also be simply described as a set of generic guiding principles common to environmental simulation methodologies. A simple graphical version of the structured ontology is presented in Figure 2. This main part of the guide also includes a series of generic tests that should be carried out during model studies to evaluate model performance. However, it does not prescribe the specific methods or algorithms (Scholten et al, 1998), or even the performance criteria, that should be used for any modelling step in any domain. Thus, while the Dutch guide provides an excellent example of guiding principle methodologies, and is comprehensive across many domains, it is not designed to serve as a practical set of guidance notes (which is what the MDBC guide is designed to deliver), nor is it sufficiently specific, detailed or pragmatic to serve as such.

The second main section of the Dutch guide documents a wide range of "sensitivities and pitfalls" for each domain. While this does provide some practical and specific (rather than generic) guidance by way of "lessons learnt", and it does utilise the structure of the adopted ontology, it does not provide the fundamental type of guidance notes as do the MDBC or UK guides. The community (in Australia at least), along with others such as inexperienced modellers, resource management agency staff and universities, seem to value such pragmatic guidance, particularly as an educational resource, but also for its detail on project management steps and model performance criteria.

Interestingly, experienced modellers are typically "under-whelmed" by pragmatic guides, and even more so by generic guides and checklists. Within the modelling industry, guidelines seem to be considered useful and necessary for modellers to learn the craft, but not sufficient for good modelling outcomes by experienced practitioners, who usually apply more advanced techniques. Nevertheless, guidelines are accepted by the water resources industry as useful for helping to ensure that study outcomes are fit for purpose, if not value for money.

It is apparent that the Dutch guide provides a substantial part of the structure and content for the HarmoniQuA project (discussed in detail later), which is developing a Modelling Support Tool (MoST) for quality assurance, project management and communication on multi-discipline modelling projects. Among the modellers with whom I discussed this issue, it is a commonly held view (which I strongly support) that generic guiding principles are important, but, in order to be practically useful, the MoST tool also needs to have more specific and pragmatic guidance content for each domain. Regarding the groundwater domain in particular, this means that the guidance content in MoST needs be substantially augmented by including appropriate information from the MDBC, UK, German and other selected guidelines, which have been proven in application over recent years.

3.2.4 UK Guidance Notes

The UK Environment Agency has a range of guidance relating to various aspects of groundwater modelling, notably:

- Framework for Groundwater Resources Conceptual and Numerical Modelling (Environment Agency, 2001); promotes a long term strategy for the use of models for a range of major regional modelling projects across England and Wales, and outlines basic technical and project management approaches.

- Groundwater Resource Modelling Guidance Notes and Template Project Brief (Environment Agency, 2002); documents the collective wisdom gained from practical experience on technical and project management aspects of regional modelling methodologies, and is intended to complement existing text books.

- Guidance on Conducting a Hydrogeological Impact Appraisal (HIA) (Environment Agency, 2003); outlines the approaches for impact assessment for use by hydrogeologists involved in abstraction licensing, which includes both regional and local scale appraisals.

- Guide to Good Practice for the Development of Conceptual Models and the Selection and Application of Mathematical Models of Contaminant Transport Processes in the Subsurface (Environment Agency, 1999).

The guidance notes (EA, 2002, 2003) were designed for use by hydrogeologists and modellers to implement the Modelling Framework (EA, 2001) by undertaking investigations, analysis and modelling projects for the purposes of abstraction licensing and the Catchment Abstraction Management Strategy (CAMS), and for resource assessment and management related to the European Union Habitats Directive and Water Framework Directive. Appendices describe case studies that illustrate the integrated use of the various guidance topics. The solute transport guidance notes (EA, 2001) are mentioned for completeness, but are not discussed in detail here, as the scope of this document does not extend beyond groundwater flow modelling.

Although they were produced independently, the methodologies described in the UK guidance notes are broadly consistent, and also specifically consistent in most areas, with the MDBC guide. Apart from the different hydrological regimes of the two countries, and thus a somewhat different focus on technical methods, the differences are mainly due to the need for more detailed guidance in the UK to manage the much greater project budgets, data availability and timeframes for their regional studies.

The project management procedures and template project brief are particularly comprehensive and detailed, probably because they are designed for application to major projects (with budgets typically in the range €100,000 to €400,000). In comparison, the scoping and project management guidance that is given in the MDBC (Australian) guidelines are much more oriented to the smaller scale of project resources available in Australia. Thus, it is recommended later in this report that content from both the UK and Australian guides be used to upgrade the content in the MoST package being developed by the EU HarmoniQuA project, as this is likely to be a tool that will be applied in future major modelling projects.

One major area of difference between the UK and MDBC guides is that, while the UK guide incorporates the need for reviews throughout the modelling study, it does not detail model review procedures, whereas the MDBC guide is much more detailed on reviews. However, on other aspects, the UK guidance notes sometimes explore the philosophies, guiding principles and/or technical methods in greater detail than does the MDBC guide. For example, the UK guide suggests that regional models should not be calibrated to short term pumping tests, and also discusses conceptual models in more detail, introducing the term "perceptual model" that was first suggested by Beven (2002). In simple terms, the perceptual model is "the qualitative model in your head", whereas "conceptual model" includes quantitative elements, such as water balance components and aquifer parameter values. Including the quantitative elements allows the model to be tested and evaluated objectively.

The HIA guidance notes are particularly useful in that they discuss some common misconceptions about groundwater abstractions and associated impacts. This provides some very good guidance, not only to inexperienced hydrogeologists, but also to non-specialists (and even experienced hydrogeologists have been known to benefit from explanation of the issues and the case studies).

There is a focus throughout all the UK guidance notes on the need to develop a conceptual model in a staged manner, starting from an initial simple representation of the essential features of the system, and developing further complexity gradually (eg. refer to Figure 1). Model refinement involves testing the concepts in qualitative and quantitative terms with water balance estimates, technical analysis and numerical modelling.

This is linked with the need to evaluate uncertainty in terms of natural hydrological variability, and of uncertainty in regard to models and data sets. The overall approach is cyclical, involving conceptualisation, simulation/analysis, review and refinement until the risk (the prediction of impacts on the environment or resource sustainability) is considered "acceptable" in relation to the study purpose. This approach is quite consistent with the approaches outlined in the MDBC guide, particularly the need to establish, at the outset of the study, the study purpose, resources available, and the appropriate model complexity, and then to assess whether the model is "fit for purpose" (including the simulation results) through specified review methods.

In summary, the HIA approach outlines that as the investigation progresses, the cost increases, as should the confidence in the model, with a resulting decrease in uncertainty, until the risk is reduced to an "acceptable" level. The "acceptable" level of risk is a subjective assessment, depending fundamentally on the purpose of the investigation, such as licensing decisions, or assessment of resource availability for sustainable development, etc. The application of the precautionary principle within an adaptive management approach can also help manage remaining risks.

These approaches are quite consistent with the MDBC guide, notably where it outlines the "**quick-cheap-good**" paradox (Section 2.1, Study Purpose and Model Complexity):

> "*The end-user can readily obtain a model with one or two of these three attributes, but not all three. If a model is required to be done quickly, it can also be done cheaply, but the results may not be good enough on which to base important resource development or management decisions. Such a simple model may be good enough for rough calculations to guide a field program, or to assess the broad impacts of a certain proposal, but would usually not be sufficient for project approval or licensing purposes. Alternatively, if a good, reliable model is required, then it is not likely to be able to be developed quickly or cheaply.*"

3.2.5 German Guidelines

The FH-DGG guideline is intended to assist clients, consultants, and regulatory officers in groundwater resources in the modelling process. The strength of the guide is illustrated by the substantial detail it provides on the assessment of the database, the choice of an adequate model approach based on the spatial scale and the data situation, as well as possible necessary revisions of the model approach. The procedures are designed to achieve an efficient approach to model development, irrespective of the individual hydrogeological situation and the posed problem.

The guideline is based on the idea of a "Hydrogeological Model" (HGM), which aims to provide a consistent framework for the transfer of the complexities of hydrogeological nature into a quantified model. The eventual HGM is intended to enhance the understanding of hydrogeological systems and to provide quantified predictions of their behaviour through analytical or numerical calculations. Thus, all the quantifiable elements are involved before the final HGM is achieved (ie. not just water balance components and aquifer parameters, but also model calibration and validation steps). There is acknowledgement of the possible need for process feedback to refine the HGM concept throughout the process, consistent with other best practice guides. The final HGM could be described in simple terms as a calibrated and validated model with a sound conceptual basis, or a predictive scientific tool that has been shown to replicate the historical behaviour of the system.

The initial focus of the FH-DGG guide is the development of a HGM concept (this is broadly synonymous with the "conceptual model" of most other guides, although there are significant differences when details are considered). The HGM concept simplifies nature adequately with respect to the problem to be solved and the relevant dominant hydrogeologic features. The final HGM is only achieved if the HGM concept is proved by a sound evaluation, consisting of an analysis of quantified model results with respect to non-uniqueness, accuracy, sensitivity, and model application range as well as a quantification of uncertainty.

The FH-DGG guideline is not considered to be a strict "recipe" on modelling, but rather provides a systematic framework for the generation of HGM concepts. The FH-DGG guide is quite detailed on some practical issues like problem specification, commissioning of work, efficient workflow, and structural quality assurance. In addition to the content on analysing the hydrogeological system and devising a suitable modelling approach, the project management content would add value to the HarmoniQuA/MoST guidance, especially the commissioning process and the process of review at specific milestones jointly by clients, consultants, and regulatory officers.

3.3 MDBC GROUNDWATER FLOW MODELLING GUIDELINE

3.3.1 Background

The following introductory comments (taken from the MDBC Groundwater Flow Modelling Guideline, 2001) describe the purpose and applicability of the de facto Australian best practice guide:

> *The aim of most guidelines is to reduce and reveal model uncertainty for the users of modelling studies, including resource management decision makers and the community. This is achieved by promoting transparency in modelling methodologies and encouraging innovation, consistency, and best practice. Guidance should be provided to project managers and the community (ie. non-specialist modellers) by outlining the steps involved in scoping, managing, and evaluating the results of groundwater modelling studies. The MDBC guidelines also serve modelling specialists*

by providing a baseline set of ideas and procedures from which they can innovate, and detailing model performance criteria, which are used by auditors/reviewers of models.

The MDBC guidelines are intended for use in raising the minimum standard of modelling practice and allowing appropriate flexibility, without limiting necessary creativity or rigidly specifying standard methods. The guidelines also should not limit the ability of modellers to use simple or advanced techniques, appropriate for the study purpose. Techniques recommended in the guidelines may be omitted, altered, or enhanced, subject to the modeller providing a satisfactory explanation for the change and negotiation with the client and/or regulator as required. Not all aspects of the guidelines would necessarily be applicable to every study. It also is acknowledged that standardization of modelling methods will not preclude the need for subjective judgment during the model development process.

The guidelines are to be applied to new groundwater flow modelling studies and reviews of existing models. The guidelines should be seen as a best practice reference point for framing modelling projects, assessing model performance, and providing clients with the ability to manage contracts and understand the strengths and limitations of models across a wide range of studies (scopes, objectives, budgets) at various scales in various hydrogeological settings. The intention is not to provide a prescriptive step-by-step guidance, as the site-specific nature of each modelling study renders this impossible, but to provide overall guidance and to help make the reader aware of the complexities of models, and how they may be managed.

3.3.2 Need for an Update?

Throughout the Churchill Fellowship, I reviewed and tested the structure and content of the MDBC guide in relation to the structure and content of other leading guidelines, and in relation to developing trends and protocols likely to define future best practice modelling. While there are various elements in other guides and papers that describe aspects of modelling philosophy or the procedural steps more comprehensively or more elegantly, there are very few areas where the MDBC guide needs updating. The three notable areas where updates may be worthwhile are in the model review checklists, parameter optimisation methods, and the notes on leading modelling codes (especially regarding integrated surface-groundwater modelling). Each of these area are discussed further in Sections 3.3.3 to 3.3.5.

Fundamentally, I found that the MDBC guide is still very much fit for its purpose.

More importantly, with the HarmoniQuA project set to release MoST at the end of 2005, it would be prudent to wait and see the final configuration of that tool, and review its guidance content, before considering the need to update the MDBC guide.

3.3.3 Model Review

The review methodology in the MDBC guide is acknowledged as unique among other guidelines (Hill et al, 2004). Its value was further confirmed by the HarmoniQuA project adapting it for use in MoST. The schedule of review questions have also proven their value in Australia, as they have been adapted by hydrological (surface water) modellers to devise audit schedules for river basin (allocation) models, and I have also adapted them for application to "landscape impact" type models (eg. for the review of the SIMRAT/SIMPACT GIS tools that include analytical models to estimate the salinity and water balance impacts of irrigation allocation and transfer proposals).

Middlemis and Merrick (2004) describe some of the lessons learned over the last few years from application of the review methodology and schedules, and conclude that the review protocols are still valid, although the framing of the questions could be updated. The questions are currently configured in “closed” or “checklist” form, which tends to encourage yes/no type answers, whereas audit protocols should ask “open” type questions. This would encourage more explanatory answers, and should require documentation of “objective evidence” of compliance (ie. references to reports). It is recommended that the MDBC consider the need to collate and review existing model audit schedules, and to revise them where necessary to provide more “open-ended” questions, with clear performance criteria, and to make them available on their website for universal application.

Middlemis and Merrick (2004) also find that there is not an extensive list of experienced modelling specialists in Australia who meet the requirements for a model reviewer/auditor as outlined in Appendix C of the MDBC guide. Model auditors should be selected from experienced professionals who are actively engaged on practical modelling projects, so that they understand the strengths and limitations of model methodologies, conceptual understanding and data availability, and thus can undertake a realistic review of whether a model is fit for its intended purpose (ie. "inside knowledge" is needed). Thus, any list of model reviewers needs to provide project principals with the capability of selecting a knowledgeable model auditor who is well-informed yet is not compromised in terms of potential conflicts of interest. It is recommended that the MDBC consider the need for a panel of nationally-recognised modelling audit specialists, which can be considered by project principals for inclusion on project steering committees as advisers and auditors.

3.3.4 Parameter Optimisation and Uncertainty Assessment Methods

The MDBC guidance on parameter optimisation and uncertainty assessment methods and their application is not detailed, and there has been substantial development in this field in recent years. However, the guidance in MoST is very good in this regard, and the papers on these topics in conferences demonstrates that best practice is still developing and is not yet mature (unlike many other aspects of modelling). Thus, while the MDBC guide could still be considered fit for its purpose, it is recommended that the final configuration of MoST be reviewed in late 2005 or early 2006, to confirm whether it constitutes a suitable guide for parameter optimisation and uncertainty assessment. At that time it may be appropriate to devise action plans for updating the MDBC guide.

In addition, the JUPITER API (Joint Universal Parameter IdenTification and Evaluation of Reliability Application Programming Interface) is due for release in 2005. The IGWMC October 2004 newsletter (http://typhoon.mines.edu/news/) indicates that the Jupiter project and API "strives to energise the areas of sensitivity analysis, data assessment, calibration, and uncertainty analysis of groundwater models." Applications developed using the Jupiter API will provide the opportunity for users to readily:

- experiment with a number of techniques for generating conceptual models (eg. geologic process, geostatistics, upscaling)
- compare alternative algorithms for the same task
- evaluate and evolve/refine conceptual models through ranking and multi-model inferential analysis
- assess data needs to improve calibration in light of prediction results and uncertainty

The outcomes of the Jupiter project should move best practice forward in the area of parameter identification and optimisation.

3.3.5 Leading Modelling Codes

The MDBC guide has an appendix that summarises the key capabilities of a number of numerical modelling codes, but this information is now out of date. An update is currently being undertaken as part of a project for the MDBC, and it is recommended that the outcomes of the update be made available for download from the MDBC web site. This issue is also currently being discussed with the International Groundwater Modeling Centre (http://www.mines.edu/igwmc/), with a view to providing a list on their website.

In the meantime, the next section provides some summary comments on some of the latest integrated modelling tools.

4 INTEGRATED MODELLING

4.1 ISSUE

A question that is often asked by project managers and the community is: "What modelling package should I use for this study?". One of the aims of the MDBC modelling guide was to provide some information on the capabilities of the leading packages. This section provides an update to that information. The update is provided from the perspective of the need for integrated surface-groundwater modelling tools to address conjunctive and integrated catchment management issues, rather than just from a groundwater modelling viewpoint.

Integrated catchment management (ICM) has been identified as a clear need for improved natural resources management, and this is echoed in policy settings in Australia and internationally. For ICM to be delivered, there is a need for comprehensive/multi-disciplinary integrated modelling tools, rather than the simplistic approaches that have been historically applied (eg. surface water models with a very simple groundwater capability, and vice-versa). To the best of my knowledge, Australian best practice on model development and applications is broadly consistent with approaches in the UK at least, if not much of Europe, although it is lagging behind the USA in this field. That is, best practice Australian (and UK) projects on surface-groundwater modelling have tended to take either a surface water or a groundwater focus, with the non-primary domain represented adequately, but in less detail. These approaches are simplistic in that they do not represent the dynamic interaction (ie. within the one software package) between the key elements of entire hydrological cycle, from rainfall-runoff and river/channel flow to infiltration through the unsaturated zone and flow in aquifer systems. From a groundwater modelling point of view, rivers and wetlands are typically represented using specified water levels in the river, or in less common cases using specified flows (with river water levels calculated by the model, using Manning's equation). However, in the USA, more comprehensive integrated modelling approaches have been applied, notably on projects in Florida (where wetlands are prominent in the landscape) and California. This is an area where best practice in Australia needs to improve in future, and the following section outlines information on some of the tools that should be suitable for application.

4.2 INTEGRATED MODEL SELECTION CRITERIA

There has been substantial development in recent years of detailed integrated tools for surface-groundwater modelling, much of it in the USA. There has also been substantial application on real-world projects, again, mostly in the USA, presumably because of the resources that can be or need to be applied to meet environmental management or compliance requirements.

Table 1 provides summary information on the capabilities and limitations of the integrated modelling packages that are either currently available, or in the final stages of testing and development, which would be suitable for application in Australia. This is a rapidly changing field, and the reader is cautioned to check that the information below is up-to-date, and to confirm that the model details listed are correct and applicable to their study purpose, before committing the substantial sums involved in software purchases, and in the modeller's learning curve.

The selection criteria for this shortlist comprises:

- 3D groundwater flow with/without unsaturated zone simulation, with detailed packages for two-way surface-groundwater interaction, evapotranspiration, drainage, recharge, wells, etc.
- 1D open channel flow and two-way interaction with aquifers and the unsaturated zone, preferably with dynamic elevation-area-volume-leakage relationships for channels, storages and wetlands, and representation of the surface channel/wetland geometry specifically (as opposed to simple grid representation only)
- preferably with capability to model (continuous) rainfall-runoff processes, dynamically model hydraulic control structures and on/off triggers such as weirs, gates, pumps, etc.
- comprehensive and accurate water budget analysis
- well-documented code validation, and reasonable model set-up and execution times
- graphical user interface for pre- and post-processing, preferably with good GIS linkages.

Most of the information for the assessment of model capabilities was drawn from two key evaluation reports (Camp Dresser and McKee (2001), and South Florida Water Management District (2002)), plus information from the MDBC Groundwater Flow Modelling Guideline (2001), discussions with various professionals during this Churchill Fellowship, and Aquaterra project experience). The CDM report considered 75 models, of which the nine models (listed below) met the integrated modelling criteria. The SFWMD evaluation report considered 15 models and combinations of models, which can be summarised into the 13 models listed in Table 1.

TABLE 1 INTEGRATED SURFACE-GROUNDWATER MODEL SUMMARY

Camp Dresser and McKee (2001)	South Florida Water Management District (2002)
HSPF (process-based watershed model supported by US EPA and USGS; often integrated with groundwater models)	BASINS 3.0 (incl. HSPF, FEQ)
HMS (process-based watershed model developed by US Army Corps of Engineers)	HEC-HMS, and HEC-RAS with UNET
MIKE-SHE	MIKE-SHE / MIKE-11
MODBRANCH (coupled MODFLOW with 1D open channel and interactive leakage)	MODBRANCH (coupled MODFLOW with 1D open channel and interactive leakage)
MODFLOW (limited surface water)	MODFLOW (limited surface water)
SWMM (urban; limited groundwater)	EPA SWMM; XP-SWMM2000
DYNFLOW (MODFLOW equivalent)	
SWATMOD (uses MODFLOW)	
FHM-FIPR (incl. MODFLOW and HSPF)	ISGW (origins in FHM; incl. MODFLOW & HSPF)
	IHM (origins in FHM; incl. MODFLOW & HSPF)
	WASH123D
	MODHMS (incl. MODFLOW and HMS)
	MODNET (incl. MODFLOW and UNET)
	HSM (South Florida Hydrologic Systems Model, comprising a range of innovative approaches)
	AdICPR

From the modelling packages listed in Table 1, the shortlist in Table 2 includes those models with the more comprehensive capability, notably regarding rainfall runoff, 1D channel flow, unsaturated zone and 3D groundwater flow, and operating within a unified software package with a good graphical interface. In addition, some newly developed models, not considered by the CDM (2001) and SFWMD (2002) reports, have been identified and added to Table 2. Some other models, which were not included in the above reports, were considered but not included because of known issues with regard their capabilities (eg. see LaBolle et al, 2003 regarding the IGSM model).

In summary, the Mike-SHE model is regarded as the most comprehensive package for physically based, spatially distributed, integrated modelling, with excellent data management and visualisation capabilities, and a proven track record on integrated modelling studies (but not yet in Australia). This is closely followed by MODHMS, which has been applied by practitioners in Australia, although mainly from a groundwater point of view. Both of these packages are being used on integrated projects in Australia at present by several consultants, including Aquaterra. The other packages listed in Table 2 represent fairly recent developments, but their capabilities are either more limited than these two codes or are they are still being developed or evaluated, and a watching brief should be applied to them.

TABLE 2 BEST PRACTICE INTEGRATED SURFACE-GROUNDWATER MODEL SUMMARY

Package	Strengths	Limitations
MIKE-SHE	• 3D variably saturated groundwater flow, fully integrated with rainfall, ET, overland, channel flow and pipe flow; dynamic control of hydraulic structures; comprehensive capabilities • three levels of groundwater models from simple lumped parameter to complex 3D • excellent graphical interface, strong links with ESRI (ARCGIS etc); applied extensively and successfully in USA for integrated studies • DHI is non-profit organisation, applies substantial funds to code development	• river channels linked directly to saturated 3D model without unsaturated zone interaction • reputation of being "data-hungry" persists, but to lesser extent than before 2000; still needs substantial study resources • somewhat expensive, no access to source code, and no ability for variable grids • Modflow capability still being developed (due in 2005)
MODHMS	• full Modflow capability, plus density coupled flow and solute transport, and variably-saturated capability • most useful at meso-scale, currently being tested at watershed scale	• limited capability for continuous rainfall-runoff • surface water components not widely applied to watershed scale; no controlling dynamic hydraulic structures (only per stress period) • graphical interface developing • no access to source code
HYDROSPHERE	• 3D variably-saturated, finite element counterpart to MODHMS, includes density coupled flow and solute transport using FRAC3DVS	• only recently developed and still being tested • graphical interfaces still being developed (GMS and others) • no access to source code
IHM	• 3D variably-saturated finite element flow and solute transport, with overland and channel flow and water budget for the full hydrologic cycle • groundwater includes porous media, fractures, conduits, macro-pores and perched water tables • access to source code	• recently developed and still being tested; unknown evapotranspiration capability • no hydraulic structures • graphical interfaces still being developed (GMS and others)
SFWMD	• 3D Modflow-based, includes specially developed physically-based wetland, diversion and ET-recharge packages • access to source code	• no unified graphical interface • designed for regional scale, not for local scale decision support
SFRSM (South Florida Regional Simulation Model)	• object-oriented surface and 3D groundwater model, plus overland and channel flow, unsaturated zone and wetlands, complex water management operational rules • access to source code	• still being developed (esp. water management module and graphical interface) • designed for regional scale, not for local scale decision support
WASH123D	• public domain, physically based, distributed parameter, finite element, 3D variably-saturated, integrated surface water and groundwater model, and groundwater quality • can simulate dynamic hydraulic structure operating rules	• GMS graphical interface adequate but still being developed • Long run times and needs parallel super computers to run integrated models at catchment scale; still being tested

Note: website details for the models in Table 2 are listed at the end of the Bibliography.

4.3 IQQM-MODFLOW

The Integrated Quantity and Quality Model (IQQM) is a generic river basin simulation package, including in-stream water quality components such as salinity and nutrients, that is used for all surface water management modelling in Queensland and New South Wales. The MDBC has recently adopted IQQM as their preferred river basin model for strategic resources studies, as distinct from a river operations modelling tool (A. Close, pers. comm.).

Although IQQM has a "groundwater" component, it is a non-calibratable node within the model that is usually used as a storage term for stream losses to assist in the stream-flow calibration. In the early 1990's, the Queensland Department of Natural Resources Management (DNRM) developed linkages between IQQM and MODFLOW, and also linked IQQM to a simple groundwater storage model. Initial trials of the linked modelling system proved successful, but further development and more rigorous trials are required before it would form a suitable tool for widespread application.

Once it is fully developed, the coupling of a detailed river basin resource assessment, allocation and management model, such as IQQM, with a groundwater model with similarly detailed features for simulating aquifer systems, such as Modflow, would provide the type of tool that is required for comprehensive conjunctive resources assessment and management in much of Australia. Modelling tools such as these (refer to Tables 1 and 2) can be quite expensive (in terms of purchase cost, staff training time and data collection and input requirements), but the IQQM-Modflow approach would provide substantial cost-savings potential. Cost savings could accrue because the IQQM tool is an Australian product, and there are a number of trained staff in agencies in NSW and Queensland, with capacity-building in progress in the MDBC. Similarly, there are a large number of professionals in agencies and consultancies that are well-trained and experienced in Modflow studies.

It is strongly recommended that appropriate resources be applied in a sustained programme to develop the IQQM-MODFLOW package into a comprehensive tool for analysing and modelling conjunctive resource assessment and management.

The other river basin models in common usage in Australia are REALM (applied in Victoria) and WaterCress (applied in South Australia). These models were developed from a water supply perspective, in which a user defines the supply and demand distribution system, using nodes such as reservoirs, demand centres such as cities, and so on, which are connected by conveyances which can be rivers or pipes. The node of interest within the context of stream-aquifer interaction is the groundwater node, which basically represents a groundwater storage that can be recharged or depleted. Groundwater is thus simulated in a manner similar to a surface water reservoir volume, but the groundwater flow system is not modelled as such. However, since the aquifer is treated as a depletable storage component that can be connected through rivers, the surface water and groundwater are treated as a single resource. The development of IQQM-Modflow may inspire these jurisdictions to develop similar linkages for REALM and WaterCress, or it may be considered more appropriate to adopt the IQQM-Modflow or other technology (eg. Mike-SHE).

4.4 ZOOMQ3D

During my Churchill Fellowship, and with support from my employer (Aquaterra), I took the opportunity to attend a training course on the ZoomQ3D model. It is one of the first applications of object-oriented groundwater modelling in the world, and thus represents a significant future direction in this field. The ZOOMQ3D model was officially released at the course, along with its associated object-oriented recharge model, ZOOBRM, although it had been "launched" previously. I also took the opportunity to visit the British Geological Survey (BGS) offices in Wallingford to discuss plans for further development of the Zoom models, particularly the need for the development of integrated (surface-groundwater) modelling tools.

ZOOMQ3D has been developed over many years by a partnership of the University of Birmingham, the Environment Agency and the BGS. The model code is quite consistent with MODFLOW (the industry-standard groundwater flow model), in terms of its modularity and functionality (eg. surface-groundwater interaction features). Its additional features include local grid refinement (LGR), the ability to represent model features (such as rivers) independently of the grid, and a variable hydraulic conductivity with depth mechanism (VKD). The particle tracking implementation is consistent with VKD by simulating variable velocity with depth within the one layer. A recharge model has also been developed that is compatible with ZOOMQ3D and which uses GIS based data directly. Both models have been applied by the BGS to a variety of regional groundwater flow studies on Chalk and Permo-Triassic aquifers in the UK.

4.4.1 Local Grid Refinement

Fine grids are required in modelling for two main reasons:

a. to represent physical mechanisms which only operate at a relatively small scale, such as surface-groundwater interaction at wetlands or rivers, which is particularly relevant to integrated surface-groundwater modelling

b. to reduce numerical errors potentially inherent with larger grid spacings (eg. in areas of rapidly changing hydraulic gradients; and when undertaking solute transport and/or particle tracking).

One of the major innovative aspects of ZoomQ3D is the Local Grid Refinement (LGR) capability. Using the object-oriented and grid-independent data approach, coupled with the LGR methods, one can have several local-scale model features (eg. river reaches, wetlands, contaminant sites) scattered within the same regional model without having to build several separate local models to investigate in detail the local scale effects. With the LGR method, one regional model can be devised, which is operating in some places at the local scale. In simple terms the LGR gives Zoom a kind of finite element scale-functionality within a finite difference model.

The following paragraphs on LGR issues are edited from a set of notes provided by Paul Hulme (UK Environment Agency) regarding discussions from the groundwater modellers' group he established.

The discussion centred on a paper on telescopic mesh refinement (TMR) techniques in the MODFLOW2003 conference by Steffen Mehl and Mary Hill entitled "Local Grid Refinement Methods for MODFLOW: The Good, the Bad, and the Ugly". The paper compared three techniques, 1) simply refining the grid ("ugly"), 2) "standard" TMR (potentially "bad"), and 3) using LGR ("good") which is what ZOOM does.

Mehl and Hill concluded that simply refining the grid is "ugly" and that cells at the edge of the mesh can have numerical errors due to the high aspect ratio, but the argument was considered somewhat weak, with no known serious documentation.

The second technique (TMR) was found to be a problem by Mehl and Hill because, while TMR provides the link from the original regional model to the boundary conditions of the local (TMR) simulation, there is no feedback loop from the local model back to the regional model. This is only a problem if the local model is being (re)calibrated, or if predictions are run which could feed back to the regional setting. Most people use TMR to make predictions at a finer scale, so the lack of an active link between the two models is not usually a problem.

Paul Hulme illustrated the problem by pointing out that stresses applied within the local grid can significantly affect its boundary conditions (eg. adding a pumping well to remove contaminant). Thus, the regional model should be re-run with the pumping well added, and the TMR boundary conditions adjusted appropriately. This could be solved by coupling the local and regional grids, which would involve applying the iterative coupling method of Mehl and Hill, or the intrinsically coupled ZOOM method, or the traditional LGR available through finite element methods.

The two real advantages with the third approach (LGR) are i) issues of boundaries are resolved (although, as noted above this can be resolved in other ways in most cases), and ii) that it saves memory. Since computer memory is not usually limiting these days, this is not usually much of an issue. There may be a computational advantage to LGR but benchmark studies have not been completed, and the new generation of Multigrid Solvers may show that there is little computation difference. It is interesting to note that the test cases presented by Mehl and Hill show a small measurable error in head/flux for LGR that is not present in the case of simply refining the grid. In summary, the group tended to disagree with the conclusions of Mehl and Hill, and concluded that, until some case studies become available to test the approaches, simply refining the grid in the traditional manner may be the best approach.

4.4.2 Other ZoomQ3D Functionality Needs

There are plans to enhance the current horizontal LGR by extending it to the vertical grid. The combination of horizontal and vertical LGR would make the Zoom models much better than existing techniques for investigating the detailed flow process in wetland areas and near rivers. It would also help address the dimensionality problem that affects saline intrusion models in particular, where multiple layers and a very fine grid are required in coastal areas to accurately simulate these complex flow systems and avoid numerical errors, but such a fine grid is not necessarily needed in other areas of the model. For saline intrusion models, a density-coupled mathematical formulation of ZOOMQ3D would also be useful.

One major issue that is limiting wider application of Zoom in general practice is that there is currently no graphical user interface, and so Zoom models must be developed using a combination of spreadsheets and small utility programmes. While the BGS are interested in mathematical model development, they are not in the business of developing user interfaces, although it is understood that they have had discussions with organisations such as the Danish Hydraulic Institute regarding their interface software.

I believe that Zoom would not be taken up in general practice unless and until it is available with reasonable interface software.

Another limitation is that, despite its unique features, Zoom does not have all the functionality of Modflow in terms of surface-groundwater interaction and other hydrological packages. This is also one of the main reasons why Modflow in particular tends to be preferred in general practice (ie. it tends to have the best capability in terms of surface-groundwater interaction), despite its finite difference limitations compared to finite element model approaches. The local grid refinement capability of Zoom, combined with its intrinsically coupled approach (automated feedbacks of flow influences between the local and regional grids), lends itself to model applications for wetlands and rivers (ie. to address the WFD and Habitats Directive), but improved hydrological packages are needed before Zoom can properly fulfil this role.

In summary, Zoom demonstrates an innovation in groundwater modelling, with its object-oriented formulation and local grid refinement. However, it is my view that there are a number of limitations in its current form that will need to be addressed before it would be regarded as a useful addition to the modellers' toolbox.

5 CONCLUSIONS

- This study has found that the existing MDBC Groundwater Flow Modelling Guidelines are very well regarded internationally, and are fundamentally still relevant and fit for purpose as a best practice guide for groundwater flow studies. Further, the MDBC guide has influenced the development of best practice guidance internationally, including the next generation of combined quality assurance and guidance software, the Modelling Support Tool (MoST) from the European HarmoniQuA project.

- The Modelling Support Tool (MoST) from the HarmoniQuA project should provide a useful tool for the international community (when it is completed by December 2005) for recording the assumptions, data sources, model results and decisions throughout a study. This tool provides both a means to communicate approaches and progress, and for quality assurance and an audit trail, and it should thus help deliver "transparency", a much sought-after attribute of model-based studies. MoST may well be suitable as a replacement for the existing MDBC guide, especially for integrated and multi-disciplinary studies (surface water, groundwater, ecology, etc), subject to review of its final configuration.

- The current (unfinished) content of MoST draws heavily from the Dutch Good Modelling Practice Handbook, with some aspects also drawn from the MDBC guide, and the Bay-Delta Modeling Forum. To encourage widespread uptake of MoST as a practical tool for improved QA, project management and communication, its current somewhat generic guidance content should be improved, especially for the groundwater domain, by including more specific information from the MDBC, UK, German and other selected guidelines, which have been proven to be effective.

- There are, and always have been, leading edge modelling technologies being developed that extend beyond best practice methods for devising decision support tools for natural resources management. For groundwater modelling, these leading techniques generally relate to devising and testing alternative model “realisations” by using parameter optimisation methods and undertaking uncertainty assessment. While the existing MDBC guide is considered adequate for its purpose in regard to outlining the basic technologies, it should be reviewed and updated regularly, as the “advanced” practices become more mainstream. The first review should occur in 2006, after the MoST tool and the JUPITER API, become available, as improved guidance on 0these methodologies is being provided through these systems.

- Other leading edge technologies relate to the current spate in developing and applying numerical modelling tools for integrated surface-groundwater modelling, which seems to be focused on the USA. This study found that the packages with the most capability/promise are Mike-SHE and MODHMS. In addition, IQQM-Modflow has substantial unrealised potential.

6 RECOMMENDATIONS

- Content from the UK, German and Australian guides should be used to upgrade the content in the HarmoniQuA project’s MoST package, as this is likely to be a tool that will be applied in future major modelling projects for quality assurance, project management and web-based communication of project progress and status

- The final configuration of the HarmoniQuA project’s MoST package should be reviewed in early 2006, to confirm whether that tool delivers suitable guidance in addition to the QA and project management capabilities, especially for parameter optimisation and uncertainty assessment methods. Appropriate action plans may then need to be devised for updating the MDBC guide.

- The MDBC should consider the need to collate and review existing model audit schedules (“checklists”), and revise them where necessary to provide more “open-ended” questions, with clear performance criteria, and consider making them available on their website for universal application. In addition, it is recommended that the MDBC compile a panel of nationally-recognised modelling specialists, who can be considered as adviser and auditor candidates by project principals for inclusion on project steering committees.

- The MDBC should initiate a project to finalise development of the IQQM-Modflow model as a comprehensive tool for integrated surface-groundwater modelling, to support the current investment in IQQM (surface water modelling) and Australian-based technologies.

7 BIBLIOGRAPHY

Aly, A., M. Ross, J. Geurink and Z. Hosseinipour (2003). A field-scale application of the Integrated Hydrologic Model. In Proceedings, Modflow and More 2003: Understanding through Modeling. Poeter, Zheng, Hill and Doherty, eds. www.mines.edu/igwmc/

Anderson, M.P. and W.W. Woessner (1992). The role of postaudit in model validation. Advances in Water Resources. 1992;15:167-73.

Anderson M.P. and W.W. Woessner (1992). Applied Groundwater Modeling: Simulation of Flow and Advective Transport. Academic Press. San Diego, 381pp.

Arbeitskreis "Hydrogeologische Modelle" der Fachsektion Hydrogeologie (1999) Hydrogeologische Modelle – Ein Leitfaden für Auftraggeber, Ingenieurbüros und Fachbehörden. Schriftenreihe der Deutschen Geologischen Gesellschaft; Heft 10; English translation by Johannes Riegger, "Hydrogeological models. Guidelines for clients, consultants, and regulatory officers. German Guidelines." www.fh-dgg.de/ ; www.tu-darmstadt.de/fb/geo/fhdgg/arbeitskreis-frame.htm, 2004, 36p.

Bay-Delta Modeling Forum – see CWEMF and www.cwemf.org/

Beven K.J. (2002). Towards a coherent philosophy for modelling the environment. Proceedings Royal Society of London, A., 458, 2465-2484.

Bredehoeft J.D. (2003). From models to performance assessment: the conceptualization problem. Ground Water, 41, 5, pp. 571-577.

California Water and Environmental Modeling Forum (2000). Protocols for Water and Environmental Modeling. Bay-Delta Modeling Forum. Ad hoc Modeling Protocols Committee. www.cwemf.org/

Camp Dresser McKee Inc. (2001). Evaluation of integrated surface water and ground water modeling tools. February 2001.

Environment Agency (2001). Guide to Good Practice for the Development of Conceptual Models and the Selection and Application of Mathematical Models of Contaminant Transport Processes in the Subsurface. National Groundwater and Contaminated Land Centre Report NC/99/38/2. July 2001. 105p plus appendices.

Environment Agency (2001). Environment Agency Framework for Groundwater Conceptual and Numerical Modelling. National Groundwater and Contaminated Land Centre. R&D Technical Report W214. 41p plus appendices.

Environment Agency (2002). Eds. Hulme, P. et al. Groundwater Resources Modelling Guidance Notes & Template Project Brief (Version 1). R&D Technical Report W213. Environment Agency. www.environment-agency.gov.uk/science/454158/110943/871689/?version=1&lang=_e

Environment Agency (2003). Risk-based Decision Making for Water Resources Licensing. Phase 2: Guidance on Conducting a Hydrogeological Impact Appraisal. National Groundwater and Contaminated Land Centre. R&D Technical Report W6-071/TR2. 66p plus appendices. May 2003.

Hassan, A.E. (2004a) Validation of numerical ground water models used to guide decision making. Ground Water, 42, 2, pp 277-290. March-April 2004.

Hassan, A.E. (2004b). A methodology for validating numerical ground water models. Ground Water, 42, 3, pp 347-362. May-June 2004.

Hill, M. C. (1998), Methods and guidelines for effective model calibration: U.S. Geological Survey Water-Resources Investigations Report 98-4005, 90p. http://pubs.water.usgs.gov/wri984005/

Hill M.C., Hugh Middlemis, Paul Hulme, Eileen Poeter, Johannes Riegger, Shlomo Neumann, Howard Williams, Mary Anderson. (2004). Brief overview of selected groundwater modeling guidelines. FEM-Modflow; International Conference on Finite Element Models, Modflow and More, 2004, Karlovy Vary, Czech Republic, 13-16 September, 2004.

Hulme, P., D. Johnson, S. Fletcher (2003). Challenging our misunderstandings through modelling - lessons learned from the Environment Agency national groundwater modelling programme. In MODFLOW 2003, Proceedings, International Groundwater Modeling Center (IGWMC), September 16-19, 2003.

LaBolle E.M., A.A. Ahmed, G.E. Fogg (2003). Review of the Integrated Groundwater and Surface-Water Model (IGSM). Ground Water, 41, 2, March-April 2003. 238-246.

Merrick, N.P, H. Middlemis, and J.M. Ross (2002). Groundwater modelling guidelines for Australia - the review process. International Association of Hydrogeologists Groundwater Conference “Balancing the Groundwater Budget”, Darwin, Australia, May, 2002.

Middlemis, H., N.P. Merrick, J.M. Ross and K.L. Rozlapa (2002). Groundwater modelling guidelines for Australia - an overview of the need for and use of the guidelines. International Association of Hydrogeologists Groundwater Conference “Balancing the Groundwater Budget”, Darwin, Australia, May, 2002.

Middlemis H., and N.P. Merrick. (2004). Australian groundwater modelling guidelines – lessons learned from the review process. FEM-Modflow; International Conference on Finite Element Models, Modflow and More, 2004, Karlovy Vary, Czech Republic, 13-16 September, 2004.

Murray-Darling Basin Commission (2001). Groundwater Flow Modelling Guideline. Report by Aquaterra, January, 2001. Download from:
www.mdbc.gov.au/publications/si_and_e_reports.htm

Pattle Delamore Partners (2002). Groundwater Model Audit Guidelines. Report submitted to New Zealand Ministry for the Environment, 73p plus 4 Appendices.

Refsgaard and Henriksen (2004) Modelling Guidelines - terminology and guiding principles. Advances in Water Research, 27, pp71-82. see also: http://harmoniqua.wau.nl/ or www.harmoniqua.org

Refsgaard, J. C., H. J. Henriksen, B. Harrar, H. Scholten, and A. Kassahun. (2004, in press). Quality assurance in model based water management - review of existing practice and outline of new approaches. Accepted by Environmental Modelling and Software.

Reilly, T.E. and Harbaugh, A.W. (2004), Guidelines for evaluating ground-water flow models: USGS Scientific Investigations Report 2004-5038.

Scholten H., R.H. Van Waveren, S. Groot, F. Van Geer, H. Wösten, R. Koeze, J. Noort. (2000). Good Modelling Practice in water management. In Proceedings, HydroInformatics 2000, 23-27 July, 2000, Cedar Rapids, Indiana, USA.

Scholten H., A. Kassahun, J.C. Refsgaard, Kargas, Gavardinas and Beulens. (2004, in press). The HarmoniQuA methodology to support multi-disciplinary model-based water management. Submitted to Advances in Water Research.

South Florida Water Management District (2002). Final Model Evaluation Report. EAA Storage Reservoirs – Phase 1. Central and Southern Florida Project. Comprehensive Everglades Restoration Plan. November 2002.

Van Waveren, R.H., S. Groot, H. Scholten, F. Van Geer, H. Wösten, R. Koeze, J. Noort. (1999). Good Modelling Practice Handbook, STOWA, Utrecht, The Netherlands (in Dutch). English version from www.landandwater.tudelft.nl/Downloads/GMP-uk.pdf

Web Sites for Selected Modelling Packages:

MIKE-SHE www.dhisoftware.com/mikeshe/

MODHMS www.modhms.com

(HMS www.hec.usace.army.mil/software/hec-hms/hechms-hechms.html)

Hydrosphere www.modhms.com

IHM www.intera.com/techology_ihm.php (see also Aly, 2003)

SFWMD and SFHSM www.sfwmd.gov/org/pld/hsm/hsm.html

GMS (Wash123D, Hydrosphere development) www.gms.watermodeling.org/

BRIEF OVERVIEW OF SELECTED GROUNDWATER MODELING GUIDELINES

Mary C. Hill

U.S. Geological Survey, Boulder, Colorado, USA, mchill@usgs.gov

Hugh Middlemis

Aquaterra, Australia, hugh.middlemis@aquaterra.com.au

Paul Hulme

Environmental Agency Science Group – Air, Land, Water; paul.hulme@environment-agency.gov.uk

Eileen Poeter

Colorado School of Mines, International Ground Water Modeling Center, Golden, Colorado, USA, epoeter@mines.edu

Johannes Riegger

University of Stuttgart, Stuttgart, Germany, riegger@iws.uni-stuttgart.de

Shlomo P. Neuman

University of Arizona, Tucson, Arizona, USA, neuman@hwr.arizona.edu

Howard Williams

Pattle Delamore Partners Ltd, P.O. Box 389, Christchurch, New Zealand, howard.williams@pdp.co.nz

Mary Anderson

University of Wisconsin, Madison, Wisconsin, andy@geology.wisc.edu

Abstract This paper briefly reviews 12 sets of guidelines for groundwater modeling. The guidelines originate from 8 countries on three continents. After reviewing terminology differences, shared, unique, and conflicting aspects of selected sets of guidelines are presented.

Key words: groundwater model guidelines; model calibration; inverse modelling; conceptual models; perceptual models.

FEM-Modflow; International Conference on Finite Element Models, Modflow and More, 2004, Karlovy Vary, Czech Republic, 13-16 September, 2004.

1. Introduction

Hydrogeological investigations and groundwater modeling are dynamic and inexact. They are dynamic in the sense that (1) the state of any hydrological system changes with time, (2) new scientific techniques with which to evaluate these systems are continually developed and (3) new data challenge previously held concepts about the systems. They are inexact in the sense that groundwater systems are complicated and are largely inaccessible, so we cannot evaluate them comprehensively in detail, and we invariably do not have sufficient data to do so (even if we had the ability).

Over the last 70 years, many ideas and procedures have been introduced to address hydrogeological investigations, including groundwater modeling. Beginning with the four-part series of articles concluded by Freeze et al. (1992), attempts have been made to present these ideas and procedures more comprehensively, and these efforts have evolved such that now there exist many sets of guidelines for the development of groundwater models.

This paper seeks to increase communication and coordination between guideline developers. This goal is thought to be advantageous for two reasons. First, these developers are some of the most active groundwater modelers in the world and it is likely that increased communication will improve all sets of guidelines. Through this effort it is hoped that groundwater modeling will mature more quickly into a method that can be used reliably to investigate and manage groundwater systems. Second, differences in the sets of guidelines are likely to cause confusion for governmental agencies that depend on the guidelines. Such confusion is likely, for example, when managing groundwater systems that cross national boundaries or when groundwater models are used in litigation. Awareness of differences in the sets of guidelines by the modeling community can help address the consequences of inconsistencies.

This paper identifies and presents information about selected sets of these guidelines. There are many sets of guidelines and many of the sets of guidelines are described in manuscripts of considerable length. It is not the intent of this work to consider all sets of guidelines or to comment comprehensively on the sets of guidelines considered. Instead, this work is intended to provide enough information about selected sets of guidelines to encourage communication between guideline developers and users.

Cumulatively, the sets of guidelines cover a wide range of the modeling process, such as determining the scope and objectives, conceptualizing the system, data management, model development, sensitivity analysis, simulating predictions, evaluating prediction uncertainty, documentation and reports, and review. An individual set of guidelines may have a narrower range. This paper includes information about the entire range of the modeling process.

After this introduction, the paper consists of sections 2 through 6. Section 2 lists the selected sets of guidelines considered here and the major people and institutions involved in their development and use. Section 3 lists major differences in terminology used in the different sets of guidelines and identifies the terminology used in this work. Sections 4, 5, and 6 categorize major ideas and procedures presented in the guidelines as (1) shared by most of the sets of guidelines, (2) unique, or (3) currently in conflict. Short discussions are included to clarify the ideas and procedures involved.

2. Sets of existing guidelines

The sets of guidelines included in this work were chosen mostly based on their level of development, prominence, inclusion of new or controversial ideas, or some combination of these considerations. At a minimum, the sets of guidelines included in this work were required to cover a broad range of the problems encountered when simulating groundwater systems. Most are in whole or in part applicable to many other types of systems as well. Table 1 lists the sets of guidelines, regulatory agencies that use the guidelines, and references. Table 2 lists contacts for the guidelines.

Table 1 Selected sets of groundwater modeling guidelines

[Governmental organizations: FH-DGG, Hydrogeology Section of the German Geological Society; USGS, United States Geological Survey; USNRC, United States Nuclear Regulatory Commission.]

Name for this report	Used or officially adopted by regulators?	Reference and Comments
AUS	Murray-Darling Basin Commission (MDBC); informally adopted by State water agencies	Middlemis, 2001. Download at: www.mdbc.gov.au/publications/si_and_e_reports.htm. Includes very helpful annotated bibliography and comparison of codes and GUIs.
UK	Environment Agency (England & Wales)	Environment Agency, 2002
GLUE	Ideas used in UK guidelines	Beven 2001, 2004
FH-DGG	Not officially adopted, but widely used	Arbeitskreis "Hydrogeologische Modelle" FH-DGG 1999, 2002; Riegger, 2004
A&W	Not officially adopted, but widely used	Anderson and Woessner 1992
USNRC	Hydrogeologic modeling strategy in contractor report is being evaluated and used by USNRC staff as a technical resource	Neuman and Wierenga 2003, NUREG/CR-6805 <http://www.nrc.gov/reading-rm/doc-collections/nuregs/contract/cr6805/> Includes some USGS guidelines.
USGS	Parts used by the US-NRC guidelines	Hill 1998, Hill et al. 2001, Tiedeman et al. 2003, Reilly and Harbaugh, 2004.
ASTM	Used by some at USEPA. Funded by USEPA, USGS, and Dept. of the Navy. Some reference in litigation and RFPs (Requests for Proposals).	Documented in standards. For example: D5447-93e1 , apply a model to a site D5490-93, compare model to data D5611-94, sensitivity analysis D5718-95, document application D5880-95, flow & transport modeling D5979-96, conceptualization D5981-96, calibration D6000-96, water level reporting D6170-97e1 , select code
DK	Danish EPA	Hans Jørgen Henricksen, GEUS
HarmonIT Harmoniqua Harmoni-ca Harmoni-rib	European Union	Generic Framework papers: http://www.genericframework.org/uk/techdocu.htm http://harmoniqua.wau.nl/Summary.htm quality assurance http://www.harmoni-ca.info/HarmoniCA/Public/index.php Software guidelines
NZ	Ministry for the Environment	Guidelines for audit and review of models. http://www.pdp.co.nz
USEPA	Draft being considered for significant future use by USEPA	Guidance for Environmental Models of the Agency's Council on Regulatory Environmental Modeling (CREM). http://www.epa.gov/osp/crem/library/CREM%20Guidance%20Draft%2012_03.pdf

Table 2 Contact information for the selected sets of groundwater modeling guidelines

[--, not applicable; ASTM, American Society for Testing and Materials; PNNL, Pacific Northwest National Laboratory; FH-DGG, .]

Name for this report	Web access	Contact person	Professional affiliation, email	Country
AUS	www.mdbc.gov.au/publications/si_and_e_reports.htm	Hugh Middlemis Noel Merrick	Aquaterra Simulations, hugh.middlemis@aquaterra.com.au University of Technology Sydney, nmerrick@uts.edu.au	Australia
UK	--	Paul Hulme Mark Whiteman	Environment Agency: Science Group, paul.hulme@environment-agency.gov.uk … Policy & Process (Hydrogeology), mark.whiteman@environment-agency.gov.uk	United Kingdom
GLUE	--	Keith Beven	University of Lancaster, k.beven@lancaster.ac.uk	United Kingdom
FH-DGG	www.fh-dgg/ak-hgm	Johannes Riegger	University of Stuttgart, riegger@iws.uni-stuttgart.de	Germany
A&W	--	Mary Anderson Bill Woessner	University of Wisconsin, andy@geology.wisc.edu University of Montana. gl_www@selway.umt.edu	USA
USNRC	--	Shlomo Neuman Phil Meyer	University of Arizona, neuman@hwr.arizona.edu PNNL, philip.meyer@pnl.gov	USA
USGS	http://pubs.water.usgs.gov/wri984005/	Mary Hill	U.S. Geological Survey, mchill@usgs.gov	USA
ASTM	www.astm.org	Jim Rumbaugh	Environmental Systems, Inc jrumbaugh@groundwatermodels.com	USA
DK	--	Hans Jørgen Henricksen	GEUS hjh@geus.dk	Den-mark
HarmonIT	http://www.harmonit.org	Mainly DHI, Delft Hydraulics, Wallingford. +11 European organizations	Gregersen jbg@dhi.dk harmonit@harmonit.org http://www.genericframework.org/uk/partners.htm	Europe
Harmon-iqua	http://harmoniqua.wau.nl/	Huub Scholten	Wageningen University, huub.scholten@wur.nl	Nether-lands
NZ	http://www.pdp.co.nz	Howard Williams	Pattle Delamore Partners Ltd, Howard.williams@pdp.co.nz	New Zealand
USEPA	http://www.epa.gov/osp/crem.htm	Pasky Pascual	USEPA, pascual.pasky@epa.gov	USA

3. Terminology

Table 3 lists inconsistencies in terminology encountered in the sets of guidelines considered and the terms used in this paper. Part I of the table lists terms for which the differences are just a matter of style or spelling. Part II lists terms for which the differences reflect fundamental concepts and ideas.

Table 3 Terms that are used differently in different sets of guidelines, and the term used in this paper.

[Terms used in this paper are listed in **bold**. For topics not covered in this paper, a term is not selected. Selected terms were determined in part by the opinions of the authors, which are identified in the third column using the following initials: ma, Mary Anderson; mh, Mary Hill; ph, Paul Hulme; hm, Hugh Middlemis; sn, Shlomo Neuman; ep, Eileen Poeter; jr, Johannes Riegger, hw, Howard Williams. ASTM, American Society for Testing and Materials.]

Meaning to be conveyed	Terms used in guidelines	Author opinions
Pat I. Superficial differences		
Subsurface water[1]	**Groundwater** Ground water Don't care or conflicted	sn mh, hm, ma, ph, ep, hw
Simulating, simulated, one who simulates	**modeling/modeled/modeler** modelling/modelled/modeller Don't care	 hm, mh, ep, ph, hw
Quantities measured in the laboratory or field	**Measurements** Observations	mh, ma, hw hm, ma
Measurements or values derived from measurements that are compared with simulated dependent variables	**Observations** Targets	sn, mh, ph, hw hm, ma
The calculated values that are compared to the [observations, targets]	**Simulated equivalents or simulated values** Predictions	mh, ep, hm, ph, hw
Calculated system state for future or hypothesized conditions	**Predictions** Alternative unknown	mh, ep, hm, ph, hw
Part II. Indicative of fundamental differences		
Qualitative description of the hydrology and hydrogeology to be represented in the groundwater flow model.	**Conceptual model** Hydrogeologic Model (HGM) Perceptual model[2]	mh jr ph
Quantitative description of the hydrology and hydrogeology to be represented in the groundwater flow model, including definition of hydrogeologic units, recharge distribution, surface water bodies, and so on -- everything but the values of parameters.	Hydrogeologic Framework Model (HFM)	
Quantitative description of the hydrogeology to be represented in the groundwater flow model, including definition of hydrogeologic units.	**Hydrogeologic Framework Model (HFM)** Hydrogeologic Model Concept Conceptual model[2]	mh, hm, ep jr ph, hw
Quantitative description of hydrogeology and the related processes including respective model calibration and evaluation; ready to use	Hydrogeologic Model	jr
Quantitative description of all other aspects of the system being represented	Hydrologic Model	mch
The degree to which a model application resembles or is designed to reproduce the details of the hydrogeological system.	Complexity Fidelity	hm, mh, ep, ph, hw ASTM, jr
The correspondence between the prediction of interest and the level of model complexity	Fidelity	ASTM, jr

Meaning to be conveyed	Terms used in guidelines	Author opinions
A test of the model application by checking if simulated values reasonably match a reserved data set that was excluded during model calibration.	Verification Validation Test	hm, hw FH-DGG mh, ep
Analysis of model results with respect to uniqueness, accuracy, sensitivity and model application range; quantification of uncertainty	Evaluation	jr, hw
Process of comparing simulated and measured values and changing the model to address inconsistencies.	Calibration Model refinement Parameter identification	mh, sn, ep ph, hm, hw
Process of adjusting parameter values to reduce inconsistencies in model fit.	Parameter estimation Calibration	mh, hm ph, hw
Spatial assignment of parameters for a unique and accurate relation between measurements and not directly observable values (volume / mass /energy flows) under hydrogeologic constraints	Calibration	jr
Use of optimization methods to adjust parameter values.	Inverse modeling Parameter estimation using optimization Parameter optimization	 mh, hm ph

1. In Great Britain and New Zealand, only 'groundwater' is used. Both options are used in Europe and North America.
2. Perceptual and conceptual models as presented in the UK guidelines are discussed in section 5.2.

4. Shared Ideas and Procedures

4.1 Overall perspective

The following statement, which is modified slightly from the AUS guidelines, states some basic ideas that are shared by all the sets of guidelines considered:

> *The aim of most guidelines is to reduce and reveal model uncertainty for the users of modeling studies, including resource management decision makers and the community. This is achieved by promoting transparency in modeling methodologies and encouraging innovation, consistency, and best practice. Guidance is provided to non-specialist modelers and auditors or reviewers of models by outlining the steps involved in scoping, managing, and evaluating the results of groundwater modeling studies. The guidelines serve modeling specialists by providing a baseline set of ideas and procedures from which they can innovate.*
>
> *The guidelines are intended for use in raising the minimum standard of modeling practice and allowing appropriate flexibility, without limiting necessary creativity or rigidly specifying standard methods. The guidelines also should not limit the ability of modelers to use simple or advanced techniques, appropriate for the study purpose. Techniques recommended in the guidelines may be omitted, altered, or enhanced, subject to the modeler providing a satisfactory explanation for the change and negotiation with the client and/or regulator as required. Not all aspects of the guidelines would necessarily be applicable to every study. It also is acknowledged that standardization of modeling methods will not preclude the need for subjective judgment during the model development process.*
>
> *The guidelines are to be applied to new groundwater flow modeling studies and reviews of existing models. The guidelines should be seen as a best practice reference point for framing modeling projects, assessing model performance, and providing clients with the ability to manage contracts and understand the strengths and limitations of models across a wide range of studies (scopes, objectives, budgets) at various scales in various hydrogeological settings. The intention*

is not to provide a prescriptive step-by-step guidance, as the site-specific nature of each modeling study renders this impossible, but to provide overall guidance and to help make the reader aware of the complexities of models, and how they may be managed.

4.2 Model Transparency

The goal of model transparency mentioned in section 4.1 is stressed in many guidelines and, indeed, is a major reason for the guidelines to be developed. Transparency means that the ideas and assumptions used to build a model application are clearly stated and can be tested. The complexity of the systems and the model applications and the tools used to develop groundwater-model applications (including guidelines, visualization software, database software, and so on) rarely result in applications that are completely transparent. However, the goal of transparency is important.

4.3 Valid conceptual models that start simple and build complexity as needed are crucial and fundamental

All sets of guidelines stress the importance of valid conceptual models. Most guidelines suggest some form of parsimony. For example, the UK guideline outlines how conceptual models and the model applications should be continually updated/refined from an initial “appropriately simple” approach. Model updates and refinements arise as ongoing modeling studies and more data on system responses to natural and imposed stresses produce improved understanding of the system processes and interactions. The UK guidelines note that “The first (conceptual) model is not the best and it is not the last”. In the FH-DGG guidelines, this is expressed in their “Hydrogeological Model” HGM as “Model Maintenance.”

A related concept is the step-wise method of model development: Refine the conceptualization (more simple or more complex) and(or) add more parameters as needed to obtain model fit.

Some guidelines also stress consideration of the information provided by the observations. The USGS guidelines present sensitivity measures of the information provided by observations, as discussed in section 5.

The UK guideline outlines a comprehensive modeling approach based on the need to develop understanding of the studied systems. Conceptual models are developed as quantitative descriptions of the real system using observed field values. These are then tested using a variety of methods including lumped water balances, purely investigative numerical models, and during the development of historical numerical models.

> *Investigative modeling* refers to building numerical models of alternative conceptual models to test hypotheses and to define key processes. Trial simulations are run to explore initial understanding without necessarily “calibrating” any model. In the FH-DGG guidelines, this is called the “scenario technique.”
>
> *Historical modeling* refers to building numerical models that adequately represent the historical behaviour and specifically the key flow mechanisms of the real system. Field data is used both as model input and to compare against the model outputs.

4.4 Consider predictions of interest in model development

All guidelines suggest the importance of considering the model and the calibration in light of predictions, but they differ on how this is accomplished. Some new procedures have been introduced in some sets of guidelines and are described in Section 5.

4.5 Use hydrogeologic data to constrain the model

The hydrogeologic data of concern generally includes the time-invariant data such as stratigraphy, layer elevations and extents, hydraulic conductivity data, and so on. It also can include some time-variable data such as recharge, stream-aquifer interaction, abstraction configuration and stresses, and so on. The emphasis placed on hydrogeologic data differs between the sets of guidelines. Especially, the proper role of hydraulic-conductivity measurements is in contention, as discussed in Section 6.

4.6 Use least-squares objective functions as one measure of model performance

All of the methods encourage the use of some type of least-squares or maximum-likelihood objective function to quantify how well the model fits the observations. These two are the same for a given regression if, as is common, the statistical parameters are known. Alternatives, such as the sum of absolute values, have been used, but rarely and mostly in research papers.

The types of least-squares objective functions commonly used can be classified based on how the quantities are included in the objective function (observations, prior information, or regularization) are weighted. Possibilities include: simple least-squares objective functions that have no weighting or equal weighting, weighted least-squares objective functions that have a diagonal weight matrix, and generalized least-squares objective functions that have a full weight matrix.

An alternative to a single least-squares objective function including all the observations is to divide the observations to create multiple objective functions (a recent reference of this approach is Vrugt and others (2003). The multi-objective functions are each least-squares objective functions, so the agreement cited in this section applies.

4.7 Use other measures of model performance

All guidelines make numerous suggestions about how to use evaluate models. Most performance measures are based on comparing observations to simulated values. Here are a few of the suggestions shared by most sets of guidelines considered.

Use more than hydraulic heads as observations

Hydraulic heads are the most commonly available hydraulic data in most systems. Many of the guidelines directly address the advantage of having other types of hydraulic data, such as flows, advective transport derived from concentrations, concentrations used directly, temperature, and so on. The procedures suggested for including these different types of observations vary; some of the ideas are presented in Section 6.

Use more than fit to observations to judge a model

To assist the end-user to assess whether model performance is acceptable and meets the level of complexity required, qualitative and quantitative model performance measures are proposed in many guidelines. For example, in the UK 30-year time-variant simulations are common and modeled results are routinely compared to long-term trends and seasonal behavior. The FH-DGG guidelines stress that consideration of the uniqueness and accuracy of model fit is essential. A way to determine these quantitatively is to display the objective function based on least squares versus parameter combinations. The USGS guidelines stress the normality, randomness, and magnitudes of the weighted residuals, and note that very good fits can result from undesirable fitting of observations errors.

Prescriptive performance measures should not be applied blindly, as model performance can only be gauged against observations that are usually imperfect and incomplete, and the model must replicate processes that might be poorly understood or inadequately measured.

The utility of some qualitative comparisons are in dispute. For example the issue of using contoured hydraulic heads is discussed in section 6.5.

4.8 Consider alternative models

All sets of guidelines stress the importance of considering alternative models because system dynamics are rarely clearly defined. The methods used to generate alternative models and evaluate the results vary. Both deterministic and stochastic methods are considered. Generation of alternative models is a very interesting problem that has not been addressed thoroughly.

All methods evaluate alternative models through comparisons with field data and eliminate or reduce emphasis on models that reproduce field data poorly. This idea was originally stressed by the GLUE developers as part of considering only models that adequately fit the observations, and is now widely accepted.

5. Unique Ideas and Procedures

Below are short statements from the authors of six of the sets of guidelines describing briefly what they see as unique about their guidelines.

5.1 AUS

Scoping a Modeling Study

The scoping process is a key initial step in a model study, with the outcomes being specific study objectives, model complexity, and the required/available resources of time, budget, data, and technical expertise.

Detailed information is provided in the AUS guide including, for a range of complexity, the broad data requirements, timeframes for model development, broad budget requirements, and examples of specific objectives, for use by project managers in scoping their project.

Model Complexity

The ASTM guides proposed the term “fidelity”, which was adapted to “complexity” for the AUS guide. In simple terms, model complexity can be described by the “quick-cheap-good” paradox. The end-user can readily obtain a model with one or two of these three attributes, but not all three. If a model is required to be done quickly, it also can be done cheaply, but the results may not be good enough on which to base important resource development or management decisions. Alternatively, if a good, reliable model is required, then it is not likely to be able to be developed quickly or cheaply. Thus, it is crucial to establish at the scoping stage the specific details of the study objectives, the water resources issues/scenarios, the model purpose, the development stages, and resources.

In less simple terms, the “quick-cheap-good” attributes are better defined in terms of a hierarchical scale of model complexity. The level of model complexity needs to be discussed and agreed upon by the end-user and the modeler to ensure that it suits the study purpose, objectives, and resources available for each study, including long term staged development and technology transfer.

Water managers also should be included in the scoping and model design process (if they are not already part of the project team), as they will use the model results to allocate water resources and/or to assess the impacts of proposed developments and/or to implement resource management policies. It is important for the overall project objectives that potential fatal flaws in the modeling approach are identified and rectified at an early stage, rather than presenting government agencies with the results of a study that may not be regarded as scientifically sound.

Model Reviews

The AUS guide proposes a unique model review framework and detailed checklists, with reviews recommended at all stages throughout the study, consistent with the objectives, scope, scale, and budget of the project. A model review provides a process by which the end-user can check that a model meets the project objectives. It also provides the model developer with a specification against which the modeling study will be evaluated. The level of review undertaken will depend on the nature of the project. Less complex models require less detailed reviews. Reviews necessarily add expense to the modeling process. The client and contractor must be clear at the outset as to which party is to bear the cost of each review. The reviews included in the AUS guidelines are listed in Table 4.

Table 4. Reviews prescribed by the AUS guidelines.

Type of review	Parts of model reviewed	Procedure provided in AUS[1]	Suggested reviewers
Appraisal	Report.	36 questions; App. E	Representative of stakeholders
Peer review	Report.	200 questions; App. E, plus 10 pass/fail criteria; App.G	Other professional modeler.[2]
Audit	Report; model data files, simulations, and output.	200 questions; App. E, plus 10 pass/fail criteria; App.G	Other professional modeler.[2]
Post-audit	Report model data files, simulations, and re-runs with actual stresses.	200 questions; App. E, plus 10 pass/fail criteria; App.G	Professional modeler.[2]

1. The listed appendices are in Middlemis (2001).
2. Attributes of suitable experienced model reviewers are summarized in Item 11 of Appendix C of the AUS guide

5.2 UK

The Environment Agency's Guidance Notes distil the practical experience of more than 20 modelers who have worked on groundwater resources projects over the last 30 years. Each of the 31 two or three page topics was written by an invited specialist and included modelers from consultancies, academics, and the regulator. Each topic is directly relevant to the Agency's operational use of regional groundwater modeling for water-resources management and the guidelines do not deal in any detail with source protection or contaminant transport modeling. Two appendices describe case studies that illustrate the integrated use of the various topics.

The UK guidelines use the terms *perceptual model* and *conceptual model*. Most guidelines refer to a conceptual model as some simplified understanding of the real system. However, in the various guidelines the term is used to cover both a qualitative and a quantitative understanding. In the UK these are distinguished. Beven (2002) points out that much more complexity is recognized than can be represented in a mathematical model. He refers to what we know about a system as the *perceptual model* and the mathematical representation as the *conceptual model.* The Environment Agency in the UK makes the same distinction between the qualitative understanding of a system – the model in your head – and the quantitative description of that understanding in the conceptual model (Hulme *et al*, 2003). This leads to a conceptual model that can be properly tested because it is described using numbers.

The UK guidelines are unique in their emphasis on post-project appraisal as an essential separate stage. The guidelines were written after a comprehensive review of all the Agency's time-variant regional models. The Agency has reviewed three recent models in adjacent chalk catchments and this has promoted debate on the scientific issues raised, for example, the estimation of time series of recharge (Environment Agency, 2004).

The guidelines include a Template Project Brief which provides an example specification of the purpose, approach, and outputs for each major task in a groundwater modeling study. These are presented not as a strict procedure but as a resource to be adapted accordingly.

5.3 FH-DGG

The FH-DGG guideline is based on the idea of a "Hydrogeological Model" (HGM), which aims to provide a consistent framework for the transfer of complex hydrogeological nature into a model. Its main focus is on the creation of a hydrogeological model concept that simplifies nature adequately with respect to the problem to be solved and the respective dominant hydrogeologic features. As the HGM is intended to enhance understanding of hydrogeological systems and the predictions of their behavior as well as to serve as a basis for analytical or numerical calculations, the calibration, evaluation, and possible re-iteration of the hydrogeological model concept prior to application is included. Thus, in the FH-DGG guideline, the terminus “Hydrogeological Model” is only chosen if the hydrogeological model concept is proofed by a sound evaluation consisting of an analysis of model results with respect to uniqueness, accuracy, sensitivity, and model application range as well as a quantification of uncertainty. That means the HGM is ready to use.

The FH-DGG guideline is not considered to be a strict recipe on modeling, but rather to provide a systematic framework for the generation of hydrogeological model concepts as well as for practical issues like problem specification, commissioning of work, efficient work flow, and structural quality assurance. The intention of the Guideline is to assist clients, consultants, and regulatory officers in groundwater resources in the assessment of the database, the choice of an adequate model approach based on the spatial scale and the data situation, as well as possible necessary revisions of the model approach.

The guideline also proposes a working and communication scheme throughout all fundamental steps in the construction and application of the HGM. By following this procedure, an efficient approach to model development should be guaranteed, irrespective of the individual hydrogeological situation and the posed problem. Particular emphasis is put on communication at specific milestones, where for quality assurance purposes a common work base must be formed jointly by clients, consultants, and regulatory officers. Thus, unnecessary iterations are avoided and necessary iterations emanating from an inadequate model approach are clearly identified.

5.4 USNRC

A comprehensive strategy for hydrogeologic modeling and uncertainty analysis has been developed by a U.S. Nuclear Regulatory Commission (USNRC) research contractor (Neuman and Wierenga, 2003). The strategy recommends that alternative models be evaluated using maximum-likelihood Bayesian averaging, and suggests rating different models using Kashyap’s criterion. This method includes the variance-covariance matrix on the parameters in model discrimination – models with parameters that are more precisely estimated are preferred, given a similar match to observations. The strategy uses Kashyap’s criterion to weight predictions from the different models to obtain a probability distribution for predictions that accounts for model structure uncertainty. The same weighting procedure can be used with other criteria, such as the AIC and BIC statistics (defined in most statistics textbooks and in Hill, 1998). USNRC staff are evaluating this strategy for use in reviewing performance assessments of nuclear facilities and sites.

5.5 USGS

The USGS guidelines are unique in the statistics presented for sensitivity analysis. These statistics can be used to measure the information provided by observations for parameters and predictions, and the importance of parameters to predictions. These statistics provide an effective way to improve the utility of the model for resource managers. For example, the model can be used to clearly indicate the value of additional field data and whether new data justify recalibration of a model.

The new statistics developed as part of the USGS guidelines are independent of model fit, which makes them useful even if the model is not yet calibrated. They also are well suited to evaluation of potential new data, for which the observation is not yet known.

The USGS guidelines provide a unique perspective on weighting of observations, prior information, and regularization. Using statistical theory, the guidelines discuss the importance of using weights to account for observation error in the model development effort. The weighting approach suggested provides a systematic way to include different kinds of observations in a single objective function, or in multiple objective functions.

5.6 New Zealand

The New Zealand groundwater model audit guidelines reflect a need to provide non-experts with tools to audit and review groundwater models. The guidelines include explanations of how models work and which models suit specific problems. Procedures with accompanying checklists of pertinent questions are listed such that auditors of models may determine whether there are any modeling errors, whether the results are meaningful in the context of the particular question being asked of the model, and whether model uncertainty is a result of parameter variability and measurement errors, or model assumptions.

6. Conflicting Ideas and Procedures

The following are active known disagreements reflected directly in the guidelines considered here, or simmering just beneath the surface within the broader community of groundwater modelers. A brief description of the positions is stated, and proponents are noted for positions expressed in sets of guidelines presented in this paper. Positions without a proponent listed are not expressed in the sets of guidelines considered here but are included so that opinions from elsewhere in the groundwater modeling community are expressed.

6.1 Proper role of hydraulic-conductivity measurements

Position 1: Hydraulic-conductivity measurements based on laboratory tests of field samples, slug tests, aquifer tests, and so on are not relevant enough to use in ground-water models of a significantly larger scale to support many defined hydraulic conductivity parameters. (USGS)

Position 2: It depends on the scale. Small-scale hydraulic conductivity measurements can be useful for macroscale structured models. Position 2 is true for regional models, where the effective parameters are found between the harmonic and the arithmetic mean of the local measurements depending on the anisotropy and the direction of flow with respect to the anisotropy axes. (FH-DGG)

Position 3: Hydraulic-conductivity measurements based on laboratory tests of field samples, slug tests, aquifer tests, and so on can be used in large-scale ground-water models to support many defined hydraulic-conductivity parameters. The differences in scale between the measurements and the model do not cause problems that make the relevance of the measurements questionable.

6.2 Accounting for observation errors

Position 1: Use weighting to account for random errors in the observations. This also provides a way to normalize observations that may have different units, and therefore cannot be accumulated directly into a single objective function. (USGS)

Position 2: Weights cannot be determined well enough with available data to be useful.

6.3 Proper role for pumping test head-change data in calibrating regional models

Position 1: All models, regional or local scale, should be calibrated to pumping test head-change data (for example, drawdown data).

Greenfields sites (where there is only limited and/or short term data available), commonly have data on the short term pumping of test boreholes, and the measurement of aquifer responses in observation boreholes. This position holds that even short term pumping test head-change data is suitable for model calibration, and the implication is that calibration to such data renders a model valid for regional scale simulations and/or impact assessment purposes. The corollary (ie. if Position 2 of section 6.3 prevails over Position 1), raises the question of to what extent is the common approach of undertaking these field investigations justified in the short or long term in relation to the substantial expense involved and the argument of Position 2 that the short term data is of limited value.

Position 2: Regional models should not be expected to be accurately calibrated to pumping test data, but should be (eventually) well-calibrated to large scale stresses and long term monitoring. (AUS, FH-DGG) A regional model is designed for regional scale investigations, and its setup typically involves non-homogeneous (and sometimes non-isotropic) conditions, as well as boundary inflows/outflows, recharge, and stream-aquifer interaction features that impose regional- and local-scale gradients and curvature on the water table. However, pumping tests are usually analyzed with a range of assumptions that include homogeneity, isotropy, infinite and/or fully penetrating boundaries, and other assumptions that are not consistent with the regional model setup. A short term and/or local scale pumping test does not stress the aquifer adequately to invoke regional-scale aquifer responses, which is what the regional-scale model is designed to investigate. Therefore, a regional model should not be expected to accurately reproduce local-scale changes in head in response to pumpage, and also be expected to be suitable for its prime purpose of regional-scale investigations. Monitoring data from large- scale and long-term pumping schemes (ie. flows, water levels, water quality, etc), however, is highly valued for calibrating models. The calibrated model should, however, be developed with parameter values that are consistent with the values obtained from any pumping tests, to help address model non-uniqueness issues.

6.4 Use of contoured head data

Position 1: While heads form a quantitative calibration target, subjective assessment of the goodness of fit between model and measured groundwater level contours also is important. (AUS)

Position 2: Trying to match head contours that do not reflect conservation of mass considerations generally is not helpful. (USGS)

Position 3: Contouring head data helps to interpret local hydrogeologic conditions, including definition of parameter distributions and boundaries. Yet as interpolation has no physical background for calibration, only measured point values should be used. (FH-DGG)

6.5 Establish specific goals for model performance measures

An example of such a goal is that the largest discrepancy between observed and simulated heads needs to be less than a specified amount.

Position 1: Set goals for model performance measures at the scoping stage of model development (AUS, A&W). Propose staged development with coarse initial targets to be met before invoking more accurate targets with each successive stage of refinement as understanding improves.

Position 2: It is unclear how to establish such goals and how relevant they are. The resource manager would probably be able to best suggest goals based on tolerable prediction uncertainty, but translating that into goals applicable to model calibration is not straightforward, and perhaps not even possible. (USGS)

6.6 Use optimization methods to estimate parameter values.

Position 1: Using optimization methods to estimate parameter values helps to enhance understanding of the system. Either gradient or global search methods can be used. This does not indicate that only one model is to be produced, only that intense investigation of one of several models can be informative. (USGS, US-NRC, ASTM)

Position 2: The use of optimization methods is misguided because it emphasizes a single model. It is better to generate random samples of the possible sets of parameter values, use the related simulations to calculate weighted least-squares objective functions, and use dotty plots to investigate the results. Models that match the data too poorly are eliminated from the analysis. (GLUE, UK)

Position 3: Optimization methods are useful in calibration to estimate selected parameter values. However, as the model results depend on other influences (aquifer geometry, boundaries, initial states, etc.), a proper evaluation is needed to quantify the uncertainties of the assumptions and their influence on the results. (FH-DGG)

7. Conclusions

The many sets of guidelines developed for groundwater models reflects the importance of groundwater to people, communities, nations, and the world. Working together will help groundwater modeling reliability improve. This paper introduces the conflicting ideas and procedures so that the community can begin to work together to understand and possibly resolve differences. Development of computer programs that make it easy to experiment with different approaches will facilitate joint experiments that will help to mature groundwater modeling.

8. References

Anderson, M.P. and W.W. Woessner, 1992, Applied Groundwater Modeling, Academic Press, San Diego, 381p.

Arbeitskreis "Hydrogeologische Modelle" der Fachsektion Hydrogeologie (1999) mHydrogeologische Modelle – Ein Leitfaden für Auftraggeber, Ingenieurbüros und Fachbehörden. *Schriftenreihe der Deutschen Geologischen Gesellschaft; Heft 10;* English translation by Johannes Riegger, „Hydrogeological models -- Guidelines for clients, consultants, and regulatory officers", http://www.fh-dgg/ak-hgm , 2004, 36p.

Arbeitskreis "Hydrogeologische Modelle und Grundwassermanagement" der Fachsektion Hydrogeologie, 2002, Hydrogeologische Modell – Ein Leitfaden mit Fallbeispielen. *Schriftenreihe der Deutschen Geologischen Gesellschaft; Heft 24*

Beven, K.J., 2001, Rainfall-Runoff modeling, the primer: West Sussex, England, John Wiley, 355p.

Beven, K., 2002, Towards a coherent philosophy for modelling the environment: Proceedings Royal Society of London, A., 458, 2465-2484.

Beven, K.J., 2004, A manifesto for the equifinality thesis: unpublished manuscript, 38p.

Environment Agency, 2002. Eds. Hulme, P. et al. Groundwater Resources Modelling: Guidance Notes & Template Project Brief. R&D Technical Report W213. Environment Agency (England & Wales).

Environment Agency, 2004, A comparison of Chalk Groundwater Models in and around the River Test catchment. *Environment Agency Science Project NC/03/05, Environment Agency (England & Wales)*

Freeze, R.A., Joel Massman, Leslie Smith, Tony Sperling, and Bruce James, 1992, Hydrogeological decision analysis, 4, The concept of data worth and its use in the development of site investigative strategies: Ground Water, 30(4):574-588.

Hill, M. C., 1998, Methods and guidelines for effective model calibration: U.S. Geological Survey Water-Resources Investigations Report 98-4005, 90p. http://pubs.water.usgs.gov/wri984005/

Hill, M.C., Ely, M.D., Tiedeman, C.R., D'Agnese, F.A. Faunt, C.C., and O'Brien, B.A., 2001, Preliminary evaluation of the importance of existing hydraulic-head observation locations to advective-transport predictions, Death Valley regional flow system, California and Nevada: U.S. Geological Survey Water-Resources Investigations Report 00-4282, 82p, http://water.usgs.gov/pubs/wri/wri004282/ .

Hulme, P., Johnson, D., Fletcher, S., 2003, Challenging our misunderstandings through modelling - lessons learned from the Environment Agency national groundwater modelling programme. In MODFLOW 2003 - Proceedings, International Groundwater Modeling Center (IGWMC), September 16-19, 2003.

Middlemis, Hugh, 2001, Groundwater flow modelling guideline: Murray-Darling Basin Commission, Perth, Australia, Project No. 125, 133p.

Pattle Delamore Partners, 2002. Groundwater Model Audit Guidelines. Report submitted to New Zealand Ministry for the Environment, 73p plus 4 Appendices.

Neuman, S.P. and Wierenga, P.J., 2003, A comprehensive strategy of hydrogeologic modeling and uncertainty analysis for nuclear facilities and sites:, 2003, NU-REG/CR-6805, U.S. Nuclear Regulatory Commission, Washington, D.C.

Reilly, T.E. and Harbaugh, A.W., 2004, Guidelines for evaluating ground-water flow models: USGS Scientific Investigations Report 2004-5038.

Riegger, Johannes, 2004, Hydrogeologic models - A guideline for clients, consultants and regulatory officers: unpublished manuscript.

Tiedeman, C.R., Hill, M.C.**,** D'Agnese, F.A., and Faunt, C.C., 2003, Methods for using groundwater model predictions to guide hydrogeologic data collection, with application to the Death Valley regional ground-water flow system: Water Resources Research, 39(1): 5-1 to 5-17, 10.1029/2001WR001255.

Vrugt J. A., H. V. Gupta, L. A. Bastidas, W. Bouten, S. Sorooshian, 2003, Effective and efficient algorithm for multiobjective optimization of hydrologic models, Water Resoures Research, 39 (8), 1214, doi:10.1029/2002WR001746.

APPENDIX B - Critical Review of Modelling Support Tool (MoST)

OVERVIEW

The HarmoniQuA project (www.harmoniqua.org) aims to provide guidance for multi-disciplinary teams working in model-based water management, supporting the work of these teams throughout the modelling lifecycle by means of a software tool, which uses XML to facilitate web-delivery. A simplistic description of the Modelling Support Tool (MoST) is that it is a workflow management tool. It outlines the overall process (flowchart), with descriptions of the five main steps, 50 tasks and many more activities (sub-tasks), plus outlines of best practice methods and associated guidance for specific activities. MoST is more than a simple workflow management tool, however, as it provides a means of recording the assumptions, data sources, model results and decisions throughout a study, with QA steps, in a manner that provides both a means of communicating progress and providing an audit trail, which helps deliver "transparency" (a much sought after attribute of model-based studies). The version of the software that was tested showed excellent functionality, which augurs well for the usefulness of the final tool.

Please note that the comments below are the outcomes of a critical review focussed on the groundwater (GW) domain, and do not necessarily consider or address every detailed aspect of the MoST system. The aim was to identify significant areas where the tool can be improved. Overall, this reviewer found that MoST is an excellent tool, and the developers should be proud of what they have achieved. Please note also that this draft version of the tool was reviewed in September 2004, and it is not due for completion until the end of 2005. It should also be noted that the preliminary comments were discussed in person with Huub Scholten and Ayalew Kassahun of the HarmoniQuA team in a process that proved to be of value to all parties.

REVIEW COMMENTS

1 Global spell check required (eg. "decompoese" should be decompose). Examples of other (less trivial) suggested changes:

- "modeling" with one "l" (not consistent with the commendable adoption of the Queen’s English as the standard by the HarmoniQuA team)
- "solution" definition only covers solute mixed with solvent, not problem solution.
- "upscaling" is not defined
- "zonation" is not defined, but the opposite of it is ("regionalisation"), both important in "parameterisation"; these terms also need synonyms from the German guide
- "Actors" (Task 1.1, Activities) should be better described as "Parties"
- "DPSIR framework" needs a definition (referred in Task 1.1 Activities)
- quotes ("") seem to be replaced by question marks (?) in the glossary in some cases
- Task 2.4 should say "head-dependent FLUX conditions"
- "Fall" is hypertext-linked to its definition of a drop in level, even when it is used in a generic (eg. “seasonal”) sense. This probably applies to other words. Another example is "current".
- Problems also occur with hyperlinked appearing text in the bibliographies, when they should not have hyperlinks.

2 The overall steps and tasks of the modelling process in the HarmoniQuA knowledge base are comprehensive. Perhaps I am being a little over-critical, but there could be more emphasis on the review stages at the end of Step 2 and Step 5. Specifically, there needs to be a detailed review of the Model Design Report (as I like to call it), after the data analysis and conceptualisation is undertaken, and again when the final report is submitted for any project.

This final review step specifically provides for a post-audit (or post-project appraisal in the terms of the UK Guidance Notes), which is actually another loop within the continuous improvement / quality assurance cycle of “plan-do-check-act”. In modelling terms, this becomes “conceptualise-simulate-review-refine”, with the implication that any model realisation is neither the best nor the last, but the latest representation of our understanding of the system with appropriate complexity. The post-audit is crucial in this long term model development/refinement process, and should have adequate emphasis. I note that there is a loop back from the end of a project, back though each stage of the process, which is very useful. Perhaps if I could suggest a discrete loop back from the very end of Step 5 to the very start of Step 1, appropriately labeled and described consistent with a “post-audit” approach.

3 Task 1.4 Determine Requirements - "quality" of the model study needs more definition/description. eg. discussion of the terms complexity and/or fidelity would be useful (ie. the degree to which the model is designed or able to represent reality), with content drawn from the AU, UK or DE guides

4 Task 1.5 Terms of Reference – while it can be argued that the required model accuracy cannot be determined exactly at this stage, a target accuracy could be outlined for later discussion and agreement. In my view, the Data Analysis task must be completed first to confirm a proper definition. The alternative would be to propose a nominal accuracy and allow for it to be discussed, revised and agreed at later stages. The last para of "Temporal and Spatial Scales" of Task 1.1 Activities begins to describe the interactive approach between the modeller and client to discuss and agree issues (what I refer to as the "Scoping" part of a study), but the issue needs to appear much earlier (in my view), as a major component of Step 1, not buried deep in the notes on space/time.

5 Step 1 Model Study Plan. In my view, what is called a Model Study Plan in Step 1 is really a Model Study BRIEF. There is confusion in MoST between Model Study Plan (first established at Task 1.7), and the revised draft Model Study Plan (Task 2.11) and the revised and approved Model Study Plan (Task 2.12). There does not seem to be adequate explanation of the overall principle that the Model Study Plan is updated at various times throughout the various Tasks. I believe that it should be given distinct titles, such as Model Study Brief (Step 1), Model Study/Design Plan (Step 2), Model Setup Report (Step 3), Model Calibration Report (Step 4), and Model Prediction Report (Step 4). These stages and titles are logical, and intuitively meaningful, whereas the adopted approach requires clarification of which “study plan version” is being referenced. There is a problem also with the Model Study Plan definition, which does not include the crucial component of the conceptual model, and yet the conceptual model is included in the activities for Task 2.11, which result in the revised model study plan. If the conceptual model becomes part of the model study plan (ie. at the end of Step 2 under my system, when the Model Study/Design Plan is devised), then the definition needs to reflect that approach. In my view the end of Step 2 is the appropriate place for a Model Study Plan (provided that conceptual model is included in the definition), or possibly a better title would be the Model Design Plan, as it sets the context and detail for Step 3 Model Setup. At the end of Step 2, the report on the model design is actually the outcome, in preparation for Step 3 Model Setup. The Model Design Plan eventually forms the first stages of the Study Report that is refined throughout the rest of the Study. The end of Step 1 is really a "Model Study Brief", in my view. The outcome of Step 1 could possibly be described as a "model study plan", but without known details on the complexity or accuracy required, the conceptual model details, the split sample data set to be used for calibration and validation (against the specified criteria, which are also not yet known in detail), which would usually be a subset of the available data set. However, I believe that a better term for this document is the Model Study Brief, as it provides the information for the tender process, which results in the tender outlining what the modeller proposes for their modelling approach, but the modeller would not spend the time and money to do all those things outlined in Step 1 without payment. In almost all cases, certainly in Australia, the client does not pay the modeller to prepare a proposal to address the Brief. If the client were to pay, then the process outlined in Step 1 of MoST would be acceptable, but the issue of confusion over "model study plan" definitions with/without conceptual model and complexity issues still needs to be resolved. There is one time where full definition of the Model Study Plan is possible at the start, and that is where the project is to update a model that has already been developed. In this case, a lot more is known about the system, and the procedures outlined in MoST should be appropriate.

6 Step 1 Model Study Plan. There is a related problem to item 4, and that has to do with improving the definition of the roles for this stage of the project. It is not clear from the content in this version that the tasks and activities in Step 1 need to be undertaken by the water manager, with advice from the Model Auditor on specialist issues. The UK, DE and AU guides are all much more explicit and detailed in this regard, and appropriate content should be incorporated into MoST from those sources.

7 Some Activities are repeated (eg. Problem Context for several different domains in Task 1.1), and one must look hard to ensure that one is considering the proper guidance for their study (eg. GW only or whatever). I think this may be just a familiarisation thing, and not a serious problem, and I believe that the final tool should filter the task/activity content for the profile of the user/domain etc. When navigating between screens on the guideline content (Steps, Tasks, Activities, Methods and Full Information), then the position within one menu should be maintained, rather than going back to the start of the menu.

8 Task 1.1. Conceptual model needs a lot more description for the GW domain. Time series data and System data also need more explanation for the GW domain. Communities and project managers (who are not modellers) are always asking for guidance on what data is required for a model. In particular, detailed data on groundwater pumping is typically not good quality, and yet it has a major influence on the model performance and underlying uncertainty. The information in Task 2.1 "Data to setup and run model" is probably required at the Task 1.1 stage. However, under the detailed content, MoST is confusing the framework data that does not change with time and the hydrological data that does change with time. This is an important concept that helps everyone understand the data needs, but it may be only universally applicable to the GW domain (eg. I can see that sediment transport in rivers is a process where the physical framework changes with time). I am not sure how this fits in with the MoST approach of separating system data, time series data and process data (there is no similar problem with the initial and boundary conditions data as currently described in MoST). This begins to be described in Task 2.2 Process raw data, but the concept needs to be introduced earlier, in my view.

9 Step 1 Scenarios - It is typical for only broad scenarios to be known at this stage, so I think it is impossible to be absolutely definitive about the scenarios required. However, it should be possible to outline the broad scenarios in terms of numbers of scenarios and probable configuration. The prediction scenarios should be defined exactly at the reporting stage (ie. after Task 4.11 Define Model Scope).

10 Task 2.9 “Assess soundness of conceptualisation” indicates that the modeller and/or the modeller's organisation should take a step back and review the suitability of the conceptual model. That is fine, but it needs to involve hydrogeologists and hydrologists and engineers and others from the project team, not just the modeller. Preferably, it should also involve the model auditor (working on behalf of the client) to put the conceptual model under critical scientific review. In fact, I believe that all the DECISION diamonds in the procedure should involve the project manager and the auditor, not just the modeller, as is the case for all the decision nodes as far as I can see.

11 Task 2.10; in the reference list, the website is defunct for Benchmarking Models for Water (BMW) Framework directive.

12 Task 3.2 Test Runs Completed has a Method indicated as Evaluation/Recommendation, which includes reference to a model journal (meaning a modeller’s journal, I believe; eg. a table of model run numbers and details for each run in terms of inputs and outputs; a tool such as InfoWorks from Wallingford Software is such a product). This should be introduced at Task 3.1, when the model setup starts. In my view, it is very important for the MoST tool to include a modeller’s journal tool, but I could not find one. A relatively poor example is given in the Australian Guidelines, but it provides some inspiration. This is not to be confused with the “model journal”, which I understand to be a report on the assumptions, data, approaches and decisions made throughout the project.

13 Task 3.3 Target Criteria has an Activity of Selecting observation data sets (GW domain), which I believe should be part of the data analysis in Step 2, which would be documented in the model study plan report at the end of Step 2. Assessment of observation data is properly carried out in Step 2, and a logical outcome is that split samples would be defined for calibration and validation.

14 Task 3.5 Review Model; I disagree strongly with the suggestion that only comprehensive (complex) models need independent review. Any level of complexity of model from Basic upwards may need review if the results are used as the basis for important decisions on resource management. In some cases (eg. Basic model), documentation of QA internally to the project team may be adequate, but there should be indication within the guide of the need for review at many stages through the project, provided that the project manager can over-ride the need on specific projects.

15 Task 4.6 Qualitative Performance Measures has no reference to the Australian Guidelines, which document some measures for the GW domain under discussion of calibration performance measures and assessment.

16 Post-Project Appraisal is a step that is not included in the model study closure task/activity (this is a key requirement from UK, although their guidance is also limited in this regard). In Australia, it is very common for models to be independently reviewed at the end of a project. This differs from a post-project appraisal in that it is usually done immediately at the end of the model study, whereas a post-project appraisal is usually done after some time has passed, and additional data is available to review the predictive performance of the model.

17 Two minor problems with the flow chart:

17.1 Task 3.1 has two "negative" answer loop backs, and they need to be labeled "terrible" and "not OK" (both are currently labeled "terrible") even though a hover on the link shows the status ("terrible" and "no") and the source.

17.2 Task 4.5 could possibly be labeled "Parameter optimisation and evaluate model performance" (as was the case with an earlier version of MoST). The later Task 4.7 of "Assess soundness of calibration" (and 4.9 assess soundness of validation) should remain.

18 Although the MoST tool is designed as to contain just text, to facilitate simple XML delivery, it would benefit from some graphics, especially to explain issues such as conceptual model and model performance assessment.

19 Task 4.11 has the aim of defining the model “scope”, but I find this term misleading in this context. I believe that the task is really defining the model capability in terms of process simulation within its “domain of applicability” to quote Refsgaard and Henriksen (2004). I believe that the use of the term “scope” may cause confusion with some of the activities that are carried out at the pre-commissioning stage, when the scope of the project is defined in terms of the study objectives, model complexity and resources available.

20 The ability to iterate back through steps and tasks is an excellent innovation in MoST. I believe that it could be further improved by providing some sort of iteration or loop count, so that the user is aware of how many times this particular task/activity has been visited. This should perhaps use some form of a hierarchical counter to indicate the “version number” and/or a date indicator.

21 MoST has a number of review and decision points, which could be hold points in the process. It is suggested that these hold points could be clarified by “traffic light” indicators (ie. red or yellow or green), so that the users can see in an instant the status of the hold point. For example, a red light could indicate that a review has not yet been completed by the auditor, a yellow light could indicate that the review has been completed but no fatal flaws were identified, so the model team can proceed cautiously, and a green light could mean that the review is completed, along with consultation, and the water manager has made appropriate decisions to progress the study.