Groundwater Modelling Guidelines for Australia: The Review Process

N. P. Merrick[a], H. Middlemis[b], and J. B. Ross[c]

[a] *National Centre for Groundwater Management, University of Technology, Sydney (nmerrick@uts.edu.au)*

[b] *Aquaterra Consulting, Perth (hugh.middlemis@aquaterra.com.au)*

[c] *PPK Environment and Infrastructure, Sydney (jross@ppk.com.au)*

Abstract: In a project sponsored by the Murray-Darling Basin Commission (MDBC), best practice guidelines have been developed for application to groundwater flow modelling projects. This paper discusses the detailed model review methodologies that form part of the guidelines, while a companion paper presents an overview of the guidelines. Solute transport methods and unsaturated zone modelling are not within the scope. Best practice modelling involves the peer review of a model at various stages throughout its development, to a level of detail consistent with the model study objectives, scope, scale, budget, and the model complexity. The lower the complexity of a model, the less detailed a review is required. The Australian guide proposes that reviews need to range from simple *model appraisals* using a checklist for simple models, through to more comprehensive *peer reviews* and complete *model audits* for more challenging complex models. An appraisal and a peer review would usually involve a review of a modelling study report, while an audit would also require an in-depth review of the model data files, simulations and outputs. Model reviewing provides a process by which the end-user can check consistently that a model meets the project objectives. Checklists and model performance evaluation criteria (documented in the guide) provide the model developer with specifications against which the modelling study will be reviewed and evaluated. The undertaking of a review necessarily adds expense to the modelling process. The client and contractor must be clear at the outset as to which party is to bear the cost of each review. Guidance is also provided regarding suitable attributes for model reviewers.

Keywords: Groundwater; Guidelines; Modelling; Review; Audit

1. INTRODUCTION

Groundwater flow modelling guidelines, developed for the Murray-Darling Basin Commission (MDBC), have widespread application for managing and undertaking groundwater flow modelling projects across Australia (Middlemis et al., 2000 and 2001). In order to promulgate the guidelines more widely, the Agriculture and Resource Management Council of Australia and New Zealand (ARMCANZ) is considering endorsing the guidelines for national adoption. MDBC is planning to distribute 5000 copies of a 20-page plain English summary, and 500 to 1000 copies of the 169-page technical document. The guidelines can be downloaded from the MDBC web site (www.mdbc.gov.au).

The guidelines are to be applied to new groundwater flow modelling studies and reviews of existing models. Solute transport and unsaturated zone modelling were not within the scope of this project. A national workshop was held to review the draft guide, and achieve consensus regarding practical and implementable guidelines. The guide is designed to be applied with flexibility to simple, small scale, small budget groundwater flow modelling jobs, as well as much larger and more complex regional modelling studies with substantial resource management implications.

A model review framework is a key element of the guide, with reviews recommended at the key stages throughout a study: conceptualisation, model plan and design, calibration/verification, prediction, and report completion. Of course, the amount of reviewing must be consistent with the objectives, scope, scale and particularly the budget of the project. The undertaking of a review necessarily adds expense to the modelling process, not only in having the review done, but also in the preparation of documents/files by the modeller for the reviewer. The client and contractor must be clear at the outset as to which party is to bear the cost of each review.

A model review provides a process by which the end-user can check consistently that a model meets the project objectives. It also provides the model developer with a specification against which the modelling study will be evaluated. The level of review undertaken will depend on the nature of the project. The lower the complexity of a model, the less detailed a review is required.

2. MODEL REVIEWS

2.1 Approaches

There appears to be very little published in support of the review process, apart from a useful checklist by Spitz and Moreno [1996].

The Australian guide proposes a hierarchy of reviews, guided by *pro forma* checklists. Model reviews may be made at any of four levels: model appraisal, peer review, model audit, and post-audit. An appraisal is less technical than a peer review, which in turn is not as detailed as an audit. The nominal difference between a peer review and an audit is that a peer review would usually involve a detailed review of a modelling study report, while an audit would also require an in-depth review of the model data files, simulations and outputs.

A model appraisal is made by a professional person, not necessarily with modelling skills, who represents the contractor's clientele (e.g. a government agency or the community). It might be possible with some training for a community representative to undertake an appraisal directly, or for the appraisal to be completed by group consensus.

A peer review or a model audit should only be done by an experienced groundwater modeller, different from the person who has developed the model. The conceptualisation stage of model development should be reviewed only by a competent hydrogeologist with local knowledge.

A post-audit is usually performed by the person who originally developed the model, but it could be done by a different professional modeller who has access to the model software and archived files.

2.2 Reviewer Attributes

It is difficult to be prescriptive as to the skills required of an appraiser/reviewer/auditor. For a single study, several reviewers might be required to cover the full range of tasks covered by the model. The selection of a reviewer (or reviewers) is best determined at the outset by mutual agreement between the contracting parties.

Criteria which could be used to identify suitable experienced model reviewers would include:

- Level of local hydrogeological knowledge (or access to such knowledge);
- General experience as a modelling specialist, and experience as a modelling team leader;
- Numbers of models developed of various degrees of complexity;
- Expert skills in specific modelling packages (especially the one to be used in the study) and/or specific model types (e.g. finite difference/finite element; 3D/quasi-3D/2D; flow/solute transport/heat/density coupled);
- Experience of modelling a range of hydrological and hydrogeological conditions (e.g. arid, tropical, temperate, irrigation, mine dewatering, dryland salinity, complex river-aquifer interaction).

3. MODEL COMPLIANCE

In order to flag serious model deficiencies, a checklist has been designed for Model Compliance. This consists of 10 critical questions with a PASS or FAIL response:

- Are the objectives of the modelling study stated clearly?
- Are the objectives satisfied?
- Is the conceptual model consistent with project objectives and agreed model complexity?
- Is the conceptualisation based on the full data set and a competent analysis of available data, and presented clearly?
- Has the conceptualisation been developed, endorsed or reviewed by a competent hydrogeologist (and revised if necessary)?
- Does model design/implementation conform with best practice?
- Is model calibration satisfactory?
- Are calibrated aquifer property values plausible?

- Does model prediction/application conform with best practice?
- Is there an excessive number of "Missing" or "Deficient" task performances marked on the Model Appraisal or Model Review Checklists?

The reviewer or appraiser can use this document to highlight any corrective action which must be undertaken before the model is deemed to be acceptable. This Model Compliance Statement could be disclosed to the modeller and to the client or, alternatively, used in isolation to provide a rapid overview appraisal of a model.

The Model Compliance Statement is best viewed as a summary document to be completed after a more exhaustive model appraisal or peer review.

4. MODEL APPRAISAL

An appraisal will often be undertaken by non-modellers (e.g. the client or project manager). To facilitate the appraisal process, and to encourage consistency between appraisers and between models, the guide encourages a systematic appraisal by posing 36 questions in a checklist.

The checklist asks the most important questions of a model for each of a number of categories: (1) The report; (2) data analysis; (3) conceptualisation; (4) model design; (5) calibration; (6) verification; (7) prediction; (8) sensitivity analysis; and (9) uncertainty analysis.

The checklist is limited to groundwater flow models and the applicability of a question will depend on the level of complexity of the model. It will highlight items for subsequent discussion between the client/principal and the contractor.

For each question in the checklist, the appraiser is asked to score the performance of the model or the modeller on a scale of 0 to 5. Some answers are of the YES/NO type, but others require judgement of the degree of effort or compliance (e.g. DEFICIENT, ADEQUATE, VERY GOOD). Items which are MISSING attract a zero score, unless the question is NOT APPLICABLE or UNKNOWN (in which case the question is voided). Appraisers are requested to mark the answer which best satisfies the question, and enter the appropriate score in the SCORE column. The maximum score for a question is 5 except for a question which is NOT APPLICABLE, in which case the maximum score is 0. For some less important questions, the maximum score is 3. (The appraiser is free to adjust the maximum score for any question, given its relevance to the model under consideration.) When the checklist is completed, the appraiser should record the TOTAL SCORE and TOTAL MAXIMUM SCORE, and report the performance as a percentage.

The appraisal can be done in part at any stage of the modelling process, or in full at the completion of the study.

Guideline:

"To encourage consistency of approach between appraisers and between models, for models of any complexity, a model appraisal should be conducted using a checklist of questions on (1) the report, (2) data analysis, (3) conceptualisation, (4) model design, (5) calibration, (6) verification, (7) prediction, (8) sensitivity analysis, and (9) uncertainty analysis. A guideline checklist for model appraisals is presented in Table E1 in Appendix E [Middlemis et al., 2000]. The appraisal could be undertaken by a trained community representative, by community group consensus, or by a professional person different from the person who developed the model."

5. PEER REVIEW

At present, peer reviews are sometimes but not always undertaken, at times internally to the model development team, and at other times externally by an independent reviewer. Sometimes they are costed "up front" in the proposal but often they are an after-thought, in which case there can be conflict as to who bears the cost.

A peer review is best done progressively through the modelling process at key milestones (conceptualisation, end of calibration, end of prediction, and after report completion). The experience of an expert reviewer will provide valuable feedback to the model's development, and will ensure a more reliable and useful product at the completion of the study. If left to the end of the study, there is a danger that a mistake could have been made early in the modelling process, which might invalidate subsequent work.

There is a shortage of highly experienced groundwater modellers in Australia who undertake peer reviews. While each has their own approach to a review, there is a case for standardising the approach in order to bring consistency into the review process, and to reduce the level of subjectivity. To this end, a checklist has been devised which poses questions for each of a number of categories: (1) The report; (2) data analysis; (3) conceptualisation; (4) model design; (5) calibration; (6) verification; (7) prediction; (8) sensitivity analysis; and (9) uncertainty analysis.

A checklist of 120 questions is provided in the guide. The model compliance and model appraisal checklists discussed in the previous sections are subsets of the peer review checklist. The same scoring system applies. The checklist is limited to groundwater flow models.

The checklist is designed for a high complexity model. For a model which is deliberately lower in complexity, the reviewer must be conscious that many of the questions will be NOT APPLICABLE and should not be scored.

Answers to the questions in the peer review checklist will encourage focus and balance in the reviewer's report. The reviewer cannot always assess the accuracy of model outcomes, but can offer an opinion on the plausibility of reported results. The peer review report should follow a similar structure to the model report, as outlined in the guide [Middlemis et al., 2000]. The *Introduction* should include a clear statement on what documents and other materials were provided for the review.

It is envisaged that the full scorecard will not be disclosed by the reviewer, as some reported deficiencies might be due to the complexity of the model being developed rather than poor performance on the part of the modeller. To avoid mis-interpretations by third parties, it is better for the reviewer to use the checklist as a systematic evaluation tool which can guide his/her review report, and ensure fair treatment and consistency across different reviews.

Of their nature, systematic reviews can easily tend to be negative as "no stone is left unturned". It remains to be seen whether modellers will embrace this review process or be overly defensive.

Guideline:

"To encourage consistency of approach between reviewers and between models, for models of medium to high complexity, a peer review should be conducted using a checklist of questions on (1) the report, (2) data analysis, (3) conceptualisation, (4) model design, (5) calibration, (6) verification, (7) prediction, (8) sensitivity analysis, and (9) uncertainty analysis. A guideline checklist for peer reviews of high complexity models is presented in Table F1 in Appendix F [Middlemis et al., 2000]. The review could be undertaken by an experienced modeller, different from the person who developed the model."

6. MODEL AUDIT

A model audit is conducted by a person other than the original modeller. The audit should consist of all aspects of the peer review in addition to the issues discussed below. Model audits are rarely done, except in-house as part of a quality control system. An internal audit is best done progressively through the modelling process in order to capture any inadvertent mistakes which might invalidate subsequent work. An external audit is likely to be done only after model completion (if at all).

A complete set of datafiles for at least one representative simulation should be provided to the auditor, so that he/she can verify that the model structure is as reported and runs successfully without numerical errors or mass imbalances.

The construction of a model using a graphical user interface (GUI) will make a model more open to audit. The auditor should proceed systematically through each GUI menu to check that the digital representation of the model matches the information provided in the report, or in working documents if the audit is internal. The auditor should check that all processes identified in the conceptual model are in fact activated and populated with data. For large Modflow (or other) models, stress package input files are often provided externally of the model GUI. In that case, the auditor would require documentation on the pre-processing software used to generate the data files. An auditor cannot vouch for the absolute accuracy of all datasets, and all model outcomes, but should apply plausibility criteria in all cases.

It is not possible to present in a model report the full detail of a model – particularly the spatial distributions of aquifer properties and layer elevations, and the temporal distributions of applied stresses. The auditor should pay particular attention to unreported features of the model. The auditor should also scrutinise the settings of switches or options in model packages or process algorithms, to ensure that the process is being simulated in the manner intended by the modeller. For example, some codes provide an option for either rate-limited or unlimited stream leakage when groundwater level drops below the stream bed. The auditor should comment on whether the representation of a particular model feature (e.g.

lateral inflow) would be better handled by an alternative mechanism (e.g. prescribed flow, fixed heads, or general head boundary).

As the occurrence of dewatered cells can cause some codes to become unstable, or erroneous, the auditor should pay close attention to their existence and evolution, and to the way in which the modeller has handled the affected cells.

Rough estimates of the components of the water balance should be made as a check on the values produced by the model. In some cases, fixed head cells might be responsible for an unrealistic inexhaustible supply of water.

An external audit is usually done at "arm's length", but there is a case for discussion between the auditor and the modeller before the audit report is written, in the event of unusual data handling or use of an innovative approach outside the experience of the auditor. An opportunity for communication should not be stifled.

Guideline:

"For medium and high complexity models, an internal model audit should be carried out progressively as part of an in-house quality control programme. An external audit would be warranted only in the event of an adverse peer review, or when a model is central to a matter destined for litigation."

7. POST-AUDIT

A post-audit describes the process of revisiting the modelling study several years after it is completed, to assess the accuracy of model predictions. This is essentially an alternative method of model verification, or of assessing model uncertainty, but it can be carried out only in hindsight, and therefore is not immediately useful for every modelling study. Before post-audits may be carried out, sufficient time must be allowed to gather data on the actual climatic/hydrological conditions and pumping regimes that have occurred. This will preferably include distinct hydrological conditions compared to the data set used for model calibration, which will also allow assessment of model non-uniqueness.

There have been few post-audits reported in scientific literature [Anderson and Woessner, 1992], and those that have generally showed that the model did not accurately predict the future, for two main reasons:

- inaccurate predictions resulted from poor guesswork in relation to future stresses (notably pumping rates and climatic regimes);
- inaccurate predictions were partly caused by errors in the conceptual model.

The first issue can be addressed at the post-audit stage by a "blind verification", by re-running the "prediction" in hindsight, but using the actual stresses that occurred. A valid model will produce a system response that closely matches the measured data, and calibration performance measures can be used to assess model accuracy. If a good match is achieved only after some parameters have been adjusted, then the original calibration simulation should also be re-run and re-assessed.

This is effectively a repeat of the process of verification, although it also helps address the non-uniqueness issue. It also adds value to the overall project by using the model in management mode (i.e. continual development with new data), rather than just crisis mode (i.e. to "answer" a question and then shelve the model).

The fundamental lesson from reported post-audits is that a valid and complete conceptual model is essential for making accurate predictions. This means that the model must include all the essential features of the hydrogeological system to an adequate level of detail. For example, if the available data or hydrogeological understanding has not identified the importance of leakage from a feature such as an overlying clay unit, then long term release of water from storage in the clays will not be accounted for, and the model prediction of pumping impacts will always be in error. This provides further justification for the emphasis that has been given in the guide to the need for developing a valid and robust conceptual model as an essential first step in model development.

Guideline:

"For medium and high complexity models, a post-audit should be carried out several years after original development, as part of the ongoing use of the model as a management tool. Reviews of and adjustments to the conceptual model and the model calibration may be required, which relies on the model archive produced at the end of the original study."

8. CONCLUSION

Best practice guidelines, specifically developed for application within the Murray-Darling Basin, are likely to be adopted for groundwater flow modelling Australia-wide. The guidelines address the processes of model design, calibration, prediction, uncertainty management, reporting and review. There has been a particular emphasis on producing best practice guidelines that provide support to modellers, but are still meaningful and useful to the community. A national workshop process was used to develop consensus, and a "plain English" summary guide is being developed for use by the community.

The review process is a particular feature of the guidelines. It ranges from model appraisal using a checklist for simple models, through to more comprehensive peer reviews and complete model audits for more complex models. An appraisal and a peer review would usually involve a review of a modelling study report, while an audit would also require an in-depth review of the model data files, simulations and outputs. A model appraisal is made by a professional person, not necessarily with modelling skills, but preferably with training and/or experience in undertaking reviews, who represents the contractor's clientele. A peer review or a model audit should only be done by an experienced groundwater modeller, different from the person who has developed the model. A post-audit is usually performed by the person who originally developed the model, but it could be done by a different professional modeller who has access to the model software and archived files.

It is hoped that the review guidelines and checklists in the Australian guide will inject consistency into the review process, so as to reduce the level of subjectivity and ensure fair treatment of a model and a modeller. However, the review process must always acknowledge the non-prescriptive and creative aspects typical of groundwater modelling. Good modelling practice cannot be decomposed into a set of rigid rules.

It remains to be seen whether modellers will embrace this review process or be overly defensive.

9. ACKNOWLEDGEMENTS

This project has been funded by the Murray-Darling Basin Commission. The authors wish to thank in particular the members of the Groundwater Technical Reference Group, and the many stakeholders who participated in the national workshop.

10. REFERENCES

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

Middlemis, H., N.P. Merrick, and J.B. Ross, Groundwater Flow Modelling Guideline, Aquaterra Consulting Pty Ltd, Consultancy Report, November 2000.

Middlemis, H., N.P. Merrick, J.B. Ross, and K.L. Rozlapa, Groundwater Modelling Guidelines for Australia – An Overview, Proceedings, MODSIM 2001 International Congress on Modelling and Simulation, Canberra, December 2001.

Spitz, K. and J. Moreno, *A Practical Guide to Groundwater and Solute Transport Modeling*, Wiley, New York, 416pp, 1996.