Available online at www.sciencedirect.com
SCIENCE DIRECT
Environmental Modelling & Software 21 (2006) 602–614
Environmental Modelling & Software
www.elsevier.com/locate/envsoft

Position Paper[1]

Ten iterative steps in development and evaluation of environmental models

A.J. Jakeman [a,b,*], R.A. Letcher [a,c], J.P. Norton [a,c]

[a] *Integrated Catchment Assessment and Management Centre, Building 48A, The Australian National University, Canberra, ACT 0200, Australia*
[b] *Centre for Resource and Environmental Studies, The Australian National University, Canberra, ACT 0200, Australia*
[c] *Department of Mathematics, The Australian National University, Canberra, ACT 0200, Australia*

Received 5 January 2005; accepted 10 January 2006
Available online 20 March 2006

Abstract

Models are increasingly being relied upon to inform and support natural resource management. They are incorporating an ever broader range of disciplines and now often confront people without strong quantitative or model-building backgrounds. These trends imply a need for wider awareness of what constitutes good model-development practice, including reporting of models to users and sceptical review of models by users. To this end the paper outlines ten basic steps of good, disciplined model practice. The aim is to develop purposeful, credible models from data and prior knowledge, in consort with end-users, with every stage open to critical review and revision. Best practice entails identifying clearly the clients and objectives of the modelling exercise; documenting the nature (quantity, quality, limitations) of the data used to construct and test the model; providing a strong rationale for the choice of model family and features (encompassing review of alternative approaches); justifying the techniques used to calibrate the model; serious analysis, testing and discussion of model performance; and making a resultant statement of model assumptions, utility, accuracy, limitations, and scope for improvement. In natural resource management applications, these steps will be a learning process, even a partnership, between model developers, clients and other interested parties.

Keywords: Model testing; Verification; Uncertainty; Sensitivity; Integrated assessment; System identification

1. Motivation

The pursuit of good practice in model development and application deserves thorough and sustained attention, whatever the field. Good practice increases the credibility and impact of the information and insight that modelling aims to generate. It is crucial for model acceptance and is a necessity for long-term, systematic accrual of a good knowledge base for both science and decision-making. The complexity and uncertainty inherent in management for better sustainability outcomes make the pursuit of good practice especially important, in spite of limited time and resources. Natural resource management confronts a complex set of issues, usually with environmental, social and economic trade-offs. These trade-offs are characterised by interactions at many scales and often by scarcity of good observed data. Thus natural resource managers commonly have to trade uncertain outcomes to achieve equitable results for various social groups, across spatial and temporal scales and across disciplinary boundaries. This must be achieved on the basis of information that varies in relevance, completeness and quality.

The complexity of these situations has led to model-based approaches for examining their components and interactions, and for predicting management outcomes. There is wide agreement on the potential of models for revealing the

* Corresponding author. Integrated Catchment Assessment and Management Centre, Building 48A, The Australian National University, Canberra, ACT 0200, Australia.
E-mail address: tony.jakeman@anu.edu.au (A.J. Jakeman).
[1] Position papers aim to synthesise some key aspect of the knowledge platform for environmental modelling and software issues. The review process is twofold – a normal external review process followed by extensive review by EMS Board members. See the Editorial in this issue.

1364-8152/$ - see front matter
doi:10.1016/j.envsoft.2006.01.004

implications of assumptions, estimating the impact of interactions, changes and uncertainties on outcomes, and enhancing communication between researchers from different backgrounds and between researchers and the broader community.

Managers and interest groups can also potentially benefit from use of a model to define the scope of a problem, to make assumptions explicit, to examine what is known and what is not, and to explore possible outcomes beyond the obvious ones. If models are accessible enough, they can act as a medium for wider participation in environmental management. However, the pressing need to use models in managing complex situations, rather than in sharply defined areas of research, has resulted in people with little modelling or quantitative background having to rely on models, while not being in a position to judge their quality or appropriateness. Caminiti (2004) provides a resource manager's perspective on the difficulties of choosing the best modelling approach for catchment management, concluding that ''[m]odellers can help by trying to understand the needs and expectations of the resource manager, who may not have the technical knowledge or language to express them.'' Managers may also not initially understand their own needs fully, so modelling must be an iterative learning process between modeller and manager.

The uses of models by managers and interest groups, as well as by modellers, bring dangers. It is easy for a poorly informed non-modeller to remain unaware of limitations, uncertainties, omissions and subjective choices in models. The risk is then that too much is read into the outputs and/or predictions of the model. There is also a danger that a model is used for purposes different from those intended, making invalid conclusions very likely. Taking a longer-term perspective, such inadvertent abuses detract from and distort the understanding on which science and decision-making are built.

The only way to mitigate these risks is to generate wider awareness of what the whole modelling process entails, what choices are made, what constitutes good practice for testing and applying models, how the results of using models should be viewed, and what sorts of questions users should be asking of modellers. This amounts to specifying good model practice, in terms of development, reporting and critical review of models.

As a move in that direction, this paper outlines ten steps in model development, then discusses minimum standards for model development and reporting. The wide range of model types and potential applications makes such an enterprise prone to both over-generalisation and failure to cover all cases. So the intention is to name the main steps and give examples of what each includes, without attempting the impossible task of compiling a comprehensive checklist or map of the model-development process. Such checklists have been developed within certain modelling communities where particular paradigms are dominant. Thus the Good Modelling Practice Handbook (STOWA/RIZA, 1999), financed by the Dutch government and executed by Wageningen University, has a well developed checklist for deterministic, numerical models. The guidelines for modelling groundwater flow developed by the Murray-Darling Basin Commission (2000) in Australia provide another example. Our purpose, by contrast, is to point to considerations and practices that apply in a broad range of natural resource modelling situations.

It is hoped that this paper will prompt modellers to codify their practices and to be more creative in their examination of alternatives and rigorous in their model testing. It is intended to provide a synoptic view for model builders and model users, applying to both integrated models and models within distinct disciplines. It does not deal with the surrounding issue of the appropriate development and use of environmental decision support systems (e.g. Denzer, 2005), which in addition involve issues of user interfacing, software usability and software and data integration. The paper discusses good practice in construction, testing and use of models, not in their imbedding and use in decision support systems or with software interfaces more widely.

As already indicated, the idea of guidelines for good model practice is not new. Parker et al. (2002) call for the development of guidelines for situations where formal analysis and testing of a model may be difficult or unfeasible. They state that ''the essential, contemporary questions one would like to have answered when seeking to evaluate a model (are):

i) Has the model been constructed of approved materials i.e., approved constituent hypotheses (in scientific terms)?
ii) Does its behaviour approximate well that observed in respect of the real thing?
iii) Does it work i.e. does it fulfil its designated task, or serve its intended purpose?''

Risbey et al. (1996) call for the establishment of quality-control measures in the development of Integrated Assessment (IA) models for climate change, and suggest several features that must be considered:

- a clear statement of assumptions and their implications;
- a review of 'anchored' or commonly accepted results and the assumptions that created them;
- transparent testing and reporting of the adequacy of the whole model, not only each of the component parts;
- inclusion of the broadest possible range of diverse perspectives in IA development;
- supply of instructions to model end-users on the appropriate and inappropriate use of results and insights from the analysis;
- 'A place for dirty laundry', that is, for open discussion of problems experienced in constructing complex integrative modelling, in order for solutions to these problems to be found, and to facilitate the appropriate level of trust in model results.

Ravetz (1997), considering integrated models, argues for validation (or evaluation) of the process of development rather than the product, stating that in such circumstances ''the inherently more difficult path of testing of the process may actually be more practical''. Ravetz finds that in general ''the quality of a model is assured only by the quality of its production''.