QUALITY ASSURANCE IN THE DEVELOPMENT AND APPLICATION OF GROUNDWATER MODELS

P. K. M. VAN DER HEIJDE
International Ground Water Modeling Center
Holcomb Research Institute
Butler University
Indianapolis, IN 46208, USA

ABSTRACT Because increasingly wide-reaching decisions in groundwater management are based in part on the results of modeling-based analysis, it is crucial that the development and application of models be subject to rigorous quality assurance procedures. Before models are used, their credentials must be established through systematic testing and evaluation of the code's performance characteristics. In model testing, the commonly used graphic performance evaluation approaches should be augmented by quantitative, objective measures. This paper outlines quality assurance approaches to model development and application, discusses model testing approaches and presents the requirements for quantitative performance evaluation measures.

INTRODUCTION

Groundwater modeling has become an important methodology in support of the planning and decision-making processes involved in groundwater management. Generally, three phases in the modeling of real-world systems can be distinguished: development of the modeling methodology, developing of the modeling tools (i.e. model development), and the application of methodology and tools to real-world problems. Model development consists of a research component dealing with the quantitative description of the system under study, a software development component and model testing. Model development often is driven by the short-term and sometimes by the long-term needs of management.

To develop and apply effective software in analyzing alternative solutions to groundwater problems requires a number of steps, each of which should be taken conscientiously and reviewed carefully. A systematic, well-defined and controlled approach to all steps of the model development and application process is essential for its successful use in management. Quality Assurance (QA) provides the mechanisms and framework to ensure that decisions are based on the best available data and modeling-based analyses.

Recently, QA in groundwater modeling has been the focus of discussion by the Committee on Ground Water Modeling Assessment, Water Science and Technology Board, National Research Council (NRC, 1989). Furthermore, a comprehensive set of QA procedures has been adopted by the International Ground Water Modeling Center (IGWMC) (van der Heijde, 1989). This paper highlights some of the findings of the NRC committee regarding QA and discusses the model test procedures developed at the IGWMC as part of a general QA methodology.
QUALITY ASSURANCE DEFINITIONS

Quality assurance in groundwater modeling is the procedural and operational framework put in place by the organization managing the modeling study, to assure technically and scientifically adequate execution of all tasks included in the study, and to assure that all modeling-based analysis is verifiable and defensible. QA in groundwater modeling is crucial to both model development and application, and should be an integral part of project planning and be applied to all phases of the modeling process.

The two major elements of quality assurance are quality control (QC) and quality assessment. Quality control refers to the procedures that ensure the quality of the final product. These procedures include the use of appropriate methodology, adequate validation, and proper usage of the selected methods and models. To monitor the quality control procedures and to evaluate the quality of the studies, quality assessment is applied (van der Heijde, 1987). Each project should have a QA plan, listing the measures planned to meet the project's quality objectives (van der Heijde, 1989).

Many modeling studies are performed without adequate QA. A formal QA plan is often lacking and extensive QA assessment is frequently postponed until the project reaches its final stage. This is especially true for studies where models are applied to solve site-specific problems. In contrast, policies based on modeling assessments and affecting large constituencies are reviewed more critically before they are adopted. Increasingly, however, financial and criminal liability require modelers to apply rigorous QA procedures in all stages of their projects.

QA IN MODEL DEVELOPMENT

The objective of a groundwater modeling project is a prototype system containing selected elements of a real-world multielement groundwater system (van der Heijde et al., 1988) (Fig. 1). The conceptual model of the selected groundwater system forms the basis for determining the causal relationships among various components of the system and its environment. These relationships are defined mathematically, resulting in a mathematical model. If the solution of the mathematical equations is complex or when many repetitious calculations are required, it is necessary to write a computer code. The conceptual formulations, mathematical descriptions and the computer coding together constitute the prototype model (Fig. 1).

To determine the overall validity of a prototype model, answers to the following questions are required:

(a) Is the conceptual model valid for the prototype system it represents?
(b) Does the mathematical model truly represent the conceptual model, the processes involved and the stresses present for the various design conditions?
(c) Does the code accurately represent the mathematical framework of the model?

Determining the correctness of a model is basically part of the scientific discovery process and as such is a rather subjective process. The acceptance of a model, or for that matter a theory, by the scientific community is a gradual process. Two approaches are generally applied to acquire such acceptance: (1) the rather qualitative evaluation or review process, and code testing which includes quantitative evaluation of test results using performance measures (van der Heijde, 1989).

An important aspect of the acceptance process is adherence to good QA procedures for code development and maintenance. These include (van der Heijde, 1987):

(a) documentation of code development (record keeping);
(b) verification of program structure and coding;
(c) validation of complete software product (model validation);
(d) documentation of code characteristics, capabilities and use;
(e) scientific and technical reviews;
(f) administrative auditing.

In the course of its use, the code gains confidence by proving its reliability and applicability. This can be further improved by provisions for continuing support and maintenance.

FIG. 1 Model development concepts (van der Heijde et al., 1988).

MODEL TESTING

For the purpose of this discussion one should distinguish between model testing and code testing. Code testing is limited to establishing the correctness of the computer code with respect to the criteria and requirements for which it has been designed. Model testing is more inclusive (and often more elusive) than code testing, as it represents the final step in determining the validity of the quantitative relationships derived for the real-world prototype system the model is designed to simulate (Fig. 1).

To evaluate groundwater models in a systematic and consistent manner, the International Ground Water Modeling Center (IGWMC) has developed a model review, verification and validation procedure (van der Heijde et al., 1985). IGWMC model testing takes place on three levels:
(a) level 1: verification by testing against known closed-form (analytical) solutions for highly simplified systems;
(b) level 2: verification by using complex, synthetic problems specially designed for testing specific types of models;
(c) level 3: validation against experimental (field or laboratory) results.
Model validation

The term "validation" is defined according to the discipline in which it is used. In terms of software engineering such as adopted by the National Bureau of Standards, code validation is defined as "the determination of the correctness of the final software product with respect to user needs and requirements" (Adrion et al., 1986). Applying this definition to groundwater modeling, the objective of model validation is to determine how well a model's theoretical foundation and computer implementation describe actual system behavior in terms of the "degree of correlation" between model calculations and independently derived observations of the cause-and-effect responses of the prototype groundwater system (Fig. 1).

How "true" the model is in a scientific sense, and how reliable its predictions, depend on how closely the system conceptualization and quantification of the system processes follow reality, on the degree of uncertainty in the data, and on the thoroughness of data input in the model formulation phase.

Various methods are used to quantify or describe qualitatively this degree of correlation (e.g. Koch & Link, 1980). It should be noted that the actual measured data of both model input and system response are samples of the real system and inherently incorporate errors. Thus, model validity established in this manner must always be subjective.

To be rigorous, model validation requires testing over the full range of conditions for which the model is designed. Whether a model should be used for a particular application can be assessed by applying predetermined performance criteria, sometimes called validation or acceptance criteria.

For many types of groundwater models, extensive high-quality datasets are lacking. Furthermore, available datasets are often limited with respect to the variety of conditions and stresses that occur in the sampled system. Therefore, model testing is generally limited to verification, using existing analytical solutions and synthetic benchmarks, to code intercomparison and to in-depth post-audits of model applications (NRC, 1989).

Verification of program structure and code

In software engineering, verification is the process of demonstrating consistency, completeness and correctness of the software (Adrion et al., 1986). ASTM (1984) defines verification as the examination of the numerical technique in the computer code to ascertain that it truly represents the conceptual model and that there are no inherent problems with obtaining a solution.

In applying these definitions to groundwater modeling software, the objective of the code verification process is twofold: to check the logic, correctness and computational accuracy of the algorithms used to solve the governing equations, and to ensure that the computer code is fully operational (has no programming errors).

The computational algorithms embedded in the code are first tested using a set of problems selected to ensure that the code's main program and most of its subroutines, including all of those frequently called, are being used in the testing. The effectiveness of a verification exercise can be further enhanced by using so-called walk-throughs, a step-by-step analysis of the computer program's operation, using the data of the verification problem. However, walk-throughs are tedious and labour-intensive.

In a frequently used verification approach, the test problem describes a system for which the stresses and parameters are defined in every point of its space and time domains, and the response of such a system is described by a closed-form solution to the governing partial differential equation (IGWMC test level 1). The numerical model to be tested provides solutions to the same equation at a limited number of discrete
points in space and time. Assuming that the coding is correct, the differences between the system responses described by the analytical solution and the numerical model are due primarily to the approximate nature of the numerical method involved and to limitations in computer accuracy, and are generally not randomly distributed. In many instances, the magnitude of these differences is related to the resolution in the discretization used in the computational scheme (Lapidus & Pinder, 1982). Theoretically, if the resolution increases such that the spatial and temporal step sizes approach zero, the differences between the numerical and the closed-form solutions should disappear. An additional issue for testing on this level is the numerical consistency and stability of the model.

Another important verification approach is the use of synthetic datasets (IGWMC test level 2). Here, hypothetical problems are used to test special computational features that are not represented in simple analytical models, as in testing for irregular boundary geometries, varying boundary conditions, specific spatial and temporal behavior of sources and sinks, or certain heterogeneous and anisotropic aquifer properties. Such hypothetical problems might be based on the features of a single field site or a generic mixture of features of various sites. The resulting dataset is referred to as a synthetic test dataset.

The system used for such a test is defined by its system parameters and system stresses. However, information regarding the system response obtained independently from a numerical model is not available. Testing takes place either by evaluating model behavior regarding such aspects as numerical consistency and stability, or by comparing the discretized predictions obtained from a new numerical model with those obtained from a well-established numerical model using a high-resolution spatial and temporal discretization scheme. As the absolute "truth" for these hypothetical problems is unknown, only a comparative verification of a model can be obtained. Using this approach provides a relative benchmark.

If the simulation results in such a code intercomparison do not deviate significantly from each other, a "relative" or "comparative" validity is established (van der Heijde et al., 1988). However, if significant differences occur, an in-depth analysis of the results of simulation runs performed with both codes is called for.

The International Ground Water Modeling Center is developing a series of two- and three-dimensional synthetic benchmarks for transient flow and solute transport in saturated and unsaturated porous media. Various well-established two- and three-dimensional models will be used to derive the benchmark predictions. The three-dimensional models to be selected will represent the major mathematical solution techniques used in groundwater modeling, including the finite-element method, the finite-difference method and the random walk method. The design of the benchmark problems will draw on known characteristics of the various numerical methods used in solving the linear and nonlinear groundwater flow and transport equations.

Finally, the verification stage can be used to evaluate the operational characteristics of the code, such as the sensitivity of the code to grid design, to various dominant processes and to a wide selection of parameter values.

Verification/validation criteria

An important aspect of model verification and validation is the definition of informative and efficient measures for use as evaluation or performance criteria. Such measures need to characterize accuracy and stability of the solution derived with the numerical model over total space and time domains appropriate for the model and for the full range of parameter values that might be encountered in the systems the model has been designed to simulate.

Thus far, acceptance of verification and validation tests has been based primarily on
visual inspection of graphical representation of the prediction variable (Fig. 2). Such
inspection, and for that matter any set of quantitative measures to be derived, should
involve the overall goodness-of-fit, including maximum deviation, average deviation,
and spreading in deviation, both in absolute terms and relative to the value of the
prediction variable. Furthermore, the model's ability to follow closely the shape of
complex graphs should be considered, even if for part of the domain under
consideration some offset occurs (i.e. in a time-based graph phase change) (Fig. 3).

Most of the graphical goodness-of-fit determinations use single-dimension graphs, e.g.
head versus time or head versus distance. Thus far, multidimensional graphs have been
used primarily in validation tests, e.g. in the form of contoured deviations. Extending
these multi-dimensional graphical techniques to the verification process might
significantly enhance our ability to judge goodness-of-fit (Fig. 4).

Each of the three testing levels adopted by the IGWMC requires a different ap­
proach to the definition of these measures. At the first two levels of testing, the veri­
fication phase, statistical measures are improper, but statistical-like measures might
prove useful. For example, ASTM (1984) presents some expressions for a deviation
coefficient DC, e.g.

$$DC = \frac{(x - y)}{x} \times 100\% \quad (1)$$

where $x =$ predicted value and $y =$ measured value and

$$DC = \sum_{i=1}^{n} \left[\frac{(x_i - y_i)}{x_i}\right] \times 100\% \quad (2)$$

with $n$ the total number of computed deviations.

FIG. 2 Example of a comparison between analytical and various numerical solutions (from Beljin, 1988).
DISCUSSION

Literally, quality assurance assures the quality of the product (model) or activity of concern (modeling). A more workable description is that QA in modeling guarantees that the quality of the model-based analysis and advice (to decision makers) satisfies quantitative quality criteria or measures. The principle ideas behind QA are accountability and communication.

QA cannot always assure acceptable quality of the model or modeling study; however, adequate QA can provide safeguards against faulty models or improper modeling. Regulators and decision makers should understand that there is no way to
guarantee that modeling-based advice is entirely correct, or that the model used (or any scientific model or theory, for that matter) can ever be proven, verified or validated in the strictest sense of these terms. Rather, a model can only be invalidated by disagreement of its predictions with independently derived observations of real systems.

ACKNOWLEDGEMENTS The research described in this publication has been funded in part by the U.S. Environmental Protection Agency through Cooperative Agreement #CR-815363 with Butler University, Indianapolis, Indiana. Additional funding has been made available by the Holcomb Research Institute of Butler University.

REFERENCES


