# APPROACHES TO MODEL ABSTRACTION FOR PERFORMANCE ASSESSMENT

Richard B. Codell
U.S. Nuclear Regulatory Commission
Washington DC 20555
(301)415-8167

Budhi Sagar
Center for Nuclear Waste
Regulatory Analyses
San Antonio TX
(210)522-5252

## ABSTRACT

Performance assessments for waste repositories require a large number of evaluations of complicated models. Generally, models must be abstracted to their essence because of limitations on computer and time resources. Abstraction of models with acceptable results is possible by a number of methods. Advances in computer performance will reduce the level of abstraction needed.

## I. INTRODUCTION

Three aspects of assessing performance of the repository system make it a complex process: (i) coupling between engineered and natural (geologic) components, (ii) large uncertainty in characterization of geologic component, and (iii) large time and space scales leading to a significant number of future states (scenarios) for consideration. Performance assessments (PA) must incorporate parameter and model uncertainties, spatial and temporal variabilities, and even design alternatives. Typically, a Monte Carlo sampling approach is used to propagate uncertainty in parameter values over a range of potential scenarios. This approach requires that large number of system simulations be performed. To accommodate all system components into the simulation with their inherent uncertainties and develop estimates of long-term performance within practical bounds of computational resources, some abstraction of the complex models becomes necessary. Published Total System Performance Assessments[1,2,3,4] indicate the abstractions made, but their justifications are not always clear. We discuss in this paper several aspects of the abstraction process, with the goal of developing bases for such abstraction.

## II. METHODS FOR ABSTRACTION

While, model abstraction is used in nearly all PA analyses, a cogent and systematic description of abstraction methods is not available in the literature. The following are among the many approaches that are currently reported, several of which have been used in some form in recent Nuclear Regulatory Commission (NRC) PA's[2,5]:

- Intuitive simplification – Most PA's use models that are simple at the outset, based on constraints of computational efficiency and budget. The distinction between intuitive simplification and other forms of model abstraction is that the former is accomplished at the level of the model conceptualization, and is not necessarily based on the results of a more complete simulation. This technique is justified if the phenomenon can be treated in a very conservative fashion and still demonstrate acceptable repository performance. For example, some phenomena may be entirely omitted from the model or a lower dimensionality model adopted.

- Empirical models based on full models - Response surfaces or look-up tables can be developed from the results of the prototype models, and then used in the PA's. Empirical abstraction may not work well where the sub-models include strong nonlinear couplings.

- Equivalent parameters - A simpler model can be justified by appropriate choices of coefficients; e.g., (1) the dispersion coefficient to represent transport on a scale smaller than the model scale; (2) the characteristic curves of an equivalent porous medium exhibiting both fracture and matrix properties[6], and (3) renormalization, which develops equivalent model parameters on a course grid that give nearly the same results as a much finer grid[7].

• Direct propagation of variability and uncertainty – Uncertainty and variability of model parameters are usually propagated through models using Monte Carlo Techniques. There have been some limited successes with the direct propagation in one simulation of spatial variability[8] and uncertainty[9] through the model equations.

• Integration - Time and space integration during the formulation of models may result in acceptable abstraction. For example, cumulative release at the accessible environment over 10,000 years (as required by the EPA standard[10]) can be calculated directly, then it would be possible to avoid difficult transient solutions. Integration could also be performed over similar system components, as in the case of a repository with tens of thousands of waste packages. The repository could be broken down into a number of zones, within which there would be a "representative" waste package and environment, chosen so that the behavior of that waste package represented the average behavior of all waste packages in the zone.

• Dimensional analysis – The equations representing the conservation of mass, momentum and energy in the repository can be formulated in terms of dimensionless groups[11], based on the fact that if theoretical equations exist among the variables affecting a process, they must be dimensionally homogeneous. These dimensionless groups can often reduce the number of variables in the uncertainty analysis. Once dimensionless groups are created and screened, they can be used to develop an abstracted model by correlation; e..g., a response surface.

• Abstracted Models as Limits of Detailed Models - This has rarely been done in PA literature. It would be a formal method of abstraction in which the abstracted model is derived based on taking limits (in the mathematical sense) of a more detailed model; e.g., the derivation of sorption coefficients from a surface-complexation model[12].

## III. DEGREE OF ABSTRACTION

Abstracted models should adequately represent significant processes impacting performance. They must be capable of simulating the system to a level of detail that can distinguished compliance from non-compliance. Abstracted models that are overly simplified may not be useful for determining sensitivity to those parameters and mechanisms discarded. Abstraction of complicated models without significant loss of crucial information leads to refinement of performance estimates by allowing more Monte Carlo realizations, or more scenarios. However, there is a tradeoff between the gains from increased computational efficiency and increased resources needed for abstracted model development and testing.

The abstracted models can be tested against the detailed process models directly. It is primarily a matter of devoting the necessary computational resources to perform this task in order to ensure its dependability. If detailed process models already exit, the abstracted models should be tested against them. The demonstration of the technical defensibility of the abstracted models is related in turn to the defensibility of the detailed process models.

## IV. FUTURE NEEDS FOR ABSTRACTION

For the present, there will be a need for model abstraction to work within reasonable limits of computer resources. Increasing computer speed, lower cost and better numerical methods will reduce the need for abstraction. Very fast and inexpensive workstations arranged in clusters and massively parallel computers are ideal for Monte Carlo calculations often used in PA.

In order to estimate the potential for speed-up and the need for abstraction, we compare the computational effort from a recent PA to speeds available on current and projected computers. Actual speed comparisons must take into account factors such as the degree of vectorization (for vector computers, e.g., Cray), scalability on parallel processors, and limitations of storage devices, so we caution that these are only order of magnitude estimates. The recent NRC PA for the Yucca Mountain repository[5] required on the order of 100 CPU hours (single processor) on a Cray XMP-24 computer, even with models that were highly simplified. Using the speed from the LINPACK benchmark[13], we estimate that the NRC PA equaled roughly $10^{13}$ floating point operations (FLOPS) per second. Since so much of the computational effort was related to repetitive Monte Carlo calculations, we would expect the run times to scale linearly to the number of processors on a parallel computer. Recent calculations on massively parallel computers routinely have sustained speed in excess of $5 \times 10^9$ FLOPS per second[14], which would be equivalent to 2000 seconds for the NRC PA problem. High-end speeds for presently available machines exceed $10^{11}$ FLOPS per second[15], or 100 seconds for the NRC problem. Expected performance in the next few years is expected to top $10^{12}$ FLOPS per second[15], or 10 seconds for the NRC problem.

In addition to faster computers, we expect that better algorithms to evaluate the equations would increase computation speed substantially. For example, substituting an advanced conjugate gradient solver for an iterative

solver in the gas-flow module of the NRC PA reduced the run time for this calculation about a factor of 10. Similar improvements in other parts of the PA are likely to be found.

## V. CONCLUSIONS

We have discussed a number of ways that models can be abstracted. Used carefully, these abstractions can save significant computer resources for repository PA's. We expect that faster, cheaper computers and improved algorithms will reduce the need for abstraction and allow more repetitions in the Monte Carlo solutions. Estimated speed-up for the recently completed NRC Phase 2 PA on Yucca Mountain could be in the range of 4 to 5 orders of magnitude within the decade. Much of the possible speed-up could be consumed by allowing less-abstracted models into the PA calculations. The degree of abstraction will be balanced against the resources allocated to the problem.

## VI. DISCLAIMER

The opinions expressed in this paper are solely those of the authors, and do not necessarily reflect the official positions of the Nuclear Regulatory Commission or the Center for Nuclear Waste Regulatory Analyses.

## REFERENCES


1. Author, "Preliminary Performance Assessment for the Waste Isolation Pilot Plant" SAND92-0700/1-4 UC-721, Sandia National Laboratories (1993).

2. R.B. Codell, et.al. "Initial Demonstration of the NRC's Capability to Conduct a Performance Assessment for a High-Level Waste Repository", U.S. Nuclear Regulatory Commission, Washington D.C. NUREG-1327 (1992).

3. P.W. Eslinger, et al., "Preliminary Total-System Analysis of a Potential High-Level Nuclear Waste Repository at Yucca Mountain", Pacific Northwest Laboratory PNL-8444 UC-814 (1993).

4. R.W. Barnard, et al., "TSPA 1991: An Initial Total-System Performance Assessment for Yucca Mountain", Sandia National Laboratories SNL91-7295 (1992).

5. R.G. Wescott et.al., "NRC Iterative Performance Assessment Phase 2: Development of capabilities for review of a performance assessment for a high-level waste repository", U.S. Nuclear Regulatory Commission, Washington D.C.,(1995 in press).

6. E.A. Klavetter and R.R. Peters, "Fluid Flow in a Fractured Rock Mass", Sandia National Laboratories, SAND85-0855, Albuquerque NM (1986)

7. S. Mohanty, A.C. Bagtzoglou, "Effective unsaturated hydraulic properties determination with the renormalization group approach: Methodology" High-Level Radioactive Waste Management pp 2660-2668, American Nuclear Society, La Grange Park, Ill (1994).

8. S.P. Neuman and C. Loeven, "Exact conditional moment expressions for transient unsaturated flow in randomly heterogeneous soils with scalable moisture retention characteristics", Department of Hydrology and Water Resources, The University of Arizona (1994).

9. P.C. Robinson and D.P. Hodgkinson, "Exact solutions for radionuclide transport in the presence of parameter uncertainty", Radioactive Waste Management and the Nuclear Fuel Cycle, 87(4), pp 283-311 (1987).

10. EPA "40CFR191: Environmental Radiation Protection Standards for the Management and Disposal of Spent Nuclear Fuel, High-Level and Transuranic Radioactive Wastes", Environmental Protection Agency, Washington D.C. (1993).

11. R.D. Manteufel and M.W. Powell, "Thermosyphon analysis of a repository: A simplified model for vapor flow and heat transfer" High-Level Radioactive Waste Management pp 2207-2216, American Nuclear Society, La Grange Park, Ill, 1994

12. D.R. Turner, "A Uniform Approach to Surface Complexation modeling of Radionuclide systems", CNWRA 95 xxx, Center for Nuclear Waste Regulatory Analyses, San Antonio TX (1995)

13. J.J. Dongarra, "Performance of Various Computers Using Standard Linear Equation Software", CS-89-85, Mathematical Sciences Section, Oak Ridge National Laboratory (1995).

14. J.J. Dongarra, H.W. Meuer, E. Strohmaier, "TOP-500 Supercomputer Sites", RUM 40/94, Mathemati-

cal Sciences Section, Oak Ridge National Laboratory, (1994).

15. Office of Science and Technology Policy, "High Performance Computing and Communications", Federal Coordinating Council for Science, Engineering and Technology (1994).