AN INTEGRATED TESTING PROGRAMME FOR A NEAR FIELD RADIONUCLIDE MIGRATION CODE

P. BINKS[1], A. FAIRHURST[2], D. HOWARTH[2], P. N. HUMPHREYS[3]*, T. JOHNSTONE[3], M. JONES[3], E. KELLY[2], A. LEE[3], R. MCGARRY[3], M. RANDALL[3], D. SMALLEY[3], G. THOMSON[3], D. P. TRIVEDI[3].
[1]Company Research Laboratories, BNFL, Springfields Works, Salwick, Preston, UK. [2]Sellafield Research and Technology, BNFL, Sellafield, Cumbria, UK. [3]Environmental Assessments, BNFL, Risley, Warrington, UK. *Corresponding author.

ABSTRACT

The DRINK code is a 2D, biogeochemical transport code developed as a research tool to simulate the long term evolution of near surface LLW disposal sites and to generate gaseous and liquid source terms for far field studies. The code was recently upgraded to provide a more generic modelling tool with wider application to radionuclide migration scenarios. During the development of this code, the Generalised Repository Model (GRM), an integrated strategy has been employed to ensure the production of a fully tested, verified and quality assured product. This strategy is based around a code development protocol with three main components: quality assurance and documentation, verification and realism testing. Realism testing includes both peer review and model testing, with the latter including: experimental test cases; natural and anthropogenic analogues; field observations and finally uncertainty and sensitivity analysis. This paper describes the successful application of the protocol to the development and testing of the GRM code with specific emphasis upon verification and realism testing.

INTRODUCTION

In recent years BNFL have developed a suite of geochemical and biogeochemical transport codes to provide a mechanistic understanding of radionuclide release and transport. Principle amongst these is DRINK [1,2,3], a 2D, biogeochemical transport code developed as a research tool to simulate the long term evolution of near surface, LLW disposal sites and to generate gaseous and liquid source terms for far field studies. DRINK incorporates a number of functional units which model geochemistry, sorption, corrosion, microbiology, radionuclide decay, colloid facilitated transport, mineral precipitation/dissolution and gaseous release. The combination of these processes within one code allows the evolving geochemistry of a LLW disposal site and its subsequent impact upon radionuclide transport and release to be simulated. The Generalised Repository Model (GRM) is an upgrade of the DRINK code, and was developed as a more generic modelling tool with application to a wider range of radionuclide migration scenarios. During the development of the GRM it was recognised that there was a requirement for a structured approach to the development and testing of multi-functional software tools. In particular, issues such as verification and validation, fitness for purpose, peer review and realism, all need addressing in a transparent and structured manner. In view of this an integrated strategy was developed around a code development protocol to ensure the production of a fully tested, verified and documented product.

This paper outlines some of the major issues addressed during the development of the modelling protocol, details the application of the protocol to the GRM code and gives specific examples of successful test cases. Particular emphasis is given to the approach adopted to verification and realism testing.

Mat. Res. Soc. Symp. Proc. Vol. 506 © 1998 Materials Research Society

THE GRM CODE

GRM is a 2D, contaminant transport code which employs standard advection and dispersion equations to describe the transport of dissolved species and colloidal particulates. The 2D, areal flow field is either user defined or generated by MODFLOW[4]. Geochemical calculations are carried out using a modified version of the PHREEQE[5,6] speciation code with mineral precipitation and dissolution being determined by a fully automated reaction path model. Mineral equilibration can either be instantaneous and mediated by PHREEQE, or kinetically controlled. Redox reactions involving the electron donors, organic molecules and hydrogen, and the electron acceptors oxygen, nitrate, ferric iron and sulphate, are assumed to be microbially mediated and are modelled via the microbial growth equations. The redox potential within the model is determined from the most oxidising redox couple, using standard Eh functions [7]. Step changes from one couple to the next occur through microbial action removing the most oxidising couple. This allows the modelled redox potential to progress through the biologically mediated redox sequence [8]. The GRM microbial module is based on a number of existing models [9,10,11] and currently contains seven broad microbial groups carrying out a range of redox reactions [3]. The primary electron donors are hydrogen resulting from both corrosion and microbial activity and organic carbon originating from the hydrolysis of polymeric organic substrates classified as either cellulose, protein or fat. Corrosion is partitioned between the unsaturated and saturated zones with both aerobic and anaerobic corrosion modelled[1]. Sorption is taken to be rapid with respect to the solute residence times, consequently instantaneous equilibrium is assumed. Accounting for radionuclide decay within GRM is complicated by the presence of solute, sorbed and solid phases. On decay the daughter is redistributed between solute, sorbed and mineral phases according to the ambient bulk chemistry.

THE ISSUES OF VALIDATION AND TESTING

During the development of the modelling protocol two main areas for consideration were identified, firstly the issue of validation and secondly the approach to code testing. It was noted that there has been significant scientific debate as to whether it is actually possible to validate contaminant transport models [12,13,14], principally on the grounds that the future cannot be interrogated now. Much of the confusion surrounding validation and verification has arisen from the fact that in the nuclear industry mathematical models of reactor and fuel performance have tended to set the definitions in the context of computer modelling. Verification has been taken to mean: *checks that the numerical model faithfully reflects the mathematical model,* and validation to mean: *checks that the numerical model gives the level of predictive accuracy required.*

In the context of reactor modelling and other forms of modelling such as heat transfer and stress analysis, genuine blind testing of numerical models can be undertaken since a large number of repetitive tests can be undertaken and documented. Leaving aside the fact that radionuclide migration models are addressing much more complex and less well understood processes, the multiple blind testing approach cannot be done meaningfully and a consensus opinion has developed that validation is actually impossible [12,13,14]. It is also significant that in the more established discipline of petroleum reservoir modelling, extrapolation into the future is only seen as supportable for a period of time equal to about twice as long as the period for which the model has been history matched or calibrated. This is obviously impossible when codes are to be run over the time periods of interest to the nuclear industry.

THE MODELLING PROTOCOL

The discussion above raises the issue that if models cannot be validated then what should be done instead? Authoritative opinion has developed that validation should be replaced by a good modelling protocol which should include [13]: a complete description of model design; thorough assessment of model calibration; and uncertainty analysis. A somewhat similar viewpoint has been put forward by Jackson et al. (1992) [14] where a formal framework for addressing issues relevant to validation has been proposed. More radically, Konikow and Bredehoeft[12] recommend that we move away from the false confidence in model predictions which is encouraged by nominal validation programmes and instead move towards an understanding of the complex systems with which we are dealing.

The approach adopted here is broadly in line with these approaches with the adoption of a formal modelling protocol. To avoid confusion, the term validation is replaced with the concept of building confidence in model performance, within known boundaries, through the application of a rigid modelling protocol which is outlined in Figure 1. The foundations of this protocol are a stringent set of software development, quality assurance (QA) procedures which provide the verification and supporting documentation aspects. These are then complemented by realism testing which includes both peer review and model testing.

Quality Assurance

The software QA standards employed [15] require a considerable amount of effort to be invested in in-depth documentation of the code. This provides a comprehensive, transparent description of the code's components, structure, modelling approaches and the assumptions applied, as well as providing an audit trail for all code testing and the basis of the peer review process.

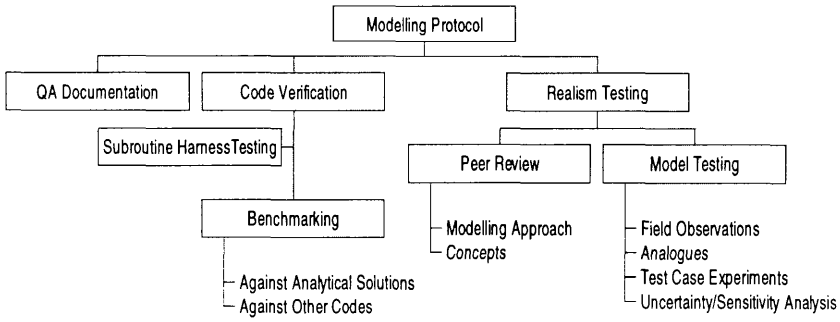

Figure 1. Model Testing Protocol

Verification and Benchmarking

Verification of the code is carried out in terms of checking that the numerical model faithfully reflects the mathematical model. This is carried out through appropriate subroutine harness testing, comparisons against relevant analytical solutions and through inter-code comparisons. Subroutine harness testing involves using harness code to isolate a subroutine to allow testing

against hand calculations etc. Benchmarking involves the running of the code against appropriate analytical solutions and runs from other similar codes. Benchmarking against analytical solutions provides insight into the impact of the inherent approximations of the numerical approaches employed. Benchmarking against other codes on the other hand provides insight into how a code performs when compared with its international 'peers'. To date extensive use has been made of analytical solutions [16,17] and international inter-code comparison exercises [18,19]. In addition more specific comparisons have been made focusing on particular capabilities of the code such as cation exchange [20] and microbial degradation of organics [21] with good results in both cases (Figures 2 and 3).

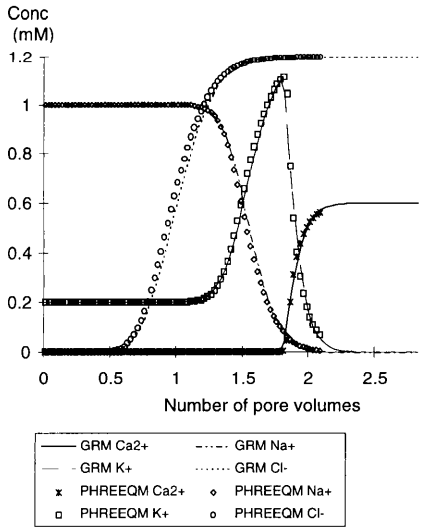

Figure 2. Modelling cation exchange using GRM and PHREEQM

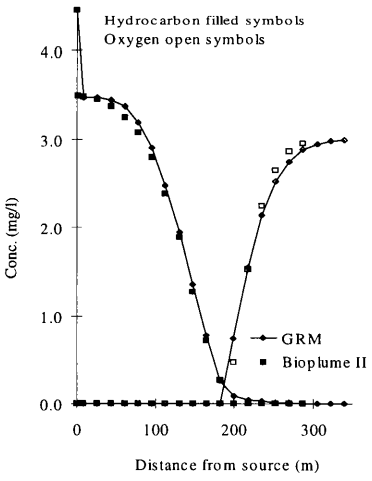

Figure 3. Comparison between GRM and BIOPLUME II

Realism Testing

The realism testing splits into two main components, peer review and model testing. The aim of realism testing is to ensure that the codes are sufficiently detailed, both conceptually and numerically, to adequately represent the situations modelled. This contrasts with the aims of code verification which determines that the numerical model (the coding) faithfully reflects the mathematical model. It is entirely possible that a fully verified code is of no practical use, because the mathematical model does not adequately represent the situation it was intended to model. Realism testing is intended to provide confidence that this is not the case through expert peer review and testing against experimental data where appropriate.

The aim of peer review is twofold, firstly to improve confidence in the work and secondly to ensure and demonstrate a high technical standard. The peer review component tests, through expert elicitation, that: the assumptions made in model formulation are supportable; the

mathematical model is fit for purpose; the techniques used to reflect the mathematical model into computer code are fit for purpose.

Model testing has four major components: experimental test cases; natural and anthropogenic analogues; field trials; and finally uncertainty and sensitivity analysis. The use of experimental test cases and analogues will be discussed here, field trials are site specific and cannot therefore be discussed in general terms and work on uncertainty and sensitivity analysis is currently being initiated. The general approach to the experimental test cases is a structured one which begins with highly characterised experiments and then moves on to more complex experiments which merge processes being investigated. In parallel with these well characterised short term experiments larger scale, less well characterised, trials are included. These trials are best suited to whole code testing since at larger scale it is more difficult to control experimental conditions. The time and physical scale of the tests are further increased within the protocol by including analogue studies.

This graded approach to experimental test cases can be clearly seen in the gas generation tests currently being investigated. These begin with simple, single substrate tests where microbiology and no chemistry is modelled [22] (Figure 4a), these then move onto more complex tests involving cellulose degradation [23], here a series of microbial groups are involved, the pH is fixed but carbon dioxide solubility is modelled (Figure 4b). The final set (Figure 4c) involve gas generation from simulated LLW [24] where microbiology, corrosion and an evolving chemistry are modelled simultaneously. All these tests are on a small scale (<10 litres), to increase the scale a set of LLW tests at a larger scale (250 litres) have recently been commissioned but no data is available to date. Experimental test cases are not only confined to microbial aspects of the code, a series of column studies investigating transport and sorption are underway (Figure 4d) and test cases such as the reactive barrier case described by Fryer and Schwartz [25] have also been employed (Figure 4e).

Work on analogue studies has been confined to examples available in the scientific literature. Although a number of natural analogues have been extensively studied by the nuclear community [26,27] anthropogenic analogues have proved to be more useful for model testing GRM. This is because there is a better fit with the processes of interest, anthropogenic analogues are more dynamic both spatially and temporally and the data set quality is generally better. Analogues currently under investigation include the Savannah River F-Area [28] which is an abandoned seepage basin previously used for the disposal of low-activity liquid wastes from the chemical separation areas at the U.S. Dept of Energy's Savannah River site. Initial modelling work has concentrated upon uranium transport and sorption. The nitrate reduction case described by Postma et al. [29] has also been used as a test case (Figure 4f). This test case is based on an aquifer contaminated with nitrate emanating from agricultural areas. A sharp redox boundary is seen in the aquifer where oxygen and nitrate are reduced through pyrite oxidation. This analogue allows GRM's ability to model microbial oxidation of reduced minerals to be tested.

This combination of well characterised short term experiments, large scale experiments and analogues studies provides a complete spectrum of tests. In the small scale short term experiments there is a high degree of certainty on the initial conditions and a full data set can be collected. Successful testing against these experiments provides confidence that the code can accurately represent the processes involved. These systems are, however, idealised representations of the "real" world and the short time scales may lead to an emphasis on modelling short term phenomena which are not relevant over a longer time scale. There are none of these problems when analogues are used for model testing. Here there is the opportunity to test the code against a real system. The time scales are longer and nearer to the time scales of interest and all the problems of heterogeneity are present. They allow a true test of the fitness for purpose of the

code. There are problems, however, which arise from the fact that with analogues it is almost impossible to get a full data set. Often only snap shots of the site at various time intervals are available. These uncertainties mean that the goodness of fit which is taken to be acceptable may be relatively poor. The combination of intensive, idealised, short duration experiments and the longer time scales of analogue studies overcomes the realism problems inherent in the laboratory work and the data set quality aspects of the larger scale, more realistic tests.

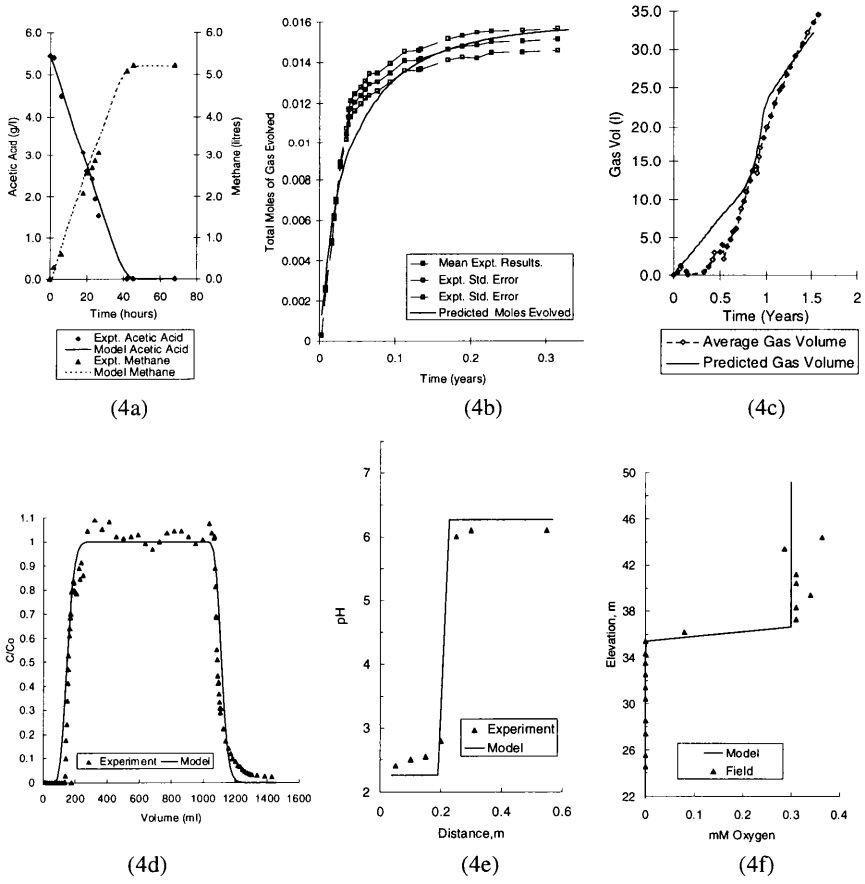

Figure 4. (a) Methane generation during acetic acid degradation, (b) Gas generation during anaerobic cellulose degradation, (c) Gas generation during LLW degradation, (d) Europium sorption and transport in a column of washed sand at pH 4.5, (e) Changes in pH down a calcite column receiving an acidic solution of $Fe(ClO_4)_3$, (f) The oxygen profile within a nitrate contaminated aquifer.

CLOSING COMMENTS

As mathematical models become increasingly complex a comprehensive approach to their quality assurance, testing and development becomes essential. It is particularly important that issues of validation are addressed in a transparent manner and that the false confidence inherent in nominal validation programmes is avoided. It should be recognised that the future cannot be interrogated now and consequently any modelling extrapolation into the future represent an approximation based upon current understanding. In order to avoid some of these pitfalls the approach adopted during the development and testing of the GRM was one of building confidence within known boundaries. Whilst recognising that once the code is applied outside these boundaries of either processes modelled or length of simulation, the degree of confidence which can be ascribed to the results has to be adjusted accordingly.

To allow this confidence building exercise to proceed a model testing protocol was developed. This protocol provides a rigid, well ordered, framework within which the development and testing of complex software tools can take place. It combined quality assurance to ensure a consistent quality of product, verification to determine how the code compares with its peers in the modelling world and realism testing to determine to what extent the code is fit for purpose. By combining within realism testing short term, well characterised experiments, with larger scale tests and analogue studies it is possible to both target specific, individual processes modelled and how the code models the whole multi-process system which occur in reality.

This approach has been applied to the development and testing of the GRM with encouraging results to date and a range of examples are provided in this paper. Work is still ongoing, with continued emphasis on analogue studies, peer review and uncertainty and sensitivity analysis. The ultimate aim is to develop a large portfolio of test cases which can be routinely applied to this and future codes. This approach will allow the envelope within which there is confidence in the code's capabilities to grow as the code is applied to an increasing number of scenarios and test cases.

REFERENCES

[1] P. N. Humphreys, T. Johnstone, D. Trivedi and A. Hoffmann, Mat. Res. Soc. Symp. Proc. **353**, p. 211, (1995).

[2] S. Manton, T. Johnstone, D. P. Trivedi, A. Hoffmann and P. N. Humphreys, Radiochimica Acta, **68**, p. 75, (1995).

[3] P. Humphreys, R. McGarry, A. Hoffmann and P. Binks, FEMS Microbiological Reviews, In Press.

[4]. M. G. McDonald, and A. W. Harbaugh, A Modular, *Techniques of Water-Resources Investigations of the United States Geological Survey*, **Book 6**, Modelling Techniques, (1988).

[5] L. Parkhurst, D.C. Thorstenson, and L. N. Plummer, *PHREEQE-A Computer Program for Geochemical Calculations*, (U. S. Geological Survey.) (1988).

[6] S. Manton, T. Johnstone, D. Trivedi, A. Hoffmann, and P. N. Humphreys in *Spectrum 92, Nuclear and Hazardous Waste Management*, International Topical Meeting, Boise, Idaho, USA, **2**, p. 1463, (1992).

[7]W.Stumm and J.J. Morgan, *Aquatic Chemistry*, 2nd Ed. (Wiley International Publications, New York, 1981).

[8] A. J. B. Zehnder and W. Stumm in *Biology of Anaerobic Microorganisms,* edited by A. J. B. Zehnder, (John Wiley and Sons, New York, 1988), p 1.

[9] A. Young in *Landfill Microbiology, R&D Workshop*, edited by P. Lawson and Y.R. Alston (Harwell laboratories UKAEA), p. 62.
[10] M. El-Fadel, A. N. Findikakis, J. O. Leckie. Waste Man. Res. **7**, 31 (1989).
[11] J. S. Kindred and M. Celia, Water Resour. Res. **25**, p. 1149 (1989).
[12] L. F. Konikow, and J. D. Bredehoeft, Adv.Wat. Res. **15**, 75, (1992).
[13] M. P. Anderson, and W. W. Woessner, Adv.Wat. Res. **15**, 167, (1992).
[14] C. P. Jackson, D. A. Lever, and P. J. Sumner, Adv. Wat. Res. **15**, 33, (1992).
[15] British Standards Institute, *Quality Systems: Model for Quality Assurance in Design, Development, Production, Installation and Servicing*. BS EN ISO 9001. Sterling Press. (1994).
[16]R. W. Cleary, M. J. Ungs, *PRINCE: Princeton Analytical Models of Flow and Mass Transport,* Waterloo Hydrogeologic Inc. Waterloo, Ontario, Canada (1994).
[17] M. S. Beljin, P. K. M. van der Heijde, *SOLUTE: Analytical Models for Solute Transport in Groundwater,* IGWC, Colorado School of Mines, Golden, CO, USA (1995).
[18] D. Read, *CHEMVAL Project: Report on Stages 3 and 4 on the Testing of the Coupled Chemical Transport Models.* CEC Report No. EUR 136.75. (1992).
[19] INTRACOIN. *INTRACOIN final report level 1, Code Verification.* SKI 84:3. Swedish Nuclear Power Inspectorate, Stockholm, Sweden (1984).
[20] C. A. J. Appelo D. and Postma, *Geochemistry and Groundwater Pollution.* A. A. Balkema, Netherlands. (1993).
[21] H. S., Rifai, P. B. Bedient, R. C. Borden and J. F. Haasbeck, *BIOPLUME II v1.1. Computer Model of Two-Dimensional Contaminant Transport Under the Influence of Oxygen Limited Biodegradation in Groundwater.* Version 1.1. National Centre for Ground Water Research, Rice University, Huston, Texas, USA. (1989).
[22] W. L. Bolle, J. van Breugel, G. C. van Eybergen, N. W. F. Kossen, and W. van Gils, Biotechnology and Bioengineering: **XXVIII**, p. 542, (1986).
[23] N. V. Ashley, S. Y. R. Pugh, C. J. Banks, P. N. Humphreys, *Microbial Treatment of Low Level Radioactive Waste.* HMIP/RR/92/06. UK DoE, London, UK. (1992).
[24] R. G. G. Holmes, and a. Rosevear, in *Gas Generation and Release from Radioactive Waste Repositories*, Proceeding of a Workshop Organised by NEA in Co-operation with ANDRA, Aix-en-Provence, 23-26 Sept 1991. OECD. (1992).
[25] A. E. Fryer, F. W. Schwartz, Wat. Res. Resour. **30** (12), p. 3455 (1994).
[26] J. E. Cross, D. S. Gabriel, A. Howarth, I, Neretnieks, S. M. Sharland, C. J. Tweed, Modelling of Redox Front and Uranium Movement in a Uranium Mine at Pocos de Caldas, Brazil. NSS/R252, UK NIREX ltd, UK (1991).
[27] D. Read, W. S. Atkins, *Geochemical Modelling of the Broubster Natural Analogue Site Caithness, Scotland.* EUR-12342, CEC Luxembourg, (1989).
[28] D. Trivedi, M. Randall, T. Johnstone, S. Serkiz, D. R. Boltz, D. I. Kaplan, AGU Transactions, **77** (46), F217. (1996).
[29] D. Postma, C. Boesen, H. Kritstiansen, F. Larsen, Wat. Res. Resour. 27(8). p. 2027 (1991).