MODELLING OF SOLUTE TRANSPORT IN GROUNDWATER: THE IMPACT OF SUPERCOMPUTERS

D. A. Barry

ABSTRACT

Supercomputers can expand greatly the possibilities for modelling subsurface solute transport. The theory relating to a specific application, viz. multicomponent solute transport, is presented. It is demonstrated that multidimensional realworld problems involving a series of chemical reactions requires significant computational power. It is demonstrated, however, that obtaining true supercomputer performance depends on algorithm selection. Other supercomputer applications are discussed, including visualization and effects of scale-dependent aquifer heterogeneity.

INTRODUCTION

It might be said, somewhat facetiously, that groundwater modelling has provided a popular pastime for computationally oriented hydrologists from the earliest days of computers. The theory of flow in porous media is extremely broad, covering a wide spectrum of engineering and scientific disciplines. For example, models developed for particular applications may draw upon fundamental results from physics, chemistry, mechanics and biology.

Models, and modelling approaches, can vary from simple to complex, but computer-based simulations are necessary for most cases where processes are to be quantified. The more complicated models of solute transport can involve literally scores of governing equations in the form of nonlinear partial differential equations over domains in which the transport parameters are spatially and temporally variable, sometimes to large magnitude. One result is that supercomputers will be called upon to handle bigger and bigger problems. It is always possible to imagine a simulation that is beyond the capacity of the largest available machine. Note that "bigger" problems can lead to important gains in knowledge. More computing capacity allows for (requires, in fact) the reexamination of basic assumptions underlying generally accepted models. As stated by *Wallis* [1987], "problems are being routinely defined that could not even have been conceived without the existence of computers". One example concerns the effects of variability of hydraulic conductivity on solute dispersion in aquifers, an area of vigorous research activity in past decade. Analytical theories have been developed for rather specialized conditions, so the question arises of situations that are not covered by the theory. Supercomputer simulations have provided insights that would probably not have been discovered as easily otherwise.

More routine problems which may push the capacity of less powerful computers are also very suited to supercomputers. A particular simulation that usually takes two hours CPU-time might only take a few minutes on a supercomputer. With a supercomputer, many more different scenarios can be examined in a very short time. If the hydraulic (or other) properties of a heterogeneous aquifer are assumed to random, then a study of the aquifer could employ Monte-Carlo computer simulation methods, meaning that many hundreds of simulations might be necessary to characterize the statistical properties of the aquifer response. Again, with a three-dimensional flow domain, a supercomputer is the most appropriate tool.

A supercomputer is defined as one in the fastest class of available machines. These machines not only have the fastest processor rates, but also convey information much faster than ordinary machines. In some machines the distances between sites within a computer are minimized, as demonstrated by the circular configuration of CRAY machines. It has been suggested that superconductors might be used to speed up transmission rates, since the hardware limitations of present technology are now recognized [*Seitz and Matisoo,* 1984]. Fast processors alone do not guarantee that groups of computations are performed faster. The efficient organization of information so that it can be processed effectively is an onerous task, particularly on multiprocessor machines.

If it is true that the technological limits of presently available computer hardware are in sight, then increases in speed can occur through increasing the number of calculations carried out in a given time by increasing the number of processors, either by operating on vectors rather than scalars, as in a vector machine, or by increasing the number of processors available, as in a multiprocessor machine. This available processing power can only be utilized by increasing the concurrency [*Sharp,* 1987] of the algorithms used. It is apparent that algorithms developed for scalar machines will not be effective on supercomputers because these algorithms are inherently sequential. In fact, the early CRAY machines were reputedly popular because they had a vary fast scalar processing rate, even though this rate was two orders of magnitude less than the maximum vector processing rate.

A good example of the effect of the lack of concurrency of an algorithm is contained in Amdahl's law [*Amdahl,* 1967; *Emmen,* 1987]. Say it takes N_0 operations or, equivalently, a time of N_0 to complete a given algorithm. The minimum possible time to complete the algorithm on P processes is N_0/P. The speedup, S, is defined by:

$$S = \frac{\text{time on sequential machine}}{\text{time on concurrent machine}} \qquad (1)$$

so S = P in this case. This is true only if the individual calculations are independent. If they are not independent, then there exists a sequential path in the algorithm, taking time (or operation count) R. The minimum time in which the algorithm can be executed on the P-processor machine is:

$$S = \frac{N_0}{R + \frac{N_0 - R}{P}} \qquad (2)$$

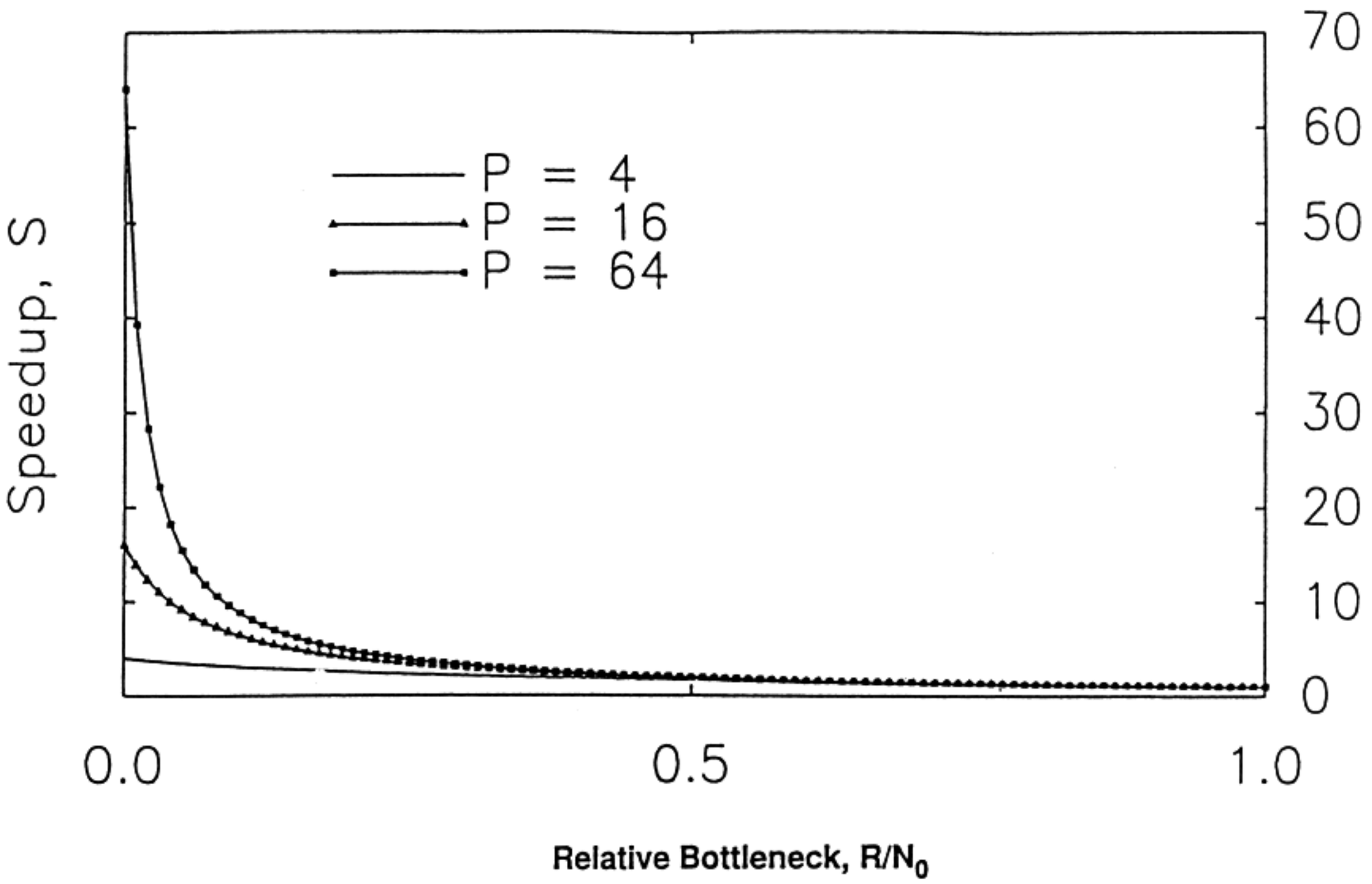

Fig. 1. Depiction of Amdahl's Law, showing the detrimental effect of bottlenecks on algorithm speedup on multiprocessor machines.

Equation (2) is based on the assumption that all the other operations in the algorithm can be performed independently. Even if $P \rightarrow \infty$, S is bounded from below by N_0/R. In other words, the minimum real time to complete the algorithm is R. Clearly, the longest sequential path in the algorithm is a serious bottleneck. In Fig. 1 some examples of (2) are displayed. If $R/N_0 = 0$ then the full speedup is achieved. The full impact of Amdahl's Law can be seen in that a rapid decline from this maximum is apparent. If one considers that for an algorithm of N_0 operations it is reasonable to take R as roughly $N_0/2$, then it is not possible to utilize fully the available processors.

Now suppose a different algorithm taking N_0^2 operations is used to perform the same task. The difference in this case is that the longest sequential part of the algorithm is reduced to, say, R/10. On a machine with a large number of processors, this second algorithm will perform much quicker in real time because the bottleneck has been removed. An algorithm that would be discounted immediately on a scalar machine suddenly becomes preferable on a multiprocessor machine. *Denning* [1988] discusses Amdahl's Law as it applies to parallel machines in more detail.

This same discussion also applies to vector machines, the only difference being that one considers the percentage of the code that is vectorizable, and the percentage that must be performed at the machine's scalar speed [*Barry,* 1990]. The scalar speed of the machine becomes the bottleneck which adversely affects the completion time of the algorithm. Although Amdahl's Law over-simplifies algorithm timing, it serves to illustrate that one needs to consider carefully algorithms used when using different machines.

MULTICOMPONENT SOLUTE TRANSPORT

Isothermal transport of a single solute species in groundwater is modelled by two governing equations, one for the solute - the transport equation - and one for the fluid - the flow equation. In many cases these equations are treated independently, even though they are coupled through concentration-dependent density changes. This approximation is appropriate for dilute solutes.

The situation is more involved for modelling the transport of reactive solutes, where several different chemical species may need be considered. For simplicity we assume isothermal conditions, ignore precipitation and dissolution of minerals, and instantaneous reactions, i.e. reaction kinetics need not be considered. The first step in the model development is to determine the chemical reactions and compose a set of components in which the reactions can be written. In other words, the set of components spans the chemical species space [*Brown and Page,* 1970, §3.1]. From the chemical reaction equations the law of mass action [*Tan,* 1982, §3.5] is used to quantify the equilibrium concentration of each species and component in each reaction via the equilibrium reaction constant. In this step a nonlinear algebraic system of equations results, the solution of which gives the equilibrium speciation of the system [*Coves and Sposito,* 1986, §1]. Many geochemical codes have been written to calculate the equilibrium speciation of natural waters [e.g. *Morel and Morgan,* 1972; *Truesdell and Jones,* 1974; *van Gaans,* 1989; *Wolery,* 1983]. The main distinguishing elements between various codes are the thermodynamic data bases that quantify individual reactions and the numerical techniques used to solve the nonlinear algebraic systems of equations. Transport of several solute species requires a transport equation for each chemical component. The basic equations governing the flow of N_c chemical components in a flowing system are [*Bear,* 1972; *Bryant et al.,* 1986]:

$$\frac{\partial(\theta c_i^T)}{\partial t} + \mathbf{L}(u_i) = 0, \qquad i = 1,...,N_c, \tag{3}$$

where c_i^T is the <u>total</u> concentration of the i^{th} component in any form whatsoever in the system and u_i is the concentration of the same component in the <u>flowing</u> (aqueous) solution. The right side of (3) is nonzero if production terms need be included, or if irreversible dissolution or precipitation occurs. The operator **L** is defined by [*Kirkner and Reeves,* 1988]:

$$\mathbf{L}(u_i) = \nabla.(u_i\mathbf{q} - \theta\mathbf{D}.\nabla u_i). \tag{4}$$

Now, the total concentration of the i^{th} component in the system is, by definition,

$$c_i^T = c_i + s_i + \sum_{j=1}^{N_x} A_{ij}x_j \tag{5}$$

Total Concentration	Component Concentration	Sorbed Concentration	Complexed Concentration

The latter two terms on the right side of (5) have not yet been discussed. The summation term refers to soluble complexation, so that the chemical reactions are of the generic form:

$$\sum_{j=1}^{N_x} A_{ij}\hat{c}_j \rightleftarrows \hat{x}_i; \qquad K_i,\ i = 1,...,N_x \tag{6}$$

where K_i is the equilibrium constant for the reaction and the circumflex indicates a chemical formula, not a concentration. Upon using the law of mass action, one obtains from (6):

$$x_i = K_i \prod_{j=1}^{N_c} c_j^{A_{ij}}, \qquad i = 1,...,N_x. \tag{7}$$

If the concentrations in (6) are not dilute, then each concentration variable must be replaced by the species activity. The total soluble concentration of the i^{th} component, u_i, is defined by:

$$u_i = c_i + \sum_{j=1}^{N_x} A_{ji}K_j \prod_{k=1}^{N_c} c_k^{A_{jk}}, \qquad i = 1,...,N_c. \tag{8}$$

Equation (8) shows that the soluble concentration can be specified as a nonlinear function of the component concentration variables alone. Observe that (8) defines N_c unknowns, u_i, in terms of N_c (presumed) known variables, c_i. In principle then, (8) can be inverted to give each c_i as a function of the set of N_c u_i variables, i.e. [*Kirkner and Reeves,* 1988]:

$$c_i = f_i^{cu}(u_1,...,u_{N_c}), \qquad i = 1,...,N_C. \tag{9}$$

We now consider the sorbed concentration term in (5). Sorption consists of surface adsorption and solute absorption within the soil grains. Unlike precipitation/dissolution reactions the medium porosity is unchanged by sorption. The main point to note here is that it is always possible to write the adsorbed concentration, s_i, in terms of the fluid phase solute concentration, i.e.

$$s_i = f_i^{sc}(c_1,...,c_{N_c},x_1,...,x_{N_x}). \tag{10}$$

Equation (10) becomes a function of the component concentrations, c_i, if the complexed concentrations are eliminated using (7).

Using (5) through (8), (3) can be written in terms of, say, the component concentrations, c_i. The resulting N_c partial differential equations are then coupled. To complete the model of the system an additional governing equation is needed to describe the water flow. There are many forms of the flow equation, but they all are derived from [*Bear,* 1972, §6.2.1]:

$$\nabla.(\rho\mathbf{q}) + \frac{\partial(\rho\theta)}{\partial t} = 0. \tag{11}$$

Equation (11) is coupled to the transport equations since both the density, ρ, and the flux, $\mathbf{q}$, vary with changes in concentration in the fluid.

This brief derivation of a governing model of multicomponent solute transport demonstrates the possible magnitude of a realistic chemical modelling effort. In such a case, not only are there numerous governing equations, but these equations must be solved simultaneously. The number of (more-or-less) general purpose models developed is low. In the main, the existing codes are for unidimensional flow and transport, and most are, in fact, the result of combining existing flow and geochemical codes.

An example of a "big" computer implementation of a multicomponent transport model is the multidimensional code called DYNAMIX that has been developed by *Narasimhan et al.* [1986]. DYNAMIX, which has been the subject of continued development, [*Liu and Narasimhan*, 1989a] couples PHREEQE [*Parkhurst et al.*, 1980] with the multidimensional transport code TRUMP [*Edwards*, 1972]. It has now been extended to include acid-base reactions, aqueous complexation, redox reactions and precipitation/dissolution reactions.

DYNAMIX uses a two-step solution approach. As the numerical solution steps through time, the transport model is solved by explicit finite differences. Then, the chemical submodel computes the chemical distribution under dynamic partial equilibrium conditions. Example problems solved by DYNAMIX are given by *Liu and Narasimhan* [1989b]. After favorable comparison of DYNAMIX to the one-dimensional codes of *Walsh et al.* [1984] and *Carnahan* [1986], *Liu and Narasimhan* [1989b] performed a 30-year simulation of solute transport from a contamination site. The two-dimensional contamination problem involved selenium and arsenic, but included also 16 aqueous components, 26 minerals and over 76 aqueous complexes. They found that arsenic was strongly retarded by precipitation as As_2S_3. On the other hand, selenium was relatively more mobile, being transported in the aqueous phase. Not surprisingly, a Cray XMP/14 supercomputer was used to perform the calculations. The reactive trace elements selenium and arsenic were included. The total CPU time needed was 3.4 hours on a CRAY X-MP/14. The code does not take advantage of the CRAY's vector speed capabilities. Rewriting computationally intensive portions of the code to allow vectorization would likely lead to substantial improvements in the CPU time used.

HETEROGENEITY OF THE SUBSURFACE ENVIRONMENT

Natural porous media are characterized by variability. This variability is reflected in the transport and hydraulic properties measured in the field. The effect of this variability on water and solute movement is a topic of significant theoretical and practical interest.

There are different ways in which variability can be modelled. For example, *Wheatcraft and Tyler* [1988] explain scale-dependent dispersion of a solute plume in groundwater using concepts from fractal theory. Thus, they model the medium as being self-similar, at least over a range scales. In the absence of an upper scale bounding the self-similar description of the medium, the solute plume dispersivity will increase without bound.

A more thoroughly investigated model of variability is based on statistical models of the variability of the hydraulic properties of the medium. The most common approach has been to model hydraulic conductivity as being random. In the most complete theories the hydraulic conductivity is assumed to be lognormally distributed, with an exponential decay in the spatial correlation. The statistical behaviour of the fluid movement can then be determined from the governing flow equation, which has become a stochastic differential equation due to the random conductivity term.

It has been demonstrated that variability in the groundwater flow has a much greater effect on solute spread than local-scale diffusion or dispersion. The cumulative effect on the spread of solute due to variability of the groundwater flow field has been termed the macrodispersivity. Since the low-order statistics of the groundwater flow field can be calculated, this information can be used to calculate the macrodispersivity of a solute plume in the case of a dilute solute. Expositions of the analytical results provided by this theory have been presented by *Dagan* [1986], *Gelhar* [1986] and *Sposito and Barry* [1987].

The analytical results are based on fairly simple conditions, viz. infinite spatial domain, second-order statistical stationarity of the hydraulic conductivity field, uncoupled governing equations of steady groundwater flow and solute transport and constant mean flow direction. Even so, these analytical predictions have had some success when compared with data from field experiments [*Freyberg,* 1986; *Barry et al.,* 1988; *Barry and Sposito,* 1990].

There is a limit, of course, to the results that can be obtained analytically. Researchers have also performed numerical studies into heterogeneity. There are two major approaches used, particle tracking and continuum models. In common with the purely analytical investigations, one of the aims of the numerical analyses is to identify the governing equations for flow and transport at macroscales, based on parameter variability at local scales. One problem encountered with the numerical approaches is that a large number of nodes might be needed to account for detail in the flow domain. At the sub-grid scale variability cannot be represented.

Particle Tracking

This technique is used to overcome some of the perceived difficulties in using continuum models based on partial differential equations. The theoretical basis of the method is the Langevin equation [*Gardiner,* 1985]:

$$\frac{d\mathbf{X}}{dt} = \mathbf{A}(\mathbf{X},t) + \mathbf{B}(\mathbf{X},t).\xi(t), \qquad (12)$$

where **X** denotes the position of the particle, **A** represents the sure changes upon the particle, $\xi(t)$ is stochastic forcing function idealized as white noise and **B** gives directional characteristics of the random component of motion. As it stands, (12) has no meaning because white noise is delta-function correlated [*van Kampen,* 1981, Chap. VIII], and infinite variances are physically impossible. If, however, we follow the Îto interpretation, then, over a time span of Δt, (12) is equivalent to:

$$\mathbf{X}(t+\Delta t) - \mathbf{X}(t) = \mathbf{A}[\mathbf{X}(t),t]\Delta t + \mathbf{B}[\mathbf{X}(t),t]. \int_{t}^{t+\Delta t} \xi(t')\, dt'. \qquad (13)$$

Since **A** is smooth it can be approximated in the time interval, Δt. The integrand of the second term varies rapidly so it cannot be treated in the same manner. In deed, (13) has already incorporated the to assumption that **B** is nonanticipating, allowing it to be removed from the integrand. This assumption is useful since it means that **B** and the integral of ξ are uncorrelated, and so **B** can be calculated from knowledge of **X**(t) [*Haken,* 1983]. Clearly, different realizations of $\xi(t)$ will influence the location of the particle under consideration. Under suitable conditions, (13) is equivalent to a Fokker-Plank, or Kolmogorov forward equation [*Cox and Miller,* 1965, Chap. 5]. The analogy to solute transport is complete when we recognize that the governing equation for solute transport is also a Fokker-Plank equation.

To complete the analogy, a mass is assigned to each particle. The movement of many (an infinite number) particles will then explore all possible realizations of the stochastic process. If the total mass is normalized to unity, then the distribution of the particles will approximate the probability density function of particle movement. This, in turn, is proportional to solute concentration. The Fokker-Plank equation corresponding (12) is [Tompson *et al.*, 1987]:

$$\frac{\partial(\rho\omega\theta)}{\partial t} + \nabla.\left\{\left[\mathbf{A} - \frac{1}{2}\nabla.(\mathbf{B}.\mathbf{B}^T)\right]\rho\omega\theta\right\} = \frac{1}{2}\nabla.[(\mathbf{B}.\mathbf{B}^T).\nabla(\rho\omega\theta)], \qquad (14)$$

where ρ is the solute density and ω is the solute mass fraction. For diffusive transport in a saturated domain and a constant velocity we would take:

$$\mathbf{v}(\mathbf{X},t) = \mathbf{A}(\mathbf{X},t), \qquad (15a)$$

and

$$\mathbf{B} = \sqrt{2D\mathbf{I}}. \qquad (15b)$$

The combination of (14) and (15) yields:

$$\frac{\partial(\rho\omega)}{\partial t} + \nabla.(\mathbf{v}\rho\omega) = D\nabla^2(\rho\omega), \qquad (16)$$

which is the usual transport equation.

Solutions to (14) can be obtained by using a Monte-Carlo approach applied to the Langevin equation. Each particle in the simulations moves independently, so the computation is well suited to evaluation by a supercomputer. Indeed, the degree of parallelism can be at least the number of particles in the simulation. Computationally, particle tracking is algebraically uncomplicated and extremely repetitive making for an easily vectorized code.

A common problem with particle tracking methods is the noisy solute distributions resulting from the simulations involving too few particles. Distributions are derived by using the assigned mass of each particle and averaging over a specified domain size. The noise is proportional to $1/\sqrt{N'}$, where N' is the number of particles. The most straightforward way to reduce noise is to increase N', although smoothing or filtering of the concentration predictions may be more economical computationally.

A recent study using particle tracking is that reported by *Tompson et al.* [1987]. These authors review in detail the theory outlined above, including the forms of the functions **A** and **B** in (14) which are appropriate for simulations of solute transport for both saturated and variably saturated flows. The simulations of *Tompson et al.* [1987] were aimed at evaluating stochastic theories. To that end, flow in a heterogeneous domain was generated using the large-scale flow simulator of *Ababou et al.* [1988], developed for the CRAY-2 supercomputer. *Ababou et al.* [1988] used the strongly implicit procedure, a less efficient forerunner of the preconditioned conjugate gradient method, to solve their linear systems of equations [*Meyer et al.*, 1989]. *Tompson et al.* [1988] did not discuss the specific algorithm design or computer used in their simulations. Their code, however, is designed to track many thousands of particles on grids of up to 10^7 nodes. Simulations of this kind are practicable only on supercomputers.

Continuum Model Solutions

Conventional solutions based on the governing equations of flow and solute transport are more common. Most applications use the assumption of a dilute solute, so that the governing equations may be treated separately.

A model that makes use of supercomputer characteristics in algorithm selection was developed by *Chiang et al.* [1989]. The code was developed to model pumping-induced flow and contaminant transport in a two-dimensional heterogeneous domain. Example simulations included transport in domains having a random hydraulic conductivity field, realizations of which were developed using the turning bands method [*Mantoglou and Wilson,* 1982]. *Chiang et al.* [1989] selected the preconditioned conjugate gradient algorithm to provide the necessary numerical solution to the linear system of equations. It was concluded that the code developed was efficient, accurate, and capable of modeling transport in systems with rapidly varying hydraulic conductivity fields.

Frind et al. [1989] report a three-dimensional simulation of transport of organic solutes undergoing aerobic biodegradation. They model a saturated groundwater system in which the solute transport process is described by a nonlinear, coupled set of equations, one for the contaminant and one for dissolved oxygen. The dual-Monod equation [*Borden and Bedient,* 1986] is used to model the growth of the microbial population. By formulating the numerical model to produce a symmetric coefficient matrix, the preconditioned conjugate gradient solution method could be used effectively. A 24 x 10^3 node simulation over 155 time steps (300 days) took about 2-h CPU time on a CRAY X-MP/24. The authors concluded that comparable two-dimensional simulations were to be avoided as oxygen availability was markedly affected by problem dimensionality.

Sudicky [1989*a*] considered the transport of a dilute solute in a steady, nonuniform, two-dimensional flow field. The spatial derivatives were discretized using the Galerkin finite element procedure. Instead of using a finite difference procedure for the temporal derivative, Laplace transforms were used. The Laplace-transformed problem was shown to not suffer from the numerical dispersion problems associated with other methods. The *Crump* [1976] algorithm for numerical inversion of the Laplace-transformed solution can be used to calculate the solution time-domain solution from a single set of Laplace-domain solution values.

Sudicky [1989*b*] extended the application of the method to the case of fractured porous formations. In the case of the traditional finite-difference approximation of the temporal derivative, numerical difficulties arise due to the large differences between the rapid advection-dominated transport in the fractures and much slower diffusional flux into the porous matrix. *Sudicky* [1989*b*] combined the Laplace-transform method with ORTHOMIN [*Behie and Vinsome,* 1982] acceleration for the rapid solution of the algebraic system of equations derived from application of the Galerkin procedure. The combined procedure was shown to be very efficient and accurate. A 30,651-node simulation of downward seepage of solute in a medium containing several thousand fractures took 72 seconds on a CRAY X-MP/24 to obtain solute distributions valid for all times of interest. By using algorithms especially suited to vector supercomputers, Sudicky has developed a code capable of simulating very large, field scale transport problems.

The data base from the Borden [*Roberts and Mackay,* 1986] aquifer transport experiment has been used to evaluate the predictions of stochastic solute transport theory. Both *Freyberg* [1986] and *Naff et al.* [1988] compared theoretical predictions of the vertically averaged tracer plume dispersivity with those estimated from experimental data and found excellent agreement. Following the success of the

two-dimensional *Dagan* [1984] model in predicting the spread of the vertically averaged tracer plume at the Borden site, *Barry et al.* [1988] checked whether the actual plume concentrations conformed with ensemble-mean concentrations. The theory was used to predict solute plumes using the measured plume from a number of different samplings as initial conditions, then comparing the predicted and measured plume at later times. The data analysis and generation of the predicted plumes, which involved integrating the initial condition numerically, were carried out using a vectorized code on a CRAY X-MP/48. It was found that the model predicted the measured plume quite accurately on relatively short time scales, becoming progressively worse with increasing elapsed time. However, the assumptions necessary in the data analysis made equivocal any comparisons of model-predicted and measured plumes.

The ensemble approach has led to elegant analytical results but, as discussed by *Dagan* [1987], they cannot be applied unless certain conditions are met. A demonstration of the possible differences between separate plume realizations is given by *Frind et al.* [1987]. These authors use a two-dimensional vertical-slice model of an aquifer with a hydraulic conductivity field that reproduced the estimated low-order statistics of the Borden aquifer. Two simulations were performed; one of which used a spatial grid of 106 elements. It was clearly demonstrated that local variations in the hydraulic conductivity dominate plume movement, particularly at early times. The short term behavior can have persistent, long-term effects. The simulation of *Frind et al.* [1987], which used 10-h of CPU time on a CRAY X-MP/48, revealed the formation of two subplumes due to the initial placement of the solute source in a region of rapidly varying conductivity. Other, similar simulations reported by *Sudicky et al.* [1990] demonstrate that in many cases less dramatic results are obtained. For example, *Sudicky et al.* [1990] discuss simulations in which the solute plume approaches a Gaussian distribution, as predicted by stochastic theory.

Sudicky et al. [1990] go on to consider the transport of a biodegradable solute in a heterogeneous porous medium. In this case there are three nonlinear governing equations, one each for the solute, available oxygen, and the microbial population [*MacQuarrie et al.*, 1990]. The behavior of the plume is quite different. If the biodegradation rate is large enough, then the plume will tend in decrease in size. For this reason, the plume will encounter, and be influenced by, only a small portion of the groundwater velocity field. Under these conditions, plumes resulting from different realizations will not necessarily converge very rapidly to those predicted by ensemble averaging of the governing equations. More comprehensive simulations presented by *MacQuarrie and Sudicky* [1990] show in detail that biodegradable solute plumes evolve differently than plumes composed of nonreactive solutes. In particular, it is suggested that the scale dependence of the plume dispersivity, that is well documented for field scale transport of tracer solutes [*Gelhar,* 1986, Fig. 12], does not apply in the case of biodegradable, organic solute plumes.

VISUALIZATION

Improvements in data visualization have kept pace with developments in computing hardware. Plotters and associated software are standard equipment. Much specialized hardware and software is available also. A powerful combination for real-time simulation is the combination of supercomputer computational rates and a dedicated graphics workstation. An example utilizing this technology was undertaken as part of a more general study on solute transport. Three-dimensional visualizations of the bromide tracer plume from the Borden site experiment [*Roberts and Mackay,* 1986] were generated using facilities at the National Center for Supercomputing Applications (NCSA). Briefly, the raw data, consisting of solute concentration measurements at given spatial locations for the 14 plume

samplings, were fitted to an interpolation function using the NCSA's CRAY X-MP/48 computer following a cross-validation procedure. Full details can be found in *Barry and Sposito* [1989]. Dense (4 - 8 x 10^4 points), regular grids of interpolated concentration data were produced and surface contours - three-dimensional versions of contour lines - were computed using the "marching cubes" algorithm [*Lorensen and Cline,* 1987]. The contour surfaces are allowed to be semi-transparent, and so allow the passage of light. Since contours are formed at regular concentration increments, domains of higher concentration are surrounded by more surfaces than areas of lower concentrations. When (mathematical) light rays are passed through the surfaces they are diffused. Domains containing relatively higher concentrations appear as being relatively more opaque. There is complete freedom in how the final images are viewed due to variations in the light source and observation positions. Solute plume images resulting from this procedure can be found in Plate 1 of *Barry* [1990].

LINEAR ALGEBRA: THE PRECONDITIONED CONJUGATE GRADIENT METHOD

The key computational element of many transport related algorithms is the solution of linear systems of equations. One of the fastest methods for supercomputer use is the preconditioned conjugate gradient (PCG) algorithm. The conjugate gradient (CG) method is based on the fact that the solution, **y**, of $\mathbf{Cy} = \mathbf{b}$ minimizes the quadratic functional $\mathbf{y}^T\mathbf{Cy} - 2\mathbf{b}^T\mathbf{y}$ if **C** is symmetric and positive definite [*Hager*, 1988, §7-5] and the superscript T indicates transposition. It is this functional which is minimized to obtain the unknown vector y. In the absence of rounding error, the method converges to the exact solution for **y** in dim(**C**) steps. It has been observed that, for some matrix types, a reasonable approximation to the solution is obtained after relatively few steps, leading to the use of the method as an acceleration technique. The convergence rate of CG depends on the condition number of **A**. Preconditioning is therefore used to increase the convergence rate [*Hager*, 1988, §7-5]. *Concus et al.* [1976] note the following advantages of the method: (i) not all of **C** needs to be stored in memory at one time, (ii) parameters used by the algorithm are automatically determined, and (iii) some restrictions on **C** required by other methods can be relaxed. A vectorized PCG algorithm was given by *Kershaw* [1982]. If **C** is not symmetric and positive definite but is invertible, then premultiplying the equation $\mathbf{Cy} = \mathbf{b}$ by $\mathbf{C}^T$ gives a system in the required form. The PCG algorithm then involves more computations, however. An alternative is the ORTHOMIN algorithm [*Behie and Vinsome,* 1982].

Meyer et al. [1989] consider linear systems of equations encountered in subsurface flow problems. The choice of preconditioner is related to the type of computer used. A number of preconditioning schemes implemented on the CRAY X-MP/48, a vector machine, and the Alliant FX/8, an eight-processor vector-parallel machine were compared. Overall, low-order polynomial preconditioning was found to give the best results. The largest problem solved by *Meyer et al.* [1989] involved 980,000 unknowns and took approximately one hour CPU time on the Alliant, with a factor 15 speedup expected for the CRAY. *Ababou et al.* [1988] used a CRAY-2 to perform a large-scale three-dimensional flow simulation, but did not use the PCG method. For this case, *Meyer et al.* [1989] estimated that using polynomial preconditioning with CG would have reduced the computation time by an order of magnitude. This example highlights the importance of algorithm selection for large-scale problems.

CONCLUSION

Although supercomputers give the opportunity to explore new computational territories, it is fitting to conclude on a cautionary note. Each increase in computational power, gives rise to the possibility that using supercomputers wholly to carry out larger and larger magnitude simulations, and so becoming the raison d'être of scientific endeavour. This point was discussed by *Cushman and Wobber* [1986] who "argue that the understanding of complex chemical and physical interactions in the subsurface and the ability to estimate parameters is so poor that the large supercomputer models may in fact be totally incorrect. In this instance, supercomputer modelling efforts may turn out to be a greater obstacle to real progress than no model at all". This argument may be elaborated upon to include the evolution of transport theory. Simulation of the effect of scale changes on plume evolution in a heterogeneous domain, for example, necessarily involves assumptions concerning the nature of heterogeneity, i.e. a model of scale-dependent heterogeneity must be invoked. The derivation of such models is certainly not a new question, but having supercomputers available adds emphasis to lack of knowledge of scale effects on solute transport, amongst other things.

ACKNOWLEDGEMENTS

Support from the Western Australian Water Authority is gratefully acknowledged. This is Centre for Water Research document ED539DB.

NOTATION

A_{ij} stoichiometric coefficient

A deterministic portion of the particle motion, LT^{-1}

b vector

B second-order tensor representing the directional coefficient of the random portion of the particle motion, $LT^{-1/2}$

c_i^T total concentration of the i^{th} component, ML^{-3}

C matrix

D diffusion coefficient, L^2T^{-1}

D hydrodynamic dispersion tensor, L^2T^{-1}

I identity tensor

K_i equilibrium constant

L second-order spatial differential operator, T^{-1}

N_0 number of arithmetic operations to complete an algorithm

N_c number of chemical components

N_x number of chemical complexes

N' number of particles

P number of processors

q Darcy flux, LT^{-1}

R number of operations in the longest sequential path in an algorithm

S speedup

s_i sorbed concentration, L^{-3}

t time, T

u_i total concentration of the i^{th} component in the aqueous phase, ML^{-3}

v mean solute velocity, LT^{-1}

x_j concentration of the j^{th} complex, ML^{-3}

X particle position, L

y vector

θ porosity

ξ stochastic forcing function, $T^{-1/2}$

ρ density, ML^{-3}

ω mass fraction

$\hat{\cdot}$ chemical formula of $\cdot$

$\cdot^T$ transpose of $\cdot$

REFERENCES

Ababou, R., L. W. Gelhar, and D. McLaughlin, Three-dimensional flow in random porous media, 2 vols., *Rep. No. 318,* Ralph M. Parsons Lab., Civil Engineering Dept., M.I.T., Cambridge, Mass., 1988.

Amdahl, G. M., Validity of the single processor approach to achieving large scale computing capabilities, *AFIPS Conf. Proc., 30,* 483-485,1967.

Barry, D. A., Supercomputers and their use in modeling subsurface solute transport, *Rev. Geophys., 28*(3), 277-295, 1990.

Barry, D. A., J. Coves, and G. Sposito, On the Dagan model of solute transport in groundwater: Application to the Borden site, *Water Resour. Res., 24*(10), 1805-1817, 1988.

Barry, D. A., and G. Sposito, Three-dimensional statistical moment analysis of the Stanford/Waterloo Borden tracer data, *Water Resour. Res., 26*(8), 1735-1748, 1990.

Bear, J., *Dynamics of Fluids in Porous Media,* American Elsevier, New York, 1972.

Behie, A., and P. K. W. Vinsome, Block iterative methods for fully implicit reservoir simulation, *Soc. Petrol. Eng. J., 22,* 659-668, 1982.

Borden, R. C., and P. B. Bedient, Transport of dissolved hydrocarbons influenced by oxygen-limited biodegradation, 1, Theoretical development, *Water Resour. Res., 22*(13), 1973-1982, 1986.

Brown, A. L., and A. Page, *Elements of Functional Analysis,* Van Nostrand Reinhold, New York, 1972.

Bryant, S. L., R. S. Schechter and L. W. Lake, Interactions of precipitation/dissolution waves and ion exchange in flow through permeable media, *Am. Institut. Chem. Eng. J., 32*(5), 751-764, 1986.

Chiang, C. Y., M. F. Wheeler, and P. B. Bedient, A modified method of characteristics technique and mixed finite elements method for simulation of groundwater solute transport, *Water Resour. Res., 25*(7), 1541-1549, 1989.

Concus, P., G. H. Golub, and D. P. O'Leary, A generalized conjugate gradient method for the numerical solution of elliptic partial differential equations, in *Sparse Matrix Computations,* edited by J. R. Bunch and D. J. Rose, Academic Press, New York, 1976.

Coves, J., and G. Sposito, *MICROQL-UCR: A surface chemical adaption of the speciation program MICROQL,* Soil Env. Sci. Dept., Univ. Calif., Riverside, California, 1986.

Cox, D. R., and H. D. Miller, *The Theory of Stochastic Processes,* Chapman and Hall, London, 1965.

Crump, K. S., Numerical inversion of Laplace transforms using a Fourier series approximation, *J. Assoc. Comput. Mach., 23*(1), 89-96, 1976.

Cushman, J. H., and F. J. Wobber, Keeping supercomputers under control, *Nature, 321*(6005), 21, 1986.

Dagan, G., Solute transport in heterogeneous porous formations, *J. Fluid Mech, 145,* 151-177, 1984.

Dagan, G., Statistical theory of groundwater flow and transport: pore to laboratory, laboratory to formation and formation to regional scale, *Water Resour. Res., 22*(9), 120S-134S, 1986.

Dagan, G., Theory of solute transport by groundwater, *Ann. Rev. Fluid Mech., 19,* 183-215, 1987.

Denning, P. J., Speeding up parallel processing, *Am. Sci., 76*(4), 347-349, 1988.

Edwards, A. L., *TRUMP: A Computer Program for Transient and Steady State Temperature Distributions in Multi-Dimensional Systems,* Rep. 14654, Rev. 3, Lawrence Livermore National Laboratory, Livermore, California, 1973.

Emmen, A. H. L., A survey of vector and parallel processors, in *Algorithms and Applications on Vector and Parallel Computers,* edited by H. J. J. te Riele, Th. J. Dekker and H. A. van der Vorst, pp. 151-154, Elsevier Science Pubs. B. V., Amsterdam, 1987.

Freyberg, D. L., A natural gradient experiment on solute transport in a sand aquifer, 2, Spatial moments and the advection and dispersion of nonreactive tracers, *Water Resour. Res., 22*(13), 2031-2046,1986.

Frind, E. O., E. A. Sudicky, and J. W. Molson, Three-dimensional simulation of organic transport with aerobic biodegradation, *Proc. IAHS 3rd Scientific Assembly,* Baltimore, Maryland, May 10-19, 1989.

Frind, E. O., E. A. Sudicky, and S. L. Schellenberg, Micro-scale modelling in the study of plume evolution in heterogeneous media, *Stochastic Hydrol. Hydraul., 1*(4), 263-279, 1987.

Gardiner, C. W., *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences,* 442 pp., Springer-Verlag, Berlin, 1983.

Gelhar, L. W., Stochastic subsurface hydrology from theory to applications, *Water Resour. Res., 22*(9), 135S-145S, 1986.

Hager, W. W., *Applied Numerical Linear Algebra,* Prentice-Hall, London, *1988.*

Haken, H., *Advanced Synergetics,* Springer Verlag, Berlin, 1983.

Kershaw, D., Solution of single tridiagonal linear systems and vectorization of the ICCG algorithm on the CRAY-1, in *Parallel Computations,* edited by G. Rodrigue, Academic Press, New York, 1982.

Kirkner, *D. J.,* and H. W. Reeves, Multicomponent mass transport with homogeneous and heterogeneous chemical reactions: Effect of the chemistry on the choice of numerical algorithm.
1. Theory, *Water Resour. Res., 24*(10), 1719-1729, 1988.

Liu, C. W., and T. N. Narasimhan, Redox-controlled multiple-species reactive chemical transport, 1, Model development, *Water Resour. Res., 25*(5), 869-882, 1989*a.*

Liu, C. W., and T. N. Narasimhan, Redox-controlled multiple-species reactive chemical transport, 2, Verification and application, *Water Resour. Res., 25*(5), 883-910, 1989*b*.

Lorensen, W. E., and H. E. Cline, Marching cubes: A high resolution 3D surface construction algorithm, *Comput. Graphics, 21*(4), 163-169, 1987.

MacQuarrie, K. T. B., and E. A. Sudicky, Simulation of biodegradable organic contaminants in groundwater, 2, Plume behaviour in uniform and random flow fields, *Water Resour. Res., 26*(2), 223-240, 1990.

MacQuarrie, K. T. B., E. A. Sudicky, and E. O. Frind, Simulation of biodegradable organic contaminants in groundwater, 1, Numerical formulation in principal directions, *Water Resour. Res., 26*(2), 207-222, 1990.

Mantoglou, A., and J. Wilson, The turning bands method for simulation of random fields using line generation by a spectral method, *Water Resour. Res., 18*(5), 1379-1394, 1982.

Meyer, P. D., A. J. Valocchi, S. F. Ashby, and P. E. Saylor, A numerical investigation of the conjugate gradient method as applied to three-dimensional groundwater flow problems in randomly heterogeneous porous media, *Water Resour. Res., 25*(6), 1440-1446, 1989.

Morel, F., and J. Morgan, A numerical method for computing equilibria in aqueous chemical systems, *Environ. Sci. Tech., 6,* 58-67, 1972.

Naff, R. L., T.-C. J. Yeh, and M. W. Kemblowski, A note on the recent natural gradient tracer test at the Borden site, *Water Resour. Res., 24*(12), 2099-2103, 1988.

Narasimhan, T. N., A. F. White, and T. Tokunaga, Groundwater contamination from an inactive uranium mill tailings pile, 2, Application of a dynamic mixing model, *Water Resour. Res., 22*(13), 1820-1834, 1986.

Parkhurst, D. L., D. C. Thorstenson, and L. M. Plummer, *U.S. Geological Survey Water Resources Investigation 80-96,* NTIS Accession No. PB8I 167801, November, 1980

Roberts, P. V., and D. M. Mackay (Eds.), A natural gradient experiment on solute transport in a sand aquifer, *Tech. Rep. No. 292,* 229 pp., Civil Engineering Dept., Stanford University, Stanford, Calif., 1986.

Seitz, C. L., and J. Matisoo, Engineering limits on computer performance, *Phys. Today, 37*(5), 38-45, 1984.

Sharp, J. A., *An Introduction to Distributed and Parallel Processing,* Blackwell Scientific, Oxford, 1987.

Sposito, G., and D. A. Barry, On the Dagan model of solute transport in groundwater: foundational aspects, Water *Resour. Res., 23*(10), 1867-1875, 1987.

Sudicky, E. A., The Laplace transform Galerkin technique: A time-continuous finite element theory and application to mass transport in groundwater, *Water Resour. Res., 25*(8), 1833-1846, 1989*a*.

Sudicky, E. A., The Laplace transform Galerkin technique for large-scale simulation of mass transport in discretely fractured porous formations, *EOS Trans., Am. Geophys. Union, 70*(43), 1082, 1989*b*.

Sudicky, E. A., S. L. Schellenberg, and K. T. B. MacQuarrie, Assessment of the behaviour of conservative and biodegradable solutes in heterogeneous porous media, in *Dynamics of Fluids in Hierarchical Porous Formations,* edited by J. H. Cushman, Academic, London, 1990.

Tan, K. H., *Principles of Soil Chemistry,* Marcel Dekker, New York, 1982.

Tompson, A. F. B., E. G. Vomvoris, and L. W. Gelhar, *Numerical simulation of solute transport in randomly heterogeneous porous media: Motivation, model development, and application,* Rep. No. UCID-21281, Lawrence Livermore Lab., Livermore, California, 1987.

Truesdell, A. H., and B. F. Jones, WATEQ, a computer program for calculating chemical equilibria of natural waters, *J. Res. U.S. Geol. Surv., 2*(2), 233-248, 1974.

van Gaans, P. F. M., WATEQX - A restructured, generalized, and extended FORTRAN 77 computer code and database format for the WATEQ aqueous chemical model for element speciation and mineral saturation, for use on personal computers or mainframes, *Comput. Geosci., 15*(6), 843-887, 1989.

Wallis, J. R., Hydrology -- the computer revolution continues, *Rev. Geophys., 25*(2), 101-105, 1987.

Wheatcraft, S. W., and S. W. Tyler, An explanation of scale-dependent dispersivity in heterogeneous aquifers using concepts from fractal geometry, *Water Resour. Res., 24*(4), 566-578, 1988.

Wolery, T. J., *EQ3NR: A Computer Program for Geochemical Aqueous Speciation-Solubility Calculations, User's Guide and Documentation,* Rep. UCRL-53414, Lawrence Livermore Lab., Univ. Calif., Livermore, California, 1983.

van Kampen, N. G., *Stochastic Processes in Physics and Chemistry,* North-Holland, Amsterdam, 1981.

Author:

D. A. Barry
Centre for Water Research
University of Western Australia, Nedlands WA 6009