WATER RESOURCES RESEARCH, VOL. 17, NO. 1, PAGES 149–161, FEBRUARY 1981

First Order Analysis of Uncertainty in Numerical Models of Groundwater Flow Part 1. Mathematical Development

MICHAEL D. DETTINGER[1] AND JOHN L. WILSON

Department of Civil Engineering, Massachusetts Institute of Technology Cambridge, Massachusetts 02139

In part 1 of this paper, first and second order analysis of uncertainty is applied to numerical models of groundwater flow. The models are cast in state-space form, with boundary conditions and inputs that are functionally dependent, but statistically independent, of time. Using a compact matrix calculus notation, first and second order Taylor series expansions of the model equations are derived and used to estimate the mean and variance-covariance properties of piezometric head predictions, given corresponding statistics for aquifer parameters: material properties, initial conditions, boundary conditions, and inputs. The mathematical results demonstrate that the prediction uncertainty is a function of the magnitude of the parameter uncertainty, and sensitivity of the predictions to the parameters. Furthermore, the first order estimate of the piezometric head is identical to the deterministic result. Part 2 of this paper, to be presented later, will illustrate these and other results through numerous applications of the methodology.

INTRODUCTION

A basic objective of the study of groundwater hydrology is the prediction of aquifer behavior under natural or stressed conditions. Aquifer behavior depends on the spatial and temporal variability of aquifer properties and other parameters, such as boundary values and the strength and location of sources and sinks. This parameter variability can, in theory, be described in terms of a recognized, deterministic distributed field of the parameters in time and space, or in terms of the field statistics. The ability to predict aquifer behavior depends on our knowledge of the variability. That is, do we know the true parameter field? Do we know the statistics of the field? Or do we estimate each from a few observations? The presence of this variability, coupled with incomplete information about its description, leads to a description of aquifer behavior in probabilistic terms. This paper describes a general probabilistic approach for the numerical solutions of groundwater flow problems. The method employed, first and second order analysis, is used to treat systems with spatially distributed and varying properties, complicated initial or boundary conditions, and complicated one-, two-, and three-dimensional flows in bounded domains. Temporal probabilistic variations are not covered.

Part 1 of this paper, presented here, includes a discussion of the nature and sources of uncertainty encountered in groundwater flow models, with particular attention given to numerical models, and a description of techniques available for use in the analysis of the effects of parameter uncertainty on piezometric head predictions. A brief review of first and second order analysis for the scalar case follows. Finally, the extension of these analyses to a general form of numerical models is presented, with details of the analysis available in three appendices.

Part 2 of this paper, to be presented later, contains a series of numerical examples illustrating the effects of parameter uncertainty (material properties, boundary conditions, and initial conditions) on predictions of heads for one- and two-dimensional aquifer flows.

UNCERTAINTY CONCEPTS IN GROUNDWATER SYSTEMS

Uncertainty in groundwater systems may be divided into two classes: intrinsic uncertainty and information uncertainty. The first class derives from the variability of certain natural properties or processes and is an irreducible uncertainty inherent to the system. The second class is the result of 'noisy' or incomplete information about the system and may be reduced by various strategies, notably further measurements and analysis.

The spatial and temporal variation of parameters, such as the recharge rate, and spatial variability of properties, such as hydraulic conductivity, are extremely complicated and in general defy exact description. For example, the relatively simple or large scale spatial changes of properties, such as the change of hydraulic conductivity between geological units, can be included in deterministic descriptions. However, smaller scale spatial fluctuations are superimposed on the large 'deterministic' scale variation. The small scale fluctuations, in turn, contain even smaller scale fluctuations and so on until the property itself ceases to have meaning. Since there is reason to believe that this local variability will affect aquifer behavior, several authors have developed methods to include it in aquifer analysis. The general approach [see, e.g., *Freeze*, 1975; *Bakr et al.*, 1978; *Sagar*, 1978; *Dagan*, 1979; *Smith and Freeze*, 1979] has been to characterize the variations as spatial stochastic processes or phenomena. Acknowledging the fact that this variability is highly complex and irreducible, these techniques do not attempt to manipulate the actual variation of properties; instead they attempt to preserve the spatial statistical properties of the variation. This approach provides a probabilistic description of the magnitude, spatial extent and nature of the effects that the possible range of property variations can have on aquifer behavior, particularly piezometric head. Hence these stochastic property variations are described as a form of intrinsic uncertainty.

Information uncertainty represents the lack, in quantity or quality, of information concerning the aquifer system. Estimates of various properties or descriptive parameters of the

[1] Currently with CDM–Water Resources Engineers, Walnut Creek, California.

Paper number 80W1092.
0043-1397/81/080W-1092$01.00

system will generally contain many inaccuracies, both small and large. These errors arise from a number of sources, such as noisy or scarce measurements, and may be statistical or conceptual in nature. Thus information uncertainty involves errors or the possible range of errors in an analysis of a groundwater system, in contrast to intrinsic uncertainty, which is a concept describing the intrinsic variability of aquifer parameters. Another distinction between the two forms of uncertainty is that information uncertainty is generally reducible through measurements while intrinsic uncertainty is a physical variability undiminished by observations. Finally, the term information uncertainty can refer to our estimate of the true parameter field, or it can refer to our estimate of the statistics of the spatial stochastic process. To date, all analyses of groundwater flow as a spatial stochastic phenomenon have assumed perfect knowledge of these statistics [see, e.g., *Freeze,* 1975; *Bakr et al.,* 1978; *Dagan,* 1979; *Smith and Freeze,* 1979]. This paper makes the same assumption, realizing that it can be relaxed at a later time. Thus, below, consider that information uncertainty refers to knowledge of the true parameter field and not to its statistics. There is only one true field, but it is unknown. A probabalistic description of it is required.

Intrinsic uncertainty or variability, by its very nature, is independent of information uncertainty, but not the reverse! Information uncertainty, and our knowledge of the true parameter field, depend on the description of intrinsic variability. This description can be used to decrease information uncertainty by, for example, permitting qualified inferences between widespread measurements.

As an illustration of the relationship between intrinsic and information uncertainty consider an n dimensional aquifer with hydraulic property, K, distributed in space. Suppose that this property has a known homogeneous, isotropic, two-parameter statistical description, with mean $\hat{K}$ and covariance Cov (r), where r is the distance between any two points. This description has two interpretations. First it represents the spatial stochastic process from which the true field K is drawn, which is one of many possible realizations of K, at least in theory. Here trend $\hat{K}$ and Cov (r) represent the spatial structure of K. In the second interpretation $\hat{K}$ is an estimate for the true field K, and Var [K] = Cov (0), called the estimation variance, represents the confidence or 'degree of belief' in that estimate. These two interpretations span both the concept of intrinsic uncertainty, through the idea of a stochastic process, and information uncertainty, through the Bayesian concept of degree of belief [see, e.g., *Ramsey,* 1950; *Savage,* 1972; *de Finetti,* 1974] in the estimate of the true field.

Now let m point samples of K be taken in the field, with zero sampling error. Then the exact value of K is known at these m points and, because of the spatial correlation structure, these samples also provide some information about the value of K nearby. A new estimate of the K field can be derived, with reduced information uncertainty as indicated by a lower estimation variance. Additional measurements of K, or even other parameters correlated or functionally related to K, will lead to an even better identification of the K field (i.e., even less information uncertainty). The basic statistical properties of the K spatial process is unaltered through all this, and in fact will often be used to aid in this procedure.

This illustration argues that although the information uncertainty on the true K field is reduced by sampling, the intrinsic uncertainty of the spatial stochastic process is irreducible. A particular example is the application of Kriging to the estimation of the property K, as described by *Delhomme* [1978]. In Kriging the true field is called a regionalized variable and there are two characteristic uncertainties, one represented by the variogram and one by the estimation variance. The variogram describes the random component of the spatial variability of the property K and is an irreducible descriptor of the system. The Kriging process must take this spatial variability into account in order to properly synthesize the available data into a map of the property. The estimation variance, on the other hand, is a statistical measure of the accuracy of the Kriging process. In regions where sufficient data about K are available, the Kriging process can achieve good results and therefore yields a small estimation variance; where little information is available, the variance may be large. In this illustration, the variogram represents the intrinsic uncertainty in the system, while the estimation variance is an expression of information uncertainty.

Intrinsic uncertainty and information uncertainty play different roles in the prediction of aquifer behavior. Studying the effect of the spatial variability of property K on piezometric head h leads *Bakr et al.* [1978] to interpret the predicted mean head as a trend, about which there are small scale spatial fluctuations of head described statistically by the head covariance structure. In other words, the spatial stochastic process K begets a spatial stochastic process h. A 'degree of believer' would interpret the results differently. Given that Var [K] represents information uncertainty as to the estimate of the actual distribution of K, he would interpret the predicted mean head as the best estimate of the actual distribution of h, and the calculated covariance would be a measure of the uncertainty of the estimate. This second, information uncertainty, interpretation clearly allows the incorporation of sampled data, either utilizing the underlying spatial stochastic process [*Delhomme,* 1979], or ignoring it [*Wilson et al.,* 1978].

Given that there is significant information uncertainty in natural aquifer systems, and given that the stochastic description of intrinsic uncertainty is incorporated and used profitably in an analysis of information uncertainty on aquifer behavior, it seems that the 'degree of belief' interpretation of predictions is the more practical. Besides, once this approach is selected, questions concerning the ergodicity of the stochastic spatial processes [see, e.g., *Bakr et al.,* 1978; *Dagan,* 1979], become less relevant, as the degree of believer explicitly recognizes that there is only one realization of the stochastic process, which he is trying to estimate.

The procedures described in this paper and illustrated in part 2, however, can be viewed with either interpretation, both of which deal with parameter uncertainty.

UNCERTAINTY IN NUMERICAL MODELS

Above we have presented a discussion of parameter uncertainty, uncertainty due to the variability of parameters and our knowledge of the variability. Parameters have been defined to include aquifer properties, boundary values, source/sink strength, and initial conditions. There is another source of uncertainty due to the appropriateness of our model of aquifer behavior. The model is an abstract description of the aquifer system synthesized from available data and information. This abstraction includes a general description of the governing equations, the position and nature of the boundary conditions, the form of the initial conditions and location of sources/sinks to the system, questions of aquifer homogeneity, isotropy, etc. This abstraction is called a conceptual model

and is ideally a mathematically or intuitively manageable description of the system. In most practical problems this model is solved numerically by discretizing the aquifer domain in time and space. As a result of this discretization a numerical model provides only an approximate solution of a conceptualization of the aquifer situation. More importantly the conceptual model itself may be faulty. All of these sources of information uncertainty about the model are termed model error. Often inaccuracies resulting from model error are attributed to other sources, in particular parameter uncertainty, so that conceptual and other model errors are difficult to isolate, much less quantify.

There are some sources of model error due to the numerical discretization that can be accounted for immediately, and thus eliminated. One of these is the effect of the spatial discretization on the underlying statistics of spatial stochastic processes. For finite difference and finite element models with constant material properties within a cell or an element, continuous aquifer properties of the actual domain are spatially integrated over cells or elements of the discretized domain. Thus the probabalistic description of the spatial process must also be integrated. *Smith and Freeze* [1979] and others have failed to recognize this important feature, which results in a reduction of the variance of the spatially averaged property and an increase of its correlation length. Part 2 will document and illustrate this variance-covariance reduction due to the discretization. By properly accounting for this and other effects of the discretization, model error can be thought of as primarily a description of the effects of a faulty conceptual model.

The remainder of this paper is concerned with methods to assess the quality of numerical predictions under parameter uncertainty, due to intrinsic variability and/or information uncertainty. The role of conceptual model error will be left for future research efforts.

TECHNIQUES FOR THE ANALYSIS OF UNCERTAINTY

A variety of procedures are available to treat the effects of uncertainty on groundwater flow. Some of these are more suited to the complexity and dimension of numerical models. The various approaches can be divided into two main groups: full distribution analyses and first and second moment analyses. Full distribution methods begin with a complete specification of the probabilistic properties of all nondeterministic inputs and parameters of a flow system and attempt to specify completely the probability distribution of the resulting flow. First and second moment methods assume that the first two moments of a random variable or function are sufficient to characterize it. On the basis of this assumption, they consider only the mean and variance-covariance of the piezometric heads and/or flow rates.

The two most important full distribution techniques are the method of derived distributions and Monte Carlo simulation. The derived distributions approach is an analytical method for deriving the probability distribution of a random function given the distributions of its independent variables [*Benjamin and Cornell,* 1970]. The analysis becomes prohibitively complicated unless applied to simple systems with relatively simple probabilistic properties. An example of a derived distribution analysis is *Eagleson*'s [1978] recent evaluation of infiltration due to stochastic precipitation events (temporal variability), or *Sagar and Kisiel*'s [1972] examination of parameter uncertainty for aquifer pump tests. More widely applicable is the Monte Carlo method, which employs numerous replications of flow system simulations, with the parameters and inputs of each simulation generated at random from their respective probability distributions. The results of the simulations are compiled to form estimates of the probability distribution of the aquifer flow or piezometric heads. This method is readily automated and can be applied to analytically complicated distributed parameter systems. A disadvantage of Monte Carlo simulation is that the results are never in the closed analytical form that a derived distribution study strives for and therefore results are not readily transferable to a new situation. Monte Carlo simulations have been applied to the investigation of the effects of spatial variability of physical properties in flow through porous media by several authors, including *Warren and Price* [1961], *Freeze* [1975], and *Smith and Freeze* [1979]. More will be said of Freeze's work in a later discussion of the strengths and limitations of first order analysis.

It is often difficult to obtain the probability distributions that are input to a full distribution analysis, and only slightly easier to estimate their moments. Thus the results obtained, which depend on the exact distribution selected, can be deceptive. They may reveal nothing more than an analysis conducted using only the first two moments. In fact the results of full distributional studies are often presented only in terms of these moments [see *Sagar and Kisiel,* 1972; *Freeze,* 1975; *Eagleson,* 1978].

The assumption underlying the first and second moment methods is that the important information about the random variables (or functions) of interest can be summarized with the mean representing the central or expected tendency of the variable (or function), and the variance-covariance representing the amount of scattering or variation around the mean. Unless the third moment (skewness) or higher moments of the variable are relatively large, they are generally of little interest in application. An example of a function fulfilling this assumption is one which is normally distributed. Such a function has zero skewness and other higher moments of odd order, and all even order moments can be calculated from the variance [*Benjamin and Cornell,* 1970]. Within the framework of second moment methods it is, of course, impossible to test the assumption that the mean and variance-covariance fully describe a function. Thus other methods, in particular full distribution methods, should be used when possible to check it. For example, in his paper on Monte Carlo simulation of steady state one-dimensional flow through uncorrelated porous blocks, *Freeze* [1975] assumed that the logarithm of hydraulic conductivity was normally distributed. Using this model he studied the second moment properties of the piezometric heads along the flow line, finding, among other things, that only a portion of the heads pass statistical tests of normalcy. Freeze's Figure 4 suggests that truncation of the set of possible values that the head can assume near the boundaries results in highly skewed head probability distributions in those areas. His Figure 6, on the other hand, indicates that for coefficient of variation of log K less than or equal to 0.25, the predicted heads are in fact approximately normally distributed over most of the length of flow. These two results suggest that second moment methods may be appropriate when applied to (1) systems with second moment inputs and parameters with relatively small variance, and (2) regions of the flow domain in which the truncating effects of boundary conditions

are not too important. An example of the second case is a system with uncertain boundary conditions.

First and second moment methods can be applied in a number of ways: perturbation and/or Taylor series expansion. In perturbation analyses, the partial differential equation governing the piezometric heads is perturbed slightly, yielding a new equation which governs the random or fluctuating component of the piezometric surface. Analyses based on Taylor series expansions, on the other hand, generally expand analytical or numerical solutions of the piezometric head or flow governing equation around the expected values of the solution's parameters and independent variables. These series expansions may then be used to deduce the probabilistic moments of the piezometric heads. The perturbation approach has been used in both the space [*Tang and Pinder*, 1977] and wave number [*Bakr et al.*, 1978; *Gutjahr et al.*, 1978] domains to attack the problem of spatial variability. The latter manipulates Fourier-Stieltjes integral representations of the probabilistic moments of the logarithm of hydraulic conductivity and flow equation, while the former deals with a more familiar description of groundwater flow as a function of space and time. Analysis in the wave number domain has produced analytical descriptions of the second moment properties of heads in one-, two-, and three-dimensional flows [*Bakr et al.*, 1978]. These results statistically describe flow in aquifers whose properties are continuously variable, rather than aquifers with properties that vary in discrete blocks or aquifers described by lumped parameters. This continuous description is certainly the most realistic and is the main advantage of the wave number domain approach over any other. However, since boundary conditions imply a nonstationarity of the random fields describing the flow system which has not been dealt with using current techniques, most boundary conditions cannot be adequately incorporated into the analysis. Analyses in the wave number domain produce very satisfying results, but are limited to flows of infinite extent where piezometric heads are sought. The time-space domain perturbation approach of *Tang and Pinder* [1978], on the other hand, uses a numerical model to deal with boundary value problems. Despite difficulties in their analysis [see *Gelhar et al.*, 1979] their approach is equivalent to the Taylor series approach suggested in this paper; that is to say, a first order perturbation analysis should yield the same information as a Taylor series analysis of the same order.

First and second moment methods based on Taylor series expansions have been employed by a number of authors. *Cornell* [1972] presented applications of the approach to a wide variety of simple hydrologic and water resources problems and suggested much wider application. *Wilson and Dettinger* [1980] present a number of simple analytical examples specific to groundwater flow applications. A finite element model of flow in a confined aquifer was analyzed by *Sagar* [1978] using this approach with a simple one-dimensional flow example. The method of analysis used in his presentation is the same as that used in this paper; however, in this paper the analysis will be applied to a much broader class of numerical models, which also happens to include the finite element model described by Sagar.

Review of First and Second Order Analysis

In this paper the terms first and second order analysis are often used. First order analysis will be defined to be the analysis of the mean and variance-covariance of a random function based on its first order Taylor series expansion. Second order analysis will refer to analysis of the mean based on a second order Taylor series expansion with the concurrent analysis of the variance-covariance still restricted to use of the first order series expansion. Defined in this way, the mean derived in first and second order analyses may be different; the variance-covariance will not.

First order analysis of the moments of a random function is a simple way of estimating its probabilistic properties. Let f be a generic function of some random variable x. The function is expanded in a Taylor series about the expected value of its independent variable,

$$f(x) = f(\hat{x}) + f^{(1)}(x - \hat{x}) + \frac{1}{2} f^{(2)}(x - \hat{x})^2 + \frac{1}{6} f^{(3)}(x - \hat{x})^3 + \cdots \tag{1}$$

with $f^{(k)}$ equal to the kth derivative of f with respect to x, evaluated at $\hat{x}$, the expected value of x. Second and higher order terms, i.e., terms multiplied by the factor $(x - \hat{x})^k$, $k \geq 2$, are neglected as being small compared to the first two terms. With these terms neglected, the mean or expected value of f is calculated to be

$$E[f] = \hat{f} \doteq E[f(\hat{x}) + f^{(1)}(x - \hat{x})] \tag{2}$$

with $\doteq$ denoting equals, to a first order approximation. $E[f]$ denotes the expected value of f. The expected value is a linear operation so that $E[a + b] = E[a] + E[b]$ and $E[cb] = cE[b]$, where c is a constant and a and b are two random variables.

Using these properties of the expected value,

$$\begin{aligned} \hat{f} &\doteq f(\hat{x}) + f^{(1)}E[x - \hat{x}] = f(\hat{x}) + f^{(1)}[E(x) - \hat{x}] \\ &= f(\hat{x}) \end{aligned} \tag{3}$$

since $E[x] = \hat{x}$ by definition. The second moment of f can be estimated to first order also. The second moment most often used in studying random functions is the covariance, or, in the case of a scalar function such as f, the variance which is defined to be $E[(f - \hat{f})^2] \equiv \sigma_f^2$. The variance is the second moment around the mean value. Using the first order series for f, the variance is

$$\begin{aligned} \sigma_f^2 &\doteq E[\{f(\hat{x}) + f^{(1)}(x - \hat{x}) - \hat{f}\}^2] \\ &\doteq E[\{f(\hat{x}) + f^{(1)}(x - \hat{x}) - f(\hat{x})\}^2] \\ &\doteq E[\{f^{(1)}(x - \hat{x})\}^2] = [f^{(1)}]^2 E[(x - \hat{x})^2] = (f^{(1)})^2 \sigma_x^2 \end{aligned} \tag{4}$$

A second order analysis of the mean of f is carried out analogously using a second order series expansion for the function:

$$f(x) \doteq^2 f(\hat{x}) + f^{(1)}(x - \hat{x}) + \frac{1}{2} f^{(2)}(x - \hat{x})^2 \tag{5}$$

with $\doteq^2$ denoting equals, to a second order approximation. Taking the expected value yields the second order approximation of the mean:

$$\begin{aligned} E[f] = \hat{f} &\doteq^2 f(\hat{x}) + f^{(1)}E[(x - \hat{x})] + \frac{1}{2} f^{(2)}E[(x - \hat{x})^2] \\ &= f(\hat{x}) + \frac{1}{2} f^{(2)}\sigma_x^2 \end{aligned} \tag{6}$$

This estimate of the mean is more accurate than the first order estimate, using information about the expected value and

variability of x. The second order estimate of the mean, then, is actually the expected value of f conditional on the mean and variance of x.

These results are generalized to the case where f and x are vectors in the next section of this paper, when describing numerical models of flow.

First and second order analyses, in general, can only properly be applied to nonlinear systems in which the coefficients of variation are a small fraction of one. These analyses, including the ones presented in this paper, are based on the expected values of truncated Taylor series expansions. If the truncation error is to be small, then the higher order terms in the expansions must be negligible; either the higher order derivatives or the higher moments of the variables must be small. *Cornell* [1972] has suggested that for coefficients of variation $V \leq 0.2$, the analysis method is applicable to moderately nonlinear systems. This restriction is, of course, in agreement with the limitation of second moment analyses to systems with moderate to small variances discussed earlier in connection with Freeze's work.

In this paper the methods of first and second order analysis will be applied to numerical models of groundwater flow. Because of their complexity, neither derived distributions nor wave number domain analyses are suited for use in analyzing numerical models or aquifer conditions in which numerical methods are warranted. Monte Carlo experiments may also be limited in cases requiring complicated numerical solutions. The numerous simulations required by Monte Carlo techniques can lead to an enormous computational burden when the aquifer requires detailed modeling and has complicated boundary conditions, heterogeneous parameters, and/or two- and three-dimensional flow. This computational burden will generally place serious limitations, raised both by economics and expediency, on the accuracy with which estimates of probabilistic parameters can be obtained, since the accuracy of Monte Carlo experiments is an increasing function of the number of simulations carried out. Because of these limitations on the use of other methods of anlaysis, first and second order analyses of numerical models are a natural choice and can generally be made with an accuracy consistent with the accuracy of the numerical model itself.

MATHEMATICAL DEVELOPMENT

In this section, a first and second order analysis of a class of numerical models of flow in porous media is presented. Because these numerical models are based on discrete grids, the solution obtained at each time step may be conceived of as vector of piezometric heads, and the numerical models can be manipulated in the form of matrix equations. Upon expansion of these equations in a Taylor series, predictions of the first and second moments of the piezometric head can be obtained through first order analysis. Prediction of the first moment can be further refined using second order analysis.

The State-Space Equation

The flow of water in a saturated porous medium is described by a differential equation incorporating both mass conservation and dynamic (Darcy's law) principles. For example, one-dimensional steady flow is described by

$$\frac{d}{dx}\left[K\frac{dh}{dx}\right] = 0 \tag{7}$$

where K is the hydraulic conductivity and h is piezometric head. For three-dimensional, transient flow in a deformable medium, an appropriate expression is [*Bear*, 1972]

$$S_0\frac{\partial h}{\partial t} = \nabla \cdot [\mathbf{K}\nabla h] \tag{8}$$

where $\mathbf{K}$ is the hydraulic conductivity tensor and S_0 is specific storage. Many aquifers are commonly modeled with the two-dimensional, essentially horizontal flow, 'hydraulic' equation [*Bear*, 1972]:

$$S\frac{\partial h}{\partial t} = \frac{\partial}{\partial x}\left[T_{xx}\frac{\partial h}{\partial x}\right] + \frac{\partial}{\partial y}\left[T_{yy}\frac{\partial h}{\partial y}\right] + \frac{\partial}{\partial x}\left[T_{xy}\frac{\partial h}{\partial y}\right] + \frac{\partial}{\partial y}\left[T_{yx}\frac{\partial h}{\partial x}\right] + \Sigma Q_i\,\delta(x - x_i)\,\delta(y - y_i) \tag{9}$$

where T_{xx}, T_{yy}, T_{xy}, and T_{yx} are components of the second rank transmissivity tensor, x and y are Cartesian coordinates, S is aquifer storativity, Q_i is the strength of a source/sink at location x_i, y_i, and δ is the Dirac delta function. These equations can be discretized in space and time and solved numerically using various methods, most commonly finite differences or finite elements. For situations with boundary conditions at fixed locations the numerical equations can be manipulated into the following matrix equation [see, e.g., *Pinder and Gray*, 1976], called the state-space equation [*Brockett*, 1970].

$$\frac{d\,\mathbf{h}}{dt} = \mathbf{A}\mathbf{h} + \mathbf{C}\mathbf{u} \tag{10}$$

where $\mathbf{h}$ is the $(n \times 1)$ vector of piezometric heads to be solved for and n is the number of node points in the discretized grid. The notation used throughout the paper is that boldface capital letters (e.g., $\mathbf{A}$) indicate matrices and boldface lower case letters (e.g., $\mathbf{h}$) indicate vectors. The vector $\mathbf{u}(m \times 1)$ is made up of boundary conditions, both Dirichlet and Neumann. $\mathbf{A}(n \times n)$ and $\mathbf{C}(n \times n)$ are coefficient matrices and functions of model parameters, such as the hydraulic conductivity and specific storage, or the transmissivity and storativity. As discussed in part 2, these model parameters are not necessarily idential to the aquifer properties. The exact form of $\mathbf{A}$ and the product $\mathbf{Cu}$ generally depends upon the modeling method implemented (e.g., finite difference, finite element), although cases do exist, such as some one-dimensional flows, in which the same $\mathbf{A}$ matrix is used for both finite difference and finite element models. It is also assumed for the present paper that the state-space equation is linear in terms of state. Equations (7) and (8) fit this model, but (9) requires the further specification that saturated thickness at a point remain essentially constant over time. Many aquifer situations conform to this restriction.

The state-space equation (10) specifies that the future piezometric head at a node point in the discretized aquifer domain is a linear function of the current piezometric head, the piezometric head at other node points in the domain, and the boundary conditions. Altogether, the heads are known as the state vector; they describe the state of the system. Solution of the time dependent state-space equation is generally based on a finite difference discretization of the time derivative. The result may be written in matrix notation as a state-transition equation:

$$\begin{aligned}\mathbf{h}(t + \Delta t) &= \mathbf{\Theta}(t + \Delta t, t)\mathbf{h}(t) + \mathbf{G}\mathbf{u} \\ &= \mathbf{\Phi}\mathbf{h}(t) + \mathbf{G}\mathbf{u}\end{aligned} \tag{11}$$

where $\Theta(t_2, t_1)$ is the $(n \times n)$ state transition matrix [*Brockett*, 1970] relating the state at time t_2 to the earlier state at time t_1, Φ is the particular transition matrix for the state propagation from time t to time $t + \Delta t$, so that $\Phi \equiv \Theta(t + \Delta t, t)$ and $\mathbf{Gu}(n \times 1)$ is given by [*Brockett*, 1970]

$$\mathbf{Gu} = \int_t^{t+\Delta t} \Theta(t + \Delta t, \tau)\mathbf{Cu}(\tau)\, d\tau \tag{12}$$

In this study it is assumed that the parameters and boundary conditions, and consequently **A** and **u**, are constant over Δt. This is actually an assumption of perfect temporal correlation for the boundary conditions and parameters. Another choice could conceivably have been made but only at the cost of greatly complicating the analysis. The state-transition equation (11) says that the piezometric heads at some time $t + \Delta t$ are a linear function of the heads at some earlier time t, and that a matrix function Φ, of the parameters can be formed which describes the transition of heads from time t to $t + \Delta t$. The heads at times earlier than t do not affect the head, $\mathbf{h}(t + \Delta t)$, predicted by the equation. For this reason, each time step can be studied individually and, often in the following, $\mathbf{h}(t)$ will be referred to as the initial condition for $\mathbf{h}(t + \Delta t)$. The heads $\mathbf{h}(t + \Delta t)$ are also a function of boundary conditions **u**. The relationship is assumed to be linear and described by the matrix function of parameters, **G**.

The transition matrix Φ and the matrix **G** are a function of the time integration method used. Some of the forms that may be taken are:

Explicit

$$\Phi = \mathbf{I} + \mathbf{A}\Delta t \qquad \mathbf{G} = \mathbf{C}\Delta t \tag{13a}$$

Fully implicit

$$\Phi = (\mathbf{I} - \mathbf{A}\Delta t)^{-1} \qquad \mathbf{G} = (\mathbf{I} - \mathbf{A}\Delta t)^{-1}\mathbf{C}\Delta t = \Phi\mathbf{C}\Delta t \tag{13b}$$

Exponential [*Brockett*, 1970]

$$\Phi = \left(\mathbf{I} + \mathbf{A}\Delta t + \frac{1}{2}\mathbf{A}^2\Delta t + \cdots\right) = \exp(\mathbf{A}\Delta t)$$

$$\mathbf{G} = \Delta t\left(\mathbf{I} + \frac{1}{2}\mathbf{A}\Delta t + \frac{1}{6}\mathbf{A}^2\Delta t^2 + \cdots\right)\mathbf{C}$$

$$= -\mathbf{A}^{-1}[\mathbf{I} - \exp(\mathbf{A}\Delta t)]\mathbf{C} \tag{13c}$$

where **I** is an identity matrix of the same dimension as **A** (i.e., $n \times n$). Obviously, Φ and **G** may be very nonlinear functions of **A** and **C**, and thus are generally very complicated functions of the parameters. In addition, they depend explicitly on the time discretization interval, Δt.

The steady state configuration of heads may be calculated by letting $d\mathbf{h}/dt = \mathbf{0}$ where **0** is a null vector of dimension $(n \times 1)$. The resulting matrix equation is $\mathbf{0} = \mathbf{Ah} + \mathbf{Cu}$, which yields a steady state vector of heads described by

$$\mathbf{h}_s = \mathbf{Gu} \tag{14}$$

where $\mathbf{G} = -\mathbf{A}^{-1}\mathbf{C}$. Notice that the steady state heads are a function of the parameters and boundary conditions but of course are not a function of initial conditions or time step Δt.

Taylor's Series Expansion and Vetter Calculus

The methods of predicting to second order the mean, variance, and covariance of a function require a Taylor series expansion of the function in terms of its uncertain independent variables. Because $\mathbf{h}(t + \Delta t)$ is a sum of matrix products, and the matrices are functions of the numerous model parameters, a special notation helps to conveniently express the Taylor expansion. *Vetter* [1973] has proposed a notation that allows the manipulation of derivatives of matrices with respect to matrices. Derivatives of vectors with respect to vectors, and so on, are specializations of his calculus. In addition, *Vetter* [1973] has derived the proper form of the Taylor series expansion of matrices. Because of its compact form and the completeness of the theory as derived by Vetter, his notation will be employed here. In addition, the calculus has the advantage that all arrays are either vectors or matrices and no higher dimensioned tensors are required. Because the notation will be unfamiliar to most readers, a basic introduction to the aspects of Vetter matrix calculus required in this paper is presented in Appendix A.

The Taylor expansion of a vector, derived by *Vetter* [1973] is

$$\mathbf{f}(\mathbf{y}) = \mathbf{f}(\mathbf{y}_0) + \sum_{k=1}^{\infty} \frac{1}{k!}[D_{y'^k}\mathbf{f}] \cdot (\mathbf{y} - \mathbf{y}_0)^{*k} \tag{15}$$

where **f** is a vector function of some other vector **y**, and $\mathbf{y}_0$ is the value of **y** around which the expansion takes place. $[D_{y'^k}\mathbf{f}]$ is the kth derivative of **f** with respect to the transpose of **y**, evaluated at $\mathbf{y}_0$, the prime indicates matrix transpose, and $(\mathbf{y} - \mathbf{y}_0)^{*k}$ is the result of k Kronecker products (see Appendix A) of $(\mathbf{y} - \mathbf{y}_0)$ with itself. As a result of Vetter's notation, this equation looks very much like the usual Taylor expansion of a scalar (1). Notice, however, that if **f** is an $n \times 1$ vector and **y** has dimension $p \times l$, then $D_{y'^k}\mathbf{f}$ has dimension $n \times p^k$ and $(\mathbf{y} - \mathbf{y}_0)^{*k}$ is a $p^k \times 1$ vector. Although the notation is compact, the expansion involves matrices of very large dimension. Even a second order expansion can involve matrices that are prohibitively large unless the structure of the derivative matrices are taken into account. Fortunately, the structure of many groundwater flow models is such that the derivative matrices are very sparse, and great computational savings can be obtained. In fact, second order estimates can be computed exactly through manipulation of matrices no larger than the covariance matrices involved.

To first order the Taylor series is

$$\mathbf{f}(\mathbf{y}) \stackrel{1}{=} \mathbf{f}(\mathbf{y}_0) + (\mathbf{D}_y\mathbf{f})(\mathbf{y} - \mathbf{y}_0) \tag{16}$$

where $\stackrel{1}{=}$ signifies equals, to a first order approximation. Taking the expected value of this, the mean of **f** is estimated to first order by

$$E[\mathbf{f}] = \hat{\mathbf{f}} \stackrel{1}{=} E[\mathbf{f}(\hat{\mathbf{y}})] + E[(D_y\mathbf{f})(\mathbf{y} - \hat{\mathbf{y}})] = \mathbf{f}(\hat{\mathbf{y}}) + (D_y\mathbf{f})E[\mathbf{y} - \hat{\mathbf{y}}] = \mathbf{f}(\hat{\mathbf{y}}) \tag{17}$$

where it has been assumed that the expansion is around the expected value of **y**, $\mathbf{y}_0 = E[\mathbf{y}] = \hat{\mathbf{y}}$. Notice that both $\mathbf{f}(\hat{\mathbf{y}})$ and $(D_y\mathbf{f})$ are constant. The first order estimate of the mean is exactly the value obtained through the application of a traditional deterministic approach. This estimate of the mean is the expected value of **f** conditional on the expected value of **y**.

To second order the estimate of the mean **f** is, under the same assumption,

$$E[\mathbf{f}(\mathbf{y})] = \hat{\mathbf{f}} \overset{2}{=} \mathbf{f}(\hat{\mathbf{y}}) + (D_{\mathbf{y}}\mathbf{f})E[\mathbf{y} - \hat{\mathbf{y}}] + \frac{1}{2}(D_{\mathbf{y}^2}\mathbf{f})E[(\mathbf{y} - \hat{\mathbf{y}})^{*2}]$$

$$= \mathbf{f}(\hat{\mathbf{y}}) + \frac{1}{2}(D_{\mathbf{y}^2}\mathbf{f})E[(\mathbf{y} - \hat{\mathbf{y}})^{*2}] \tag{18}$$

where $E[(\mathbf{y} - \hat{\mathbf{y}})^{*2}]$ is a vector of the elements in the covariance matrix of $\mathbf{y}$. The second order mean is the expected value of $\mathbf{f}$ conditional on the mean and covariance of $\mathbf{y}$. Thus the second order mean includes an additional contribution that depends on the amount of variation or error that may be expected to exist between the mean value of $\mathbf{y}$ and its true value. This contribution is also a function of the nonlinearity of $\mathbf{f}$ near $\hat{\mathbf{y}}$. As the function $\mathbf{f}$ becomes more nonlinear, the result of variations of $\mathbf{y}$ about the mean $\hat{\mathbf{y}}$ are likely to be more important. For a highly nonlinear function of $\mathbf{y}$, a first order or 'deterministic' model may yield misleading results even when the estimate of the independent variable $\mathbf{y}$ is subject to only small errors. When $\mathbf{y}$ is uncertain or varies about its mean, the second order estimate of $\hat{\mathbf{f}}$ will generally be more accurate than a deterministic estimator. However, caution should be exercised when the difference between first and second order estimates is significant, as this probably indicates such extreme nonlinearity that the second order estimate is inadequate as well. In such cases, Monte Carlo simulations or other full distributional techniques may be required. Generally, though, it is expected that first order estimates will be adequate.

The covariance of $\mathbf{f}$ may be approximated to first order by

$$\begin{aligned}\text{Cov}(\mathbf{f}) &= E[(\mathbf{f} - \hat{\mathbf{f}})(\mathbf{f} - \hat{\mathbf{f}})'] \\ &\overset{1}{=} E[\{\hat{\mathbf{f}} + (D_{\mathbf{y}}\mathbf{f})(\mathbf{y} - \hat{\mathbf{y}}) - \hat{\mathbf{f}}\}\{\hat{\mathbf{f}} + (D_{\mathbf{y}}\mathbf{f})(\mathbf{y} - \hat{\mathbf{y}}) - \hat{\mathbf{f}}\}'] \\ &= E[(D_{\mathbf{y}}\mathbf{f})(\mathbf{y} - \hat{\mathbf{y}})(\mathbf{y} - \hat{\mathbf{y}})'(D_{\mathbf{y}}\mathbf{f})'] \\ &= (D_{\mathbf{y}}\mathbf{f})E[(\mathbf{y} - \hat{\mathbf{y}})(\mathbf{y} - \hat{\mathbf{y}})'](D_{\mathbf{y}}\mathbf{f})' \\ &= (D_{\mathbf{y}}\mathbf{f})\,\text{Cov}(\mathbf{y})(D_{\mathbf{y}}\mathbf{f})'\end{aligned} \tag{19}$$

with Cov $(\mathbf{y})$ the covariance of $\mathbf{y}$ and $(D_{\mathbf{y}}\mathbf{f})$ evaluated at the expected value of $\mathbf{y}$. Notice that the covariance of $\mathbf{f}$ is a function of the uncertainty or variability of $\mathbf{y}$, and the sensitivity of $\mathbf{f}$ to $\mathbf{y}$ in the neighborhood of $\hat{\mathbf{y}}$. Thus any elements of $\mathbf{f}$ that are insensitive to $\mathbf{y}$ will be little affected by uncertainty in $\mathbf{y}$. Alternatively, if $\mathbf{y}$ is known very well (i.e., small covariances), then $\mathbf{f}$ will have small covariances unless $\mathbf{f}$ is very sensitive to $\mathbf{y}$.

Estimate of the First Moment of Piezometric Head for Transient Flow

Applying first and second order Taylor series expansions to numerical solution of the state-transition equation for piezometric heads in an aquifer (11) yields

$$\mathbf{h}(t + \Delta t) \overset{1}{=} \hat{\Phi}\hat{\mathbf{h}}(t) + \hat{\mathbf{G}}\hat{\mathbf{u}} + \mathbf{D}_{\mathbf{b}}(\mathbf{b} - \hat{\mathbf{b}}) + \mathbf{D}_{\mathbf{u}}(\mathbf{u} - \hat{\mathbf{u}}) + \mathbf{D}_{\mathbf{h}}[\mathbf{h}(t) - \hat{\mathbf{h}}(t)] \tag{20a}$$

and

$$\begin{aligned}\mathbf{h}(t + \Delta t) \overset{2}{=} \hat{\Phi}\hat{\mathbf{h}}(t) &+ \hat{\mathbf{G}}\hat{\mathbf{u}} + \mathbf{D}_{\mathbf{b}}(\mathbf{b} - \hat{\mathbf{b}}) + \mathbf{D}_{\mathbf{u}}(\mathbf{u} - \hat{\mathbf{u}}) \\ &+ \mathbf{D}_{\mathbf{h}}[\mathbf{h}(t) - \hat{\mathbf{h}}(t)] + \frac{1}{2}\mathbf{D}_{\mathbf{bb}}(\mathbf{b} - \hat{\mathbf{b}})^{*2} \\ &+ \frac{1}{2}\mathbf{D}_{\mathbf{bu}}[(\mathbf{b} - \hat{\mathbf{b}}) * (\mathbf{u} - \hat{\mathbf{u}})] \\ &+ \frac{1}{2}\mathbf{D}_{\mathbf{bh}}[(\mathbf{b} - \hat{\mathbf{b}}) * (\mathbf{h}(t) - \hat{\mathbf{h}}(t))] \\ &+ \frac{1}{2}\mathbf{D}_{\mathbf{uu}}(\mathbf{u} - \hat{\mathbf{u}})^{*2} + \frac{1}{2}\mathbf{D}_{\mathbf{ub}}[(\mathbf{u} - \hat{\mathbf{u}}) * (\mathbf{b} - \hat{\mathbf{b}})] \\ &+ \frac{1}{2}\mathbf{D}_{\mathbf{uh}}[(\mathbf{u} - \hat{\mathbf{u}}) * (\mathbf{h}(t) - \hat{\mathbf{h}}(t))] \\ &+ \frac{1}{2}\mathbf{D}_{\mathbf{hh}}[\mathbf{h}(t) - \hat{\mathbf{h}}(t)]^{*2} \\ &+ \frac{1}{2}\mathbf{D}_{\mathbf{hb}}[\{\mathbf{h}(t) - \hat{\mathbf{h}}(t)\} * (\mathbf{b} - \hat{\mathbf{b}})] \\ &+ \frac{1}{2}\mathbf{D}_{\mathbf{hu}}[(\mathbf{h}(t) - \hat{\mathbf{h}}(t)) * (\mathbf{u} - \hat{\mathbf{u}})]\end{aligned} \tag{20b}$$

respectively. The operator $*$ signifies a Kronecker product (see Appendix A). The $\hat{\Phi}$ and $\hat{\mathbf{G}}$ denote the matrices Φ and $\mathbf{G}$ evaluated at the expected value of the parameters. Also $\mathbf{b}$ is a vector of the uncertain parameters. An abbreviated notation has been adopted in this equation and will be continued throughout the paper; that is $\mathbf{D}_{\mathbf{b}}$, $\mathbf{D}_{\mathbf{u}}$, $\mathbf{D}_{\mathbf{h}}$ are defined to be the derivative or sensitivity matrices of $\mathbf{h}(t + \Delta t)$ with respect to the transpose of the parameters $\mathbf{b}'$, the boundary conditions $\mathbf{u}'$, and the initial conditions $\mathbf{h}'(t)$. The second derivatives of $\mathbf{h}(t + \Delta t)$ with respect to parameters, boundary conditions, and initial conditions are denoted by $\mathbf{D}$ matrices with double subscripts. All of these derivative matrices are evaluated at the expected values of $\mathbf{b}$, $\mathbf{u}$, and $\mathbf{h}(t)$, i.e., $\hat{\mathbf{b}}$, $\hat{\mathbf{u}}$, $\hat{\mathbf{h}}(t)$. A discussion of the various $\mathbf{D}$ matrices above is presented in Appendix B (first derivatives) and Appendix C (second derivatives).

Taking the expected value of (20*a*) and (20*b*) yields first and second order estimates of the mean of $\mathbf{h}(t + \Delta t)$:

$$\hat{h}(t + \Delta t) \overset{1}{=} \hat{\Phi}\,\hat{\mathbf{h}}(t) + \hat{\mathbf{G}}\,\hat{\mathbf{u}} \tag{21a}$$

and

$$\begin{aligned}\hat{h}(t + \Delta t) \overset{2}{=} \hat{\Phi}\,\hat{\mathbf{h}}(t) + \hat{\mathbf{G}}\,\hat{\mathbf{u}} + \frac{1}{2}\mathbf{D}_{\mathbf{bb}}E[(\mathbf{b} - \hat{\mathbf{b}})^{*2}] \\ + \mathbf{D}_{\mathbf{bh}}E[(\mathbf{b} - \hat{\mathbf{b}}) * (\mathbf{h}(t) - \hat{\mathbf{h}}(t))]\end{aligned} \tag{21b}$$

The second order mean above was derived under the assumption that the boundary conditions $\mathbf{u}$ and the parameters $\mathbf{b}$ are uncorrelated. It was also noticed that $\mathbf{h}(t + \Delta t)$ is a linear function of both $\mathbf{u}$ and $\mathbf{h}(t)$, and thus $\mathbf{D}_{\mathbf{hh}}$, $\mathbf{D}_{\mathbf{uu}}$, $\mathbf{D}_{\mathbf{uh}}$, and $\mathbf{D}_{\mathbf{hu}}$ are null matrices. Finally, use is made of the fact that the second order contribution

$$1/2\ \mathbf{D}_{\mathbf{bh}}E[(\mathbf{b} - \hat{\mathbf{b}}) * (\mathbf{h}(t) - \hat{\mathbf{h}}(t))]$$

is equal to

$$1/2\ \mathbf{D}_{\mathbf{hb}}E[(\mathbf{h}(t) - \hat{\mathbf{h}}(t)) * (\mathbf{b} - \hat{\mathbf{b}})]$$

to simplify the form of the second order mean; thus the fourth term in (21*b*) is the sum of the two contributions.

Examination of (21*b*) reveals that the second order contributions to the mean are only due to parameter uncertainty or variability; this is because the equations of a numerical model of groundwater flow are generally nonlinear in their parameters, such as hydraulic conductivity. For this reason a deterministic model may be expected to yield inaccurate results in the presence of parameter uncertainty, unless special measures

are taken in fitting the parameters to obtain a 'best' model performance. A simple application of expected values may be inadequate.

Variations of the boundary conditions $\mathbf{u}$ around their expected values do not affect higher order estimates of the mean heads. In applying a numerical model to a system with no uncertainty in parameters, such as transmissivities or permeabilities, a deterministic estimate of the mean piezometric heads is sufficient.

Estimate of the Second Moment of Piezometric Head for Transient Flow

The first order estimate of the covariance of heads at time $t + \Delta t$ is

$$\begin{aligned}\mathbf{P}_h(t+\Delta t) &\stackrel{1}{=} E(\{\mathbf{D}_b(\mathbf{b}-\hat{\mathbf{b}}) + \mathbf{D}_u(\mathbf{u}-\hat{\mathbf{u}}) + \mathbf{D}_h(\mathbf{h}(t)-\hat{\mathbf{h}}(t))\}\\ &\quad \{\mathbf{D}_b(\mathbf{b}-\hat{\mathbf{b}}) + \mathbf{D}_u(\mathbf{u}-\hat{\mathbf{u}}) + \mathbf{D}_h(\mathbf{h}(t)-\hat{\mathbf{h}}(t))\}')\\ &= \mathbf{D}_b\mathbf{P}_b\mathbf{D}_b' + \mathbf{D}_b\mathbf{P}_{bh}(t)\mathbf{D}_h' + \mathbf{D}_u\mathbf{P}_u\mathbf{D}_u' \\ &\quad + \mathbf{D}_u\mathbf{P}_{uh}(t)\mathbf{D}_h' + \mathbf{D}_h\mathbf{P}_{hb}(t)\,\mathbf{D}_b' + \mathbf{D}_h\mathbf{P}_{hu}(t)\mathbf{D}_u' \\ &\quad + \mathbf{D}_h\mathbf{P}_h(t)\mathbf{D}_h'\end{aligned} \tag{22}$$

with $\mathbf{P}_y(t)$ defined to be the covariance matrix of the vector $\mathbf{y}$ at time t and $\mathbf{P}_{yz}(t)$ equal to the cross-covariance matrix of $\mathbf{y}$ and $\mathbf{z}$. By definition, $\mathbf{P}_{yz} = \mathbf{P}_{zy}'$. The derivative matrices $\mathbf{D}_h$ and $\mathbf{D}_u$ are equal to $\hat{\mathbf{\Phi}}$ and $\hat{\mathbf{G}}$, respectively, as shown in Appendix B. To keep this expression for $\mathbf{P}_h(t + \Delta t)$ simple, it has been assumed that the parameters and boundary conditions are uncorrelated; i.e.,

$$\mathbf{P}_{bu} = E[(\mathbf{b}-\hat{\mathbf{b}})(\mathbf{u}-\hat{\mathbf{u}})'] = \mathbf{0}$$

is a null matrix. Because the matrix product $\mathbf{DD}'$ for any real matrix $\mathbf{D}$ is positive semidefinite, the first order estimate of $\mathbf{P}_h(t + \Delta t)$ always meets the requirement that all covariance matrices be positive semidefinite.

$\mathbf{DD}'$ is a diagonal matrix for only a small class of matrices $\mathbf{D}$, suggesting that the covariance matrix $\mathbf{P}_h(t + \Delta t)$ will not generally be a diagonal matrix even when $\mathbf{P}_u$, $\mathbf{P}_b$, etc., are; that is, the elements of $\mathbf{h}(t + \Delta t)$ will often be correlated even when $\mathbf{u}$, $\mathbf{b}$, and $\mathbf{h}(t)$ are not. On the other hand, if the $\mathbf{D}$ matrices are diagonal, then $\mathbf{P}_h(t + \Delta t)$ assumes a correlation structure that is similar to the correlation of the uncertain properties at time t. Thus the covariance matrix $\mathbf{P}_h(t + \Delta t)$ is affected by two important length scales: (1) the correlation length of the uncertain properties at time t and (2) the distance over which the heads are sensitive to spatially distributed parameters. The two 'length scales' interact so that generally $\mathbf{P}_h(t + \Delta t)$ is correlated over longer distances than the uncertain parameters, etc.

Experience shows that in most cases, $\mathbf{P}_h(t + \Delta t)$ is greater than $\mathbf{P}_h(t)$, in the absence of measurements. Thus as time proceeds, uncertainty regarding the piezometric heads in an aquifer grows at a rate which depends on the sensitivity of $\mathbf{h}(t + \Delta t)$ to the parameters, boundary conditions, and initial conditions, and depends on the amount of uncertainty in each. When the derivative matrices or covariance matrices contain large numbers (high sensitivities or uncertainties), the increase in uncertainty at each time step can be significant.

The cross-covariance matrices in (22) propagate with each time step as follows

$$\begin{aligned}\mathbf{P}_{hb}(t+\Delta t) &= E[(\mathbf{h}(t)-\hat{\mathbf{h}}(t))(\mathbf{b}-\hat{\mathbf{b}})'] \\ &\stackrel{1}{=} E[\{\mathbf{D}_b(\mathbf{b}-\hat{\mathbf{b}}) + \mathbf{D}_u(\mathbf{u}-\hat{\mathbf{u}}) + \mathbf{D}_h(\mathbf{h}(t)-\hat{\mathbf{h}}(t))\} \\ &\qquad \{\mathbf{b}-\hat{\mathbf{b}}\}'] \\ &= \mathbf{D}_b\mathbf{P}_b + \mathbf{D}_h\mathbf{P}_{hb}(t)\end{aligned} \tag{23}$$

and

$$\mathbf{P}_{hu}(t+\Delta t) = \mathbf{D}_u\mathbf{P}_u + \mathbf{D}_h\mathbf{P}_{hu}(t) \tag{24}$$

where $\mathbf{P}_{ub}$ is again assumed equal to a null matrix. The parameter and boundary condition covariances are not a function of time as a result of the assumption throughout this paper that $\mathbf{b}$ and $\mathbf{u}$ are constant in time.

Estimates of First and Second Moments for Steady Flow

If the boundary conditions are specified such that the expected value of the predicted heads approaches a steady state value, then so too will the covariance matrix. The steady state approached by the heads is given to first and second order by

$$\mathbf{h}_s \stackrel{1}{=} \hat{\mathbf{G}}\hat{\mathbf{u}} + \mathbf{D}_b(\mathbf{b}-\hat{\mathbf{b}}) + \mathbf{D}_u(\mathbf{u}-\hat{\mathbf{u}}) \tag{25a}$$

and

$$\mathbf{h}_s \stackrel{2}{=} \hat{\mathbf{G}}\hat{\mathbf{u}} + \mathbf{D}_b(\mathbf{b}-\hat{\mathbf{b}}) + \mathbf{D}_u(\mathbf{u}-\hat{\mathbf{u}}) + \frac{1}{2}\mathbf{D}_{bu}[(\mathbf{b}-\hat{\mathbf{b}})*(\mathbf{u}-\hat{\mathbf{u}})]$$

$$+ \frac{1}{2}\mathbf{D}_{ub}[(\mathbf{u}-\hat{\mathbf{u}})*(\mathbf{b}-\hat{\mathbf{b}})] + \frac{1}{2}\mathbf{D}_{bb}[(\mathbf{b}-\hat{\mathbf{b}})^{*2}] \tag{25b}$$

where the $\mathbf{D}$ matrices, in this case, denote derivatives of $\mathbf{h}_s$ and where it has been noted that $\mathbf{D}_{uu}$ is a null matrix. Taking the expected value of these two equations yields first and second order estimates of the steady state mean:

$$\hat{\mathbf{h}}_s = E[\mathbf{h}_s] \stackrel{1}{=} \hat{\mathbf{G}}\hat{\mathbf{u}} + \mathbf{D}_bE[\mathbf{b}-\hat{\mathbf{b}}] + \mathbf{D}_uE[\mathbf{u}-\hat{\mathbf{u}}] = \hat{\mathbf{G}}\hat{\mathbf{u}} \tag{26a}$$

and

$$\hat{\mathbf{h}}_s \stackrel{2}{=} \hat{\mathbf{G}}\hat{\mathbf{u}} + \frac{1}{2}\mathbf{D}_{bb}E[(\mathbf{b}-\hat{\mathbf{b}})^{*2}] \tag{26b}$$

when it is assumed that $\mathbf{b}$ and $\mathbf{u}$ are uncorrelated. Notice that, as in the case of transient heads $\mathbf{h}(t + \Delta t)$, the first order mean is the same as a deterministic prediction, while the second order mean includes contributions due to parameter uncertainty. The steady state value that the covariance of heads approaches is given to first order by

$$\begin{aligned}\text{Cov}(\mathbf{h}_s) = \mathbf{P}_h(s) &= E[(\mathbf{h}_s-\hat{\mathbf{h}}_s)\cdot(\mathbf{h}_s-\hat{\mathbf{h}}_s)'] \\ &\stackrel{1}{=} \mathbf{D}_b\mathbf{P}_b\mathbf{D}_b' + \mathbf{D}_u\mathbf{P}_u\mathbf{D}_u'\end{aligned} \tag{27}$$

The steady state covariance of piezometric heads is a function only of uncertainty in parameters and boundary conditions and does not depend on uncertainties in the head prior to achieving the steady state. This suggests that the contribution of initial condition uncertainty becomes less important as time passes. The steady state covariances can be obtained by allowing the second terms in (23) and (24) to vanish as the initial conditions cease to contribute to head uncertainty.

Comments

As long as the numerical solution method for steady state of transient piezometric heads can be conceived in a state-space

form, $\mathbf{A}$ and $\mathbf{\Phi}$ matrices can be derived or approximated. A Vetter-Taylor series expansion based on exact derivatives can be used to yield first and second order head estimates. Although computationally burdensome, the derivative matrices can also be found through numerical methods if exact methods are prohibited by an excessively complicated model structure. Numerical derivatives should be avoided if possible since the advantage of the well-ordered sparcity of the derivative matrices (especially the second derivatives) would be difficult to capture in an approach based on approximate derivatives.

The results derived in this section are quite general, since no specification has been made of the structure of $\mathbf{A}$, $\mathbf{C}$, and $\mathbf{u}$, or the time integration scheme used, i.e., the form of the $\mathbf{\Phi}$ and $\mathbf{G}$ matrices. In Appendices B and C, equations for the derivatives of $\mathbf{h}(t + \Delta t)$ required in the above analysis are presented. Although several of the formulas in those appendices depend on the precise model formulation, this does not indicate any real limitation of the analysis described above.

COMPUTATIONAL BURDEN OF FIRST AND SECOND ORDER ANALYSES

Computationally, the analysis described above is considerably more burdensome than a simple deterministic solution for the piezometric heads. If a full matrix inversion is performed to solve for the steady state heads (14), then the number of multiplications and divisions is of order n^3 (n equal to the number of nodes) with most of the burden arising in the formation of the $\mathbf{G}$ matrix. If the banded nature of the coefficient matrices is taken into account, great computational savings can be obtained, reducing the complete deterministic solution to $O(n^2)$ multiplications for a square grid [*Dahlquist and Bjorck*, 1974]. The matrix inversion must be carried out if a first or second order analysis is to be applied. Equation (27) for the head covariance represents $O(2np^2 + 2nr^2)$ additional operations, where p is the number of uncertain parameters describing aquifer properties and r is the number of uncertain boundary conditions. In a one-dimensional flow problem, $n \simeq p$ and $n \gg r$, implying $O(2n^3)$ additional operations, or $O(3n^3)$ total operations. For two-dimensional flows $p \cong kn$, where $k = 2$ for the link hydraulic conductivities in the finite difference solution of part 2, and $k < 1$ for some finite element models; thus operations total $O[(2k^2 + 1)n^3]$. A second order analysis of the mean can be enormously burdensome. However, if the simple and sparse structure of the $\mathbf{D}_{bb}$ matrix is taken into account in (26*b*), the operations may add only an additional $O(3np^2)$ steps.

Compare this to a Monte Carlo simulation which requires the solution of a large number of matrix equations (14), thus requiring from $O(mn^2)$ to $O(mn^3)$ computational steps, where m is the number of simulations (usually in the hundreds) required to obtain statistically meaningful results. An important additional burden is the generation of the random parameters. The procedure, outlined by *Smith and Freeze* [1979] or *Wilson* [1979] for generating spatially correlated parameters appears to require at least one inversion or decomposition of a matrix describing the interrelationship of the parameters. The matrix chosen by Smith and Freeze is relatively simple, but more realistic correlation structures require full covariance matrices to be decomposed [*Wilson*, 1979]. The result of this decomposition must then be multiplied by one vector of p random deviates for each of the m simulations. These two steps require $O(p^3)$ and $O(mp^2)$ multiplications, respectively, bringing the total computational requirements for Monte Carlo simulation to $O(mn^2 + p^3 + mp^2)$ or $O(mn^3 + p^3 + mp^2)$ depending on the deterministic solution scheme employed.

On the basis of these rough estimates of computational requirements, first order analysis will usually be much less expensive than Monte Carlo simulation for large problems where analytical solutions do not exist. If it is assumed that both methods employ the same matrix inversion scheme, then the ratio of first order analysis to Monte Carlo simulation computational requirements is roughly

$$\frac{(2k^2 + 1)n^3}{k^3n^3 + mk^2n^2 + mn^3}$$

with $p \cong kn$ as above. When $k \cong 1$ and m is small compared to n, this ratio is $3/(1 + m)$, which is a very small number. When $m \simeq n$, the ratio becomes $3/(2 + m)$, still a small number. On the other hand, if all of the matrix inversions are carried out taking advantage of the band structure of the coefficient matrices, the ratio can be recomputed, leaving Monte Carlo simulation in a better position for most grid systems. The ratio may be estimated at approximately

$$\frac{2k^2n^3}{mn^2 + k^3n^3 + mn^2k^2}$$

which reduces to $2n/(2m + n)$ when $k \cong 1$. Thus the ratio is 2 when $m \ll n$ and 2/3 when $m \cong n$.

Comparison of the computational requirements of first order analysis and Monte Carlo simulation is very tricky since the two methods may be designed to return very different information and since the computational burden is distributed differently. For instance, the estimates above all assume that the first order analysis returns the full head covariance matrix. Implicitly it has been assumed that the Monte Carlo simulations returned only the mean and variances, having neglected the additional $O(mn^2)$ operations required to evaluate the correlations of the Monte Carlo head results. If only the head variances were required of a first order analysis, nearly half of the multiplications in (27) could be avoided. This would shift all of the ratios above even more in first order analyses favor. More importantly the number of multiplications in the first order analysis can be greatly reduced by taking advantage of the correlation structure, assuming zero correlation for parameters poorly correlated, and banding the covariance matrix. Finally, first order analysis explicitly provides sensitivity matrices $\mathbf{D}$, which are very useful in designing data collection programs and in the solution of the inverse problem [*Wilson et al.*, 1978; *Wilson and Dettinger*, 1978].

Note that second order analysis and nonsteady state first order analysis of large numerical models are not usually as economical as the steady state first order analysis just described.

SUMMARY AND CONCLUSIONS

The first and second order analyses of numerical models presented in this paper can be used in examining a broad range of groundwater flow problems with many different grid geometries, aquifer properties and boundary conditions. Homogeneous or nonhomogeneous, isotropic or anisotropic, uniform or spatially varying parameters and boundary conditions can be studied through the proper choice of covariance matrices. Thus the analysis presents a useful alternative or complement to Monte Carlo methods, in studies of the effects of

information uncertainty and/or intrinsic variability on groundwater flow and piezometric head predictions.

The derivation presented is limited to systems with boundary conditions and aquifer properties that are statistically uncorrelated with each other and in time. As noted earlier, this does not indicate a true restriction of the technique, rather it represents a simplification by the authors in the interests of clarity. The results have also been limited to numerical models that can be conceived as state-space models, i.e., linear in the state or piezometric heads, and cast in the form of (10). Again this equation is simply chosen as a good example since it is sufficiently general to describe many numerical model structures. First and second order analyses are also applicable to other model formulations, although models that are highly nonlinear in the state must be regarded cautiously. Finally, model error is not included in the results above, although it could be if the error statistics were known.

Within the framework of these assumptions, first and second order analysis leads to a number of generalities concerning uncertainty about piezometric heads' predictions:

1. To first order, the expected value of piezometric heads will be identical to the heads predicted by deterministic methods.
2. Uncertainty in these head levels, as measured by the variance of a prediction or estimate, is a function both of the amount of uncertainty concerning aquifer parameters, and of the sensitivity of the heads to these same parameters.
3. An improved estimate of the heads, as achieved through a second order analysis, may differ from the first order estimate if there is important uncertainty concerning parameters of the aquifer that enter the predictions of head in a nonlinear fashion. Examples of such parameters are transmissivity and storativity or hydraulic conductivity and specific storage coefficient. In a state-space formulation, boundary and initial conditions enter predictions linearly and the corresponding uncertainties do not affect second and higher order estimates.
4. Uncertainty of the system state generally varies with time, increasing or decreasing as the aquifer system moves into states that are more or less sensitive to the sources of uncertainty. The assumption that the statistics of boundary conditions and aquifer properties are constant in time, eliminates the possibility that sources of uncertainty vary in time, although this assumption is not necessary to the method of analysis.
5. As the expected value of the heads approaches a steady state, the uncertainty of the head prediction approaches a steady state of its own. If the state approached by the heads is independent of any initial conditions, then the head uncertainty will be independent of initial conditions and initial condition uncertainty as well.
6. The first and second order analysis procedure can be applied to nonlinear systems with 'reasonably small' coefficients of variation. Part 2 will demonstrate that many field situations meet this constraint.
7. Two length scales of the aquifer affect piezometric head uncertainty: the length over which aquifer properties are correlated and the length at which heads are still sensitive to parameter variation or variability. These two interact so that in general heads are correlated over a considerably longer distance than the aquifer properties.
8. First and second order analysis is usually much less expensive to apply than Monte Carlo simulation. The trade-off is cost and accuracy, i.e., loss of accuracy with first and second order analysis due to the Taylor series approximation; loss of accuracy with Monte Carlo simulation due to a limited number of simulation experiments.
9. First and second order analysis explicitly provides sensitivity matrices **D** which are very useful in the design and operation of data collection systems and operations [*Wilson et al.*, 1978; *Wilson and Dettinger*, 1978].

Application of the results presented above to a particular numerical model makes it possible to develop more concrete conclusions. Such an application is the basis for the results presented in part 2.

The analysis presented here is restricted. It requires, in most valid applications, that the variability or variance of the system parameters or properties be relatively small; the coefficient of variation of the driving uncertainties should be a fraction of one. Another restriction, more practically based, may arise as a result of the computational burden, which is considerably larger than that imposed by a simple deterministic solution. Yet the added reliability information gained in this analysis strongly recommends it for application. This will be especially true as array processors and other computer technologies progress and erode the cost of matrix manipulation.

APPENDIX A

This appendix presents an introduction to Vetter calculus. Definitions of matrix derivative operations are given, as well as the identities that are most important in carrying out the analyses suggested in this paper.

Any explanation of Vetter's matrix calculus must begin with a definition of the Kronecker product. *Vetter*'s [1970, 1971, 1973] matrix calculus is related to and uses a notation that is similar in form. The Kronecker or outer product may be illustrated as follows [*Bellman*, 1960]:

$$\mathbf{A} * \mathbf{B} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & b_{13} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}b_{11} & a_{11}b_{12} & a_{11}b_{13} & a_{12}b_{11} & a_{12}b_{12} & a_{12}b_{13} \\ a_{21}b_{11} & a_{21}b_{12} & a_{21}b_{13} & a_{22}b_{11} & a_{22}b_{12} & a_{22}b_{13} \end{bmatrix}$$

where **A** is a 2×2 matrix, **B** a 1×3 matrix, and the operator $*$ signifies Kronecker product. In general, $\mathbf{A} * \mathbf{B}$, with **A** and **B** of dimension $n \times p$ and $q \times r$, respectively, is an $nq \times pr$ matrix made up of $n \times p$ submatrices, each equal to $a_{ij}\mathbf{B}$. Notice that the matrices involved in the Kronecker product need not have complementing dimensions as in the more common inner matrix product [*Bellman*, 1960; *Brockett*, 1970].

Vetter's definition of the derivative of the matrix $\mathbf{B}(q \times r)$ with respect to the matrix $\mathbf{A}(n \times p)$ is an $nq \times pr$ matrix composed of $n \times p$ submatrices, each equal to $\partial\mathbf{B}/\partial a_{ij}$. The derivative $\partial\mathbf{B}/\partial a_{ij}$ is defined in the usual way, as a $q \times r$ matrix of the partial derivatives of each element of **B** with respect to a_{ij}.

$$\frac{\partial \mathbf{B}}{\partial \mathbf{A}} = \begin{bmatrix} \frac{\partial \mathbf{B}}{\partial a_{11}} & \frac{\partial \mathbf{B}}{\partial a_{12}} & \cdots & \frac{\partial \mathbf{B}}{\partial a_{1p}} \\ \frac{\partial \mathbf{B}}{\partial a_{21}} & \frac{\partial \mathbf{B}}{\partial a_{22}} & \cdots & \cdot \\ \vdots & \vdots & & \vdots \\ \frac{\partial \mathbf{B}}{\partial a_{n1}} & \frac{\partial \mathbf{B}}{\partial a_{n2}} & \cdots & \frac{\partial \mathbf{B}}{\partial a_{np}} \end{bmatrix}$$

In this paper, Vetter's notation for the derivative is used, i.e., $\partial \mathbf{B}/\partial \mathbf{A} = \mathbf{D}_{\mathbf{A}}\mathbf{B}$.

Several identities are helpful in carrying out the first and second order analysis presented in this paper. The more important of these are listed below. Notice that all of the identities follow directly from application of the definitions of the matrix operations involved.

For the purposes of this paper, the most important identity for the Kronecker product is

$$\mathbf{AB} * \mathbf{CD} = (\mathbf{A} * \mathbf{C})(\mathbf{B} * \mathbf{D}) \tag{A1}$$

The derivative of a vector with respect to its transpose is a particularly simple relation:

$$\mathbf{D}_{\mathbf{y}}\mathbf{y}' = \mathbf{I}_n \tag{A2}$$

with $\mathbf{y}$ an $n \times 1$ vector, the prime signifying the transpose of the vector and $\mathbf{I}_n$ equal to an $n \times n$ identity matrix. This is not equal to

$$\mathbf{D}_{\mathbf{y}}\mathbf{y} = cs\,\mathbf{I}_n = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ \text{etc.} \end{bmatrix} \begin{matrix} \\ \\ \\ \\ \\ \left.\begin{matrix} \\ \\ \\ \\ \end{matrix}\right\} n \\ \left.\begin{matrix} \\ \\ \\ \\ \end{matrix}\right\} n \\ \\ \end{matrix}$$

an $n^2 \times 1$ vector made up of the columns of the identity matrix $\mathbf{I}_n$. Note that *cs* stands for the term column string. The derivatives of a matrix with respect to its transpose and with respect to itself are somewhat more complicated and can be found in *Vetter* [1973].

The derivatives of matrix inner products are

$$D_{\mathbf{B}}(\mathbf{AC}) = (D_{\mathbf{B}}\mathbf{A})(\mathbf{I}_l * \mathbf{C}) + (\mathbf{I}_k * \mathbf{A})\, D_{\mathbf{B}}\mathbf{C} \tag{A3}$$

and

$$D_{\mathbf{B}}(\mathbf{A} * \mathbf{C}) = (D_{\mathbf{B}}\mathbf{A}) * \mathbf{C} + (\mathbf{I}_k * \mathbf{E}_{m\times n}{}^{n\times m})[(D_{\mathbf{B}}\mathbf{C}) * \mathbf{A}](\mathbf{I}_l * E_{n\times p}{}^{p\times n}) \tag{A4}$$

where $\mathbf{B}$ is a $k \times l$ matrix, $\mathbf{A}$ an $m \times n$ matrix, and C an $n \times p$. The matrix $\mathbf{E}_{m\times p}{}^{p\times m}$, which also appears in Appendix C, is defined as a matrix made up of m rows and p columns of submatrices each with dimension $p \times m$. The ijth submatrix is a matrix of zeroes with a one in the (j,i)th element. Thus $\mathbf{E}_{2\times 2}{}^{2\times 2}$ is

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Notice that $\mathbf{E}_{m\times p}{}^{p\times m}$ has dimension $mp \times mp$ and that each row and column has only a single nonzero element. The derivative of a matrix inverse with respect to a vector is

$$\mathbf{D}_{\mathbf{b}'}(\mathbf{A}^{-1}) = -(\mathbf{A}^{-1})(\mathbf{D}_{\mathbf{b}'}\mathbf{A})(\mathbf{I}_p * \mathbf{A}^{-1}) \tag{A5}$$

with $\mathbf{b}'$ a $1 \times p$ vector. This relation may be proved by taking the derivative of the constant matrix, $\mathbf{I}_n = \mathbf{AA}^{-1}$ with respect to $\mathbf{b}'$, followed by some algebra.

Finally, if $\mathbf{A}(\mathbf{C})$ is a function of $\mathbf{C}$ which in turn is a function of $\mathbf{B}$, then the following matrix calculus 'chain rule' applies.

$$\mathbf{D}_{\mathbf{B}}\mathbf{A} = [\mathbf{D}_{\mathbf{B}}(cs\ \mathbf{C})' * \mathbf{I}_m](\mathbf{I}_l * \mathbf{D}_{cs\,\mathbf{C}}\mathbf{A}) \tag{A6}$$

where $\mathbf{A}$ is an $m \times n$ matrix, $\mathbf{B}$ is a $k \times l$ matrix, and

$$cs\,\mathbf{C} = \begin{bmatrix} C_{11} \\ C_{21} \\ C_{31} \\ \vdots \\ C_{12} \\ C_{22} \\ D_{23} \\ \vdots \\ C_{np} \end{bmatrix} = \begin{bmatrix} \mathbf{C}_1 \\ \mathbf{C}_2 \\ \mathbf{C}_3 \\ \vdots \\ \mathbf{C}_n \end{bmatrix}$$

which is an $np \times 1$ vector made up of the columns of the matrix $\mathbf{C}$.

With the definitions and identities in this appendix, the derivations indicated in this paper can be obtained for any state-space model of groundwater flow. Additional relationships can be arrived at through construction or found in the work of *Vetter* [1970, 1973] and *Brewer* [1977].

APPENDIX B

First derivatives of $\mathbf{h}(t + \Delta t)$ with respect to parameters and boundary and initial conditions are required in a first and second order analysis of the numerical groundwater flow model. The forms of these derivatives are presented in this appendix.

The form of the derivatives depends on the particular solution method applied. The derivatives of explicit, implicit, and steady state solutions for the heads have been derived and are the forms presented here.

Because the state-space equation for the heads is linear in terms of the initial conditions $\mathbf{h}(t)$ and the boundary conditions $\mathbf{u}$, the derivatives of $\mathbf{h}(t + \Delta t)$ with respect to $\mathbf{h}(t)$ and $\mathbf{u}$ are easily found. The derivative of $\mathbf{h}(t + \Delta t)$ with respect to the initial conditions $\mathbf{h}'(t)$ can be found using identity (A3),

$$\begin{aligned} D_{\mathbf{h}'}\mathbf{h}(t + \Delta t) = \mathbf{D}_{\mathbf{h}} &= D_{\mathbf{h}'}[\Phi\mathbf{h}(t) + \mathbf{Gu}] = D_{\mathbf{h}'}[\Phi\mathbf{h}(t)] \\ &= (D_{\mathbf{h}'}\Phi)[\mathbf{I}_n * \mathbf{h}(t)] + \Phi[D_{\mathbf{h}'}\mathbf{h}(t)] \\ &= \mathbf{0} + \Phi\,\mathbf{I}_n = \Phi \end{aligned} \tag{B1}$$

for transient solutions, with n equal to the number of elements in $\mathbf{h}$, and identically zero (a null matrix) for a steady state solution. Similarly, the derivative with respect to the boundary conditions $\mathbf{u}'$ is

$$D_{\mathbf{u}'}(t + \Delta t) = \mathbf{D}_{\mathbf{u}} = \mathbf{G} \tag{B2}$$

for transient and steady state solutions.

The matrices Φ and $\mathbf{G}$ are complicated functions of the model parameters but relatively simple functions of the coefficient matrices $\mathbf{A}$ and $\mathbf{C}$. In turn, these coefficient matrices are simple functions of the parameters. The exact relationship of Φ to $\mathbf{A}$, etc., is dependent upon the time integration scheme; consequently, the derivatives with respect to parameters must be derived separately for each scheme.

The transition matrix Φ for the explicit scheme (13*a*) is linear in $\mathbf{A}$ and thus the derivative has a simple form. If there are p uncertain or stochastic parameters, b_i, $i = 1, \cdots, p$, then the derivative of the vectors of heads at time $t + \Delta t$, $\mathbf{h}(t + \Delta t)$, with respect to the vector of parameters $\mathbf{b}'$ is, using (A3), (11), and (13*a*),

$$D_{\mathbf{b}'}\mathbf{h}(t + \Delta t) = \mathbf{D}_{\mathbf{b}} = (D_{\mathbf{b}'}\mathbf{A}\Delta t)[\mathbf{I}_p * \mathbf{h}(t)] + D_{\mathbf{b}'}(\mathbf{Cu}\Delta t) \tag{B3}$$

for the explicit formulation where $\mathbf{I}_p$ is a $p \times p$ identity matrix.

The matrices Φ and G for an implicit scheme, as well as the G matrix for a steady state prediction, are functions of inverse matrices (see (13*b*) and (14)). Applying identity (A5) leads to the derivatives

$$\mathbf{D}_b = \Phi(D_{b'}\mathbf{A}\Delta t)[\mathbf{I}_p * (\Phi\mathbf{h}(t) + \mathbf{Gu})] + \Phi(D_{b'}\mathbf{C}\mathbf{u}\Delta t)$$
$$= \Phi[(D_{b'}\mathbf{A}\Delta t)\{\mathbf{I}_p * \mathbf{h}(t + \Delta t)\} + D_{b'}\mathbf{C}\mathbf{u}\Delta t] \quad \text{(B4)}$$

for an implicit scheme for $\mathbf{h}(t + \Delta t)$, and

$$\mathbf{D}_b = \mathbf{G}(D_{b'}\mathbf{A})(\mathbf{I}_p * \mathbf{h}_s) + (-\mathbf{A}^{-1})(D_{b'}\mathbf{C}\mathbf{u}) \quad \text{(B5)}$$

for a steady state model.

Because the coefficients of matrices A and C depend on the model structure (i.e., finite difference, finite element) used to describe the aquifer, the derivative of these matrices with respect to the model parameters must be obtained for the specific numerical model, discretization and parameterization chosen in an application. Therefore a lengthy description of particular cases will not be presented here. However, an example application will be presented in part 2.

The calculation of the derivatives of A and C is generally no more difficult than calculating and filling the coefficient matrices themselves! Products of the form

$$(D_{b'}\mathbf{A}\Delta t)(\mathbf{I}_p * \mathbf{l})$$

where $\mathbf{l}$ is some vector ($n \times 1$) with n equal to the number of elements in the solution vector $\mathbf{h}(t + \Delta t)$, appear in both the first derivatives above and the second derivatives in Appendix C. These products have dimension $n \times p$, and can be filled and manipulated with considerably more ease than the separate terms $D_{b'}\mathbf{A}\Delta t$ and $(\mathbf{I}_p * \mathbf{l})$, which have dimensions $n \times np$ and $np \times n$, respectively, and which are both very sparsely filled.

If the coefficient matrices are functions of some transformation of the uncertain parameters, then the chain rule (A6) may be applied to ease the derivation. This is the case, for example, when the logarithm of the hydraulic conductivity is the parameter of interest and the model is based on the untransformed conductivities.

The derivatives presented in this appendix are exact derivatives of the aquifer model and thus as accurate as the model itself.

APPENDIX C

The second derivatives of piezometric head predictions $\mathbf{h}(t + \Delta t)$ required in the second order analysis of the mean are derived through a direct extension of the results presented in Appendices A and B. The forms of the second derivatives are listed in this appendix.

Because $\mathbf{h}(t + \Delta t)$ is a linear function of the boundary and initial conditions, the second derivatives with respect to $\mathbf{u}'$ and $\mathbf{h}'(t)$, respectively, are null matrices. The second derivative of $\mathbf{h}(t + \Delta t)$ with respect to the parameters $\mathbf{b}'$ and the initial conditions $\mathbf{h}'(t)$ is

$$\mathbf{D}_{bh} = D_{b'}\Phi \quad \text{(C1)}$$

for transient models. For an implicit time integration scheme, this expression is found from identity (A5) to be

$$\mathbf{D}_{bh} = \Phi(D_{b'}\mathbf{A}\Delta t)(\mathbf{I} * \Phi)$$

if $\mathbf{b}'$ has dimension $1 \times p$ and $\mathbf{I}_p$ is a $p \times p$ identity matrix. For an explicit model,

$$\mathbf{D}_{bh} = D_{b'}\mathbf{A}\Delta t \quad \text{(C2)}$$

The derivative with respect to parameters and boundary conditions may be derived in an analogous manner.

The second derivative of head predictions with respect to parameters $\mathbf{b}'$ is more complicated. When an explicit time integration scheme is implemented, then using (A3) the derivative is

$$D_{b'b'}(\mathbf{h}_{t+\Delta t}) = \mathbf{D}_{bb} = (D_{b'b'}\mathbf{A}\Delta t)[\mathbf{I}_{pp} * \mathbf{h}(t)] + (D_{b'b'}\mathbf{C}\mathbf{u}\Delta t) \quad \text{(C3)}$$

where $\mathbf{I}_{pp}$ is a $p^2 \times p^2$ identity matrix, $D_{b'b'}\mathbf{A}\Delta t$ is the second derivative of $\mathbf{A}\Delta t$ with respect to the parameters $\mathbf{b}'$, and $D_{b'b'}\mathbf{C}\mathbf{u}\Delta t$ is defined analogously.

The steady state and implicit scheme first and second derivatives have very similar structures. This is due to their similar dependence on inverse matrix functions of A. For an implicit scheme, the second derivatives with respect to the parameters are

$$\mathbf{D}_{bb} = \Phi\{(D_{b'}\mathbf{A}\Delta t)(\mathbf{I}_p * \mathbf{D}_b) + (D_{b'}\mathbf{A}\Delta t)\mathbf{E}_{p\times n}{}^{n\times p}(\mathbf{D}_b * \mathbf{I}_p)$$
$$+ (D_{b'b'}\mathbf{A}\Delta t)[\mathbf{I}_{pp} * \mathbf{h}(t + \Delta t)] + D_{b'b'}\mathbf{C}\mathbf{u}\Delta t\} \quad \text{(C4)}$$

and for a steady state

$$\mathbf{D}_{bb} = \mathbf{G}(D_{b'}\mathbf{A})(\mathbf{I}_p * \mathbf{D}_b) + \mathbf{G}(D_{b'}\mathbf{A})\mathbf{E}_{p\times n}{}^{n\times p}[(\mathbf{D}_b) * \mathbf{I}_p]$$
$$+ \mathbf{G}(D_{b'b'}\mathbf{A})(\mathbf{I}_{pp} * \mathbf{h}_s) - \mathbf{A}^{-1}(D_{b'b'}\mathbf{C}\mathbf{u}\Delta t) \quad \text{(C5)}$$

using identities (A3)–(A5) and where $\mathbf{E}_{p\times n}{}^{n\times p}$ is defined in Appendix A, the solution vector $\mathbf{h}(t + \Delta t)$ is $n \times 1$ dimensional and $\mathbf{b}$ is $p \times 1$.

Acknowledgments. The work presented in this paper was performed at the Ralph M. Parsons Laboratory for Water Resources and Hydrodynamics at the Massachusetts Institute of Technology. The work was supported in part by funds from the U.S. Agency for International Development through the MIT/Cairo University Technology Planning Program, under AID grant NE-C-1291.

REFERENCES

Bakr, A. A., L. W. Gelhar, A. L. Gutjahr, and J. R. MacMillan, Stochastic analysis of spatial variability in subsurface flows, 1, Comparison of one- and three-dimensional flows, *Water Resour. Res.*, *14*(2), 263–271, 1978.

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

Bellman, R., *Introduction to Matrix Analysis,* McGraw-Hill, New York, 1960.

Benjamin, J. R., and C. A. Cornell, *Probability, Statistics, and Decision for Civil Engineers,* McGraw-Hill, New York, 1970.

Brewer, J. W., The derivative of the exponential matrix with respect to a matrix, *IEEE Trans. Autom. Control, AC-22*(4), 656–657, 1977.

Brockett, R. W., *Finite Dimensional Linear Systems,* John Wiley, New York, 1970.

Cornell, C. A., First order analysis of model and parameter uncertainty, International Symposium on Uncertainties in Hydrologic and Water Resource Systems, Univ. of Ariz., Tucson, 1972.

Dagan, G., Models of groundwater flow in statistically homogeneous porous formations, *Water Resour. Res., 15*(1), 47–63, 1979.

Dahlquist, G., and A. Bjorck, *Numerical Methods,* Prentice-Hall, Englewood Cliffs, N. J., 1974.

de Finetti, B., *Theory of Probability,* John Wiley, New York, 1974.

Delhomme, J. P., Kriging in the hydrosciences, *Adv. Water Resour., 1*(5), 251–266, 1978.

Delhomme, J. P., Spatial variability and uncertainty in groundwater flow parameters: A geostatistical approach, *Water Resour. Res., 15*(2), 269–280, 1979.

Eagleson, P. S., Climate, soil, and vegetation, 5, A derived distribution of storm surface runoff, *Water Resour. Res., 14*(5), 741–748, 1978.

Freeze, R. A., A stochastic-conceptual analysis of one-dimensional

groundwater flow in nonuniform homogeneous media, *Water Resour. Res., 11*(5), 725–741, 1975.
Gelhar, L. W., J. L. Wilson, and A. L. Gutjahr, Comment on Simulation of groundwater flow and mass transport under uncertainty, *Adv. Water Resour., 2,* 101, 1979.
Gutjahr, A. L., L. W. Gelhar, A. A. Bakr, and J. R. MacMillan, Stochastic analysis of spatial variability in subsurface flows, 2, Evaluation and application, *Water Resour. Res., 14*(5), 953–959, 1978.
Pinder, G. F., and W. G. Gray, Is there a difference in the finite difference methods?, *Water Resour. Res., 12*(1), 105–107, 1976.
Ramsey, M. A., *The Foundations of Mathematics,* Humanities Press, New York, 1950.
Sagar, B., Galerkin finite element procedure for analyzing flow through random media, *Water Resour. Res., 14*(6), 1035–1044, 1978.
Sagar, B., and C. C. Kisiel, Limits of deterministic predictability of saturated flow equations, Joint Symposium on Fundamentals of Transport Phenomena in Porous Media, Univ. of Guelph, Guelph, Ont., Canada, 1972.
Savage, L. J., *The Foundations of Statistics,* Dover, New York, 1972.
Smith, L., and R. A. Freeze, Stochastic analysis of steady state groundwater flow in a bounded domain, 1, One-dimensional simulations, *Water Resour. Res., 15*(3), 521–528, 1979.
Tang, D. H., and G. F. Pinder, Simulation of groundwater flow and mass transport, *Advan. Water Res., 1*(1), 25–30, 1977.
Vetter, W. J., Derivative operations on matrices, *IEEE Trans. Autom. Control, AC-15,* 241–244, 1970.
Vetter, W. J., Corrections to: Derivative operations on matrices, *IEEE Trans. Autom. Control, AC-16,* 1971.
Vetter, W. J., Matrix calculus operations and Taylor expansions, *SIAM Rev., 15*(2), 352–369, 1973.
Warren, J. E., and H. S. Price, Flow in heterogeneous porous media, *Soc. Pet. Eng. J., 1,* 153–169, 1961.
Wilson, J. L., The synthetic generation of areal averages of a random field, Socorro Workshop on Stochastic Methods in Subsurface Hydrology, N. Mex. Inst. of Mines and Technol., Socorro, 1979.
Wilson, J. L., and M. D. Dettinger, Steady-state vs. transient parameter estimation in groundwater systems, Verification of Mathematical Models in Hydraulic Engineering, Proceedings of the 26th Annual Hydraulics Division Special Conference, Am. Soc. Civ. Eng., New York, 1978.
Wilson, J. L., and M. D. Dettinger, Elementary applications of first order analysis to groundwater hydrology, manuscript in preparation, 1980.
Wilson, J. L., P. K. Kitanidis, and M. D. Dettinger, State and parameter estimation in groundwater models, in *Applications of Kalman Filter to Hydrology, Hydraulics, and Water Resources,* edited by C. L. Chiu, Stochastic Hydraulics Program, University of Pittsburgh, Pittsburgh, Pa., 1978.

(Received April 24, 1979;
revised July 24, 1980;
accepted July 31, 1980.)