AN EFFICIENT STOCHASTIC DECOMPOSITION APPROACH FOR LARGE-SCALE SUBSURFACE FLOW PROBLEMS

DONGXIAO ZHANG[1,2], ZHIMING LU[3], AND GAISHENG LIU[1]

[1] *Mewbourne School of Petroleum and Geological Engineering, The University of Oklahoma, Norman, OK 73019, USA*
[2] *Department of Energy and Resources Engineering, College of Engineering, Peking University, Beijing 100871, China*
[3] *Hydrology, Geochemistry, and geology Group, Los Alamos National Laboratory, Los Alamos, NM 87545, USA*

ABSTRACT

Significant spatial heterogeneity and a limited number of measurements lead to uncertainty in characterization of subsurface formation properties and thus, to uncertainty in predicting fluid flow and transport in the formations. Such uncertainties add another dimension in probability space to the already large-scale subsurface problems. In this work, we develop an accurate yet efficient approach for solving flow and transport problems in large-scale heterogeneous formations. We do so by obtaining higher-order solutions of the prediction and the associated uncertainty of subsurface fluid quantities using the moment-equation approach based on Karhunen-Loéve decomposition (KLME). In the KLME approach, the log permeability (ln$k$) field is first expanded into a multiscale series in terms of orthogonal standard Gaussian random variables with their coefficients obtained from the eigen-decomposition of the ln$k$ covariance. Next, the dependent flow and transport quantities are all decomposed with perturbation expansions in which each individual term is further expanded into a polynomial series of orthogonal Gaussian random products. The coefficients associated with these series are solved recursively from low to high expansion orders. The statistical moments of flow and transport predictions can then be calculated from these coefficients using simple algebraic operations. The new approach is validated and its efficiency and accuracy are demonstrated by comparing with traditional Monte Carlo simulations in a couple of subsurface test problems.

1. INTRODUCTION

Subsurface formations are inherently heterogeneous and exhibit a high degree of variability in medium properties such as permeability (k) at a multiplicity of scales. Due to the incomplete knowledge of medium properties, it is often more suitable to approach subsurface fluid flow and transport problems in a stochastic framework. The added dimension in probability space to the already large-scale subsurface problems, however, calls for the development of new stochastic methodologies that are both accurate and computationally effective. The traditional Monte Carlo simulations (MCS) method is conceptually straightforward, but is very difficult to apply for large-scale problems under the field conditions because of its computational cost (Hassan et al., 1998). A major alternative to the

MCS method is the approach based on the moment equations (ME). In the ME approach (Graham and McLaughlin, 1989; Neuman, 1993; Zhang, 1998; Zhang and Lu, 2002) a set of deterministic partial differential equations are derived for the statistical moments of fluid quantities (such as hydraulic heads and chemical concentrations) using the method of perturbation. These equations are then solved either analytically or numerically. In general, analytical solutions of the moment equations are difficult to obtain except for some limiting cases under simplifying conditions. The computational effort required in the numerical ME approach increases rapidly with the size of model, which makes its application infeasible in most field practices. Although the ME approach has produced some success when applied to systems of large heterogeneity, it is typically restricted to problems of relatively small variability in medium properties.

To alleviate the computational burden, Zhang and Lu (2004) developed a new approach that combines Karhunen-Loéve decomposition and perturbation methods. Specifically, using the new approach, with a much lower computational cost, they were able to evaluate the steady state mean pressure head and the mean flux up to the fourth order and variances of the pressure head and flux to the sixth order with respect to the standard deviation of log permeability. The computational cost required by the KLME approach has been systematically compared with the Monte Carlo simulations and the conventional moment method (Lu and Zhang, 2004). Lu and Zhang (2006) extended the KLME approach to transient fluid flow in the three-dimensional space and demonstrated the applicability of this approach to simulating flow in large-scale heterogeneous reservoirs. In this work, we further develop the KLME approach for solving stochastic fluid flow and transport problems in large-scale heterogeneous formations.

2. METHODOLOGY

1.1 Governing Equations.

The governing equation for single-phase flow in saturated porous media under transient conditions can be written as,

$$\nabla \cdot [K_s(\boldsymbol{x}) \nabla h(\boldsymbol{x},t)] + q = S_s \frac{\partial h(\boldsymbol{x},t)}{\partial t}, \tag{1}$$

subject to appropriate initial and boundary conditions. Here $K_s(\boldsymbol{x})$ is the hydraulic conductivity defined as $k\rho_f g / \mu$ with ρ_f the fluid density, μ the fluid viscosity and g the acceleration of gravity; h is the hydraulic head; q is the flow sink/source; S_s is the specific storage; $\boldsymbol{x}$ is the vector of spatial Cartesian coordinate $(x, y, z)^{\mathrm{T}}$; t is time. In the study $K_s(\boldsymbol{x})$ (or k) is taken as a random function and all other parts of the flow model are assumed to be deterministic.

The transport of a conservative solute in three-dimensional flow is given by the advection-dispersion equation with sinks/sources,

$$\frac{\partial C(\boldsymbol{x},t)}{\partial t} = \nabla \cdot \left(D(\boldsymbol{x},t) \nabla C(\boldsymbol{x},t) \right) - \nabla \cdot \left(v(\boldsymbol{x},t) C(\boldsymbol{x},t) \right) + q_s C_s, \tag{2}$$

where C is the solute concentration; v is the velocity vector $(v_x(\boldsymbol{x},t), v_y(\boldsymbol{x},t), v_z(\boldsymbol{x},t))^{\mathrm{T}}$ that can be determined through Darcy's law, $v(\boldsymbol{x},t) = q(\boldsymbol{x},t)/\theta = -K_s(\boldsymbol{x}) \nabla h(\boldsymbol{x},t)/\theta$ after h is solved from (1); D is the hydrodynamic dispersion tensor whose values are dependent on velocities

as well as medium properties (Burnett and Frind, 1987); q_s and C_s are the flow rate and solute concentration associated with external sinks/sources, and θ is porosity The transport equation is typically solved sequentially after the solution of flow equations is obtained assuming that the changes of solute concentrations do not affect velocity fields significantly.

1.2 Karhunen-Loéve (KL) decomposition of permeability.

As only single-phase saturated fluid is considered in this paper, ρ_f and μ are fixed and do not change with space and time. Therefore, for convenience, we directly work on the hydraulic conductivity instead of permeability in the following discussions. Let $Y(\boldsymbol{x}) = \ln[K_S(\boldsymbol{x})]$, the covariance function of $Y(\boldsymbol{x})$ is $C_Y(\boldsymbol{x}_1, \boldsymbol{x}_2) =< Y'(\boldsymbol{x}_1)Y'(\boldsymbol{x}_2) >$, where perturbation $Y'(\boldsymbol{x})$ is defined as $Y'(\boldsymbol{x}) = Y(\boldsymbol{x}) - < Y(\boldsymbol{x}) >$, and $< >$ is expected mean operator. The basic idea of the KL decomposition is to decompose the positive definite covariance function of $Y(\boldsymbol{x})$ as (Courant and Hilbert, 1953),

$$C_Y(\boldsymbol{x}_1, \boldsymbol{x}_2) = \sum_{n=1}^{\infty} \lambda_n f_n(\boldsymbol{x}_1) f_n(\boldsymbol{x}_2), \tag{3}$$

where λ_n and f_n(x) are the deterministic eigenvalues and eigenfunctions that can be calculated by solving the following Fredholm equation analytically or numerically,

$$\int_D C_Y(\boldsymbol{x}_1, \boldsymbol{x}_2) f(\boldsymbol{x}_1) d\boldsymbol{x}_1 = \lambda f(\boldsymbol{x}_2). \tag{4}$$

The mean-removed term $Y'(\boldsymbol{x})$ can be expanded in terms of λ_n and $f_n(\boldsymbol{x})$,

$$Y'(\boldsymbol{x}) = \sum_{n=1}^{\infty} \xi_n \sqrt{\lambda_n} f_n(\boldsymbol{x}). \tag{5}$$

where ξ_n are the orthogonal Gaussian standard random variables. Since eigenvalues λ_n and their eigenfunctions $f_n(\boldsymbol{x})$ always appear together, in the following derivations, we define new functions $\widetilde{f}_n(\boldsymbol{x}) = \sqrt{\lambda_n} f_n(\boldsymbol{x})$ and then the tilde over f_n is dropped for simplicity.

1.3 Karhunen-Loéve expansion-based moment equations (KLME).

Because the variability of dependent variables $h(\boldsymbol{x},t)$, $v(\boldsymbol{x},t)$, $D(\boldsymbol{x},t)$ and $C(\boldsymbol{x},t)$ depends on the input variabilities, i.e., the variability of $Y'(\boldsymbol{x})$, one may express $h(\boldsymbol{x},t)$, $v(\boldsymbol{x},t)$, $D(\boldsymbol{x},t)$ and $C(\boldsymbol{x},t)$ as an infinite series,

$$\phi(\boldsymbol{x},t) = \sum_{m=0}^{\infty} \phi^{(m)}(\boldsymbol{x},t), \tag{6}$$

where ϕ stands for h, v, D or C, or $\phi^{(m)}$ are the m^{th}-order expansions with respect to the standard deviation of log hydraulic conductivity, σ_Y. Substituting (6) into the original governing equations (1) – (2) and then collecting terms at separate orders, for the transport equation, one obtains at $m = 0$,

$$\frac{\partial C^{(0)}}{\partial t} = \nabla \cdot \left(D^{(0)} \nabla C^{(0)} \right) - \nabla \cdot \left(v^{(0)} C^{(0)} \right) + q_s C_s, \tag{7}$$

and for $m \geq 1$,

$$\frac{\partial C^{(m)}}{\partial t} = \nabla \cdot \left(D^{(0)} \nabla C^{(m)} \right) - \nabla \cdot \left(v^{(0)} C^{(m)} \right) + s^{(m)}, \tag{8}$$

where

$$s^{(m)} = \nabla \cdot \sum_{i=0}^{m-1} (D^{(m-i)} \nabla C^{(i)} - v^{(m-i)} C^{(i)}). \tag{9}$$

The equations for flow can be obtained in a similar manner (Zhang and Lu, 2004). In the equations above for simplicity we have omitted the spatial and temporal indices $\boldsymbol{x}$ and t. In the conventional ME approach, the statistical moments of C and h are directly solved based on some simple manipulation of these expanded equations (Zhang and Lu, 2002). In the KLME approach, however, one further decomposes the random terms $h^{(m)}$, $v^{(m)}$, $D^{(m)}$ and $C^{(m)}$ as,

$$\phi^{(m)} = \sum_{i_1, i_2, \ldots, i_m = 1}^{\infty} (\prod_{j=1}^{m} \xi_{i_j}) \phi^{(m)}_{i_1, i_2, \ldots, i_m}, \tag{10}$$

where ξ_{i_j} are the orthogonal Gaussian random variables that are used to decompose Y' in (5); $\phi^{(m)}_{i_1, i_2, \ldots, i_m}$ are deterministic functions to be determined; $i_1, i_2, \ldots, i_m$ are referred to as modes at the m^{th} order. Substituting (10) and the KL decomposition of Y' into (8) – (9), one obtains the equations for the mode coefficients $C^{(m)}_{i_1, i_2, \ldots, i_m}$,

$$\frac{\partial C^{(m)}_{i_1, i_2, \ldots, i_m}}{\partial t} = \nabla \cdot \left(D^{(0)} \nabla C^{(m)}_{i_1, i_2, \ldots, i_m} \right) - \nabla \cdot \left(v^{(0)} C^{(m)}_{i_1, i_2, \ldots, i_m} \right) + s^{(m)}_{i_1, i_2, \ldots, i_m}, \tag{11}$$

where

$$s^{(m)}_{i_1, i_2, \ldots, i_m} = \nabla \cdot \sum_{k=0}^{m-1} [\frac{(m-k)!}{m!} \sum_{P_{i_1, i_2, \ldots, i_m}} (D^{(m-k)}_{i_{k+1}, \ldots, i_m} \nabla C^{(k)}_{i_1, \ldots, i_k} - v^{(m-k)}_{i_{k+1}, \ldots, i_m} C^{(k)}_{i_1, \ldots, i_k})]. \tag{12}$$

The equations for $h^{(m)}_{i_1, i_2, \ldots, i_m}$ and $v^{(m)}_{i_1, i_2, \ldots, i_m}$ can be found in Zhang and Lu (2004). The summation $\sum_{P_{i_1, i_2, \ldots, i_m}}$ in (12) is taken over a subset of the permutation of $\{i_1, i_2, \ldots, i_m\}$ in which repeated terms are excluded. For example, $\sum_{P_{i,j,k}} D^{(1)}_i \nabla C^{(2)}_{j,k} = D^{(1)}_i \nabla C^{(2)}_{j,k} + D^{(1)}_j \nabla C^{(2)}_{i,k} + D^{(1)}_k \nabla C^{(2)}_{i,j}$. $D^{(1)}_i \nabla C^{(2)}_{k,j}$ is identical to $D^{(1)}_i \nabla C^{(2)}_{j,k}$ and thereby excluded as $C^{(2)}_{j,k}$ calculated this way is symmetric with respect to its subscript indices. Detailed derivations of $D^{(m)}_{i_1, i_2, \ldots, i_m}$ can be found in Liu et al. (in review).

Once the deterministic coefficients $h^{(m)}_{i_1, i_2, \ldots, i_m}$, $v^{(m)}_{i_1, i_2, \ldots, i_m}$, $D^{(m)}_{i_1, i_2, \ldots, i_m}$ and $C^{(m)}_{i_1, i_2, \ldots, i_m}$ are calculated, one can easily compute their means and variances by simple algebraic operations,

$$< \phi > \approx \phi^{(0)} + \sum_{i=1}^{\infty} \phi^{(2)}_{i,i}, \tag{13}$$

where the first term on the right hand side is the zeroth-order mean solution and the second term represents the second-order correction. The variances can be estimated as,

$$\sigma^2_{\phi} \approx \sum_{i=1}^{\infty} [\phi^{(1)}_i]^2 + 2 \sum_{i,j=1}^{\infty} [\phi^{(2)}_{i,j}]^2 + 6 \sum_{i,j=1}^{\infty} [\phi^{(1)}_i \phi^{(3)}_{i,j,j}], \tag{14}$$

where the first term on the right-hand side is the variance up to the first order in σ^2_Y, and the second and third terms represent the second-order corrections.

Compared to the conventional ME approach, the KLME method has two distinctive advantages. First, unlike the conventional ME approach, the KLME method does not require solving directly the covariance equations such as head covariances and head-permeability

cross covariances whose computations are proportional to the number of grid nodes (N), thereby significantly reducing the computational efforts especially for large-scale problems. Second, as demonstrated above, the higher-order terms can be easily incorporated into the KLME method. The conventional ME approach usually approximates the covariance of flow quantities only up to the first order in σ_Y^2 because the computational burden increases drastically when higher-order corrections are considered. For instance, to obtain the hydraulic head variance up to second-order in σ_Y^2, one needs to solve equations for terms such as $< Y'(\boldsymbol{x}_1)Y'(\boldsymbol{x}_2)h'(\boldsymbol{x}_3,t) >$, which generally requires solving the partial differential equations for N^2 times. In (13) – (14) we have computed the means and variances up to the second-order. Even higher-order implementations are practical in the KLME method (Zhang and Lu, 2004). It is believed that incorporation of higher-order correction terms has enhanced the applicability of the KLME method to cases in which the heterogeneity is large.

There are two additional attractive features associated with the KLME approach. First, all equations for the deterministic coefficients $h^{(m)}_{i_1,i_2,\ldots,i_m}$ and $C^{(m)}_{i_1,i_2,\ldots,i_m}$ share exactly the same structure as the original equations, which greatly simplifies its implementation as the existing simulators/solvers can be utilized without major modifications. This feature also allows for a significant reduction in the computation effort as the coefficient matrix remains unchanged and only the right-hand-side vector needs to be updated across different orders and modes. Second, at each expansion order, the equations at different modes are independent of each other, which allows performing massively parallel computation when the physical problem dimensions are large.

3. ILLUSTRATIVE EXAMPLES

The KLME approach developed in this study is validated and its efficiency and accuracy are demonstrated with comparisons to traditional Monte Carlo simulations in a series of hypothetical numerical experiments. In this section we first present a three-dimensional test case in which the flow quantities (i.e., heads and velocities) are evaluated. In the second test case we consider a two-dimensional example where fluid transport results (i.e., chemical concentrations) are focused on. In both examples a total number of 5000 realizations are used in the MC simulations; in the KLME method, the number of modes at the first three orders is 100, 20 and 10, respectively, resulting in a total of 1 (zeroth-order) + 100 (first-order) + 210 (second-order) + 220 (third-order) = 531 model simulations. The computational effort is approximately the same for each MC realization and for each mode simulation.

1.1 Three-dimensional flow simulation.

The flow field is a three-dimensional block of 8 m long by 8 m wide by 4 m deep (Figure 1). There is no flow across the northern, southern, top and bottom boundaries; at the western and eastern boundaries the hydraulic heads are prescribed as constant at 10.5 and 10.0 m, respectively. A numerical block-centered finite-difference mesh of 21 by 20 by 10 regular cells (0.4 m on a side) is used to represent the physical domain. There is a well at the center of domain, pumping with a rate of 0.7 m^3/d. The model simulation is transient with initial heads determined from a steady-state calculation without well operation. The storage coefficient is

10^{-4}. The log hydraulic conductivity field is assumed random with zero mean (the geometric mean K_G = 1.0 m/d) and follows a separable exponential covariance function,

$$C_Y(\boldsymbol{x}_1, \boldsymbol{x}_2) = \sigma_Y^2 \exp(-|x_1 - x_2|/\eta_x - |y_1 - y_2|/\eta_y - |z_1 - z_2|/\eta_z), \quad (15)$$

where the correlation lengths $\eta_x = \eta_y = \eta_z$ = 2.0 m and variance σ_Y^2 = 1.0.

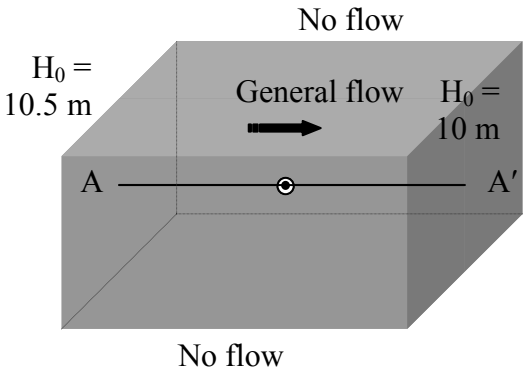

FIGURE 1. Schematic diagram of model setup in the flow test case.

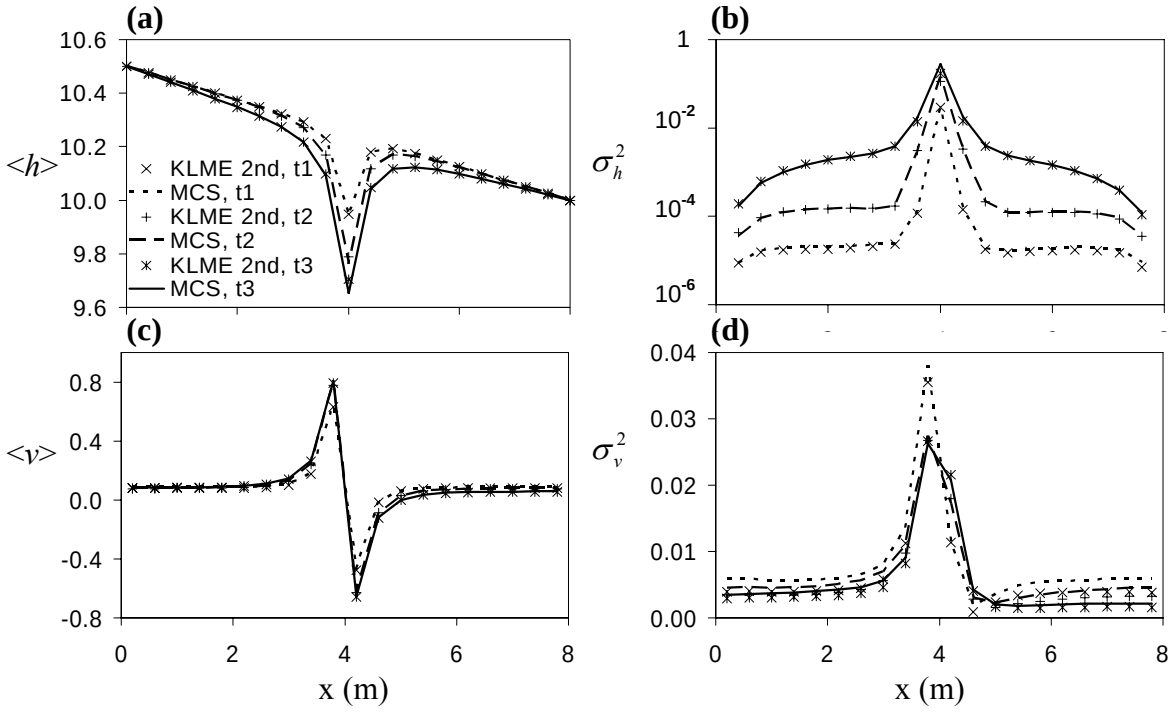

FIGURE 2. The means and variances of heads and velocities at different times after pumping.

Figure 2 displays the means and variances of heads and flow velocities along the profile A-A′ at three different times after well pumping. The legends for different curves are provided in Figure 2(a). Both the means and variances are approximated up to the second-order in the KLME method. Across different times, the KLME approach is accurate in matching the MCS solutions for all flow quantities examined. Due to pumping, there is a sharp head drop at the

well location. The head variances rise quickly near the well and stay relatively small and flat in the area away. The velocity statistics are also dominated by well pumping. The mean velocities show a large variation at the well and so do the velocity variances. From this example we can see the KLME method is accurate in providing the stochastic flow solutions. As demonstrated earlier, this is always a necessary condition before the transport problems can be appropriately solved.

1.2 Two-dimensional transport simulation.

The transport analysis is conducted in a rectangle of 30 m long by 10 m wide with no flow across the northern and southern boundaries and constant head 10.0 m at the eastern boundary. There is a specified flux across the western boundary whose value is determined such that an average hydraulic gradient of 0.001 is established in the west-east direction. The log conductivity field is assumed to be second order stationary and follows a separable exponential covariance with parameters $\eta_x = 2\,\text{m}$, $\eta_y = 1\,\text{m}$, $\sigma_Y^2 = 0.25$ and $K_G = 1.0$ m/d. The transport simulations employ a uniform effective porosity of 0.35, longitudinal and transverse dispersivities 0.2m and 0.02 m, and a molecular diffusion coefficient of 5.0×10^{-4} m^2/d. Initial mass is instantaneously injected in a single cell (Figure 3).

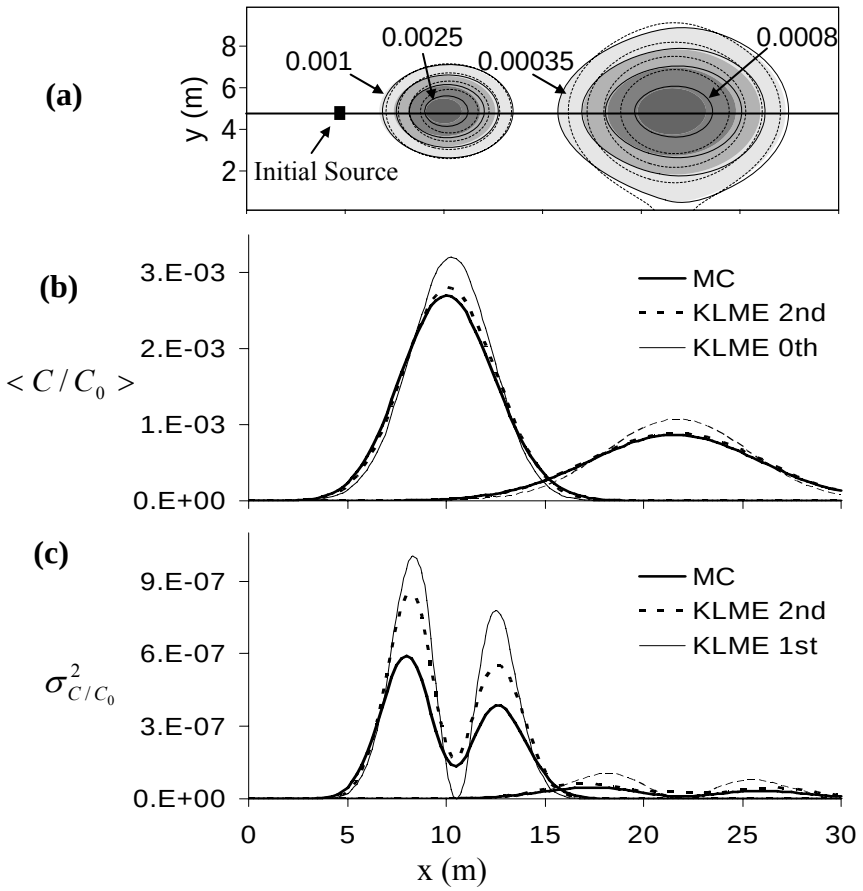

FIGURE 3. The means and variances of concentrations for the transport test case.

Figure 3 shows the mean concentrations and concentration variances computed from the KLME approach as compared to the MCS solutions. Results are presented for concentrations

normalized by the initial source concentration C_0 at two different times. Figure 3(a) contours the zeroth- and second-order mean concentrations along with the MCS solutions, indicating that the zeroth results underestimate solute spreading towards the edges of plume and show a large overestimation near the plume center. By adding the second-order corrections, the KLME approach is able to match the MCS solutions better. These observations are further corroborated by the concentration profiles along A-A′ shown in Figure 3(b). For the concentration variances shown in Figure 3(c), the first-order results overshoot the MCS solutions at the bimodal peaks while undershoot the low values in the center area. The second-order corrections help improve the agreement by rectifying both the overestimation at the peaks and the underestimation at the middle low values on the variance profiles.

4. CONCLUSIONS

This paper has presented a Karhunen-Loéve decomposition based moment-equation (KLME) approach for solving stochastic fluid flow and transport problems in large-scale heterogeneous formations. The accuracy and efficiency of this approach have been validated with comparisons to the classical Monte Carlo simulations (MCS) in two hypothetical examples. Results indicate that the KLME approach offers one order of magnitude reduction in the computation effort, and yet still provides accurate solutions comparable to the MCS.

References

Burnett, R. D. and E. O. Frind (1987), An alternating direction Galerkin technique for simulation of groundwater contaminant transport in three dimensions: 2. Dimensionality effects, *Water Resour. Res.*, 23(4), 695–705.

Courant, R. and D. Hilbert (1953), *Methods of Mathematical Physics*, Interscience, New York.

Graham, W.D. and D. McLaughlin (1989), Stochastic analysis of nonstationary subsurface solute transport, 1. Unconditional moments, *Water Resour. Res.*, 25, 215.

Hassan, A. E., J. H. Cushman and J. W. Delleur (1998), A Monte Carlo assessment of Eulerian flow and transport perturbation models, *Water Resour. Res*., *34*, 1143–1163.

Liu, G., Zhang, D. and Z. Lu, Stochastic Uncertainty Analysis for Solute Transport in Randomly Heterogeneous Media Using a Karhunen–Loéve Based Moment Equation Approach, *in review*.

Lu, Z., and D. Zhang (2004), A comparative study on quantifying uncertainty of flow in heterogeneous media using Monte Carlo simulations, the conventional and the KL-based moment-equation approaches, *SIAM J. Scientific Computing***,** 26(2), 558-577.

Lu, Z., and D. Zhang (2006), Accurate, Efficient Quantification of Uncertainty for Flow in Heterogeneous Reservoirs using the KLME Approach, *SPE Journal*, in press.

Neuman, S.P. (1993), Eulerian–Lagrangian theory of transport in space–time nonstationary velocity fields: Exact nonlocal formalism by conditional moments and weak approximations, *Water Resour. Res.*, 29, 633.

Zhang, D. (1998), Numerical Solutions to Statistical Moment Equations of Groundwater Flow in Nonstationary, Bounded Heterogeneous Media, *Water Resources Research*, Vol.34, No.3.

Zhang, D. and Z. Lu (2002), Stochastic analysis of flow in a heterogeneous unsaturated–saturated system, *Water Resour. Res.,* 38, 10.1029/2001WR000515.

Zhang, D. and Z. Lu (2004), An efficient, high-order perturbation approach for flow in random porous media via Karhunen–Loéve and polynomial expansions, *J. Comput. Phys.*, 194, 773– 794.