Advances in Water Resources 75 (2015) 53–66

Contents lists available at ScienceDirect

Advances in Water Resources

journal homepage: www.elsevier.com/locate/advwatres

A heterogeneity model comparison of highly resolved statistically anisotropic aquifers

Erica R. Siirila-Woodburn [a,b,*], Reed M. Maxwell [a]

[a] Hydrologic Science & Engineering Program, Integrated GroundWater Modeling Center (IGWMC), Department of Geology and Geological Engineering, Colorado School of Mines, Golden, CO 80401, United States

[b] Department of Geotechnical Engineering and Geosciences, Universitat Politècnica de Catalunya, UPC-BarcelonaTech, 08034 Barcelona, Spain

ARTICLE INFO

Article history:
Received 26 May 2014
Received in revised form 24 October 2014
Accepted 28 October 2014
Available online 11 November 2014

Keywords:
Heterogeneous
Geostatistics
High resolution
Uncertainty
Spatial moment
Breakthrough curve

ABSTRACT

Aquifer heterogeneity is known to affect solute characteristics such as spatial spreading, mixing, and residence time, and is often modeled geostatistically to address aquifer uncertainties. While parameter uncertainty is often considered, the model uncertainty of the heterogeneity structure is frequently ignored. In this high-resolution heterogeneity model comparison, we perform a stochastic analysis utilizing spatial moment and breakthrough curve (BTC) metrics on Gaussian (*G*), truncated Gaussian (*TG*), and non-Gaussian, or "facies" (*F*) heterogeneous domains. Three-dimensional plume behavior is rigorously assessed with meter (horizontal) and cm (vertical) scale discretization over a ten-kilometer aquifer. Model differences are quantified as a function of statistical anisotropy, ε, by varying the *x*-direction integral scale of hydraulic conductivity, *K*, from 15 to 960 (m). We demonstrate that the model is important only for certain metrics within a range of ε. For example, spreading is insensitive to the model selection at low ε, but not at high ε. In contrast, center of mass is sensitive to the model selection at low ε, and not at high ε. A conceptual model to explain these trends is proposed and validated with BTC metrics. Simulations show that *G* model effective *K*, and 1st and 2nd spatial moments are much greater than that of *TG* and *F* models. A comparison of *G* and *TG* models (which only differ in *K*-distribution tails) reveal drastically different behavior, exemplifying how accurate characterization of the *K*-distribution may be important in modeling efforts, especially in aquifers where extreme *K* values are often not measured, or inadvertently overlooked.

1. Introduction

Because the subsurface is largely unknown, geostatistical approaches utilizing the global statistics of the aquifer (for example, a mean, variance, and directional correlation lengths of hydraulic conductivity, K (m d^{-1})), are used to address aquifer uncertainties. Stochastic methods have a long history in hydrogeology (e.g. [11,40]), and are advantageous in assessing contaminant transport and the associated risks as they allow for an endpoint (such as a concentration, a flux, or through analytical equations the probability of risk) to be assessed in terms of bounds of uncertainty (e.g. [1,5,41,47]. While parameter uncertainty is often analyzed for sensitivity, the associated uncertainty of the heterogeneity model used to define the structure of the flow field is often ignored. Given the same global statistics of input parameters but different heterogeneity models, how different are important plume metrics such as arrival times and plume spreading? Moreover, how sensitive are these differences, if any, to highly uncertain aquifer parameters such as the degree of statistical anisotropy? Such an analysis is important in scenarios where site characterization is poor, leading to an un-informed heterogeneity model selection. Additionally, when the site is well characterized and one model is considered more suitable over the other, the comparison of heterogeneity models of equivalent model statistics may help to provide a type of confidence interval on the error resulting from selecting one model over another.

The number of heterogeneity model comparisons in the literature are surprisingly sparse (e.g. [18,51,54,55], especially in three dimensions [3,4,12,29] where differences in important behavior such as late-time tailing are observed. To accurately characterize plume transport and concentrations in natural settings, modeling of aquifer connectivity and potentially tortuous flow pathways with a three-dimensional, spatially correlated heterogeneity model

* Corresponding author at: Department of Geotechnical Engineering and Geosciences, Universitat Politècnica de Catalunya, UPC-BarcelonaTech, 08034 Barcelona, Spain.

E-mail address: erica.siirila@upc.edu (E.R. Siirila-Woodburn).

http://dx.doi.org/10.1016/j.advwatres.2014.10.011
0309-1708/

is imperative, especially when analyzing macrodispersion [17]. As demonstrated in previous work [38,43,44] there is a clearly linked relationship between aquifer connectivity, statistical anisotropy of K, solute residence time (and thus allowed time for a reaction to occur), and water quality assessment. The question of how to accurately model the three-dimensional structure of heterogeneous flow fields, and furthermore the sensitivity in the heterogeneity model, is an important one which deserves more analysis.

While there are a range of emerging approaches in spatial statistics, in this work we focus on the comparison of two of the most widely used representations of spatial persistence of K in hydrogeology: (1) Gaussian random fields and (2) non-Gaussian, or facies indicator approach [26]. We contend that inherent in the choice of a heterogeneity model may be a choice in aquifer connectivity (e.g. [50]). For example, in a Gaussian random field, high and low K materials are spatially isolated due to the continuous nature of the way the heterogeneity is produced in this model. In contrast, the facies approach allows for the possibility of sharp interfaces between high and low K material, potentially causing differences in how connected the simulated aquifer is. LaBolle and co-authors found the importance of neighboring strata and hydrofacies (i.e. low and high K zones) to be controlling in correctly dating post-1950, prebomb peak ^{3}H and ^{3}He water due to molecular diffusion and diffusive fractionation [27]. The idea of low-K connections in an isotropic, Gaussian random field has also been explored in the context of mobile-immobile mass transfer [55], where flow fields with nearly identical lognormal univariate conductivity distributions but different connectivity patterns of high and low conductivity regions (connected high K structures, connected intermediate K structures, and connected low K regions) result in different flow and transport. In a comprehensive analysis of facies fields, Zhang et al. [53] found that the main factor affecting non-Fickian transport is the vertical mean length of diffusion-limited layers, suggesting the importance of the structure and arrangement of high and low K material in aquifers is crucial.

To date, very few studies have analyzed the three dimensional differences in flow given Gaussian and non-Gaussian models. Most recently, Berg and Illman [3] compared the performance of hydraulic tomography with traditional approaches (kriging, effective parameters, transition probability Markov chain models, geologic models, stochastic inverse modeling), and assessed performance using observed well-drawdown data. The analysis of the traditional methods were inconsistent; that is, for some drawdown tests the error was small, and for others it was substantial. In a Lawrence Livermore National Laboratory pump test simulation utilizing transition probability indicator simulations (T-PROGS) and sequential Gaussian random fields, the T-PROGS model yielded an aquifer with greater lateral connectivity and could better reproduce simulated drawdown response behavior [29]. The simulated aquifer, however, was characterized (and conditioned) by a large number of facies samples in a highly heterogeneous aquifer (with a variance of $\ln(K) \approx 25$), potentially leading to a naturally better fit with a geologically realistic, facies model. A different study comparing a small, 6 (m) $\times$ 6 (m) $\times$ 6.2 (m), area of the macrodispersion experiment site (MADE) showed 43% of the fastest particle's paths are located within a high K zone in a Gaussian random field and as high as 69% in a T-PROGS field [4]. In this comparison preferential flow dictated transport, where fast pathways in the variogram-based methods were not necessarily through the highest K material, suggesting particle "jumps" or "leaks". To accurately characterize these interactions over long transport distances, fine cell discretization and large ensembles are needed to provide robust statistical conclusions. None of the aforementioned three-dimensional studies investigates far-field interactions, where the three analyses were restricted to sub-fifteen meter spatial extents.

To avoid these limitations, comparisons of Gaussian and non-Gaussian random fields in this analysis are (1) finely discretized (meter in the horizontal direction, cm in the vertical direction), (2) large extent (ten kilometers), and (3) composed of an ensemble of multiple realizations to minimize statistical error. To address aquifer connectivity, spatial auto-correlation lengths of K (and therefore statistical anisotropy) are also varied in this analysis, where integral scales vary from the tens of meters to the kilometer scale. With this highly resolved, large-scale numerical setup, our primary objective is to quantify and compare the effect of heterogeneity model on plume metrics such as peak and mean arrival times, aquifer connectivity, and plume center of mass and spreading. Sensitivity to model selection is assessed as a function of statistical anisotropy for a given travel distance, and through the use of numerical solutions. Section 2.1 describes the heterogeneity models used to generate the statistically anisotropic K fields and the range of parameters assessed, followed by Section 2.2 which describes the numerical setup of the flow and transport models used for all heterogeneity models. Lastly, Section 2.3 defines the metrics used to characterize plume behavior and to perform the heterogeneity model comparison.

2. Methodology

2.1. Model of heterogeneity

In the facies approach, an integer code is assigned to lithologic or hydrologic units, creating an indicator database to each facies type [8,14,20,22,26,39]. Other additional information can be included such as the volumetric proportion, mean lengths, and juxtapositional tendencies of each indicator, yielding the statistical basis for the facies model. Examples of facies models are SISIM within the GSLIB package [14] and T-PROGS [9] which simulates integer-coded categorical variables or continuous variables with an indicator database. The advantage of a facies approach is a more geologically realistic model of spatial persistence, able to represent features such as asymmetric patterns of heterogeneity, sharp interfaces, and an upward-fining feature typical of many alluvial deposits (e.g. [16]). Statistically, a Gaussian random field can be generated in an equivalent way to a facies field, and could thus be compared. Common models used to produce such geostatistical representations include GSLIB [14] and the turning-bands approach [23,30–32,48]. The turning bands algorithm generates spatially correlated random fields from a normal distribution with zero mean and specified covariance structure. The method involves the generation of a series of one-dimensional random processes along lines radiating from a coordinate origin that are projected from these lines onto a three-dimensional space. The advantage of this approach is the computational efficiency of utilizing a one-dimensional solution in three-dimensional space, but requires a large number of lines in the solution to avoid spatial distortion. In this analysis, three-dimensional Gaussian and non-Gaussian models of spatially correlated random K fields are stochastically simulated and compared. A Gaussian random field assumes a log-normal distribution of the hydrologic parameter (i.e. K), and is completely characterized by the mean, variance, and semivariogram (e.g. [40].

Here, Gaussian fields are generated using the turning bands algorithm [48]. The turning-bands algorithm enforces a semivariogram function through rotation of one-dimensional lines (or bands) through space, where each value in the random field is a weighted average of values contained within each band [48]. An exponential model is used to define spatial correlation of K via a separation distance, ξ (m), and a directional integral scale, I (m):

$$R(\xi) = \sigma^2 \exp^{-\xi/I} \qquad (1)$$