Equivalent hydraulic conductivity of an experimental stratigraphy: Implications for basin-scale flow simulations

Ye Zhang,[1] Carl W. Gable,[2] and Mark Person[3]

Received 11 November 2005; revised 10 January 2006; accepted 24 January 2006; published 3 May 2006.

[1] Critical issues facing basin-scale groundwater flow models are the estimation of representative hydraulic conductivity for the model units and the impact of nonrepresentation of within-unit conductivity heterogeneity on the model flow prediction. In this study, high-resolution, fully heterogeneous basin-scale hydraulic conductivity map is generated by scaling up an experimental stratigraphy created by physical sedimentation processes and by assuming increasing conductivity for increasing gray scale (proxy for sand content). A fully heterogeneous model is created, incorporating the complete conductivity variation. Two hydrogeologic framework models are also created, one of coarser stratigraphic division. A novel numerical up-scaling method is developed to compute an equivalent conductivity for each irregularly shaped framework model unit by conducting basin-scale flow experiments in the fully heterogeneous model. In each experiment, different boundary conditions are specified, subjecting the basin to various flow conditions. To evaluate the impact of using equivalent conductivity on the prediction of basin-scale hydraulic head and groundwater flow, the flow experiments conducted in the fully heterogenous model are repeated in the framework models. Results indicate that for most deposits, the behavior of the equivalent conductivity with increasing ln(K) variance is consistent with the prediction of an analytic-stochastic theory. The equivalent conductivity is also insensitive to the boundary condition and the number of flow experiments performed, indicating the possible emergence of an effective conductivity. Although all equivalent conductivities are full tensors, the off-diagonal term is 2–3 orders of magnitude smaller than the diagonal terms. Ignoring the off-diagonal term has minimal impact on the framework-model-predicted hydraulic head and groundwater flow paths, when compared to the impact of nonrepresentation of within-unit conductivity heterogeneity. Under certain boundary conditions, significant head deviation can develop within framework model units that contain trended or strongly stratified deposits. However, the accuracy of head prediction is improved when the length of the no-flow boundary is increased. In a topography-driven system, progressive degradation is observed in the prediction of basin-scale flow pattern, flow rate, and location of recharge/discharge, when the progressively up-scaled framework models are used. In summary, the accuracy of the framework models is controlled by the level of stratigraphic division, conductivity heterogeneity, and boundary conditions.

Citation: Zhang, Y., C. W. Gable, and M. Person (2006), Equivalent hydraulic conductivity of an experimental stratigraphy: Implications for basin-scale flow simulations, *Water Resour. Res.*, *42*, W05404, doi:10.1029/2005WR004720.

1. Introduction

[2] Environmental and petroleum studies during the past two decades have documented that spatial variation in hydraulic conductivity, or conductivity heterogeneity, is the rule rather than exception in natural sedimentary deposits. Conductivity heterogeneity is considered a major factor influencing solute migration. It is considered less crucial for estimating bulk flow characteristics. In typical regional to basin-scale groundwater studies, conductivity heterogeneity is rarely incorporated into the flow models, due to the prohibitive cost of conducting detailed sampling over large spatial scales and/or computation limit. In these studies, sedimentary deposits are commonly represented by a series of internally homogeneous hydrogeologic units [e.g., *Garven and Freeze*, 1984; *Bethke*, 1985; *Senger and Fogg*, 1987; *Person and Garven*, 1992; *Gupta and Bair*, 1997; *Castro et al.*, 1998; *Bethke et al.*, 1999; *Walvoord et al.*, 1999; *Person et al.*, 2003; *Zhang et al.*, 2005a]. An equivalent (saturated) hydraulic conductivity is assigned to each unit to relate the mean head gradient to the average groundwater fluxes. Lateral or vertical conductivity trends

[1]Department of Geological Sciences, University of Michigan, Ann Arbor, Michigan, USA.
[2]Los Alamos National Laboratory, Los Alamos, New Mexico, USA.
[3]Department of Geological Sciences, Indiana University, Bloomington, Indiana, USA.

0043-1397/06/2005WR004720

within each unit are ignored. A hydrogeologic framework model is developed to study a variety of basin-scale geological, geothermal and environmental problems impacted by groundwater flow. However, against natural heterogeneity, hydrogeologic framework models are mathematical conceptualizations. Thus the critical issue facing these studies is the accuracy of the framework model and the associated equivalent conductivity, and, more importantly, the impact of nonrepresentation of within-unit conductivity heterogeneity on model predictions of flow and transport.

[3] Numerous methods have been developed to estimate the equivalent hydraulic conductivity for heterogenous deposits (see reviews by *Desbarats* [1992], *Sánchez-Vila et al.* [1995], *Wen and Gómez-Hernández* [1996], and *Renard and de Marsily* [1997]). The equivalent conductivity differs from an effective conductivity which is defined within a stochastic context where the small-scale conductivity heterogeneity is considered a random space function (RSF) [*Dagan*, 1989; *Gelhar*, 1993; *Indelman and Dagan*, 1993]. The effective conductivity can be estimated from the spatial correlation and variability characteristics of the RSF. It is considered an intrinsic property of the RSF, thus independent of the boundary condition. On the other hand, the equivalent conductivity represents a fictitious homogenous medium that preserves the mean flux of the heterogenous deposit for a given head gradient. Compared to the effective conductivity, the estimation of the equivalent conductivity does not require restrictive assumptions on the RSF or the flow condition (e.g., stationarity, mean uniform flow). This is particularly useful for deposits that exhibit complex spatial heterogeneities or long-term correlation, i.e., the problem size is finite compared to the conductivity correlation range. However, the equivalent conductivity depends on the boundary condition, thus it is not unique. The equivalent conductivity may approach the effective conductivity when the domain size is much larger than the conductivity correlation range [*Renard and de Marsily*, 1997].

[4] In this study, a novel numerical up-scaling method is developed to calculate an equivalent conductivity for various hydrogeologic units, based on results of flow simulations conducted in a high-resolution, fully heterogeneous model of basin-scale hydraulic conductivity. The equivalent conductivity is first compared to an effective conductivity predicted by an analytic-stochastic theory. The sensitivity of the equivalent conductivity to the boundary condition is explored. The equivalent conductivity is then assigned to the hydrogeologic units to construct two framework models which employ progressively lower resolution of the geologic heterogeneity (i.e., the level of stratigraphic division). By conducting basin-scale flow simulations within the fully heterogeneous model and each of the framework models, the impact of ignoring within-unit conductivity heterogeneity on the prediction of hydraulic head and groundwater flow path is evaluated. The consequence of low-resolution representation of the geologic heterogeneity is highlighted.

[5] Compared to prior work based on synthetic data, the conductivity map of this study is not generated to satisfy any assumptions on sedimentary structure or small-scale conductivity correlation. Instead, it is created by scaling a stratigraphic image of a physical deposit created by a sediment transport experiment. Another unique aspect is that the hydrogeologic units (subjected to upscaling) are irregular in shape, more realistic considering that natural sedimentary deposits often exhibit irregular formation or lithofacies contact. Prior upscaling work was usually done for a "box" or "cube" reflecting the shape of numerical grid cells, i.e., computing the grid conductivity based on conductivity measured at smaller data support. The upscaling methodology developed in this study can be easily extended to calculate the equivalent conductivity for any natural systems, provided that detailed local conductivity measurements are available.

[6] In the remainder of the text, the fully heterogenous, basin-scale hydraulic conductivity map is introduced first. The construction of the three models is described. The upscaling methodology is presented along with the boundary conditions specified for the flow experiments. The results are presented in two parts: (1) the estimation of equivalent conductivity and (2) the impact of using equivalent conductivity. Finally, the implications for basin-scale flow modeling are discussed and future research indicated.

2. Methodology

2.1. Basin-Scale Hydraulic Conductivity Map

[7] This study is based on a stratigraphic image of a deposit created in an experimental facility where multiple sedimentary facies formed in response to a variety of depositional processes (Figure 1) [*Schneider*, 1998; *Paola*, 2000; *Paola et al.*, 2001; *Heller et al.*, 2001; *Sheets et al.*, 2002; *Cazanacli et al.*, 2002; *Strong et al.*, 2006]. In the 1996 prototype experiment, fine quartz sand and coal were used as proxies for coarse- and fine-grained (clay) sediments, respectively. During the experiment, the elevation of the standing water at the downstream side was varied in a sinusoidal manner to represent eustatic sea level fluctuations. From a single source location, sediments (50:50 mixture of sand and coal) were deposited within an accommodation space of $1.6 \times 1.0 \times 0.8$ m^3, overlying a flexible, subsiding floor. The sediment influx rate was controlled to keep pace with the subsidence rate. Fluvial, shoreline and submarine processes occurred over varying spatial and temporal scales, forming channels and floodplain and creating turbidity currents in the adjacent sea. At the distal end of the flume, coal particles settled as fine-grained clay would settle in an open ocean. Over time, changes in sea level also induced large-scale shoreline regression and transgression, and accordingly, transition from upstream fluvial deposits to downstream deepwater deposits.

[8] The prototype deposit formed in response to the combined forcing of sediment input, basement subsidence and sea level change; the dominant processes that formed the deposit were comparable to those forming a natural fluvial/deltaic system. After the deposit was produced, it was dissected along the average sediment transport direction in centimeter-scale increment. High-resolution photographs of the exposed sediment faces were taken and digitized. The dissection revealed that most of the fluvial sands and floodplain coals were laterally continuous. This is because sheet flow dominated fluvial deposition in this experiment. Thus a two-dimensional image is considered representative of the stratigraphic heterogeneity of the deposit (Figure 2).

[9] This image contains 1000 pixels in the horizontal direction and 600 pixels in the vertical direction. Associated

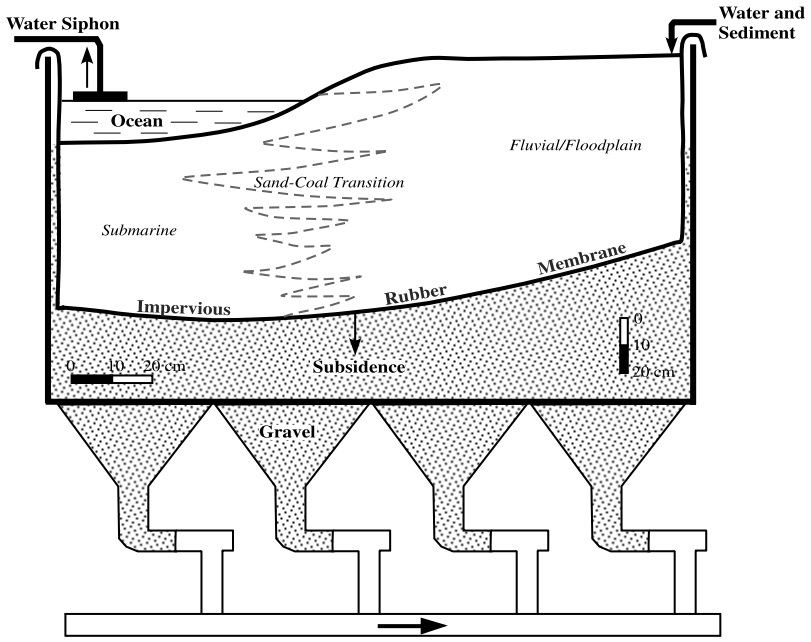

Figure 1. A schematic diagram of the prototype experiment. Beneath the rubber membrane is pea-sized gravel overlying a system of funnels which controls subsidence. Spatial and temporal continuity in subsidence is ensured by firing high-pressure water jets into the elbow pipes and knocking small volumes of gravel into an exhaust line.

with each pixel is a gray scale which ranges from 0 to 232, corresponding to the highest coal and sand fractions, respectively. This image is first scaled up to sedimentary basin dimensions by assuming that each pixel represents a homogenous and isotropic porous media or local representative elementary volume (REV) of dimensions: 100×5 m^2. The basin dimensions are chosen so that the average topographic slope is 1/100, and the basin length-to-depth ratio is 50:1, falling within the observed range for natural systems [*Belitz and Bredehoeft*, 1990]. Each REV is assigned a ln(K) value via linear interpolation of gray scale and end-member ln(K) (one for pure clay; one for pure sand). These end-members are selected based on field measurements of an unconsolidated alluvial fan deposit [*Lu et al.*, 2002]. Note that visual images of sedimentary rock faces can provide a description of the spatial structure of ln(K) [*Tidwell and Wilson*, 2002]; the end-member conductivities are used since the sand/coal used in the experiment have permeability of medium sand, the coal is also more permeable. By using end-member conductivities, lighter and sandier facies in the image are assigned higher conductivities, vice versa.

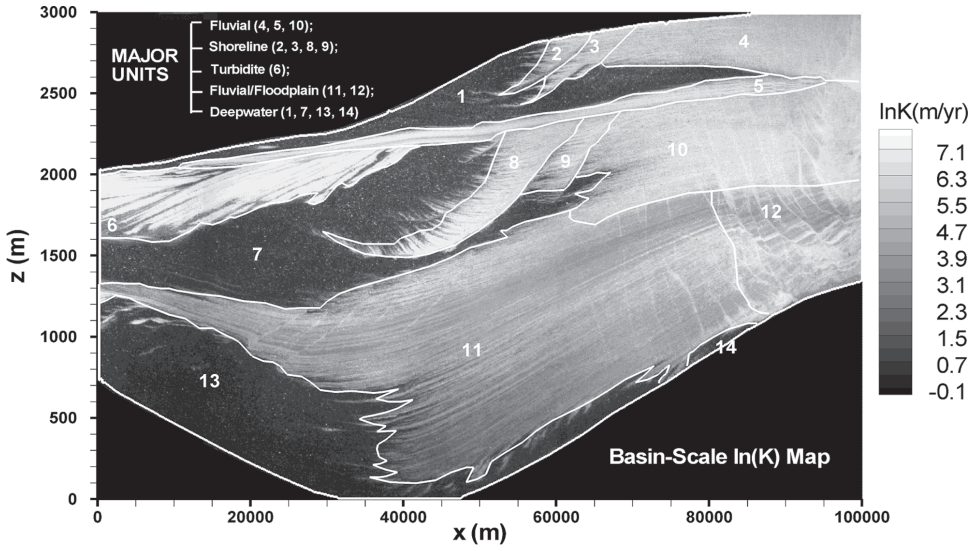

Figure 2. A basin-scale ln(K) map (in m/yr) with 14 stratigraphic units representing different depositional environments.

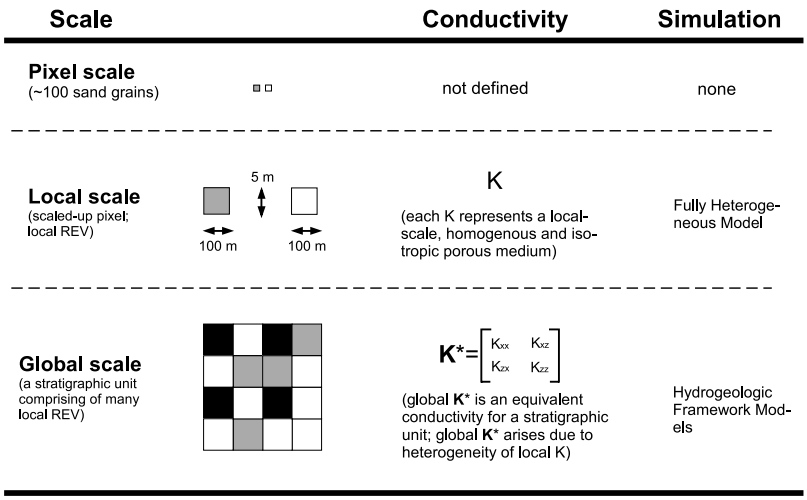

Figure 3. A schematic diagram illustrating the scale-up procedure and the up-scaling analysis. The scale-up procedure is to transform a deposit-scale image to sedimentary basin dimensions. Up-scaling occurs when a global equivalent conductivity (**K***) is computed for a heterogeneous deposit.

[10] The resulting conductivity map is geologically realistic and hydrologically definable with a mean, variance and frequency distribution comparable to those of the fan deposit [*Zhang et al.*, 2005b]. Though this conductivity map is synthetic, it is unique compared to other numerical aquifers [e.g., *Scheibe and Freyberg*, 1995]. The experimental deposit was produced under controlled conditions (sediment flux, subsidence, sea level change) by processes that represent a subset of those active in nature. These include important forms of self-organization and spontaneous pattern formation that are common in nature yet difficult to capture using simulations. By associating each pixel with a local conductivity and REV, no cutoff is employed in this map beneath which conductivity heterogeneity no longer exists. The procedure of image scale-up and the subsequent upscaling analysis to compute the equivalent conductivity is summarized in Figure 3. More detailed discussions on the experimental stratigraphy and the construction of the basin-scale conductivity map are given by *Zhang et al.* [2005b].

2.2. Model Construction

[11] On the basis of the depositional environment, the heterogenous conductivity map is divided into 14 stratigraphic units (Figure 2). These units provide the basis for the construction of two hydrogeologic framework models, one with a finer division of the geological heterogeneity. However, the shapes of the basin and the stratigraphic units are not regular. The construction of the fully heterogenous model and hydrogeologic framework models involves a number of steps:

[12] 1. A mesh generation program, Lagrit, generates a high density finite element grid (http://lagrit.lanl.gov). This grid incorporates all the pixel data, within and outside the basin (Figure 2). Each local conductivity is represented exactly by two triangular elements.

[13] 2. A mesh generation package, Argus, generates a low density finite element grid. Using a digitization tool, stratigraphic unit types (1–14) are mapped onto this grid as elemental material tag. Area outside the basin is given tags 100, 200, and 300.

[14] 3. Lagrit maps the elemental material tag of the Argus grid onto the dense grid and removes the elements and nodes associated with tag 100, 200, and 300. A new grid is created with 424,217 nodes and 845,208 elements. Each element has an associated conductivity and material tag (Figure 2).

[15] 4. All models use this high density grid. The fully heterogenous model is given the elemental conductivity. A 14-unit hydrogeologic framework model (hydrostratigraphic model) has the elemental material tag.

[16] 5. A two-unit model (lithostratigraphic model) is also created after retagging the 14-unit model: unit 1 for an aquitard unit (units 1, 7, 13, 14 of Figure 2) and unit 2 for an aquifer (all sand-rich units).

[17] Three model representations of the basin are thus created, the fully heterogeneous model is considered a reference model (ground truth) against which the performance of the two framework models is evaluated. In section 2.3, the method of estimating the equivalent conductivity for the hydrogeologic units is described.

2.3. Equivalent Conductivity Estimation

[18] To compute the equivalent conductivity, steady state, incompressible groundwater flow experiments are conducted in the fully heterogeneous model (or reference model):

$$\nabla \cdot \mathbf{q} = 0$$
$$\mathbf{q} = -K(x,z)\nabla h \tag{1}$$

where ∇ is the gradient operator, center dot is the inner product q is the Darcy flux. Note that in this model, each local conductivity $K(x, z)$ is scalar. However, because of heterogeneity, the equivalent conductivity is usually tensorial. For each unit of the hydrostratigraphic or lithostratigraphic models, the equivalent conductivity can be expressed as

$$\mathbf{K}^* = \begin{bmatrix} K_{xx} & K_{xz} \\ K_{zx} & K_{zz} \end{bmatrix} \tag{2}$$

If symmetry is assumed, the two diagonal components (K_{xx}, K_{zz}) and one off-diagonal component (K_{xz}) need to be

obtained for each unit. To uniquely solve for these three unknowns, at least two experiments are needed, under two different boundary conditions. For the first experiment and the associated boundary condition (e.g., BC 1), the mean head gradient and Darcy flux within each unit can be estimated for which a global Darcy's law can be written

$$\mathrm{BC1}: \begin{Bmatrix} \langle q_x \rangle_1 \\ \langle q_z \rangle_1 \end{Bmatrix} = - \begin{bmatrix} K_{xx} & K_{xz} \\ K_{zx} & K_{zz} \end{bmatrix} \begin{Bmatrix} \langle \partial h/\partial x \rangle_1 \\ \langle \partial h/\partial z \rangle_1 \end{Bmatrix} \tag{3}$$

where q_x, q_z are the components of the Darcy flux $\mathbf{q}$, angle brackets indicate spatial averaging. For a different boundary condition (e.g., BC 2), another set of mean gradient and flux can be estimated and Darcy's law written:

$$\mathrm{BC2}: \begin{Bmatrix} \langle q_x \rangle_2 \\ \langle q_z \rangle_2 \end{Bmatrix} = - \begin{bmatrix} K_{xx} & K_{xz} \\ K_{zx} & K_{zz} \end{bmatrix} \begin{Bmatrix} \langle \partial h/\partial x \rangle_2 \\ \langle \partial h/\partial z \rangle_2 \end{Bmatrix} \tag{4}$$

where the subindexes indicate the boundary conditions. By combining equations (3) and (4), the components of $\mathbf{K}^*$ (K_{xx}, K_{xz}, K_{zx}, K_{zz}) can be determined by solving a set of linear equations. Note that this method does not guarantee symmetry (In practice, an average of the off-diagonal terms is used to obtain K_{xz}). Moreover, an equivalent conductivity estimated from results of two experiments may exhibit dependency on the specific boundary condition. To obtain a representative equivalent conductivity or one that does not vary greatly with changing boundary condition, ideally, numerous flow experiments are needed, representing a full spectrum of likely groundwater flow scenarios and the associated boundary conditions [*Wen and Gómez-Hernández*, 1996]. An equivalent conductivity is obtained by incorporating results from all experiments:

$$\mathrm{BC1}: \begin{Bmatrix} \langle q_x \rangle_1 \\ \langle q_z \rangle_1 \end{Bmatrix} = - \begin{bmatrix} K_{xx} & K_{xz} \\ K_{zx} & K_{zz} \end{bmatrix} \begin{Bmatrix} \langle \partial h/\partial x \rangle_1 \\ \langle \partial h/\partial z \rangle_1 \end{Bmatrix}$$

$$\cdots \qquad\qquad \cdots$$

$$\mathrm{BC}m: \begin{Bmatrix} \langle q_x \rangle_m \\ \langle q_z \rangle_m \end{Bmatrix} = - \begin{bmatrix} K_{xx} & K_{xz} \\ K_{zx} & K_{zz} \end{bmatrix} \begin{Bmatrix} \langle \partial h/\partial x \rangle_m \\ \langle \partial h/\partial z \rangle_m \end{Bmatrix}$$

$$\begin{bmatrix} \langle \partial h/\partial x \rangle_1 & \langle \partial h/\partial z \rangle_1 & 0 & 0 \\ 0 & 0 & \langle \partial h/\partial x \rangle_1 & \langle \partial h/\partial z \rangle_1 \\ \cdots & \cdots & \cdots & \cdots \\ \langle \partial h/\partial x \rangle_m & \langle \partial h/\partial z \rangle_m & 0 & 0 \\ 0 & 0 & \langle \partial h/\partial x \rangle_m & \langle \partial h/\partial z \rangle_m \end{bmatrix} \begin{Bmatrix} K_{xx} \\ K_{xz} \\ K_{zx} \\ K_{zz} \end{Bmatrix} = - \begin{Bmatrix} \langle q_x \rangle_1 \\ \langle q_z \rangle_1 \\ \cdots \\ \langle q_x \rangle_m \\ \langle q_z \rangle_m \end{Bmatrix} \tag{5}$$

where m is the number of flow experiments. Note that the coefficient matrix ($\mathbf{A}$) is of dimensions $2m \times 4$, the right-hand-side vector ($\mathbf{b}$) is of dimensions $2m$, while we only have four unknowns ($\mathbf{x} = [K_{xx}, K_{xz}, K_{zx}, K_{zz}]^T$). Thus, for $m > 2$, this system of equation is overdetermined and is solved via least squares solution to minimize the 2-norm of the residue vector $\mathbf{r} = \mathbf{Ax} - \mathbf{b}$, the 2-norm is given as $|\mathbf{r}| = \sqrt{\sum_{i=1}^{2m} (r_i^2)}$.

[19] To calculate an equivalent conductivity for each stratigraphic unit, numerical flow experiments are conducted using a finite element (FE) program. A direct solver is used to compute the hydraulic head. The finite element method is chosen due to its expediency in estimating the mean head gradients and Darcy fluxes for the irregular units. For example, using linear shape functions, the FE approximation of the hydraulic head over a three-node triangular element is

$$\hat{h}(x,z) = \sum_{i=1}^{3} \psi_i(x,z) h_i \tag{6}$$

where h_i is the computed nodal hydraulic head. The spatial variation over the element is represented by the shape functions $\psi_i(x, z)$, $i = 1, 3$. To compute the elemental Darcy flux, substitute equation (6) into the Darcy's law:

$$\mathbf{q}_e = -K_e \nabla \hat{h} = -K_e \sum_{i=1}^{3} [\nabla \psi_i(x,z)] h_i \tag{7}$$

where $\mathbf{q}_e$ is the elemental flux, K_e is the elemental (scalar) hydraulic conductivity. The elemental head gradient ($\nabla \hat{h}$) can be similarly estimated. Since each element of the high-resolution grid has a material tag: hydrostratigraphic model (1–14), lithostratigraphic model (1, 2), the mean flux and head gradient for a particular unit are obtained by directly averaging among the elements with the specific tag.

[20] The above method can be applied when the local hydraulic conductivity is a full tensor, e.g., if stratification is assumed within each REV. (In this study, each local conductivity of the fully heterogeneous model is assumed a scalar, so the equivalent conductivity can be compared to an effective conductivity predicted by an analytic-stochastic theory.) This method can also be applied to any spatial dimensions, and, to any type of finite element cells and levels of approximation, i.e., the shape functions can be higher-order polynomials. It is easily implementable with other numerical schemes, e.g., finite difference (FD) or finite volume method. Note that FD flux exists on the boundary between two cells, an external layer of ghost cells has to be employed for every unit as well as making assumptions on conductivity of the ghost cells.

2.4. Numerical Flow Experiments

[21] In this study, due to the large computation demand of solving the flow equation in the fully heterogenous model, and the amount of bookkeeping required to assemble the coefficient matrix and right-hand-side vector for each stratigraphic unit, four basin-scale flow experiments ($m = 4$) are conducted in the fully heterogenous model by choosing distinct boundary conditions that result in significantly different hydraulic head and groundwater flow patterns in the basin. In each experiment, the basin groundwater was dominated in turn by (1) vertical mean flow, (2) lateral mean

flow, (3) topography-driven flow, and (4) topography-driven flow with a pair of injection/production wells. Thus, for each unit, experiment 1 ensures that the mean vertical head gradient and Darcy flux will be large (compared to the other experiments) to result in accurate determination of K_{zz}. Experiments 2–4 are dominated by lateral flows and thus accurate determination of K_{xx}.

[22] All four experiments solve the same groundwater flow equation (equation (1)); the difference among them is how the boundary condition is assigned:

2.4.1. Experiment 1

[23] The top boundary of the basin is assigned a specified head which approximately follows the top boundary elevation. For the bottom boundary, a similar head profile is assigned which follows the basement elevation. Both sides are assigned no-flow boundaries. A vertical head gradient is established in the basin where groundwater is driven to flow downward. Since the bedding plane in the basin is nearly horizontal (on average, the stratified deposits have a dip angle of around 1°), the overall flow direction is approximately perpendicular to the bedding plane. This type of flow behavior may exist in actively subsiding, geologically young basins, as rapid deposition in overlying low-permeability sediments may create overpressure driving fluids to move downward.

2.4.2. Experiment 2

[24] A lateral head gradient is assigned in the basin: the top and bottom boundaries are no-flow; the right boundary is assigned a high head (1001 m) and the left boundary is assigned a low head (1 m). Within the basin, the average lateral head gradient is 0.01. Groundwater flows from right to left, generally following the bedding plane. This type of flow behavior may occur within a set of confined strata bounded above and beneath by low conductivity deposits. The confined strata receives recharge from farther upstream and discharges farther downstream; thus groundwater flow within the strata is mostly horizontal.

2.4.3. Experiment 3

[25] A specified head that follows the top boundary elevation is assigned to the top basin. No-flow boundaries are assigned to the sides and bottom. Since there is about 1 km head drop from the highest to the lowest elevation, the average head gradient across the basin is 0.01. Beneath the recharge and discharge areas, the groundwater flow direction is subvertical (inclined to the bedding plane); between these areas along the midline basin, the flow direction is mostly lateral (parallel to bedding plane). This type of topography-driven flow can occur in uplifted, subaerial basins that are subjected to higher recharge rates along the uplifted areas. The regional topographic lows generally correspond to discharge areas.

2.4.4. Experiment 4

[26] This experiment assigns the same external boundary conditions as experiment 3; however, a set of injection and production wells is placed in the basin. The injection and production wells are both point source operating at the same flow rate. No extra fluid is thus produced or extracted from the basin. The injection well is placed in the upstream basin, within the sand-rich fluvial deposit; the production well is placed farther downstream, within the fluvial/floodplain deposit. The flow rate at the wells is selected to be sufficiently high (i.e., 10,000 m/yr) to ensure significant perturbations on the basin-scale head and flow pattern. Compared to experiment 3 that represents a naturally equilibrated system, experiment 4 represents a system that has reequilibrated after artificial recharge/discharge.

[27] In setting up these experiments, the goal is to maximize the difference in hydraulic head and groundwater flow pattern. Each flow scenario is thus hypothetical and an exact analogy to real world situations is not the purpose. The above method with the associated flow experiments is neither limited to basin-scale systems nor any particular system scales. Had we decided to scale up the image to the scales of a hydrogeologic field study (10–100 m), the same method can be applied after appropriate modifications of boundary conditions. In this study, we're interested in basin-scale since hydrogeologic framework models are most commonly employed in such studies. By scaling up the image to basin dimensions, a higher range of conductivity can be used. The three models developed can be used for future evaluations on various basin-scale flow/heat/transport processes.

2.5. Accuracy of Framework Models

[28] The emphasis of this study is to evaluate the impact of using equivalent conductivity ($\mathbf{K}^*$) by hydrogeologic framework models on the prediction of hydraulic head and groundwater flow. The simulations conducted in the fully heterogenous model to obtain $\mathbf{K}^*$ are repeated by the framework models subjected to the same boundary conditions, i.e., experiments 1–4:

$$\nabla \cdot \mathbf{q} = 0$$
$$\mathbf{q} = -\mathbf{K}^* \nabla h \tag{8}$$

In these models, each element is assigned a tensorial equivalent conductivity estimated for the particular unit to which this element belongs. The same FE program is used to compute the hydraulic head. To eliminate numerical "errors" due to grid coarsening, the same high-density grid is used. Thus deviation in framework model prediction from that of the heterogeneous model is introduced only by the difference in local conductivity assignment (scalar versus equivalent).

[29] It is straightforward to calculate an absolute head deviation predicted by the framework models everywhere in the basin: $\Delta h = h_{fw} - h_{\mathrm{ref}}$, where Δh is the nodal head deviation, h_{fw} and h_{ref} are the head computed by the framework models and the reference model, respectively. Δh is then contoured; its characteristics evaluated against the underlying heterogeneity and global boundary condition. Note that the hydraulic head is used as a measure of discrepancy since framework models are commonly used to study regional pressure anomalies. In these studies, within-unit heterogeneity or conductivity trends are ignored [e.g., *Senger and Fogg*, 1987]. It is important to understand the sources of error in head prediction introduced by such models.

[30] To further compare head deviation among experiments whereby different head magnitude is calculated, a global mean error in hydraulic head can be calculated:

$$\mathrm{ME} = \frac{1}{n} \sum_{i=1}^{n} \left| h_{fw} - h_{\mathrm{ref}} \right| \tag{9}$$

Table 1. Equivalent Hydraulic Conductivity Estimated for Each Unit of the Hydrostratigraphic Model and the Lithostratigraphic Model[a]

Unit	Depositional Environment	$E[\ln(K)]$	$\mathrm{Var}[\ln(K)]$	K_{xx}	K_{zz}	K_{xz}	$K_{\max}$	$K_{\min}$	$K_{\max}/K_{\min}$
1	deepwater	2.10	0.73	17.18	7.01	−0.25	17.18	7.01	2.45
2	shoreline	4.73	1.48	188.96	61.36	−9.73	189.70	60.62	3.13
3	shoreline	5.34	1.15	294.35	104.04	−6.90	294.60	103.79	2.84
4	fluvial	5.68	0.31	330.53	256.31	1.82	330.58	256.27	1.29
5	fluvial	5.98	0.91	554.56	224.41	−6.38	554.68	224.28	2.47
6	turbidite	6.26	1.85	871.00	151.54	10.21	871.14	151.40	5.75
7	deepwater	2.07	0.83	19.16	7.17	0.53	19.18	7.15	2.68
8	shoreline	6.00	1.25	583.76	211.36	28.33	585.91	209.21	2.80
9	shoreline	5.44	1.13	321.95	120.82	−5.15	322.08	120.69	2.67
10	fluvial	6.29	0.37	591.62	536.97	3.92	591.90	536.69	1.10
11	fluvial/floodplain	4.53	1.10	140.41	69.51	1.16	140.43	69.49	2.02
12	fluvial/floodplain	5.68	0.62	388.44	213.41	0.08	388.44	213.41	1.82
13	deepwater	1.32	0.41	5.53	3.69	0.01	5.53	3.69	1.50
14	deepwater	2.02	0.59	11.66	6.49	0.15	11.66	6.48	1.80
1	aquitard	1.76	0.77	13.28	5.41	−0.03	13.28	5.41	2.46
2	aquifer	5.29	1.54	324.61	198.49	2.09	324.65	198.46	1.64

[a]The mean and variance of $\ln(K)$ are also listed as well as the principal components ($K_{\max}$, $K_{\min}$) and anisotropy ratio ($K_{\max}/K_{\min}$) of the equivalent conductivity. For the hydrostratigraphic model, the depositional environment for each unit is listed. All conductivities are in m/yr.

n is the number of nodes. The ME is normalized by the absolute head drop across the basin for each experiment. The resulting mean relative error (MRE) is an unbiased indicator of the relative head deviation for each experiment. It can be used to compare among experiments.

[31] A velocity-based metric can also be constructed, e.g., components of the elemental Darcy flux can be used to compute absolute or relative deviations. However, besides heterogeneity, boundary conditions are found to have significant impact on the accuracy of framework-model-predicted flow path (see section 4.2). Under certain boundary conditions and depending on which model is used, the same area can be a recharge or discharge zone, a laterally dominated or vertically dominated zone. Changes in groundwater flow direction can be locally important, but difficult to capture using global averages. A velocity-based metric may be more appropriate when the mean flow direction is the same in all models [e.g., *Scheibe and Yabusaki*, 1998]. In this study, to characterize the velocity field and basin-scale flow pattern, groundwater flow paths from the recharge toward discharge areas are computed, originating from the same locations in all models. To evaluate the location and rate of recharge or discharge, groundwater fluxes are computed along selected basin boundaries. A global flux metric is computed and relative error estimated.

3. Results

3.1. Upscaling

3.1.1. Comparison to Effective Conductivity

[32] An equivalent conductivity is estimated for each unit of the framework models using results of all four experiments by least squares solution: $\mathbf{K}_4^*$ (Table 1). The subscript "$_4$" indicates the number of experiments. The mean and variance of $\ln(K)$ are also listed for each unit. We note that (1) all equivalent conductivities are full tensors with the diagonal components dominating over the off-diagonal components, consistent with the small bedding angles in most deposits; (2) the relatively sand-rich units (e.g., fluvial, fluvial/floodplain, turbidite, shoreline deposits) have higher conductivity than the clay-rich units (e.g., deepwater deposits), as expected; and (3) the anisotropy ratio ($K_{\max}/K_{\min}$) of the equivalent conductivity ranges from 1.1 to 5.8, indicating higher permeability in the lateral direction, consistent with the observed stratification.

[33] The equivalent conductivity can be compared to an effective conductivity predicted by an analytic-stochastic theory. For example, in a stationary RSF with a constant mean head gradient, a well-known result of the effective conductivity for statistically anisotropic media is given [*Zhang*, 2002, p. 143, (3.155)]:

$$K_{ii}^{ef} = K_G\left[1 + \sigma_f^2(0.5 - F_i)\right]$$
$$F_i = \frac{1}{\sigma_f^2}\int \frac{k_i^2}{k^2} S_f(\mathbf{k}) d\ \mathbf{k} \tag{10}$$

where K_{ii}^{ef} are the components of the effective conductivity along the principal axes of $\ln(K)$ anisotropy, K_G is the geometric mean of the local conductivity, σ_f^2 is the variance of $\ln(K)$, $\mathbf{k} = (k_1, \ldots, k_d)^T$ is the wave number vector, d is the number of spatial dimensions, and S_f is the spectral density of $\ln(K)$. For $d = 2$, simple relations of F_i have been derived [*Gelhar and Axness*, 1983]:

$$F_1 = \frac{e}{1+e}; \qquad F_2 = \frac{1}{1+e} \tag{11}$$

where $e = \lambda_2/\lambda_1$ ($0 < e \leq 1$) is the statistical anisotropy ratio, λ_2 and λ_1 are the $\ln(K)$ integral scales along the minor and major statistical axes of correlation, respectively. Note that when $e = 1$, $K_{11}^{ef} = K_{22}^{ef} = K_G$, a well-known solution for statistically isotropic deposits. When $e \to 0$, the deposit is highly stratified: $K_{11}^{ef} \to K_G(1 + 0.5\sigma_f^2)$, $K_{22}^{ef} \to K_G(1 - 0.5\sigma_f^2)$.

[34] A prior geostatistical analysis on selected regions of the $\ln(K)$ map indicates that the principal statistical axes of $\ln(K)$ are aligned close to the global coordinate axes [*Zhang et al.*, 2005b]. Since the equivalent conductivity is diagonally dominant, it is reasonable to compare K_{11}^{ef} with $K_{\max}$, K_{22}^{ef} with $K_{\min}$. The principal components of $\mathbf{K}_4^*$ are plotted against $\ln(K)$ variance for each stratigraphic unit (Figure 4a). K_{11}^{ef} and K_{22}^{ef} predicted by equation (10) are also plotted, for e: 0–1. All conductivities are normalized by K_G.

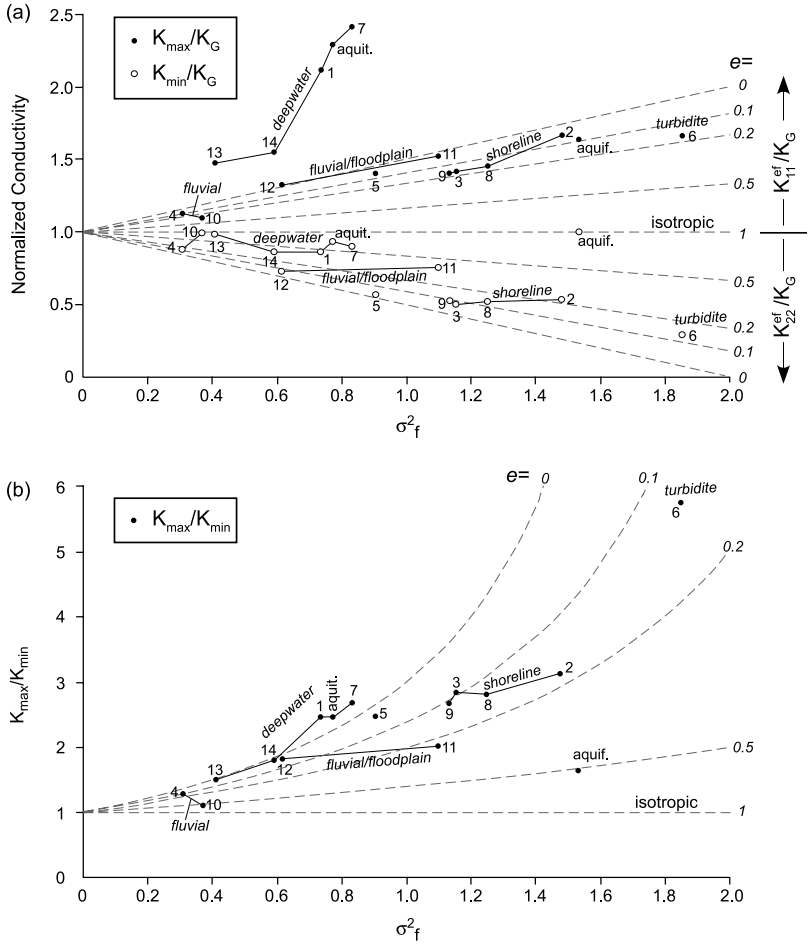

Figure 4. (a) Normalized principal components of $\mathbf{K}_4^*$ against σ_f^2 for each unit: $K_{\max}/K_G$, and $K_{\min}/K_G$ (dots). $K_{\min}/K_G$ for unit 10, 13, and aquifer is 0.992, 0.982, and 0.998, respectively. The analytical solution (K_{11}^{ef}/K_G, K_{22}^{ef}/K_G) is plotted for different e (dashed lines). (b) Anisotropy ratio of $\mathbf{K}_4^*$ against σ_f^2. The analytical K_{11}^{ef}/K_{22}^{ef} is also plotted.

The anisotropy ratio of the equivalent conductivity is then compared with the analytic prediction (Figure 4b).

[35] Several observations are made: (1) As ln(K) variance increases, the depositional environment changes from fluvial (4, 10), deepwater (13, 14, 1, 7), fluvial/floodplain (12, 11), shoreline (9, 3, 8, 2) to turbidite (6). Deposits created by similar sedimentary processes have similar characteristics in variance and $\mathbf{K}_4^*$. (2) For all units, $K_{\max} > K_G$, $K_{\min} < K_G$, reflecting statistical anisotropy due to within-unit stratification. With the exception of the aquifer unit, most sand-rich units lie within $e < 0.2$. Thus, for the given basin dimensions, an apparent average length/thickness ratio of the within-unit facies is greater than 5. The deepwater deposits are above the range predicted by equation (10), given $e \simeq 0.05$ estimated for deepwater subregions [*Zhang et al.*, 2005b]. Inclusions of minor sand lenses act to elevate the equivalent conductivity. (3) For the stratified deposits (units 12, 11, 9, 3, 8, 2, 6), the anisotropy ratio increases with variance (Figure 4b). In these units, higher variance results from higher contrast in sand/clay conductivity, thus higher anisotropy in the equivalent conductivity. (4) Compared to its constituent sand-rich units, the anisotropy ratio of the aquifer unit is relatively low, while its statistical anisotropy ratio is relatively high (Figure 4b). The exact reason for this is not clear. The aquifer unit is statistically inhomogeneous, e.g., unit 6 and 10 have higher mean conductivity than the other sand-rich units (Table 1). Direct comparison with the effective conductivity may be inappropriate.

[36] In comparing the equivalent and effective conductivity, it's important to note that the effective conductivity is defined for a RSF, or an ensemble of infinite realizations of small-scale conductivity. The equivalent conductivity is the property of a single conductivity field or one realization of the RSF. Strictly speaking, the comparison should be made between K_{ii}^{ef} and an ensemble average of the equivalent conductivity. This may be accomplished by creating a large number of basin-scale conductivity realizations (e.g., adding a small random component to each local conductivity). For each realization, a suite of equivalent conductivity can be obtained via upscaling. The ensemble conductivity is then computed by averaging results across all realizations. This, however, would require numerous flow experiments, beyond the scope of the current study.

3.1.2. Sensitivity to Boundary Condition

[37] To evaluate the sensitivity of the equivalent conductivity to boundary conditions, a $\mathbf{K}_2^*$ can be estimated using results from two flow experiments. For each unit, three sets of $\mathbf{K}_2^*$ are obtained based on experiments 1 and 2, 1 and 3,

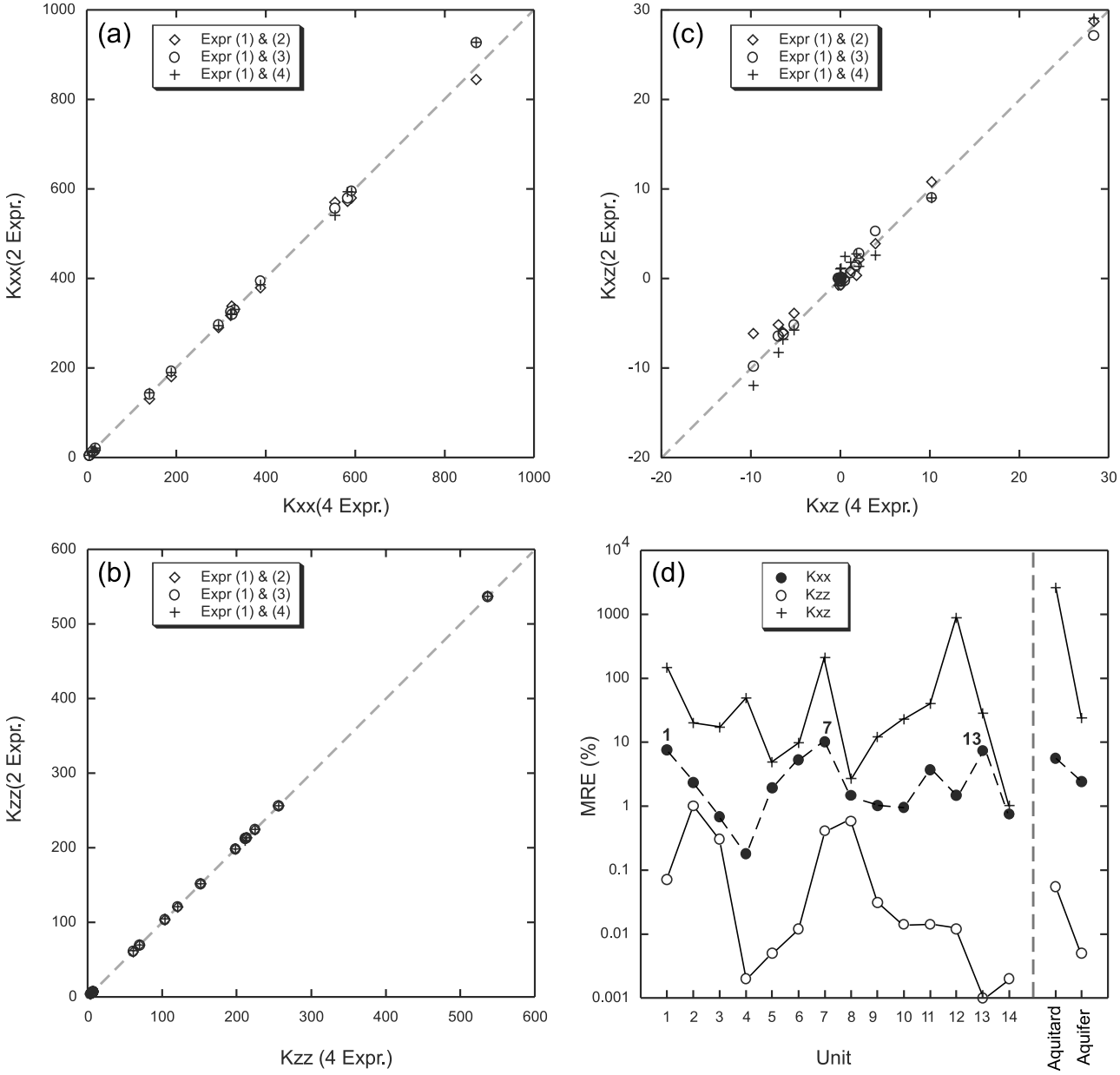

Figure 5. Components of $\mathbf{K}_2^*$ compared to those of $\mathbf{K}_4^*$: (a) K_{xx}; (b) K_{zz}; (c)K_{xz}. Conductivity is in m/yr. All 16 units of the framework models are included. (d) Mean relative error (MRE; in percent) of K_{xx}, K_{zz}, and K_{xz}.

and 1 and 4. In these cases, $\mathbf{K}_2^*$ is solved exactly (m = 2), and $\mathbf{K}_4^*$ provides the basis for comparison (Figure 5). Surprisingly, different boundary conditions do not produce appreciable trends in the deviation of $\mathbf{K}_2^*$ from $\mathbf{K}_4^*$. A mean relative error is also computed for each component of $\mathbf{K}_2^*$ against that of $\mathbf{K}_4^*$ (Figure 5d). Clearly, the largest MRE is associated with the off-diagonal terms; thus K_{xz} is the most sensitive to the boundary condition. K_{zz} has the least variation, likely due to the fact that K_{zz} is constrained by the results of experiment 1. The MRE for K_{xx} varies from 0.1% to 10%, in between that of K_{zz} and K_{xz}, with the largest deviation (MRE > 5%) found in the deepwater deposits, i.e., unit 1, 7, 13. In the lithostratigraphic model, the aquitard unit also has consistently higher MRE in all components of $\mathbf{K}_2^*$ than those of the aquifer unit. This suggests that the equivalent conductivity of the clay-rich deposits is more sensitive to boundary conditions. Overall, for the flow configurations investigated herein, the equivalent conductivity is not sensitive to the boundary conditions assigned, nor is it sensitive to the number of flow experiments performed.

3.2. Basin-Scale Flow Experiments

[38] In the second part of this study, the suite of $\mathbf{K}_4^*$ obtained from the up-scaling analysis is assigned to the appropriate stratigraphic units of the framework models (Figure 6). All conductivity maps are plotted with the same gray scale legend. Since the equivalent conductivity represents an ''averaged'' conductivity for each unit, similarity in gray scale is observed between the reference model K and framework model K_{xx}. The hydrostratigraphic model (HSM) preserves the higher mean conductivity in the upper basin,

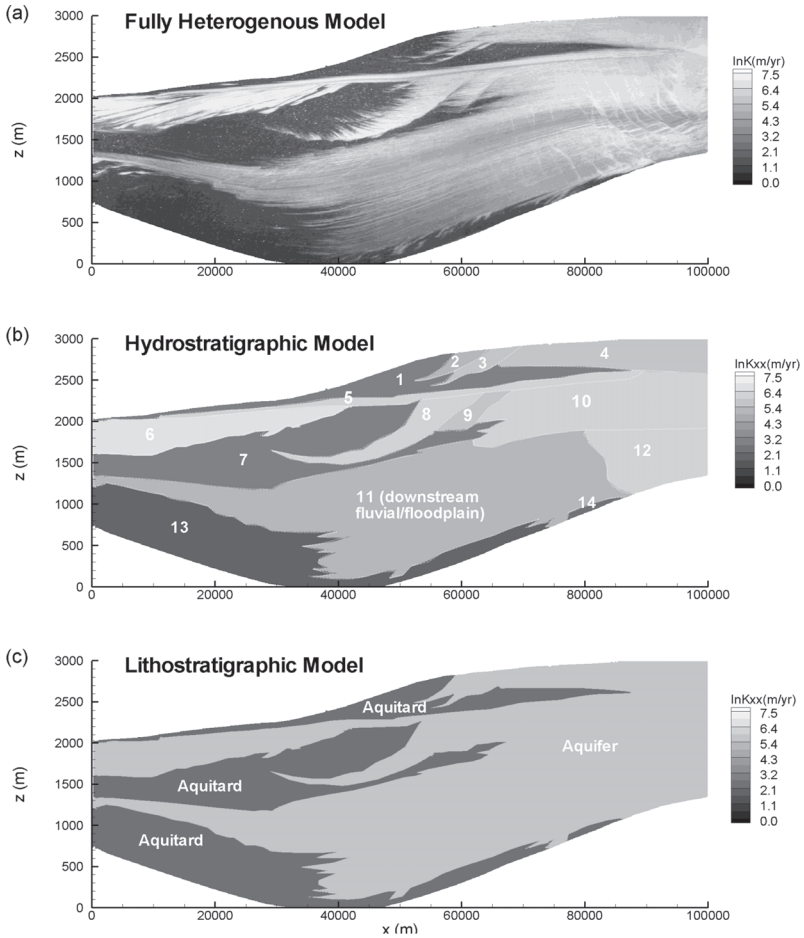

Figure 6. Three basin-scale conductivity representations: (a) a fully heterogenous model (local K), (b) a hydrostratigraphic model (K_{xx}) (with unit ID), and (c) a lithostratigraphic model (K_{xx}). The dotted lines in Figure 6b are added to emphasize the boundaries between units.

and, in the upstream, faulted fluvial/floodplain deposit of the lower basin. Such mean conductivity variation is smoothed out in the lithostratigraphic model (LSM). Note that alternatively, $\mathbf{K}_2^*$ can be assigned to each experiment, i.e., $\mathbf{K}_2^*$ by experiments 1 and 2, 1 and 2, 1 and 3 and 1 and 4 assigned to experiments 1, 2, 3, and 4, respectively. However, since the equivalent conductivity is not sensitive to the boundary condition, $\mathbf{K}_4^*$ is used for all experiments.

3.2.1. Experiment 1

[39] A vertical flow field is created with a mean downward flow direction. The steady state head distributions computed by all models observe similar features, e.g., smooth head contours, high gradient in low-K deposits (Figure 7). The smooth head contours predicted by the heterogeneous model are expected since the groundwater flow equation acts as a mathematical filter: the high wave number components (small scale within-unit heterogeneity) are filtered out. Significant head deviations ($|\Delta h| > 100$ m) are observed in large areas of the framework models (Figure 8). The LSM has a greater extent of head deviation than the HSM. The maximum and minimum Δh are also indicated, corresponding to the maximum head overestimation and underestimation in the basin, respectively. The LSM also has a greater range of head deviation ($\Delta h_{\max} - \Delta h_{\min}$).

[40] Within the downstream fluvial/floodplain deposits (unit 11 of HSM), the hydraulic head is significantly underestimated in both models, forming a symmetric deviation cell (Figure 8). The corresponding heterogenous deposit is characterized with a trend of decreasing conductivity with depth. This trend is smoothed out in the framework models: $\mathbf{K}_4^*$ assigned to unit 11 (HSM) or the aquifer unit (LSM) is lower than the mean conductivity of the upper portion of the heterogenous deposit, but higher than the mean conductivity of the lower portion of this deposit. Compared to the heterogeneous model, a higher head gradient is required for groundwater to enter this region; lower head gradient is required for groundwater to leave this region. Depending on the lateral extent of the deposit, a deviation cell forms. By similar reasoning, for the given flow direction, a downward increasing conductivity trend would result in head overestimation. In the upper basin, the turbidite (unit 6), deepwater (unit 1, 7), and upstream fluvial (unit 4) deposits have low head deviations. The corresponding heterogeneity contains no appreciable trends along the mean flow direction.

[41] To visualize the basin-scale flow pattern, a number of flow paths are placed in each framework models (Figure 9). The flow paths predicted by the heterogeneous model are plotted in the background. The flow paths predicted by all models are approximately vertical. The nonorthogonality observed in midbasin corresponds to the global facies transitions from the sand-rich fluvial deposits to the clay-rich deepwater deposits, e.g., units 7, 8, 11 of the HSM. The

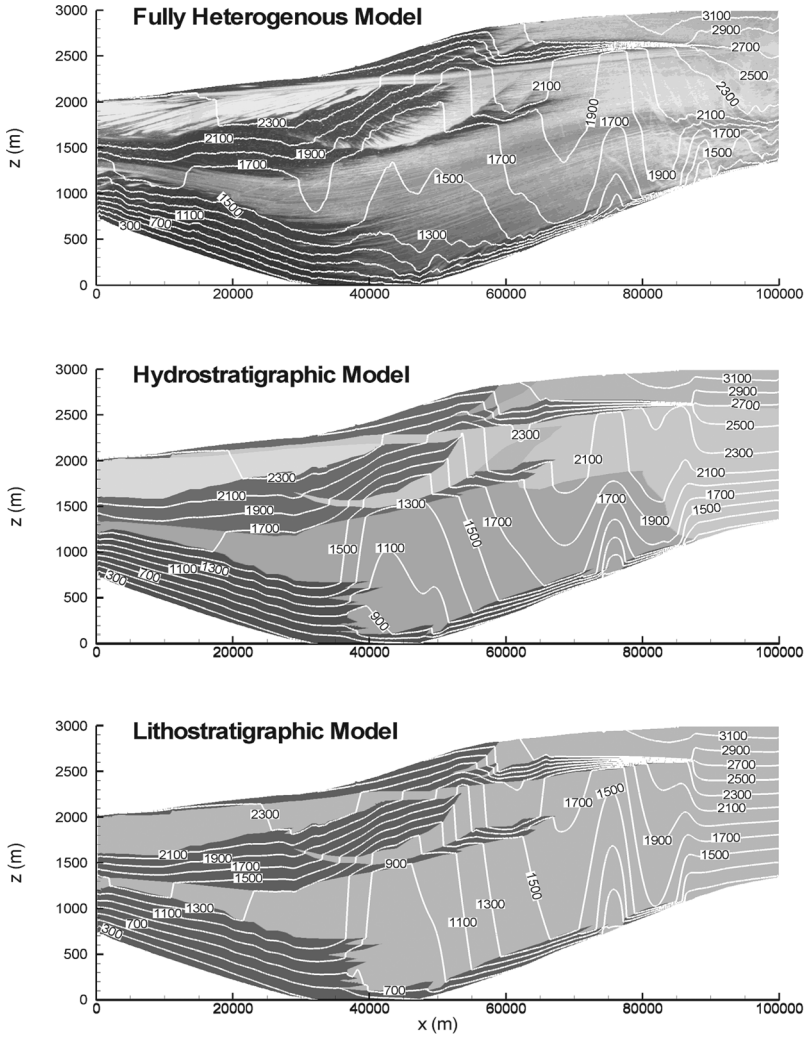

Figure 7. Hydraulic head (*m*) computed by all models in experiment 1, superimposed onto the respective conductivity field. The contour interval is 200 m.

within-unit stratification in the reference model and the anisotropy in the equivalent conductivity assigned to the framework models may also contribute. Overall, both framework models preserve the flow patterns in the basin. The maximum flow path deviation occurs in the lower basin, corresponding to the locations of the highest head deviation; the minimum flow path deviation occurs near the no-flow boundaries.

3.2.2. Experiment 2

[42] A lateral flow field is established in the basin. The head contours predicted by all models are again smooth. They are also straight, parallel to the sides of the basin where constant heads are assigned (not shown). Significant head deviations (although much smaller than those of experiment 1) are observed in the framework models; the deviation contours parallel the head contours (Figure 10). The HSM has negative deviation throughout the basin, indicating higher resistance to flow compared to the heterogenous model which contains bedding planes parallel to flow, contributing to lateral flow. In particular, the highest deviation develops in unit 6 of the HSM. The corresponding heterogeneity has the highest variance, thus large contrast in sand/clay conductivity. For the given flow direction, the sand-rich facies within this unit constitute preferential flow pathways.

[43] However, in the same region of the LSM ($x < 60$ km), positive deviation is observed, indicating overall less resistance to flow. Since the equivalent conductivity of the aquifer unit results from flow averaging across all sand-rich units of variable mean conductivity, in the downstream fluvial/floodplain region, the aquifer is more permeable than the corresponding heterogeneity. This region is also proportionally more important than the turbidite (in which the aquifer unit is less permeable), creating overall head overestimation. The general flow paths predicted by all models are again consistent (not shown): in the downstream basin where high-K deposits (units 8, 6, 11) are embedded within low-K deposits (units 1, 7, 13), flow channeling is predicted by all models. The largest flow path deviation occurs here too, corresponding to the largest head deviation.

3.2.3. Experiment 3

[44] Regional groundwater is driven by the topographic gradient along the top boundary and recharges from upland and discharges into lowland. The head contours predicted by all models are again smooth (not shown), similar to those of experiment 2. Although the same magnitude of head drop

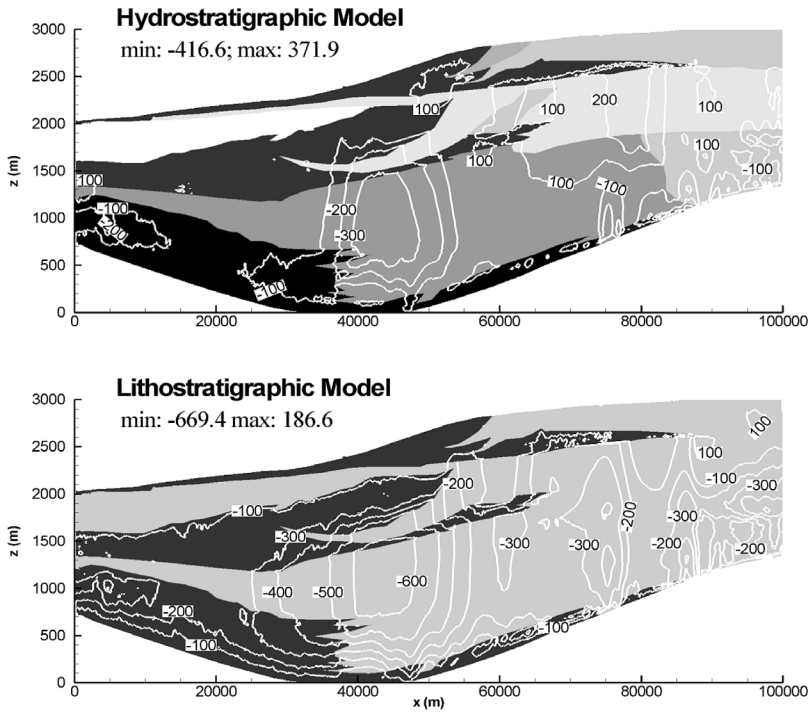

Figure 8. Head deviation (m) of the framework models in experiment 1.

is imposed along the basin length as in experiment 2, the magnitude, spatial extent, and extremes of the head deviation are much smaller (Figure 11). In experiment 3, most head deviation occurs in central basin, away from the sides where no-flow boundaries are specified. In particular, the significant underestimation (HSM) and overestimation (LSM) predicted in experiment 2 at the downstream basin ($x < 40$ km) has largely disappeared.

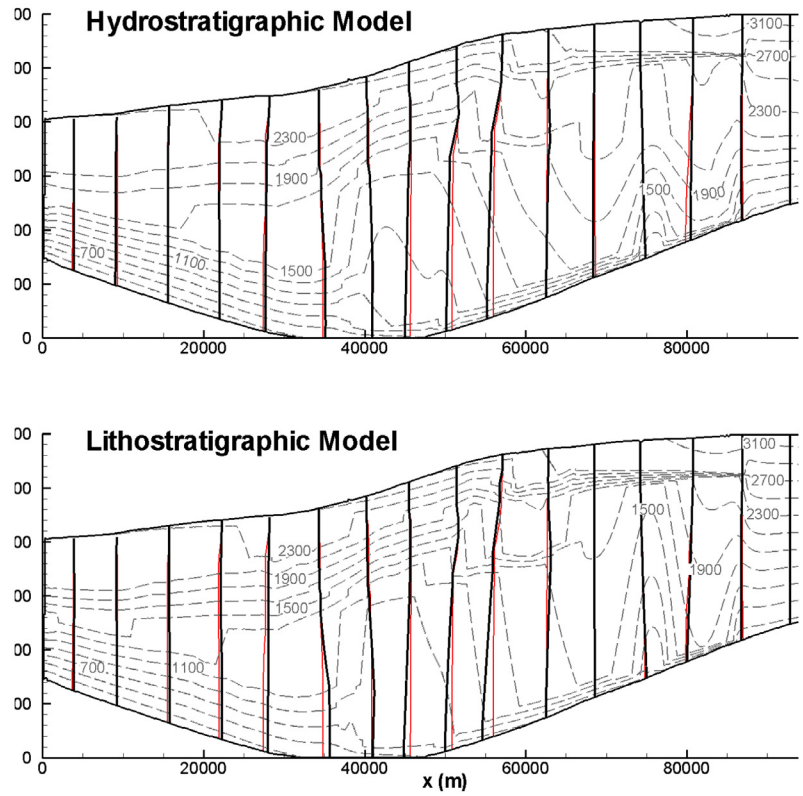

Figure 9. Flow paths predicted by the framework models in experiment 1, superimposed onto the head contours. Flow paths predicted by the heterogeneous model (red lines) are plotted in the background. The overall flow direction is downward.

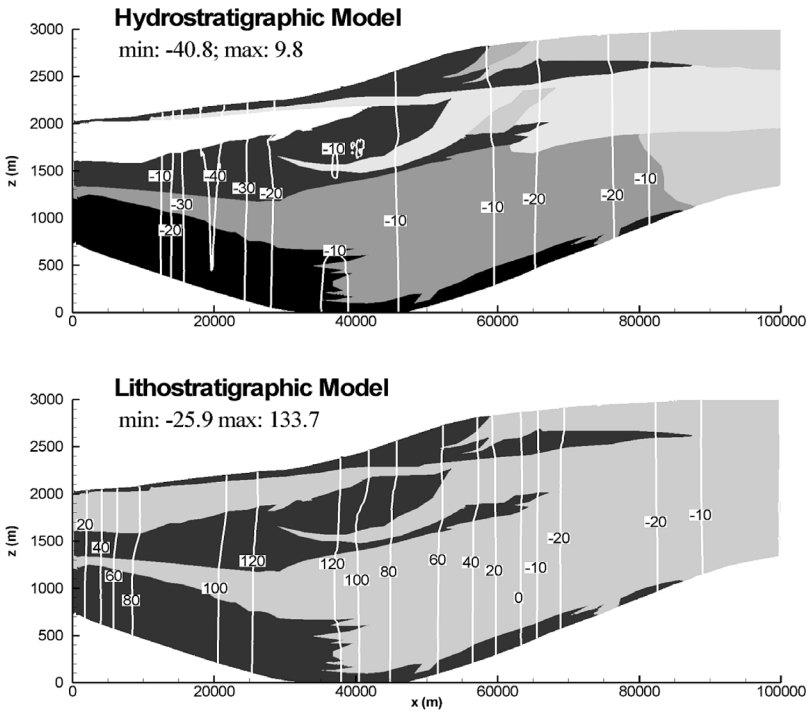

Figure 10. Head deviation (m) of the framework models in experiment 2.

[45] In terms of flow pattern, three recharge-discharge zones can be distinguished in the fully heterogenous model (Figure 12): (1) flow paths above s_1 correspond to a local circulation zone in the upland, to a depth of ~1 km. (2) Flow paths between s_1 and s_2 correspond to an intermediate zone: the recharge area is farther upstream, the discharge area is farther downstream. (3) Flow paths between s_2 and s_3 correspond to the deepest regional scale flow. Overall, the reference model predicts that groundwater flow paths are more concentrated in the upper basin, consistent with the water table gradient and the fact that hydraulic conductivity is generally decreasing with depth. The geometric alignment of the low-K unit 7 and high-K unit 8 forces flow to discharge toward the surface, also contributing to this flow pattern.

[46] In the HSM, only two recharge/discharge zones can be distinguished: (1) a shallow zone in the upper 1 km of the

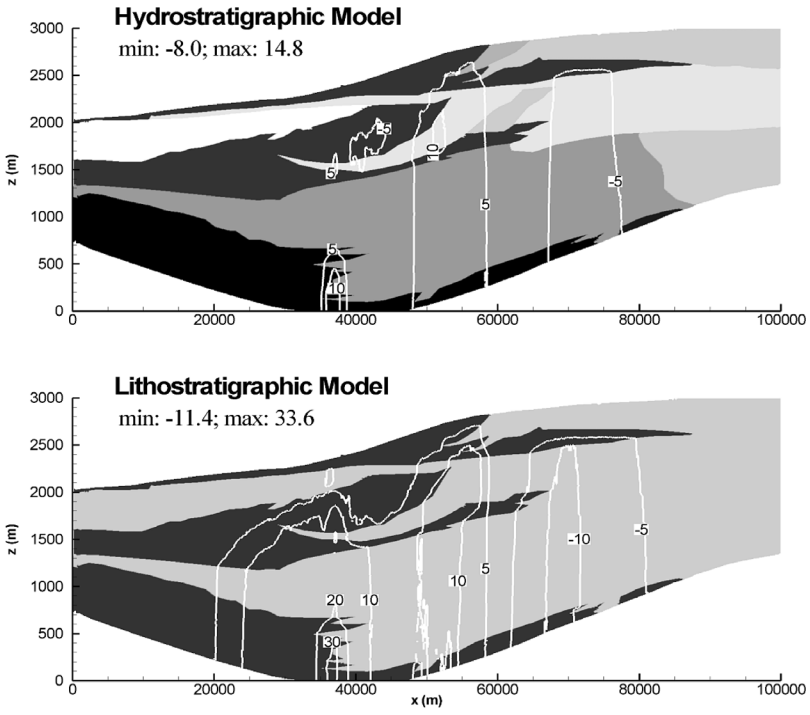

Figure 11. Head deviation (m) of the framework models in experiment 3.

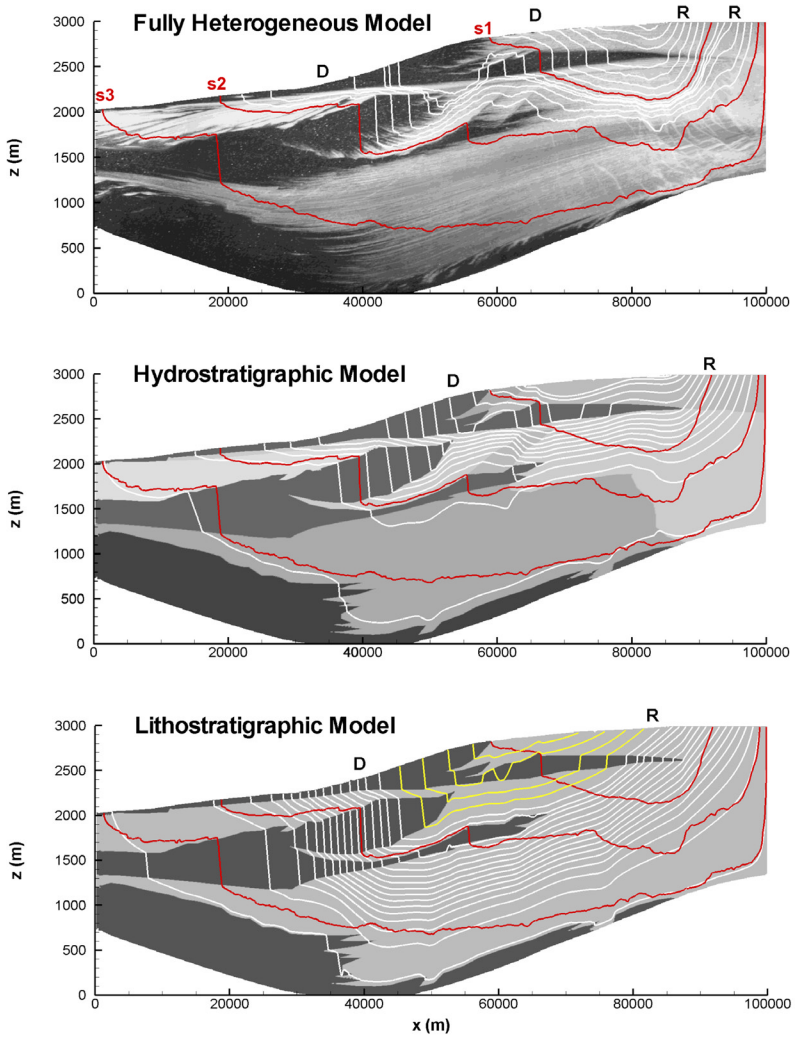

Figure 12. Flow paths predicted by all models in experiment 3, superimposed onto the respective conductivity field. Additional flow paths are plotted in the LSM (yellow lines). In the reference model, selected flow paths (red lines, s_1, s_2, s_3) divide the basin into three recharge (**R**) and discharge (**D**) zones. These flow paths are superimposed onto the framework models.

basin (its location corresponds to the top two zones of the reference model) and (2) a deeper zone of regional-scale flow (its location is approximately equivalent to the deepest zone of the reference model). However, in the lowermost unit 11, the equivalent conductivity is larger than the mean conductivity of the corresponding heterogenous deposit, s_3 is pulled down by ~500 m (a slight head overestimation is also observed; Figure 11). Although this model preserves the overall concentration of flow paths in the upper basin, the detailed flow paths in this region are significantly distorted.

[47] In the LSM, only one recharge/discharge zone is predicted: groundwater sweeps down nearly uniformly into the aquifer, reflecting the uniform equivalent conductivity assigned to this unit. Correspondingly, head is overestimated in the downstream fluvial/floodplain deposit, indicating easier passage for flow (Figure 11). The shallower, localized fluid circulation predicted by the other models has disappeared. Worse still, along an extensive area of the basin surface ($60 < x < 80$ km), discharge area predicted by the reference model is now a recharge area (yellow lines). Compared to experiments 1 and 2, despite the much smaller head deviation, considerable difference in groundwater flow pattern exists between the reference model and the framework models.

3.2.4. Experiment 4

[48] A pair of internal fluid source and sink is imposed on the topography-driven flow of experiment 3. The head contours predicted by all models are again smooth, and the magnitude and distribution of head deviation in the framework models are similar to those of experiment 3 (not shown). Since the well flow rate is sufficiently high, the basin-scale groundwater flow pattern is significantly modified (Figure 13). In all models, strategic flow paths are placed, originating from the surface (l_i and r_i lines) or the injection well (p_i lines). The overall flow pattern in the reference model can be divided based on the well location: two flow cells develop in the basin, one upstream from the injection well (marked by l_1), one downstream from the production well (marked by r_3). Between wells, there are discharge to the surface due to injection (p_i), recharge from the surface due to production (r_1, r_2), and direct flow (unmarked lines). The HSM predicts a slight fluid redistri-

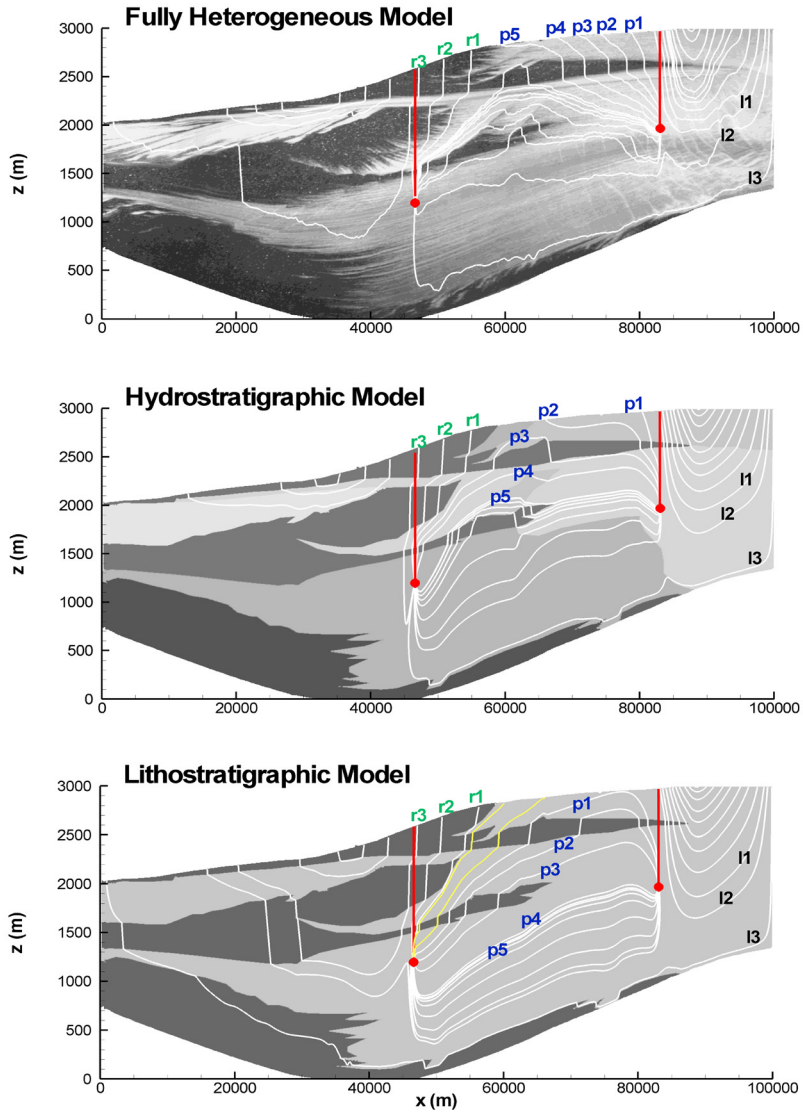

Figure 13. Flow paths predicted by all models in experiment 4. The l_1, l_2, l_3 originate in the upland basin; p_1–p_5 originate at the injection well; r_1, r_2, r_3 originate farther downstream.

bution in the basin, in particular, the injection-induced discharge flow paths (p_3, p_4, p_5) now reach the production well. The LSM is again the worse predictor, the discharge marked by p_i lines is nearly reversed, i.e., the yellow flow paths. Overall, the most significant change predicted by the framework models occurs in between wells.

[49] The above observation has implications for modeling fluid flow or transport in a basin impacted by well operations. For example, injection may be predicted by the framework models to be more efficient, i.e., low predicted leakage rate. For the given boundary condition and well rate, the framework-model-predicted target area is also deeper and spatially more extensive. In field-scale studies, this analogy may be extended to pump-and-treat operations. However, as correctly pointed out by the reviewer, the above observation may be quite sensitive to the depth of the well screen, thus may not be generalized to other situations. For example, in the case of a deep disposal well and a shallow domestic well, the characteristics of flow path predicted by the framework models may be different. A deep disposal well may be placed at the same location as the injection well (upstream basin) or the production well (downstream basin); a shallow well may be placed in unit 4 (regional recharge area) or unit 6 (regional discharge area). To fully evaluate these scenarios, a systematic study is needed, beyond the scope of the current study.

4. Discussions

4.1. Conductivity Upscaling

[50] The upscaling analysis indicates that with the exceptions of the deepwater deposits which contain sand lenses and the aquifer unit which contains deposits of variable mean conductivity, as the ln(K) variance increases, the equivalent conductivity for most deposits observes similar behavior as that predicted by the analytic-stochastic theory for an effective conductivity. This points to the possibility of obtaining the equivalent conductivity without the detailed flow simulations, if accurate statistical correlation ranges can be determined for each unit (i.e., "e" in Figure 4). For example, an effective conductivity for each unit of the HSM can be predicted by equation (10). The accuracy of this

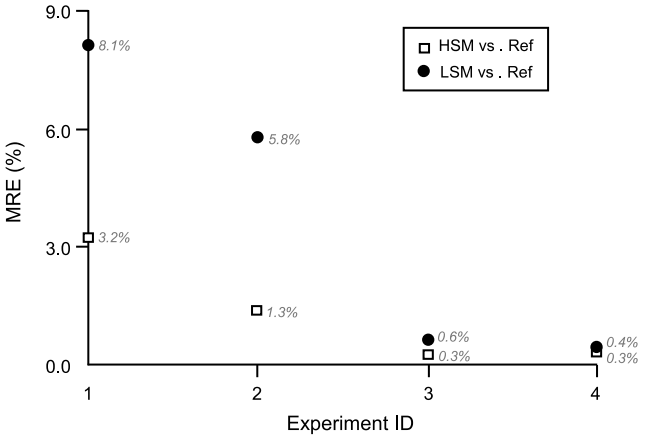

Figure 14. Global MRE of the hydraulic head predicted by the framework models for all experiments.

approach can be evaluated by comparing the flow predictions with those of the fully heterogeneous model. If the effective approach is proven useful in modeling the experimental basin, the methodology may be extended to natural systems. In that case, insights on the variability and geometry of within-unit facies can provide constraint for ln(K) variance and "e"; insights on the extent of the depositional environments can provide guidance on the development of a hydrogeologic framework model.

[51] For all hydrogeologic units and both upscaling scales (HSM versus LSM), the equivalent conductivity is insensitive to the boundary condition and the number of flow experiments. This is a significant observation. It points to the possible emergence of an effective conductivity that is independent of boundary condition. This may be a result of relatively large system scale compared to the extent of average within-unit stratification. However, future work is needed to estimate the principal correlation lengths for each unit. Moreover, such insensitivity also suggests that for such deposits, two sets of simulations may be sufficient to determine the equivalent conductivity, one ensuring a mean flow direction parallel to bedding, the other ensuring a mean flow direction perpendicular to bedding.

[52] The results further indicate that the off-diagonal term is the most sensitive to the boundary condition; however, such sensitivity should have little impact on flow since it is numerically insignificant compared to the diagonal components. To test this, additional simulations are conducted in the framework models by setting K_{xz} of $\mathbf{K}_4^*$ to 0.0. For both models and for all boundary conditions, the predictions of hydraulic head and groundwater flow paths with zero K_{xz} are extremely close to those using the full tensor (not shown), e.g., in experiment 1, the minimum, average, and maximum head deviation when K_{xz} is ignored is −11.9, 0.05, 6.1 m for the HSM and −2.3, −0.2, 6.8 m for the LSM, respectively, out of an average head drop of 2.8 km. Since most flow simulators for large-scale problems do not incorporate the off-diagonal terms, this suggests that the deviation introduced by such practice is minimum compared to the deviation as a result of nonrepresentation of within-unit heterogeneity, at least in the case when the off-diagonal term is 2 to 3 orders of magnitude smaller than the diagonal terms. This is the case in this study due to the small regional stratigraphic dip (∼1°). In situations where the bedding plane is more inclined, the significance of K_{xz} should increase. Future analysis may increase the basin depth-to-length ratio to investigate the critical bedding angle where K_{xz} becomes important.

4.2. Accuracy of Framework Models

[53] For all experiments, the LSM is less accurate than the HSM in predicting the hydraulic head and groundwater flow paths. This is expected since the HSM honors the mean conductivity variation among the depositional environments. The accuracy of the framework-model-predicted head is also affected by the boundary condition. For example, in unit 11, both framework models develop significant head deviation cells when the flow field is vertical (Figure 8). However, the same models, when subjected to different boundary conditions, predict very different pattern and magnitude of head deviation (e.g., Figures 10 and 11). This is also reflected in the global MRE of the head (Figure 14): experiments 3 and 4 have significantly smaller MRE than experiments 1 and 2. These two experiments are driven by a water table boundary with no-flow imposed on the sides and bottom, while experiments 1 and 2 only have two no-flow boundaries. Clearly, the extent of the no-flow boundary exerts an important control over the accuracy of head predicted by the framework models. The longer the no-flow boundary, the better the head prediction. Conceivably, if the extent of no-flow boundary is increased further, the MRE will approach 0.0.

[54] The accuracy of the framework-model-predicted groundwater flow pattern is also affected by boundary conditions. However, the extent of no-flow boundary has little impact in this respect: while the flow paths do not degrade appreciably for experiments 1 and 2, they degrade significantly for experiments 3 and 4. In experiment 3, the framework models fail to predict the flow pattern in the upper basin; in experiment 4, they fail to predict the flow pattern between wells.

[55] To further compare the recharge or discharge rate predicted by all models, a groundwater flux (Q; m^2/yr) is computed for every cell length along the recharge or discharge boundaries (Figure 15). For experiments 1, 3, and 4, the flux is computed along the top boundary; the cell length is 100 m. For experiment 2, the flux is computed along the right boundary; the cell length is 5 m. Q is computed via cross product of the Darcy flux with the element boundary, thus positive value indicates recharge to the basin, vice versa. A total flux is also computed along these boundaries; a relative error is estimated for each framework model. We observe the following: (1) The framework model profiles are much smoother compared to the reference model profiles. (2) The HSM profile is more closely following the mean variation of the reference profile. (3) In the HSM, the relative error in the total flux is −1.9%, 3.0%, −21.3%, and −13.5%, for experiments 1, 2, 3, and 4, respectively. Compared to experiments 1 and 2, experiments 3 and 4 have higher errors, consistent with the more severe degradation of the flow pattern. (4) In the topography-driven system, framework models locally predict opposite recharge or discharge to that of the reference model, with or without well operations, e.g., between the thin gray lines.

4.3. Implications in Basin-Scale Flow Modeling

[56] The results presented in this study based on an experimental stratigraphy have significant implications for basin-scale groundwater modeling studies. In the HSM,

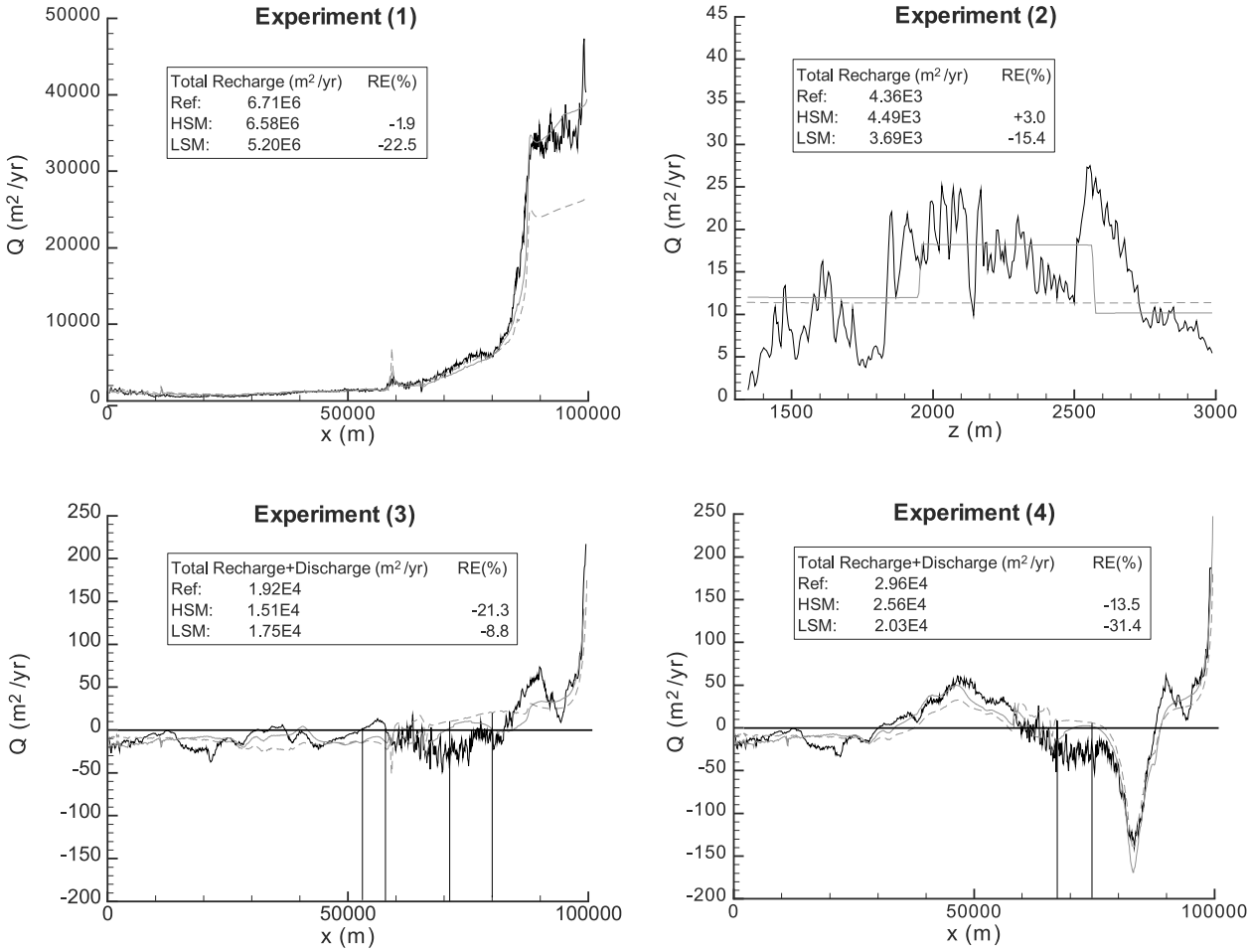

Figure 15. Groundwater flux across one cell length (Q; m^2/yr) along the recharge or recharge/discharge boundaries for all experiments: reference model (black), HSM (gray), and LSM (gray dashed). A total flux ($Q_{sum} = \sum|Q|$) is also computed; a relative error is calculated for each framework model: RE (%) = $(Q_{sum}^{fw} - Q_{sum}^{ref})/Q_{sum}^{ref} \times 100$.

even with the finer stratigraphic division and the fact that it uses an equivalent conductivity obtained via detailed numerical simulations, significant errors exist in the prediction of the hydraulic head and groundwater flow paths. For example, large head deviation can develop within a unit that contains trended deposits when the flow is parallel to the trend, or in a unit that contains stratified deposits when the flow is parallel to stratification. Thus, if a hydrogeologic unit contains a trended facies transition (as in a regressive sequence) or sand/shale juxtaposition, head prediction can be off by hundreds of meters under certain flow conditions, even if an accurate equivalent conductivity is obtained for the unit. Such facies transition is prevalent in fluvial, alluvial or intermountain basins where sediments are derived from mountain uplands and grain size decreases toward the basin interior. In such systems, if the flow direction is lateral (e.g., driven by high recharge rates in the mountains), accurate hydraulic head prediction will be impossible using the framework model approach.

[57] In attempting to fit the modeled head to the observed head, model calibration likely aggravates parameter estimation, since the appropriate comparison should be made between the modeled head and the head obtained if accurate, upscaled conductivity is known. The analysis conducted in this study indicates that for trended or stratified deposits, under certain boundary conditions, the "upscaled" head should deviate from the observed head. In such cases, using the observed head for calibration, the best fit conductivity will likely deviate significantly from the equivalent conductivity. Therefore, if head prediction is the goal of a study, e.g., in evaluating regional-scale pressure anomalies [e.g., *Senger and Fogg*, 1987; *Kaiser et al.*, 1994; *Bredehoeft et al.*, 1994], explicit incorporation of within-unit facies trend or juxtaposition is necessary.

[58] In the topography-driven system, the framework models predict progressive degradation in the basin-scale flow pattern when the level of stratigraphic division is reduced. Progressive displacement or even reversal of surface recharge or discharge are evident. As long as the framework model approach is used, the prediction of regional flow paths and the location/rate of recharge or discharge should be considered suspect. In this study, the conductivity map reflects an unconsolidated system, in mature sedimentary basins where conductivity contrast is larger, such effects may be more pronounced. Clearly, heterogeneity is important in predicting bulk flow characteristics. Using framework models to understand regional hydrodynamics can lead to significant errors.

[59] On a more positive note, as the number of stratigraphic division increases, the framework model should become more accurate in predicting head, flow path and recharge/discharge. For example, to improve the prediction of the HSM, unit 6 can be subdivided into sand-rich mass flow units and clay-rich suspended flow units. However, such division will be intractable if the unit contains numerous sand and clay facies, e.g., the trended unit 11. In such cases, geostatistical approaches may be more advantageous [e.g., *Eggleston et al.*, 1996; *Bierkens*, 1996; *Lu et al.*, 2002; *Zhou et al.*, 2003]. Future work may consider (1) increasing the level of stratigraphic division and (2) based on the current division, use geostatistical methods to generate within-unit conductivity heterogeneity, and if necessary, trended heterogeneity. Which approach is more efficient in improving model predictions of flow/transport will be of significant practical interest.

5. Conclusions

[60] A high-resolution, fully heterogeneous basin-scale hydraulic conductivity map is created by scaling up a stratigraphic image of an experimental deposit and assuming increasing conductivity for increasing gray scale (proxy for sand content). A fully heterogeneous model is created by incorporating the complete conductivity variation in the basin. On the basis of depositional environment, two hydrogeologic framework models are also created, one of coarser stratigraphic division. A novel numerical up-scaling method is developed to compute an equivalent conductivity for each irregularly shaped framework model unit by conducting basin-scale flow experiments in the fully heterogeneous model. In each experiment, distinctly different boundary conditions are specified, subjecting the basin to different flow conditions. This up-scaling methodology does not require any assumptions on the local conductivity correlation structure. It is applicable to any system scales as well as for a system with locally tensorial conductivity. It is also easily implementable with different numerical methods. Results indicate that for most deposits, the behavior of the equivalent conductivity with increasing $\ln(K)$ variance is consistent with that predicted by the analytic-stochastic theory for an effective conductivity. The equivalent conductivity is also insensitive to the boundary condition and the number of flow experiments conducted, indicating the possible emergence of an effective conductivity. Thus, for these deposits, two sets of flow experiments may be sufficient to compute a representative conductivity. Although all equivalent conductivities are full tensors, ignoring the small off-diagonal term introduces minimal error in the prediction of head and flow paths compared to that due to nonrepresentation of within-unit heterogeneity.

[61] To evaluate the impact of using equivalent conductivity on the prediction of basin-scale hydraulic head and groundwater flow, the experiments conducted in the fully heterogenous model (to obtain the equivalent conductivity) are repeated in the framework models. In each experiment, the hydraulic head and groundwater flow paths predicted by all models are compared. Results indicate that significant head deviation can develop within framework model units that contain (1) trended heterogeneity when flow direction is parallel to trend and (2) stratified heterogeneity when flow direction is parallel to stratification. The accuracy of head prediction is improved when the length of the no-flow boundary is increased. In the topography-driven flow system, progressive degradation occurs in the prediction of flow pattern, flow rate, and recharge/discharge location, when the progressively up-scale framework models are used. In summary, the accuracy of the framework models is controlled by the level of stratigraphic division, conductivity heterogeneity, and boundary conditions.

[62] It is important to note that results obtained in this study are specific to a two-dimensional system with a moderate conductivity range. In a three-dimensional system, hydraulic head and groundwater flow paths will be affected by three-dimensional structures, e.g., groundwater can flow around lenticular clay deposits.

[63] Besides what is pointed out in section 4, future work may consider a parameter estimation for the framework model units to address the nonuniqueness associated with model calibration. In light of the insights obtained in this study, a variety of basin-scale geological and environmental problems can be reevaluated [*Garven*, 1995]. Of particular interest is the impact of heterogeneity on groundwater age [e.g., *Goode*, 1996; *Park et al.*, 2002; *Bethke and Johnson*, 2002], salinity-driven flow [e.g., *Gupta and Bair*, 1997], and coupled flow and heat transfer [e.g., *Forster and Smith*, 1989; *Painter et al.*, 2003]. Note that basin-scale temperature anomalies often result from groundwater advection, temperature in addition to head data has been used to constrain formation permeability [*Deming*, 1993; *Bravo et al.*, 2002; *Zhang et al.*, 2005a]. Sand and clay have different thermal conductivities, thus both hydraulic and thermal heterogeneities exist. Although this study considers steady state flows, future work may consider transient flows subjected to temporal variations in head or recharge, as in paleohydrology and paleoclimate studies [e.g., *Senger et al.*, 1987; *Zhu*, 2000]. In these evaluations, the lower range of conductivity can be decreased to represent a more consolidated system (e.g., clay becomes shale). Finally, three-dimensional experimental stratigraphies are created based on digital reconstruction of two-dimensional images, for both sheet-dominated and channel-dominated systems. Groundwater flow, heat transfer and solute transport in the fully heterogenous three-dimensional deposits will be of significant interest.

[64] **Acknowledgments.** We wish to acknowledge three anonymous reviewers whose comments helped to improve the organization and clarity of this manuscript. The work was supported by Institute of Geophysics and Planetary Physics, Los Alamos National Laboratory.

References

Belitz, K., and J. D. Bredehoeft (1990), *Role of Confining Layers in Controlling Large-Scale Regional Groundwater Flow*, pp. 7–17, H. Heise, Hannover, Germany.

Bethke, C. M. (1985), A numerical model of compaction-driven groundwater flow and heat transfer and its application to the paleohydrology of intercratonic basins, *J. Geophys. Res.*, *90*, 6817–6828.

Bethke, C. M., and T. Johnson (2002), Paradox of groundwater age, *Geology*, *30*, 107–110.

Bethke, C. M., X. Zhao, and T. Torgersen (1999), Groundwater flow and the ^{4}He distribution in the Great Artesian Basin of Australia, *J. Geophys. Res.*, *104*, 12,999–13,011.

Bierkens, M. F. P. (1996), Modeling hydraulic conductivity of a complex confining layer at various spatial scales, *Water Resour. Res.*, *32*, 2369–2382.

Bravo, H. R., F. Jiang, and R. J. Hunt (2002), Using groundwater temperature data to constrain parameter estimation in a groundwater flow model of a wetland system, *Water Resour. Res.*, *38*(8), 1153, doi:10.1029/2000WR000172.

Bredehoeft, J., J. Wesley, and T. Fouch (1994), Simulations of the origin of fluid pressure, fracture generation, and the movement of fluids in the Uinta basin, *AAPG Bull.*, *78*, 1729–1747.

Castro, C., P. Goblet, E. Ledoux, S. Violette, and G. de Marsily (1998), Noble gases as natural tracers of water circulation in the Paris basin: 2. Calibration of a groundwater flow model using noble gas isotope data, *Water Resour. Res.*, *34*, 2467–2483.

Cazanacli, D., C. Paola, and G. Parker (2002), Experimental steep, braided flow: Application to flooding risk on fans, *J. Hydraul. Eng.*, *128*, 322–330.

Dagan, G. (1989), *Flow and Transport in Porous Formations*, Springer, New York.

Deming, D. (1993), Regional permeability estimates from investigations of coupled heat and groundwater flow, North Slope of Alaska, *J. Geophys. Res.*, *98*, 16,271–16,288.

Desbarats, A. (1992), Spatial averaging of hydraulic conductivity in three-dimensional heterogeneous porous media, *Math. Geol.*, *24*, 249–267.

Eggleston, J. R., S. A. Rojstaczer, and J. J. Peirce (1996), Identification of hydraulic conductivity structure in sand and gravel aquifers: Cape Cod data set, *Water Resour. Res*, *32*, 1209–1222.

Forster, C., and L. Smith (1989), The influence of groundwater flow on thermal regimes in mountainous terrain: A model study, *J. Geophys. Res.*, *94*, 9439–9451.

Garven, G. (1995), Continental-scale groundwater flow and geological processes, *Annu. Rev. Earth Planet. Sci.*, *23*, 89–117.

Garven, G., and R. Freeze (1984), Theoretical analysis of the role of groundwater flow in the genesis of stratabound ore deposits, 1. Mathematical and numerical model, *Am. J. Sci.*, *284*, 1085–1124.

Gelhar, L. (1993), *Stochastic Subsurface Hydrology*, Prentice-Hall, Upper Saddle River, N. J.

Gelhar, L., and C. Axness (1983), Three-dimensional stochastic analysis of macrodispersion in aquifers, *Water Resour. Res.*, *19*, 161–180.

Goode, D. (1996), Direct simulation of groundwater age, *Water Resour. Res.*, *32*, 289–296.

Gupta, E. N., and S. Bair (1997), Variable-density flow in the midcontinent basins and arches region of the United States, *Water Resour. Res.*, *33*, 1785–1802.

Heller, P., C. Paola, I. G. Hwang, B. John, and R. Steel (2001), Geomorphology and sequence stratigraphy due to slow and rapid base-level changes in an experimental subsiding basin (XES 96-1), *AAPG Bull.*, *85*, 817–838.

Indelman, P., and G. Dagan (1993), Upscaling of permeability of anisotropic heterogeneous formations: 1. The general framework, *Water Resour. Res.*, *29*, 917–924.

Kaiser, W., T. Swartz, and G. Hawkins (1994), Hydrologic framework of the Fruitland Formation, San Juan Basin, *N.M. Bur. Mines Miner. Resour. Bull.*, *146*, 133–163.

Lu, S., F. Molz, G. Fogg, and J. W. Castle (2002), Combining stochastic facies and fractal models for representing natural heterogeneity, *Hydrogeol. J.*, *10*, 475–482.

Painter, S., J. Winterle, and A. Armstrong (2003), Using temperature to test models of flow near Yucca Mountain, Nevada, *Ground Water*, *41*, 657–666.

Paola, C. (2000), Quantitative models of sedimentary basin filling, *Sedimentology*, *47*, 121–178.

Paola, C., et al. (2001), Experimental stratigraphy, *GSA Today*, *11*, 4–9.

Park, J., C. M. Bethke, T. Torgersen, and T. M. Johnson (2002), Transport modeling applied to the interpretation of groundwater ^{36}Cl age, *Water Resour. Res.*, *38*(5), 1043, doi:10.1029/2001WR000399.

Person, M., and G. Garven (1992), Hydrologic constraints on petroleum generation within continental rift basins: Theory and application to the Rhine graben, *AAPG Bull.*, *76*, 468–488.

Person, M., B. Dugan, J. Swenson, L. Urbano, C. Stott, J. Taylor, and M. Willett (2003), Pleistocene hydrogeology of the Atlantic continental shelf, New England, *Geol. Soc. Am. Bull.*, *115*, 1324–1343.

Renard, P., and G. de Marsily (1997), Calculating equivalent permeability: A review, *Adv. Water Resour.*, *20*, 253–278.

Sánchez-Vila, X., J. P. Girardi, and J. Carrera (1995), A synthesis of approaches to upscaling of hydraulic conductivities, *Water Resour. Res.*, *31*, 867–882.

Scheibe, D., and S. Yabusaki (1998), Scaling of flow and transport behavior in heterogeneous groundwater systems, *Adv. Water Resour.*, *22*, 223–238.

Scheibe, T. D., and D. L. Freyberg (1995), Use of sedimentological information for geometric simulation of natural porous media structure, *Water Resour. Res.*, *31*, 3259–3270.

Schneider, D. (1998), Tectonics in a sand box, *Sci. Am.*, *279*, 36.

Senger, R., and G. E. Fogg (1987), Regional underpressuring in deep brine aquifers, Palo Duro basin, Texas: 1. Effects of hydrostratigraphy and topography, *Water Resour. Res.*, *23*, 1481–1493.

Senger, R., C. Kreitler, and G. Fogg (1987), Regional underpressuring in deep brine aquifers, Palo Duro basin, Texas: 2. The effect of Cenozoic basin development, *Water Resour. Res.*, *23*, 1494–1504.

Sheets, B., T. Hickson, and C. Paola (2002), Assembling the stratigraphic record: Depositional patterns and time-scales in an experimental alluvial basin, *Basin Res.*, *14*, 287–301.

Strong, N., B. Sheets, T. Hickson, and C. Paola (2006), A mass-balance framework for quantifying downstream changes in fluvial architecture, *Sedimentology*, in press.

Tidwell, V. C., and J. L. Wilson (2002), Visual attributes of a rock and their relationship to permeability: A comparison of digital image and minipermeameter data, *Water Resour. Res.*, *38*(11), 1261, doi:10.1029/2001WR000932.

Walvoord, M. A., P. Pegram, F. M. Phillips, M. Person, T. L. Kieft, J. K. Fredrickson, J. P. McKinley, and J. B. Swenson (1999), Groundwater flow and geochemistry in the southeastern San Juan Basin: Implications for microbial transport and activity, *Water Resour. Res.*, *35*, 1409–1424.

Wen, X.-H., and J. Gómez-Hernández (1996), Upscaling hydraulic conductivities in heterogeneous media: An overview, *J. Hydrol.*, *183*, ix–xxxii.

Zhang, D. (2002), *Stochastic Methods for Flow in Porous Media, Coping With Uncertainties*, Elsevier, New York.

Zhang, Y., M. Person, E. Merino, and M. Szpakiewcz (2005a), Evaluation of soluble benzene migration in the Uinta basin, *Geofluids*, *5*, 106–123.

Zhang, Y., M. Person, C. Paola, C. W. Gable, X. Wen, and J. M. Davis (2005b), Geostatistical analysis of an experimental stratigraphy, *Water Resour. Res.*, *41*, W11416, doi:10.1029/2004WR003756.

Zhou, Q., H. Liu, G. S. Bodvarsson, and C. M. Oldenburg (2003), Flow and transport in unsaturated fractured rock: Effects of multiscale heterogeneity of hydrogeologic properties, *J. Contam. Hydrol.*, *60*, 1–30.

Zhu, C. (2000), Estimate of recharge from radiocarbon dating of groundwater and numerical flow and transport modeling, *Water Resour. Res.*, *36*, 2607–2620.

C. W. Gable, EES-6, MS T003, Los Alamos National Laboratory, Los Alamos, NM 87545, USA.

M. Person, Department of Geological Sciences, Indiana University, 1001 East Tenth Street, Bloomington, IN 47405, USA.

Y. Zhang, Department of Geological Sciences, University of Michigan, 2534 CC Little Bldg, 1100 North University Avenue, Ann Arbor, MI 48109, USA. (ylzhang@umich.edu)