Validity of flowmeter data in heterogeneous alluvial aquifers

Marco Bianchi[1]

[1]British Geological Survey, Environmental Science Centre, Keyworth, Nottingham, NG12 5GG, United Kingdom. marcob@bgs.ac.uk. Tel. +44 (0)155 9363136

Abstract

Numerical simulations are performed to evaluate the impact of medium-scale sedimentary architecture and small-scale heterogeneity on the validity of the borehole flowmeter test, a widely used method for measuring hydraulic conductivity (*K*) at the scale required for detailed groundwater flow and solute transport simulations. Reference data from synthetic *K* fields representing the range of structures and small-scale heterogeneity typically observed in alluvial systems are compared with estimated values from numerical simulations of flowmeter tests. Systematic errors inherent in the flowmeter *K* estimates are significant when the reference *K* field structure deviates from the hypothetical perfectly stratified conceptual model at the basis of the interpretation method of flowmeter tests. Because of these errors, the true variability of the *K* field is underestimated and the distributions of the reference *K* data and log-transformed spatial increments are also misconstrued. The presented numerical analysis shows that the validity of flowmeter based *K* data depends on measureable parameters defining the architecture of the hydrofacies, the conductivity contrasts between the hydrofacies and the sub-facies-scale *K* variability. A preliminary geological characterization is therefore essential for evaluating the optimal approach for accurate *K* field characterization.

Keywords

Hydraulic conductivity; Numerical simulation; Characterization methods; Hydrofacies; Borehole flowmeter test

1 Introduction and background

Groundwater flow and solute transport are heavily influenced by the spatial distribution of hydraulic conductivity (K) in the subsurface. Therefore, an accurate characterization of small-scale ($<10^0$ m) K variations is essential for numerical simulations at the typical scale of contaminated land remediation projects, which is in the order of few hundreds of meters. Among the available methods for measuring K at such fine scale in porous aquifers, the borehole flowmeter test is one of the most commonly applied (e.g., Maliva, 2016). It was also the preferred technique for the characterization of the K field at a number of hydrogeological research sites (Hess, 1989; Molz et al., 1989; Rehfeldt et al., 1992; Young, 1995; Vereecken et al., 2000; Riva et al. 2006; Li et al., 2008), and extensive flowmeter based K datasets have been used in myriad published studies in the fields of geostatistics, stochastic subsurface hydrology, and groundwater flow and solute transport modelling.

The borehole flowmeter test consists of measuring variations of vertical flow in a fully penetrating well and subsequently relating these variations to the vertical distribution of K along the well. The test is usually performed by pumping groundwater at a constant rate Q_w. After pseudo-steady state conditions are attained, a type of flowmeter (Hufschmied, 1986; Molz. et al., 1989; Rehfeldt et al., 1989; Young, 1995; Young and Pearson, 1995; Boman et al., 1997; Crisman et al., 2001; Newhouse et al., 2005) is lowered in the well and flow measurements are taken at regularly spaced intervals starting from near the bottom of the well (flow ≈ 0) to the top (flow $\approx Q_w$).

The key assumption for the interpretation of flowmeter data is the conceptualization of the aquifer as a perfectly stratified system (Javandel and Witherspoon, 1969). In these conditions, which in field applications are assumed to be obtained after a period of time from the beginning of pumping (i.e., pseudo-steady state conditions attained), flow in the layers is horizontal and proportional to the hydraulic conductivity of the layer. By assuming a perfectly stratified aquifer with layers of thickness Δz_i equal to the spacing between flowmeter readings, the hydraulic conductivity K_i of a layer i can then be calculated according to (Molz et al., 1989):

$$K_i = \frac{\Delta Q_i Z}{Q_w \Delta z_i} \overline{K} \qquad (1)$$

where ΔQ_i is the discharge from layer i into the well ($Q_w = \sum_i \Delta Q_i$), Z is the aquifer thickness (i.e., $Z = \sum_i \Delta z_i$), and $\overline{K}$ is an effective hydraulic conductivity value (e.g., Sanchez-Vila et al., 2006), which can be obtained from a previously performed pumping test (e.g., Molz et al., 1989; Barahona-Palomo et al., 2011; Gueting et al., 2015). An alternative interpretation approach of flowmeter data is the one proposed by Rehfeldt et al. (1989), which was used for the characterization of the K field at the Macrodispersion Experiment (MADE) site (Rehfeldt et al., 1992; Zheng et al., 2011). This approach applies an approximate solution (Cooper and Jacob, 1946) for calculating the drawdown s_i in each layer of an assumed perfectly stratified aquifer. Comparisons of K profiles calculated from flowmeter data using the approach represented by Equation 1 and the approach proposed by Rehfeldt et al. (1989) show a fairly good agreement in a relatively homogenous aquifer (Molz et al., 1989). Vertical distribution of K values based on flowmeter tests and other characterization methods have been compared in several previous studies (e.g., Molz et al. 1989; Boman et al. 1997; Whittaker and Teutsch, 1999; Zlotnik and Zurbuchen, 2003b; Butler, 2005; Tilmann et al., 2008; Illman et al. 2010; Barahona-Palomo, 2011; Bohling et al., 2012; Gueting et al., 2015; Bianchi and Zheng, 2016).

Irrespective of the approach used for interpretation, flowmeter K estimates are prone to errors from different sources. For instance, the presence of a skin region or gravel pack around the well induces vertical flow components producing biased measurements of ΔQ_i (Rehfeldt et al., 1989; Young, 1995; Boman et al., 1997). These components can negatively impact the accuracy of the K estimates especially in the presence of a skin region (Xiang, 1995; Ruud and Kabala, 1997; Riva et al., 2012). If ignored, hydraulic head losses can also generate errors in the K estimates (e.g., Rehfeldt et al., 1989) caused by unaccounted for vertical flow components (Boman et al., 1997; Dinwiddie et al., 1999; Ruud et al., 1999; Zlotnik and Zurbuchen, 2003a). It was estimated that the combined effect of hydraulic head losses and skin effects can lead to estimation errors of a factor 10 or higher (Ruud et al., 1999). Studies have evaluated the impact of variations of the borehole diameter on flowmeter K estimates (Paillet, 2004), and tested the validity of the Theis (1935) model to represent the flow dynamics in the aquifer. A this regard, it was shown numerically that the assumption of the Theis model produces systematic errors up to factor of about 1.5 (Ruud and Kabala, 1996). Additional sources of random and systematic errors are discussed in detail by Rehfeld et al. (1989).

To the best of our knowledge, no previous study has been conducted to evaluate systematic errors in flowmeter data when the assumed model of a perfectly stratified system is not representative of the actual aquifer structure. In alluvial aquifers, for instance, variations of the textural properties of the sediments and consequentially of K can occur over short lengths, not only in the vertical direction (in the order of $10^{-2} – 10^{-1}$ meters), but also in the horizontal direction (in the order of $10^{0} – 10^{1}$ meters) (e.g., Scheibe and and Freyberg 1995, Fogg et al., 1998; Heinz et al., 2003; Ramanathan et al., 2010; Bayer et al., 2011; Dell'Arciprete et al., 2012; Bianchi and Zheng, 2016). A rigorous assessment of these systematic errors is also needed because datasets of flowmeter *K* estimates from multiple wells are often considered as conditioning points in geostatistical analysis to generate three-dimensional representations of the *K* field (e.g., Rehfeldt et al., 1992; Salomon et al., 2007, Li et al., 2008; Bianchi et al., 2011). The reliability of these representations, which most often contradict the assumption of a perfectly stratified system, depends on the representativeness of the conditioning data.

Building on previous analyses of other factors affecting the validity of the flowmeter test, the main objective of this paper is to evaluate the effect of medium-scale sedimentary architecture and small-scale variability on the accuracy of the *K* estimates. As in the work of Zhang et al. (2013), a transition probability based Markov chain model is used here to represent the spatial distribution of hydrofacies in a generic alluvial aquifer, and several scenarios representing a wide range of realistic geological conditions are considered in numerical simulations of flowmeter tests. Simulated tests are then validated by comparing resulting *K* estimates with the "true" *K* values of the hydrofacies. According to the assumed conceptual model of aquifer heterogeneity, the structure of the reference *K* field is defined by the architecture of the hydrofacies, while the univariate statistical distribution of *K* is controlled by their volumetric proportions as well as by the smaller scale variability within each hydrofacies. A similar geologically consistent conceptualization of the *K* field has been applied in several previous studies showing the strong influence of the sedimentological architecture on groundwater flow and solute transport (e.g., Anderson, 1989; Webb and Anderson, 1996; Fogg, 1986; Allen-King, 1998; Fogg et al., 1998, 2000; Weissman and Fogg, 1999; Weismann et al., 1999; Lu et al., 2002; Lee et al., 2007; Riva et al., 2008; Bianchi and Zheng, 2016).

This work is also motivated by the absence of a clear explanation for the generally poor correspondence between *K* estimates from flowmeter tests and from grain-size analysis of aquifer samples (Boggs et al. 1990; Boman et al., 1997; Whitteker and Teutsch, 1999; Barahona-Palomo et al., 2011; Gueting et al., 2015; Bianchi and Zheng, 2016). Barahona-Palomo et al. (2011), for instance, compared these two types of K data at coinciding locations in an alluvial aquifer and found different values for the sample variances and correlation lengths for the two datasets. These discrepancies have been previously related to differences in support volume between the two types of measurements. However, while the support volume of estimates based on the grain-size approach can be precisely associated with the dimensions of the aquifer samples, the support volume of flowmeter based K estimates is not clearly defined (Zlotnick and Zurbuchen, 2003b). The poor correspondence between these two types of estimates might also be caused by systematic errors in the grain-size based values due to the empirical nature of the relationship between a not precisely defined effective grain diameter and K. However, modelling studies by Riva et al. (2008) and Bianchi and Zheng (2016) have shown that grain-size based K estimates are able to provide adequate information to effectively simulate transport behaviour in heterogeneous medium- to coarse-grained alluvial aquifers. The results of these studies can be considered as an indirect validation of this approach for K estimation. On the other hand, representations of the K field based on flowmeter K data do not always explain observed tracer plumes (e.g., Feehley et al., 2000; Salamon et al, 2007; Fiori et al., 2013).

2 Methods

2.1 Simulated scenarios of facies architecture

Hydrofacies architecture in a generic alluvial aquifer is simulated with the transition probability geostatistical approach (T-PROGS; Carle and Fogg, 1996, 1997; Carle, 1999). With this approach, the spatial structure of a categorical variable – here indicating different hydrofacies – is modelled with a Markov Chain model based on transition probabilities between the categories at discrete lag distances. Markov chain models can be implemented by knowing the entries in the embedded transition probability matrix, which represents the conditional probabilities of one category to occur adjacent to the others along a particular direction. The choice of a background category is also required to eliminate the need to specify corresponding entries in the embedded transition matrix since these are calculated by difference from the entries of the other categories. Implemented Markov chain models are used by a geostatistical simulation algorithm to generate stochastic realizations of the spatial distribution of the categories in either conditional or unconditional frameworks. In these realizations, the background category fills the space not occupied by the other categories.

The transition probability approach allows the generation of realistic representations of the hydrofacies architecture on the basis of a combination of measurable and interpreted geological properties including (1) the number of hydrofacies, (2) their volumetric proportions, (3) their mean lengths along specific directions, (4) the anisotropy ratios between mean lengths, and (5) the juxtapositional tendency. The latter defines the predisposition for each hydrofacies to occur adjacent to others like, for instance, when there are trends in the textural properties of the sediments along particular directions (e.g., fining upward sequences). In this work, a range of values are assigned to some of these properties to generate several realistic scenarios of aquifer heterogeneity in which to evaluate the validity of simulated flowmeter tests. However, to better fit the purpose of this study, other properties including the number of lithofacies, their volumetric proportions, and the juxtapositional tendency are kept constant in all the scenarios. The motivation of this choice will become clear after the description of the reference K fields.

Three hydrofacies are considered to represent the type of deposits typically encountered in alluvial depositional systems, including a coarser and generally more conductive hydrofacies representing, for example, fluvial channel deposits (hydrofacies “H” or High), a finer and less conductive hydrofacies representing floodplain deposits (hydrofacies “L” or “Low”), and a third hydrofacies with intermediate properties, which is meant to represents deposits of other depositional elements such as levees and crevasse splays (hydrofacies “I” or “Intermediate”). It is assumed that hydrofacies “L”, which is also considered as the background property in the Markov chain models, has the highest volumetric proportion (0.5), while hydrofacies “I” and “H” have proportions equal to 0.2 and 0.3, respectively. In the development of the Markov chain models, juxtapositional tendencies are defined by the values in the embedded transition probability matrix in each direction (Carle and Fogg, 1988; Carle, 1999). In all the scenarios, assigned embedded probabilities were chosen to favour a vertical arrangement of the hydrofacies in fining-upward successions, which are typical of fluvial deposits (Miall, 2014). Vertical embedded probabilities were also assigned to embedded probabilities in the horizontal directions (x and y), in accordance to Walther’s Law (Fogg et al., 2000).

For each developed Markov chain model, one realization of the spatial distribution of the hydrofacies is assumed as representative of the architecture for each scenario. These realizations consider a rectangular domain of dimensions equal to 241 units (u_h = 0.5 m) in the x and y directions, and 80 units ($u_v = 0.5u_h$) in the vertical (z) direction. Each cell of the 3-D grid used for geostatistical simulation has a volume of $1\ u_h \times 1\ u_h \times 1\ u_v$ for a total of more than 4.6×10^6 cells.

Different scenarios are grouped on the basis of variations of a particular property (Table 1). The first group (scenarios L_1 through L_inf) is represented by eight scenarios characterized by different isotropic horizontal mean lengths of the hydrofacies ($L_h = L_x = L_y$). For simplicity, the same L_h is assigned to hydrofacies “I” and “H”, while the mean length of the background hydrofacies “L” varies in each scenario as a function of L_h and the embedded transition probabilities of the other hydrofacies (see Carle and Fogg, 1997 for details). For the scenarios considered in this study, the horizontal mean length of the background hydrofacies is about twice the mean length of the other hydrofacies. A scenario representing a perfectly stratified architecture ($L_h = \infty$) is also included in this group. In this work, scenarios considering a perfectly stratified system represent benchmark problems to test the validity of the numerical simulations. After the first unconditional realization of hydrofacies architecture was generated considering $L_h = 1\ u_h$ (scenario L_1), the resulting vertical distribution of hydrofacies at the centre of the domain was used as conditioning data in the remaining simulations such that the vertical succession of hydrofacies at this location – which corresponds to the location of the simulated flowmeter tests – is shared by all the scenarios.

The second group of scenarios (scenarios AR_1 through AR_16) was designed to test the impact of the anisotropy ratio between the horizontal mean lengths of the lithofacies. Different hydrofacies architectures were generated from modifications of an isotropic Markov chain model with $L_h = 4\ u_h$. For each modification, the mean length in the y direction (L_y) is progressively augmented while the values of L_x and L_z remain constant. Similarly to the previous group of scenarios, the vertical succession at the centre of the domain is shared by all the scenarios.

The remaining groups in Table 1 allow the analysis of the effects of variations in K between hydrofacies (scenarios L4_K1 through L16_K4), as well as within the hydrofacies (scenarios L2_εde00 through Linf_ εde80). The first of these groups considers scenarios with similar architecture, but different K contrasts between hydrofacies "H" and "L". The effect of different architectures is also investigated, corresponding to different isotropic horizontal mean lengths ($L_h = 4\ u_h$ and $L_h = 16\ u_h$). For each L_h value, four different K ratios are considered with values ranging from 10^0 to 10^4. Different levels of intrafacies K variability, which are evaluated in the last group of scenarios, are also considered for different architectures (i.e., L_h equal to 2 u_h and infinity). The approach for simulating this small-scale variability is described in the following section.

2.2 Reference hydraulic conductivity fields

The link between simulated hydrofacies architectures and corresponding K fields is established through the relationship between hydraulic conductivity and an effective grain-size diameter d_e. In a general form, this relationship can be written as (e.g., Vuković and Soro 1992):

$$K = C\frac{g}{\nu}f(\phi)d_e^2 \qquad (2)$$

where g is the gravitational constant, ν is the kinematic viscosity, $f(\phi)$ is a function of the porosity, and C is proportionality constant expressed in consistent units. With this approach, characteristic d_e values were initially assigned to the hydrofacies and reference K values were then mapped on the basis of the simulated architectures. For the calculation of the characteristic K values of the hydrofacies, the largest and the lowest d_e values are assigned to hydrofacies "H" and to hydrofacies "L", respectively. A similar porosity value is assumed for all the hydrofacies (i.e., $f(\phi) =$ constant)

In previous studies, non-uniform K distributions within the hydrofacies have been simulated by combining stochastic facies and fractal models (Lu et al., 2002; Zhang et al., 2013), or by imposing log-normal K distributions (Frei al., 2009). Both of these approaches make assumptions on the K distribution within the hydrofacies. In this work, small-scale variability is simulated with a different approach based on modelling the non-uniform distribution of the effective diameter d_e within the hydrofacies. For each hydrofacies, the d_e distribution is modelled with a spatially correlated Gaussian field, with mean equal to $\overline{d_e}$ and standard deviation σ defined as:

$$\sigma = \frac{\varepsilon_{de}\overline{d_e}}{3} \quad (3)$$

Spatial distributions of d_e for each hydrofacies were generated with a sequential Gaussian simulation code (Deutsch and Journel, 1998) by assuming an anisotropic exponential variogram model with ranges equal to 2 u_h, 2 u_h and 1 u_v in the x, y and z directions, respectively. Different levels of intrafacies variability are generated by varying the value assigned to the parameter ε_{de}, which controls the spreading of the Gaussian distribution of d_e within each hydrofacies. When, for example, $\varepsilon_{de} = 0.2$, the 99.7% of the d_e values is within a 20% range from the characteristic mean $\overline{d_e}$. For simplicity, the same ε_{de} values are assigned to all the hydrofacies. These values were chosen to avoid superimpositions in the resulting K values of the different hydrofacies. As a result, the structure inherited from the hydrofacies architecture is preserved in the K fields, and the impact of intrafacies variability on the global K field variance is generally small (Table 1).

A fundamental feature of the synthetic K fields used in this study is that for a given K_H / K_L ratio and a given ε_{de} value, the statistical distribution of the reference K values is shared by all the scenarios notwithstanding different structures. This is the result of the choice of considering the same volumetric fractions for the hydrofacies in all scenarios. In this regard, the present study is similar in spirit with previous analyses in which K fields with identical histogram, but with different spatial structures, have been compared to understand their impact on groundwater flow and solute transport (e.g., Wen and Gomez-Hernandez, 1998; Western et al., 2001; Zinn and Harvey, 2003; Lee et al., 2007; Siirila-Woodburn and Maxwell, 2015).

2.3 Flowmeter test simulations

Flowmeter tests are simulated with a three-dimensional finite-difference code (MODFLOW-2005, Harbaugh, 2005) considering steady-state convergent flow toward a fully penetrating well in a confined aquifer. Hydraulic head (h) and specific discharge (q) satisfy the following equations:

$$\mathbf{q} = -K\nabla h \qquad \nabla \cdot \mathbf{q} + f = 0 \qquad (4)$$

where $K(x, y, z)$ is the reference K field and f is a volumetric term per unit volume of aquifer representing fluid sources (positive) or withdrawals (negative). The numerical grid is defined over a cylindrical domain of radius 120.5 u_h, and the location of the well for flowmeter test simulations is at the centre of the domain (Figure 1). A sensitivity analysis was conducted to evaluate the precision of the model outputs with respect to the resolution of the numerical grid. This analysis was performed by comparing outputs from a model with grid resolution equal to that of geostatistical grid (about 4.6 million cells) to the results of two models with horizontal resolutions that are two (about 18 million cells) and four (about 78 million cells) times higher, respectively than the resolution of the geostatistical grid. These models consider the K field of the scenario with isotropic L_h equal to 2 u_h (Group 1). No other scenarios of aquifer heterogeneity have been tested. Since comparisons did not indicate significant differences, the model with grid resolution equal to that of geostatistical grid was considered for flowmeter tests simulations. Accordingly, the flowmeter well diameter is equal to the horizontal supporting scale of the characteristic d_e and K values of the hydrofacies (1 u_h). Specified head conditions are imposed at the cells on the external surface Γ_D of the cylindrical domain, while a specified flux, simulating groundwater withdrawal at a constant rate Q_w, is applied at the cell at the top of the vertical stack of cells representing the well. No-flow boundary conditions are applied to all other boundaries. Following the approach used in previous studies (e.g., Whittaker and Teutsch, 1999; Riva et al., 2012), a significantly large K value is assigned to the cells representing the well, with result of an equilibration of the simulated heads (differences in the orders of 10^{-5}). Because of this condition, discharge into these cells depends only on the K values of the neighbouring cells. Therefore, an estimation of hydraulic conductivity for each cell can be calculated by applying Equation (1), in way similar to a flowmeter test in a real aquifer. In this numerical simulation, discharge values ΔQ_i from each layer i into the well are provided from the cell-by-cell balance of the groundwater flow model.

A series of assumptions are made in the calculation of the simulated K estimates to allow unbiased comparisons with the reference values. The first is that the measurement interval of the simulated flowmeter tests corresponds to the vertical support scale of the reference K data (1 u_v). With respect to the accuracy of the flowmeter K estimates, this represents a conservative condition because in field situations the thickness of aquifer layers with different K values is not known a priori. Moreover, studies have shown that flowmeter estimates are sensitive to the choice of the measurement interval (Rehfeldt et al., 1989; Molz et al. 1989; Boman et al. 1997; Whitteker and Teutsch, 1999). Two other conservative assumptions are that estimated K values are not affected by any source of error other than numerical approximation, and that simulated flowmeter readings are not bounded by a minimum detectable value, which studies suggest to be in the order of 0.005

l/s (Rehfeldt et al., 1989; Young and Pearson, 1995). It is also assumed that the input value in Equation (1) for the effective hydraulic conductivity of the aquifer ($\bar{K}$) is calculated with the relationship (e.g., Sanchez-Vila et al., 2006):

$$\bar{K} = \frac{Q_w}{2\pi Z(h_D - h_w)} \ln\left(\frac{r_{\Gamma_D}}{r_w}\right) \qquad (5)$$

where h_w is the average hydraulic head in the well cells, the well radius r_w is equal to 0.5 u_h, and h_D is the head imposed at the boundary Γ_D at a distance r_{Γ_D} from the well (120 u_h). Because in Equation (1) $\bar{K}$ is a constant multiplicative factor, it can be shown that the variance of the distribution of log-transformed K estimates in the well cells is independent of $\bar{K}$. Rather, different $\bar{K}$ values affect the mean of the distribution and produce shifts of the vertical profile of log-transformed K estimates toward lower or higher values.

2.4 Validation criteria

Simulated flowmeter K data are validated through comparisons with reference data based on the hydrofacies. For each scenario, the accuracy of the estimated K values is quantified by two different metrics. The first is the coefficient of determination (R^2) defined by the following:

$$R^2 = 1 - \frac{\sum_{i=1}^{N}\left(Y_{R,i} - \overline{Y_R}\right)^2}{\sum_{i=1}^{N}\left(Y_{R,i} - Y_{F,i}\right)^2} \qquad (6)$$

where $Y_{R,i}$ is the log-transformed (base 10) reference K value for the well cell i, $\overline{Y}_R$ is the mean of the distribution of $Y_{R,i}$, $Y_{R,i}$ is the corresponding log-transformed flowmeter K estimate, and $N = 80$ is total number of well cells. The coefficient R^2 is a measure of the ability of the simulated flowmeter tests to reproduce the variability of the reference log-transformed K profiles. Perfect variance recovery is indicated by an R^2 equal to 1.The second metric is the average accuracy ratio (Q) defined as:

$$Q_k = \frac{1}{M}\sum_{i=1}^{M} \text{Log}_{10}\left(\frac{K^k{}_{F,i}}{K^k{}_{R,i}}\right) \qquad (7)$$

where $K^k{}_{R,i}$ is the reference K value for hydrofacies k, $K^k{}_{F,i}$ is the corresponding flowmeter estimate, and M is the total number of reference K data for hydrofacies k. For each hydrofacies, Q quantifies the average order of magnitude of difference between reference and estimated K values. A perfect match is achieved when $Q = 0$.

3 Results

Calculated effective hydraulic conductivity ($\overline{K}$) of the reference K fields are presented in Table 1 normalized with respect to the characteristic K of hydrofacies "H" (K_H). Corresponding normalized values of the harmonic and arithmetic means of the distribution of reference K values in the well cells are also reported to provide an indication of the degree of stratification of the structure of K fields. In fact, it can be shown that these two types of means represent the effective hydraulic conductivity of 2-D perfectly stratified systems – each having the same vertical K distribution as that of the well cells – in which flow is either perpendicular (harmonic mean) or parallel (arithmetic mean) to the stratification. Calculated $\overline{K}$ values fall within these two bounds. In particular, as the horizontal mean length of the hydrofacies increases, $\overline{K}$ tends towards values more similar to the arithmetic mean indicating a corresponding increment in the degree of stratification.

The variances of the distributions of $\text{Log}_{10}(K)$ and $\ln(K)$ for each scenario allow one to compare the variability of the reference K fields to that considered in other synthetic case studies (e.g., Zhang et al., 2013) or observed in the field (e.g., Rehfeldt et al., 1992; Barahona-Palomo et al., 2011; Bohling et al., 2012). For the considered reference *K* fields, the variance of ln(*K*) varies between 1.1 and 16.3 depending on the K_H / K_L ratio and on the value of ε_{de} (Table 1). These values, which indicate a moderate to high level of heterogeneity, are within the typical range of values observed in alluvial aquifers (Zhang et al., 2013 and references therein).

3.1 Impact of the mean length of the hydrofacies

The impact of the isotropic horizontal mean length L_h of the hydrofacies on the accuracy of the simulated flowmeter tests is analysed in eight scenarios. These correspond to reference *K* fields having different structures (Figure 2a – 2d), but identical univariate statistical distribution of *K* values in the domain ($\sigma^2_{\ln K} \approx 4$) as well as identical vertical profiles of d_e and *K* along the well. The shape of vertical *K* profiles estimated with the simulated flowmeter test varies depending on the mean lengths of the hydrofacies and the degree of stratification in the reference *K* fields (Figure 3). The comparison between estimated and reference *K* data for architectures with short L_h values ($L_h \leq 4\, u_h$) shows significant discrepancies. The sign of these is both positive and negative, as shown, for instance, by the overprediction of reference *K* data between normalized elevations of 60 and 70, or by the overall underprediction of the characteristic *K* of hydrofacies "H". Moreover, the mismatch between estimated and reference *K* profiles is equally noticeable even when the arithmetic mean of the reference *K* values is used as the value assigned to $\overline{K}$ in Equation (1). This is analogous to assuming that the vertical *K* distribution along the profile is known a priori. For larger L_h values ($L_h \geq 16\, u_h$), flowmeter estimates tend to match the reference *K* values more accurately especially for the thick and horizontally continuous layers of similar conductivity that are intercepted by the testing well. However, the reference conductivity of thin intervals with vertical extent of 1 or 2 u_v can only be matched for the scenario with $L_h = \infty$. The practically perfect match between reference and simulated *K* profiles is a proof of the validity of the numerical results since this scenario represents a benchmark problem for the numerical model.

Calculated values of R^2 range from -0.184 to 1.000 corresponding to the two extreme scenarios with $L_h = 1\,u_h$ and $L_h = \infty$, respectively (Figure 4a). R^2 increases monotonically with L_h according to a relationship that can be described by the power law:

$$R^2 = 1 - aL_h^{\ b} \qquad (8)$$

where the constant *a* is positive and the exponent *b* is negative. It is noteworthy that only for scenarios with $L_h > 16$ the calculated R^2 values are higher than 0.75. From the interpretation of R^2, this result shows that the simulated flowmeter tests are able to characterize the true variability of the reference *K* field only when the *K* field structure has a high degree of stratification near the testing location. For scenarios considering imperfect stratification resulting from a more chaotic hydrofacies architecture (e.g., Figure 2a), simulated flowmeter data underpredict the true *K* variability (see also Figure 3).

Calculated *Q* values for the different structures are plotted in Figure 4b. Irrespective of the type of hydrofacies, results show that the accuracy of the estimated *K* values is strongly dependent on the mean length of the hydrofacies and therefore on the aquifer architecture. From our numerical experiment, the relationship between *Q* and L_h can also approximated by a power law in the form:

$$Q_k = a_k L_h^{\ b_k} \qquad (9)$$

where the values of the constant *a* and the exponent *b* depend on the reference *K* of hydrofacies *k*. Comparisons between *Q* values for the different hydrofacies indicates that the highest estimation errors are observed for the reference K value of the most conductive hydrofacies "H" especially for scenarios considering short L_h, while the estimation error is more uniformly distributed in more stratified architectures. Even for the scenario corresponding to $L_h = 64\, u_h$ (Figure 2c), the estimation error is still not negligible. In fact the average of the Q values for the different hydrofacies is equal to 0.135, which means that flowmeter K estimates differ from the true K values by a factor ≈1.3 on average.

3.2 Impact of the horizontal anisotropy ratio

The effect of the anisotropy ratio L_y / L_x is tested in five scenarios including an isotropic base case with $L_x = L_y = 4$ (Figure 2a) and four anisotropic scenarios in which the mean length of the hydrofacies "H" and "I" in the y direction (L_y) is progressively incremented up to a value equal to 64 u_h. Examples of the generated K fields are shown in Figures 2e and 2b. The vertical K profiles show that simulated flowmeter tests results are sensitive with respect to the ratio L_y / L_x (Figure 5), although discrepancies between corresponding K values for different scenarios appear quite random. Accordingly, the goodness of fit of the estimated K profiles does not seem to improve or decline substantially with the variation of the anisotropy ratio.

The relatively moderate impact of the horizontal anisotropy ratio of the hydrofacies on the accuracy of the simulated flowmeter tests is confirmed by the analysis of the accuracy metrics (Figure 6). An increment in R^2 from the value corresponding to the isotropic base case is observed in the first two anisotropic scenarios, which is particularly steep between the points at $L_y / L_x = 2$ and $L_y / L_x = 4$. The maximum R^2 value for the scenario with $L_y / L_x = 4$ is about 50% larger than the corresponding value in the isotropic scenario. However, this increment in R^2 is followed by a slow decline for larger anisotropy ratios as the K structure becomes more and more characterized by elongated forms (Figure 2f). Variations of Q for the different hydrofacies (Figure 6b) indicate that the increment in accuracy observed for scenarios with moderate anisotropy is particularly evident for the extreme K values (K_L and K_H), while Q values for hydrofacies "I" are generally stable for all the anisotropy ratios.

Simulations results indicate that the effective conductivity value $\bar{K}$ does not change significantly in this group of scenarios (Table 1 and Figure 5). Accordingly, the horizontal anisotropy ratio of the hydrofacies has little effect on the difference between $\bar{K}$ and the arithmetic mean of the reference K values, indicating that the overall degree of stratification in the K field structures is also relatively stable. This result can explain the moderate effect of the anisotropy ratio on the accuracy of the flowmeter test simulations. For brevity, only results for scenarios considering a short mean length of the hydrofacies in the x direction ($L_x = 4$) are presented here. However, a consistent behaviour is observed also for scenarios considering larger L_x values.

3.3 Impact of the hydraulic conductivity contrast

Accuracy metrics for scenarios with different K contrasts between hydrofacies "H" and "L" are shown in Figure 7. Results are presented for two architectures including one with a relatively short horizontal mean length of the hydrofacies ($L_h = 4\, u_h$), and a second architecture characterized by a more stratified K field structure ($L_h = 16\, u_h$). For both architectures, the analysis of the values of R^2 shows that ability of the simulated flowmeter tests to accurately represent the true K variability declines as the ratio K_H / K_L becomes larger (Figure 7a). Numerical results seem to suggest a linear relationship between R^2 and the logarithm of the K_H / K_L ratio. However, the impact of the K_H / K_L ratio on R^2 is slightly more evident for the architecture characterized by the largest L_h. On average, one order of magnitude increment in the K_H / K_L ratio for scenarios with $L_h = 16\, u_h$ corresponds to a 7% decline in R^2, while the corresponding percentage for the scenarios with $L_h = 4\, u_h$ is about 5%.

When the accuracy of the flowmeter K estimates is evaluated for the different hydrofacies (Figure 7b), results indicate that the decline in accuracy for larger K_H / K_L ratios is more apparent for architectures characterized by shorter mean lengths of the hydrofacies especially for hydrofacies “H” and “L. For scenarios considering $L_h = 4\, u_h$, for example, each order of magnitude of increment in the K_H / K_L ratio correspond to a 0.2 increment in Q. This means that the discrepancy between estimated and reference K values increases by a factor of about 1.4 for each order of magnitude of increment in the K_H / K_L ratio. When the K contrast between hydrofacies “H” and “L” is equal to 4 orders of magnitude, the average mismatch between simulated and reference K values is about one order of magnitude. On the other hand, for the scenarios characterized by a more stratified structure (i.e., $L_h = 16\, u_h$), Q values are about one half of the corresponding values for scenarios with $L_h = 4\, u_h$. As for the anisotropy ratio, the impact of the K contrast on the accuracy of the flowmeter K estimates is lower for hydrofacies with intermediate conductivity.

3.4 Impact of intrafacies variability

The degree of sub-hydrofacies scale variability in the reference K field is defined by the value of the parameter ε_{de}, which controls the range of the Gaussian distributions of the effective diameter within the hydrofacies (Figure 8). The impact of intrafacies K variability on the accuracy of the simulated flowmeter tests for different hydrofacies architectures is shown in Figure 9. In particular, the analysis of the variations of R^2 with increments of ε_{de} suggests that the ability of the flowmeter estimates to match the variability of the reference *K* is practically independent from the degree of intrafacies variability expressed by ε_{de} (Figure 9a). It is possible that this result is influenced by the approach used to model intrafacies variability, and that another approach may produce more sensitive results.

Despite having negligible effect on the values of R^2, intrafacies variability can still impact the accuracy of the simulated *K* estimates depending on type of hydrofacies architecture. Calculated values of the accuracy metric *Q* indicate that this impact is small to almost negligible for scenarios characterized by shorter L_h (e.g., Figure 9b), while it is maximum for the scenario considering a perfectly stratified *K* field (Figure 9c). In this case, simulated results indicate a linear relationship between *Q* and ε_{de}.

4 Discussion

4.1 Validity of estimated vertical *K* profiles

Simulated vertical *K* profiles are sensitive to the architecture and to some extent to the sub-unit-scale lithological heterogeneity of the hydrofacies. In particular, the magnitude of systematic errors observed in the flowmeter *K* estimates depends on the degree of deviation of the *K* field structure from the conceptual model represented by a perfectly stratified system. When the structure of the *K* field differs from the stratified model, systematic errors in the flowmeter *K* estimates are introduced due to misinterpretations of the measured discharge ΔQ_i from each measurement interval into the well. These misinterpretations are caused by the presence of vertical flow components due to the non-uniformity of the radial gradients surrounding the well (Figure 10). These components tend to increase the discharge from the less conductive hydrofacies “L” while reducing discharge from the highly-conductive lithofacies “H”. The overall effect is a volume averaging process of the true *K* values around the well, which confirms the observations of Zlotnik and Zurbuchen (2003b). Because of this process, *K* is overestimated in the lower *K* intervals and it is underestimated in the higher-*K* zones. This tendency has been suggested by previous field studies (Molz et al., 1989; Bohiling et al., 2012) and numerical analyses (Whittaker and Teutsch, 1999; Riva et al., 2012), but it was never explained relative to the characteristics of the *K* field structure. Because of this systematic under-/overprediction of extreme values, the true *K* variability along the vertical profile is underestimated.

Comparisons between scenarios with different architecture indicate that the horizontal mean length of the hydrofacies, which controls the lateral extension of the units and therefore the degree of stratification of the *K* field, has a particularly high impact on the accuracy of flowmeter *K* data. For the transition probability/Markov chain modelling approach, mean length values are indicative of spatial correlation similar to the way the variogram range relates to correlation length for variogram-based geostatistical analyses. From direct observations of aquifer outcrops or the analysis borehole data, correlation lengths of sedimentary units at the scale of hydrofacies in alluvial systems range from few meters to few tens of meters (e.g., Jussel et al., 1994; Anderson et al., 1999; Whittaker and Teutsch, 1999; Labolle and Fogg, 2001; Zappa et al., 2006; Bayer et al., 2011; Bianchi and Zheng, 2016). Given such relatively short range of lengths, it is likely that the structure of the *K* field in alluvial aquifers rarely supports the stratified structure assumption, and that systematic errors in flowmeter *K* data are therefore not negligible in most of field applications. This consideration may find confirmation in the lack of correspondence observed in several field sites between flowmeter *K* data and other *K* data measured from aquifer samples, such as those based on grain-size analysis or permeameter tests (Boggs et al. 1990; Hess et al., 1992; Whitteker and Teutsch, 1999; Barahona-Palomo et al., 2011; Gueting et al., 2015; Bianchi and Zheng, 2016).

4.2 Validity of estimated statistical distributions

So far, reference *K* data and corresponding estimates from simulated flowmeter tests have been compared for a single vertical profile. To better evaluate the impact of aquifer architecture on the univariate distribution of estimated *K* values, we now present the results of simulations of multiple flowmeter tests. For each scenario of aquifer architecture, flowmeter data are collected in 13 closely spaced locations around the centre of the domain (Figure 11). Similarly to the previously described simulations, these scenarios also consider three hydrofacies with different conductivity (i.e., K_L, K_I, and K_H). The ratio K_H / K_L in the corresponding *K* fields is assumed to be equal to 10^2, while the characteristic *K* for the hydrofacies "I" is one order of magnitude lower than K_H. Intrafacies variability of *K* is also simulated assuming ε_{de} is equal to 0.3 for all the hydrofacies. The structure of the reference *K* fields for each scenario is controlled by the aquifer architecture corresponding to a certain value of the isotropic mean length L_h. However, all the reference *K* fields share the same tri-modal univariate histogram (Figure 12a). Given the relatively short distance between the different testing locations, the effective hydraulic conductivity of the aquifer is assumed to be the value calculated in the well at the centre of the domain.

As shown in Figure 12a, reference *K* values of the well cells at the test locations for a total of more than a thousand values provide a representative sample of the *K* distribution in the entire domain. On the other hand, histograms of simulated flowmeter *K* estimates differ substantially from that of the reference K data. In particular, the tri-modal distribution of the reference K fields can be identified in the simulated flowmeter data only in the two scenarios with $L_h \geq 16\, u_h$. For architectures characterized by shorter L_h, estimated $\mathrm{Log}_{10}(K)$ values follow complex distributions with no immediately discernible shape. The variance of $\mathrm{Log}_{10}(K)$ values in these scenarios is also lower than the true variance of the reference K field. For instance, the true $\mathrm{Log}_{10}(K)$ variance is underestimated by factor ≈ 1.5 in the scenario with $L_h = 4$. Although a better match between the variances of the reference and simulated K data is obtained as L_h increases, the true variance is underestimated in all the scenarios. On the other hand, a comparison of the mean values indicates that the aquifer architecture has little impact on the mean of the flowmeter K data. This is an interesting result, which may provide an explanation for the similarity between mean values of flowmeter and grain-size based K estimates at coinciding locations in an alluvial aquifer near Tübingen (Germany), even though the variance of the flowmeter K data is lower than the corresponding grain-size based estimates (Barahona-Palomo, 2011).

Previous studies have investigated the statistical distribution of the increments (i.e., the difference of values measured at two locations separated by a spatial lag) of hydraulic conductivity (e.g., Liu and Molz, 1997; Lu and Molz, 2001; Meershaert et al., 2004; Meershaert et al., 2013, Guadagnini et al. 2013), permeability (e.g., Painter, 1996; Castle et al., 2004; Guadagnini et al., 2012; Riva et al., 2013; Siena et al., 2012), and other hydrogeological and geophysical properties (e.g., Yang et al., 2009; Guadagnini et al., 2013). Liu and Molz (1997), for example, analysed the distribution of $\ln(K)$ increments in the extensive dataset of flowmeter measurements collected at the MADE site. Results show that for a lag equal to the measurement interval (≈ 0.15 cm), the probability density function (PDF) of the increments exhibits a distinct peak around the mean equal to zero and symmetric slow decaying tails, which can be reasonably described by a Lévy-stable distribution (Samorodnitsky and Taqqu, 1994). This is a family of distributions characterized by a Lévy index α, with range 0 – 2, which controls the power-law decay of the tails of the PDF. For a Lévy index equal to 2, the Lévy distribution is Gaussian. Levy-stable or similar heavy-tailed distributions of the increments have been observed in other flowmeter K data (Hess et al.,1992; Meershaert et al., 2004), as well as in K and permeability datasets based on other measurement techniques (Painter et al., 1996; Guadagnini et al., 2013; Riva et al, 2013, among others). Lévy-stable or Lévy-like distributions provide the mathematical basis for stochastic fractal-based models of subsurface heterogeneity (see Molz et al., 2004 for a review).

PDFs of the distributions of increments of $\mathrm{Log}_{10}(K)$ for the simulated flowmeter and reference K datasets are shown in Figure 13. Increments are calculated for lags equal to 1 u_v along the vertical succession of cells in the 13 simulated flowmeter test locations, as well as for all the vertical successions of cells in the entire domain. The distribution of the increments in the reference K field is characterized by five bell-shaped peaks, and clearly reflects a number of features of the hydrofacies model and corresponding reference K field including 1) the number of hydrofacies, 2) their volumetric fractions, 3) juxtapositional tendencies between hydrofacies, 4) the K contrasts between hydrofacies, and 5) the chosen model of intrafacies variability. This reference distribution is not discernible in the distributions of $\mathrm{Log}_{10}(K)$ increments in the simulated flowmeter data. Rather, especially for scenarios with shorter L_h, the distribution of flowmeter based increments inaccurately suggest a distribution characterized by zero mean and heavy tails, which can be reasonably described by Lévy stable distributions with Lévy index α ranging from 0.677 to 1.009. It can be noted that the tails become heavier with increments of L_h. The implication of this result is twofold. First, it strongly suggests that the Lévy behaviour of the increments in log(K) observed in *K* datasets based on flowmeter measurements may be an artefact if the *K* field structure has a low degree of stratification. Lu et al. (2002) already showed that by combining facies and fractal modelling approaches it is possible to create realizations of the *K* field in which, although the distribution of the ln(*K*) increments within the facies is Gaussian, the distribution of the increments for the entire multi-facies domain follows a Lévy like distribution. In this previous study, the Lévy behaviour was indicated as a statistical artefact. On the other hand, the results of this study indicate the possibility that the Lévy behaviour frequently observed in the distribution of log(*K*) increments may be instead an artefact of the measurement method. The cause of this behaviour needs further investigation, but it is possible that the main reason is the previously described misinterpretation of the measured discharge values along the well by the traditional flowmeter data interpretation. The resulting averaging mechanism causes a smoothing effect of the true *K* contrast along the profile and an artificial enhancement of the correlation between estimated *K* values at shorter lags. The effect of this smoothing effect can be seen in the heavy tailing of the log(*K*) increments. The second implication of this result is that it suggests that it might be possible to obtain insight about of the true *K* field structure and hydrofacies architecture from the interpretation of the shape of the distribution of log(*K*) increments from flowmeter measurements. This hypothesis will be tested in a future study.

5 Conclusions

Detailed simulations of steady-state convergent groundwater flow in three-dimensional *K* fields are performed in order to evaluate the impact of medium-scale sedimentary architecture and small-scale heterogeneity on the validity of flowmeter data. In particular, we focus on systematic errors that arise when the *K* field structure differs from the hypothetical perfectly stratified conceptualization at the basis of the traditional method for flowmeter test interpretation. The general finding of this work is that flowmeter testing should not be applied in alluvial aquifers without a preliminary knowledge of the structure of the *K* field. The presented numerical analysis in fact shows that the validity of the *K* estimates strongly depends on the hydrofacies architecture of the aquifer, as well as on conductivity contrasts between and within the hydrofacies. All these features are controlled by parameters (e.g., mean lengths and volumetric fractions of the hydrofacies, grain-size distribution, etc.) that can be measured from the analysis of aquifer samples, lithological well logs, or from the application of indirect methods of subsurface characterization. The result of this preliminary characterization should be used to inform the decision regarding the identification of the best method for a fine-scale *K* field characterization.

Comparisons between simulated flowmeter estimates and reference data for scenarios considering a range of architectures, medium-scale and small-scale *K* variability also lead to the following specific conclusions.

For *K* fields having the same univariate distribution, but with different structures corresponding to a range of values of the isotropic horizontal mean length of the hydrofacies, significant systematic errors are observed for scenarios with shorter mean lengths. Numerical results suggest that the relationships between accuracy metrics of the *K* estimates and the mean length of the hydrofacies can be adequately described by power laws. When the *K* structure deviates from the hypothetical perfectly stratified system, flowmeter *K* data tend to overestimate the true *K* in low-K intervals, while they tend to underpredict reference values in high-*K* intervals. As a result, the true *K* variability is also underestimated in the profiles.

For fields having the same univariate *K* distribution, but with different structures corresponding to a range of anisotropy ratios between the horizontal mean lengths of the hydrofacies, the overall effect of the anisotropy ratio on the accuracy of the flowmeter *K* estimates is moderate.

For *K* fields sharing the same structure, but with different ranges in the *K* distributions resulting from different values of the *K* ratio between the highest and lowest conductive hydrofacies, the accuracy of the flowmeter estimates decreases linearly with the *K* ratio in a semilog scale. In particular, the highest effect on the accuracy of the *K* estimates is observed for the extremes of the reference *K* distribution, while the effect is more moderate for intermediate *K* values.

Compared to the other considered properties, the overall effect of the intrafacies *K* variability on accuracy of the flowmeter is negligible in terms of reproducibility of the actual *K* field variance. However, some level of intrafacies variability can produce non-negligible systematic errors in the estimation of the *K* profiles, even when the true medium-scale *K* structure approximates a perfectly stratified system.

The analysis of the datasets resulting from a combination of the *K* estimates from multiple simulated flowmeter tests indicate that different architectures can have a significant impact on the predicted statistical distributions of the estimated *K* values and of the increments of $\mathrm{Log}_{10}(K)$. In scenarios in which the *K* field structure deviates substantially from the perfectly stratified model, the shape of the histogram of the flowmeter log(*K*) estimates disguises the shape of the true distribution of log(*K*) values and underestimates its variance. However, the mean values of the two distributions are generally comparable irrespective of the aquifer architecture. For these scenarios, the distributions of the $\mathrm{Log}_{10}(K)$ flowmeter increments are also significantly different from the distribution of the increments in the reference *K* field. When the *K* field is characterized by imperfect stratification, the distribution of the log-transformed K increments based on flowmeter estimates tends to follow an apparent Lévy behaviour with zero mean and symmetric heavy tailing that tends to become heavier with increasing horizontal mean length of the hydrofacies.

Acknowledgments. This work was undertaken as part of the "Research Fellowship Programme" funded by the British Geological Survey (Natural Environment Research Council). Marco Bianchi publishes with the permission of the Executive Director of the British Geological Survey. Three anonymous reviewers are acknowledged for providing constructive comments that improved the paper.

References

Allen-King, R. M., R. M. Halket, D. R. Gaylord, and M. J. L. Robin (1998), Characterizing the heterogeneity and correlation of perchloroethene sorption and hydraulic conductivity using a facies-based approach, Water Resour. Res., 34(3), 385–396, doi:10.1029/97WR03496.

Anderson, M. P. (1989), Hydrogeologic facies models to delineate large-scale spatial trends in glacial and glacialfluvial sediments, Geol. Soc. Am. Bull., 101, 501–511.

Anderson, M, J. Aiken, E. Webb, and D. Mickelson (1999), Sedimentology and hydrogeology of two braided stream deposits, Sedimentary Geology, 129, 187–199.

Barahona-Palomo, M., M. Riva, X. Sanchez-Vila, E. Vazquez-Sune, and A. Guadagnini (2011), Quantitative comparison of impeller flowmeter and particle-size distribution techniques for the characterization of hydraulic conductivity variability, Hydrogeol. J., 19(3), 603–612, doi:10.1007/s10040-011-0706-5.

Bayer, P., P. Huggenberger, P. Renard, and A. Comunian (2011), Three-dimensional high resolution fluvio-glacial aquifer analog: Part 1: Field study. Journal of Hydrology, 405 (1–2) (2011), pp. 1–9.

Bianchi, M., C. Zheng, G.R. Tick, and S.M. Gorelick (2011a), Investigation of Small-Scale Preferential Flow with a Forced-Gradient Tracer Test, Groundwater, 49, 503–514, doi: 10.1111/j.1745-6584.2010.00746.

Bianchi, M., and C. Zheng (2016), A lithofacies approach for modeling non-Fickian solute transport in a heterogeneous alluvial aquifer, Water Resour. Res., 52, 552–565, doi:10.1002/2015WR018186.

Bohling, G.C., G. Liu, S.J. Knobbe, E.C. Reboulet, D.W. Hyndman, P. Dietrich, and J.J. Butler Jr. (2012), Geostatistical analysis of centimeterscale hydraulic conductivity variations at the MADE site, Water Resour. Res., 48,W02525, doi:10.1029/2011WR010791.

Boggs, J. M., S. C. Young, D. J. Benton, and Y. C. Chung (1990), Hydrogeologic characterization of the MADE site, Interim Rep. EN-6915, Electr. Power Res. Inst., Palo Alto, Calif.

Boman, G. K., Molz, F. J. and Boonec, K. D. (1997), Borehole Flowmeter Application in Fluvial Sediments: Methodology, Results, and Assessment. Ground Water, 35: 443–450. doi:10.1111/j.1745-6584.1997.tb00104.x

Butler, J. J.Jr. (2005), Hydrogeological methods for estimation of spatial variations in hydraulic conductivity, in Hydrogeophysics, Water Sci. and Technol. Libr., vol. 50, edited by Y. Rubin, and S. S. Hubbard, pp. 23–58, Springer, New York.

Castle, J. W., F. J. Molz, S. Lu, and C. L. Dinwiddie (2004), Sedimentology and fractal-based analysis of permeability data, John Henry member, Straight Cliffs formation (upper cretaceous), Utah, U.S.A., Journal of Sedimentary Research, 74(2), 270 – 284.

Carle, S.F. (1999), T-PROGS: Transition Probability Geostatistical Software, version 2.1. Davis, California: University of California.

Carle, S.F., and G.E. Fogg (1996), Transition probability-based indicator geostatistics, Math. Geol., 28(4), 453–476.

Carle S.F., and G.E. Fogg (1997), Modeling spatial variability with one and multidimensional continuous-lag Markov chains, Math. Geol., 29(7), 891–918.

Carle S.F., E.M. LaBolle, G.S. Weissmann, D. VanBrocklin, and G.E. Fogg (1998), Geostatistical simulation of hydrostratigraphic architecture: a transition probability / Markov approach, in Concepts in Hydrogeology and Environmental Geology No. 2, SEPM Special Publication, p. 147–170.

Cooper, H.H. and C.E. Jacob (1946), A generalized graphical method for evaluating formation constants and summarizing well field history, Am. Geophys. Union Trans., vol. 27, pp. 526-534.

Crisman, S. A., Molz, Fred. J., Dunn, D. L. and Sappington, F. C. (2001), Application Procedures for the Electromagnetic Borehole Flowmeter in Shallow Unconfined Aquifers. Ground Water Monitoring & Remediation, 21: 96–100. doi:10.1111/j.1745-6592.2001.tb00645.x.

Dell'Arciprete, D., R. Bersezio, F. Felletti, M. Giudici, A. Comunian, and Ph. Renard (2012), Comparison of three geostatistical methods for hydro-facies simulation: a test on alluvial sediments, Hydrogeology Journal, 20, 299-311. doi:10.1007/s10040-011-0808-0.

Deutsch, C.V., and A.G. Journel (1998), GSLIB geostatistical software library and user's guide, 2nd ed., Oxford University Press, New York, 369p.

Dinwiddie, C. L., Foley, N. A. and Molz, F. J. (1999), In-Well Hydraulics of the Electromagnetic Borehole Flowmeter, Ground Water, 37: 305–315. doi:10.1111/j.1745-6584.1999.tb00988.x

Feehley C.E., C. Zheng, and F.J. Molz (2000), A dual-domain mass transfer approach for modeling solute transport in heterogeneous porous media, application to the MADE site, Water Resour. Res., 36(9), 2501–2515, doi:10.1029/2000WR900148.

Fiori, A., G. Dagan, I. Jankovic, and A. Zarlenga (2013), The plume spreading in the MADE transport experiment: Could it be predicted by stochastic models? Water Resour. Res., 49, 2497–2507, doi:10.1002/wrcr.20128.

Frei, S., J. H. Fleckenstein, S. J. Kollet, and R. M. Maxwell (2009), Patterns and dynamics of river-aquifer exchange with variably saturated flow using a fully coupled model, J. Hydrol., 375(3–4), 383–393, doi:10.1016/j.jhydrol.2009.06.038.

Fogg, G. E. (1986), Groundwater Flow and Sand Body Interconnectedness in a Thick, Multiple-Aquifer System, Water Resour. Res., 22(5), 679–694, doi:10.1029/WR022i005p00679.

Fogg, G.E., C.D. Noyes, and S.F. Carle (1998), Geologically based model of heterogeneous hydraulic conductivity in an alluvial setting, Hydrogeology Journal, 6(1), 131–43.

Fogg, G.E., S.F. Carle, and C. Green (2000), Connected-network paradigm for the alluvial aquifer system. In: Zhang, D., C.L. Winter, eds., Theory, modeling, and field investigation in hydrogeology: A special volume in honor of Shlomo P. Neuman's 60th birthday, Geological Society of America, Special paper 348, p. 25–42.

Gómez-Hernández, J. J., J. J. Butler Jr., and A. Fiori (2016), Groundwater transport in highly heterogeneous aquifers, Eos, 97, doi:10.1029/2016EO047263.

Guadagnini, A., M. Riva, and S. P. Neuman (2012), Extended power-law scaling of heavy-tailed random air-permeability fields in fractured and sedimentary rocks, Hydrol. Earth Syst. Sci., 16, 3249–3260, doi:10.5194/hess-16-3249-2012.

Guadagnini, A., S. P. Neuman, M. G. Schaap, and M. Riva (2013), Anisotropic statistical scaling of vadose zone hydraulic property estimates near Maricopa, Arizona, Water Resour. Res., 49, 8463–8479, doi:10.1002/2013WR014286.

Gueting, N., A. Klotzsche, J. van der Kruk, J. Vanderborght, H Vereecken, and A. Englert (2015), Imaging and characterization of facies heterogeneity in an alluvial aquifer using GPR full-waveform inversion and cone penetration tests, Journal of hydrology, 524, 680 – 695, doi. 10.1016/j.jhydrol.2015.03.030.

Harbaugh, A.W. (2005), MODFLOW-2005, The U.S. Geological Survey modular ground-water model—the Ground-Water Flow Process: U.S. Geological Survey Techniques and Methods 6-A16, variously p.

Heinz, J., S. Kleineidam, G. Teutsch, and T. Aigner (2003), Heterogeneity patterns of quaternary glaciofluvial gravel bodies (SW-Germany): application to hydrogeology, Sedimentary Geology, 158 (1–2), 1–23.

Hess, K.M. (1989), Use of a borehole flowmeter to determine spatial heterogeneity of hydraulic conductivity and macrodispersion in a sand and gravel aquifer, Cape Cod, Massachusetts. In: Molz, F.J., Melville, J.G., Guven, O. (Eds.), Proceedings of the Conference on New Field Techniques for Quantifying the Physical and Chemical Properties of Heterogeneous Aquifers, Dallas, Texas, March 20–23, 1989, National Water Well Association, Dublin, OH, 497–508.

Hess, K. M., S. H. Wolf, and M. A. Celia (1992), Large-scale natural gradient tracer test in sand and gravel, Cape Cod, Massachusetts: 3. Hydraulic conductivity variability and calculated macrodispersivities, Water Resour. Res., 28(8), 2011–2027, doi:10.1029/92WR00668.

Hufschmied, P. (1986) Estimation of three-dimensional statistically anisotropic hydraulic conductivity field by means of single well pumping test combined with flowmeter measurements, Hydrogeologie, 2, 163–174.

Illman, W. A., J. Zhu, A. J. Craig, and D. Yin (2010), Comparison of aquifer characterization approaches through steady state groundwater model validation: A controlled laboratory sandbox study, Water Resour. Res., 46, W04502, doi:10.1029/2009WR007745.

Javandel, I., and P. A. Witherspoon (1969), A Method of Analyzing Transient Fluid Flow in Multilayered Aquifers, Water Resour. Res., 5(4), 856–869, doi:10.1029/WR005i004p00856.

Jussel, P., F. Stauffer, and T. Dracos (1994), Transport modeling in heterogeneous aquifers: 1. Statistical description and numerical generation of gravel deposits, Water Resour. Res., 30(6), 1803–1817, doi:10.1029/94WR00162.

Koutrouvelis, I. A. (1980), Regression-Type Estimation of the Parameters of Stable Laws, Journal of the American Statistical Association, 75(372), 918 – 928.

Koutrouvelis I. A. (1981), An Iterative Procedure for the estimation of the Parameters of Stable Law, Communications in Statistics - Simulation and Computation, 10(1), 17 – 28.

LaBolle E.M., and G.E. Fogg (2001), Role of molecular diffusion in contaminant migration and recovery in an alluvial aquifer system, Transport Porous Media, 42 (1-2), 155–179.

Lee, S.-Y., S. F. Carle, and G. E. Fogg (2007), Geologic heterogeneity and a comparison of two geostatistical models: Sequential Gaussian and transition probability-based geostatistical simulation, Adv. Water Resour., 30, 1914–1932, doi:10.1016/j.advwatres.2007.03.005.

Li, W., Englert, A., Cirpka, O. A. and Vereecken, H. (2008), Three-Dimensional Geostatistical Inversion of Flowmeter and Pumping Test Data. Ground Water, 46: 193–201. doi:10.1111/j.1745-6584.2007.00419.x.

Liu, H. H., and F. J. Molz (1997), Comment on “Evidence for non-Gaussian scaling behavior in heterogeneous sedimentary formations” by Scott Painter, Water Resour. Res., 33(4), 907–908, doi:10.1029/96WR03788.

Lu, S., and F. J. Molz (2001), How well are hydraulic conductivity variations approximated by additive stable processes? Advances in Environmental Research, 5, 39–45.

Lu, S., F. J. Molz, G. E. Fogg, and J. W. Castle (2002), Combining stochastic facies and fractal models for representing natural heterogeneity, Hydrogeol. J., 10, 475–482, doi:10.1007/s10040-002-0212-x.

Maliva, R.G. (2016), Aquifer Characterization Techniques, Schlumberger Methods in Water Resources Evaluation Series No. 4, Springer, 617 pp.

Miall, A.D (2014), Fluvial Depositional Systems. Springer International Publishing, 316 pp, doi: 10.1007/978-3-319-00666-6.

Meerschaert, M. M., T. J. Kozubowski, F. J. Molz, and S. Lu (2004), Fractional Laplace model for hydraulic conductivity, Geophys. Res. Lett., 31, L08501, doi:10.1029/2003GL019320.

Meerschaert, M. M., M. Dogan, R. L. Van Dam, D. W. Hyndman, and D. A. Benson (2013), Hydraulic conductivity fields: Gaussian or not?, Water Resour. Res., 49, 4730–4737, doi:10.1002/wrcr.20376.

Molz, F. J., R. H. Morin, A. E. Hess, J. G. Melville, and O. Güven (1989), The Impeller Meter for measuring aquifer permeability variations: Evaluation and comparison with other tests, Water Resour. Res., 25(7), 1677–1683, doi:10.1029/WR025i007p01677.

Molz, F. J., H. Rajaram, and S. Lu (2004), Stochastic fractal-based models of heterogeneity in subsurface hydrology: Origins, applications, limitations, and future research questions, Rev. Geophys., 42, RG1002, doi:10.1029/2003RG000126.

Newhouse, M.W., Izbicki, J.A. and Smith, G.A. (2005), Comparison of velocity-log data collected using impeller and electromagnetic flowmeters. Ground Water, 43: 434–438. doi:10.1111/j.1745-6584.2005.0030.x

Paillet, F.L. (2004), Borehole flowmeter applications in irregular and large diameter boreholes, Journal of Applied Geophysics 55, no.1–2: 39–59.

Painter, S. (1996), Evidence for Non-Gaussian Scaling Behavior in Heterogeneous Sedimentary Formations, Water Resour. Res., 32(5), 1183–1195, doi:10.1029/96WR00286.

Ramanathan, R., R. W. Ritzi Jr., and R. M. Allen-King (2010), Linking hierarchical stratal architecture to plume spreading in a Lagrangian-based transport model: 2. Evaluation using new data from the Borden site, Water Resour. Res., 46, W01510, doi:10.1029/2009WR007810.

Rehfeldt, K.R., Hufschmied, P., Gelhar, L.W., Schaefer, M.E. (1989), The borehole flowmeter technique for measuring hydraulic conductivity variability, Report No. EN 6511, Electric Power Research Institute, Palo Alto, CA.

Rehfeldt, K. R., J. M. Boggs, and L. W. Gelhar (1992), Field study of dispersion in a heterogeneous aquifer: 3. Geostatistical analysis of hydraulic conductivity, Water Resour. Res., 28(12), 3309–3324, doi:10.1029/92WR01758.

Riva, M., L. Guadagnini, A. Guadagnini, T. Ptak, and E. Martac (2006), Probabilistic study of well capture zones distribution at the Lauswiesen field site J. Contam. Hydrol., 88 (1–2), 92–118.

Riva, M., A. Guadagnini, D. Fernandez-Garcia, X. Sánchez-Vila, and T. Ptak (2008), Relative importance of geostatistical and transport models in describing heavily tailed breakthrough curves at the Lauswiesen site, J. Contam. Hydrol., 101, 1–13, doi:10.1016/j.jconhyd.2008.07.004.

Riva, M., Ackerer, P., and Guadagnini, A. (2012), Interpretation of flowmeter data in heterogeneous layered aquifers, Journal of Hydrology, 452–453, 76-82.

Riva, M., S.P. Neuman, and A. Guadagnini (2013), Sub-Gaussian model of processes with heavy tailed distributions applied to permeabilities of fractured tuff. Stoch. Environ. Res. Risk Assess., 27, 195–207.

Ruud, N. C., and Z. J. Kabala (1996), Numerical Evaluation of Flowmeter Test Interpretation Methodologies, Water Resour. Res., 32(4), 845–852, doi:10.1029/96WR00004.

Ruud, N. C., and Z. J. Kabala (1997), Numerical evaluation of the flowmeter test in a layerd aquifer with a skin zone. Interpretation Methodologies, Water Resour. Res., 32(4), 845–852, doi:10.1029/96WR00004.

Ruud, N.C. , Z.J. Kabala, and F.J. Molz (1999), Evaluation of flowmeter-head loss effects in the flowmeter test, J. Hydrol., 224, 55–63.

Salamon, P., D. Fernandez-Garcia, and J. J. Gómez-Hernández (2007), Modeling tracer transport at the MADE site: The importance of heterogeneity, Water Resour. Res., 43, W08404, doi:10.1029/2006WR005522.

Samorodnitsky, G., and Taqqu, M.S. (1994), Stable non-Gaussian Random Processes, Chapman & Hall, New York.

Sanchez-Vila, X., A. Guadagnini, and J. Carrera (2006), Representative hydraulic conductivities in saturated groundwater flow, Rev. Geophys., 44, RG3002, doi:10.1029/2005RG000169.

Scheibe, T. D., and D. L. Freyberg (1995), Use of sedimentological information for geometric simulation of natural porous media structure, Water Resour. Res., 31(12), 3259–3270, doi:10.1029/95WR02570.

Siena, M., A. Guadagnini, M. Riva, and S. P. Neuman (2012), Extended power-law scaling of air permeabilities measured on a block of tuff, Hydrol. Earth Syst. Sci., 16, 29–42, doi:10.5194/hess-16-29-2012.

Siirila-Woodburn, E. R., and R. M. Maxwell (2015), A heterogeneity model comparison of highly resolved statistically anisotropic aquifers, Adv. Water Resour., 75, 53–66, doi:10.1016/j.advwatres.2014.10.011.

Theis, C.V. (1935), The relation between the lowering of the piezometric surface and the rate and duration of discharge of a well using groundwater storage, Am. Geophys. Union Trans., vol. 16, pp. 519-524.

Tillmann, A., A. Englert, Z. Nyari, I. Fejes, J. Vanderborght, and H. Vereecken (2008), Characterization of subsoil heterogeneity, estimation of grain size distribution and hydraulic conductivity at the Krauthausen test site using cone penetration test, J. Contam. Hydrol. 95 (1–2), 57–75.

Vereecken, H., U. Döring, H. Hardelauf, U. Jaekel, U. Hashagen, O. Neuendorf, H. Schwarze, and R. Seidemann (2000), Analysis of solute transport in heterogeneous aquifer: The Krauthausen field experiment. Journal of Contaminant Hydrology 45, no. 3–4: 329–358.

Vuković, M., and A. Soro (1992), Determination of Hydraulic Conductivity of Porous Media From Grain-Size Composition, Water Resour. Publ., Highlands Ranch, Colorado (USA).

Weissmann, G. S., and G. E. Fogg (1999), Multi-scale alluvial fan heterogeneity modeled with transition probability geostatistics in a sequence stratigraphic framework, J. Hydrol., 226, 48–65.

Weissmann, G.S., S.F. Carle and G.E. Fogg (1999). Three-dimensional hydrofacies modeling based on soil surveys and transition probability Geostatistics. Water Resour. Res., 35(6), 1761-1770.

Wen X.H., and J. Gomez-Hernandez (1998), Numerical modeling of macrodispersion in heterogeneous media: a comparison of multi-Gaussian and nonmulti-Gaussian models. J. Contaminant Hydrol., 30(1–2), 129–56.

Western, A. W., G. Blöschl, and R. B. Grayson (2001), Toward capturing hydrologically significant connectivity in spatial patterns, Water Resour. Res., 37(1), 83–97, doi:10.1029/2000WR900241.

Whittaker, J., and G. Teutsch (1999), Numerical simulation of subsurface characterization methods: Application to a natural aquifer analogue, Adv. Water Resour., 22(8), 819–829.

Xiang, J. (1995), The evaluation of the flowmeter test in three-layer aquifers and the influence of disturbed zones, J. Hydrol., 166, 127–145.

Yang, C.-Y., K.-C. Hsu, and K.-C. Chen, (2009), The use of the Levy-stable distribution for geophysical data analysis, Hydrogeology Journal, 17, 1265–1273, doi:10.1007/s10040-008-0411-1.

Young, S. C. (1995), Characterization of High-K Pathways by Borehole Flowmeter and Tracer Tests. Ground Water, 33: 311–318. doi:10.1111/j.1745-6584.1995.tb00286.x.

Young, S. C. and Pearson, H. S. (1995), The Electromagnetic Borehole Flowmeter: Description and Application. Groundwater Monitoring & Remediation, 15: 138–147. doi: 10.1111/j.1745-6592.1995.tb00561.x.

Zappa, G., R. Bersezio, F. Felletti, and M. Giudici (2006), Modeling aquifer heterogeneity at the facies scale in gravel-sand braided stream deposits. Journal of Hydrology, 325, 1 – 4, 134 – 153.

Zhang, Y., C.T. Green, and G. Fogg (2013), The impact of medium architecture of alluvial settings on non-Fickian transport, Adv. Water Resour., 54, 78–99.

Zheng, C., M. Bianchi, and S.M. Gorelick (2011), Lessons learned from 25 years of research at the MADE site, Groundwater, 49, 649–662, doi:10.1111/j.1745–6584.2010.00753.x.

Zinn, B., and C. F. Harvey (2003), When good statistical models of aquifer heterogeneity go bad: A comparison of flow, dispersion, and mass transfer in connected and multivariate Gaussian hydraulic conductivity fields, Water Resour. Res., 39, 1051, doi:10.1029/2001WR001146, 3.

Zlotnik, V.A., and B.R. Zurbuchen (2003a), Estimation of hydraulic conductivity from borehole flowmeter tests considering head losses. Journal of Hydrology 281, no. 1–2: 115–128.

Zlotnik, V. A., and B. R. Zurbuchen (2003b), Field study of hydraulic conductivity in a heterogeneous aquifer: Comparison of single-borehole measurements using different instruments, Water Resour. Res., 39(4), 1101, doi:10.1029/2002WR001415.

Table 1. Scenarios of aquifer architecture and small-scale heterogeneity. The mean thickness L_z of the hydrofacies is equal to 2 u_v for all scenarios. HM: harmonic mean; AM: arithmetic mean. K_H and K_L are the characteristic K values of hydrofacies “H” and “L”, respectively. The characteristic K of hydrofacies “I” is one order of magnitude smaller than K_H.

Scenario ID	L_x [u_h]	L_y / L_x	ε_{de}	K_H / K_L	HM/K_H	$\overline{K}$ /K_H	AM/K_H	$\sigma_{\ln K}$	σ_{Log10K}
Group 1: Variable isotropic horizontal mean length									
L_1	1	1	0.0	10^2	0.017	0.098	0.331	4.034	0.761
L_2	2	1	0.0	10^2	0.017	0.125	0.331	4.034	0.761
L_4	4	1	0.0	10^2	0.017	0.164	0.331	4.034	0.761
L_8	8	1	0.0	10^2	0.017	0.221	0.331	4.034	0.761
L_16	16	1	0.0	10^2	0.017	0.254	0.331	4.034	0.761
L_32	32	1	0.0	10^2	0.017	0.261	0.331	4.034	0.761
L_64	64	1	0.0	10^2	0.017	0.301	0.331	4.034	0.760
L_inf	∞	1	0.0	10^2	0.017	0.323	0.331	4.034	0.813
Group 2: Anisotropy ratio L_y / L_x									
AR_1	4	1	0.0	10^2	0.017	0.169	0.331	4.034	0.761
AR _2	4	2	0.0	10^2	0.017	0.205	0.331	4.034	0.761
AR _4	4	4	0.0	10^2	0.017	0.204	0.331	4.034	0.761
AR _8	4	8	0.0	10^2	0.017	0.196	0.331	4.034	0.761
AR _16	4	16	0.0	10^2	0.017	0.211	0.331	4.034	0.761
Group 3: K_H / K_I ratio									
L4_K1	4	1	0.0	10^1	0.139	0.261	0.381	1.115	0.210
L4_K2	4	1	0.0	10^2	0.017	0.165	0.331	4.031	0.760
L4_K3	4	1	0.0	10^3	0.002	0.142	0.326	9.598	1.810
L4_K4	4	1	0.0	10^4	0.0002	0.138	0.325	17.816	3.360
L16_K1	16	1	0.0	10^1	0.139	0.339	0.381	1.115	0.210
L16_K2	16	1	0.0	10^2	0.017	0.255	0.331	4.031	0.760
L16_K3	16	1	0.0	10^3	0.002	0.223	0.326	9.598	1.811
L16_K4	16	1	0.0	10^4	0.0002	0.215	0.325	17.816	3.361
Group 4: Intrafacies K variability									
L2_εde00	2	1	0.0	10^4	0.0002	0.066	0.314	16.121	3.041
L2_εde20	2	1	0.2	10^4	0.0002	0.066	0.321	16.139	3.044
L2_εde40	2	1	0.4	10^4	0.0002	0.066	0.331	16.196	3.055
L2_εde60	2	1	0.6	10^4	0.0002	0.065	0.344	16.301	3.074
L2_εde80	2	1	0.8	10^4	0.0002	0.063	0.361	16.488	3.110
Linf_εde0	∞	1	0.0	10^4	0.0002	0.259	0.265	14.684	2.770
Linf_εde20	∞	1	0.2	10^4	0.0002	0.263	0.285	14.707	2.774
Linf_εde40	∞	1	0.4	10^4	0.0002	0.266	0.308	14.765	2.785
Linf_εde60	∞	1	0.6	10^4	0.0002	0.268	0.334	14.875	2.805
Linf_εde80	∞	1	0.8	10^4	0.0002	0.268	0.364	15.062	2.841

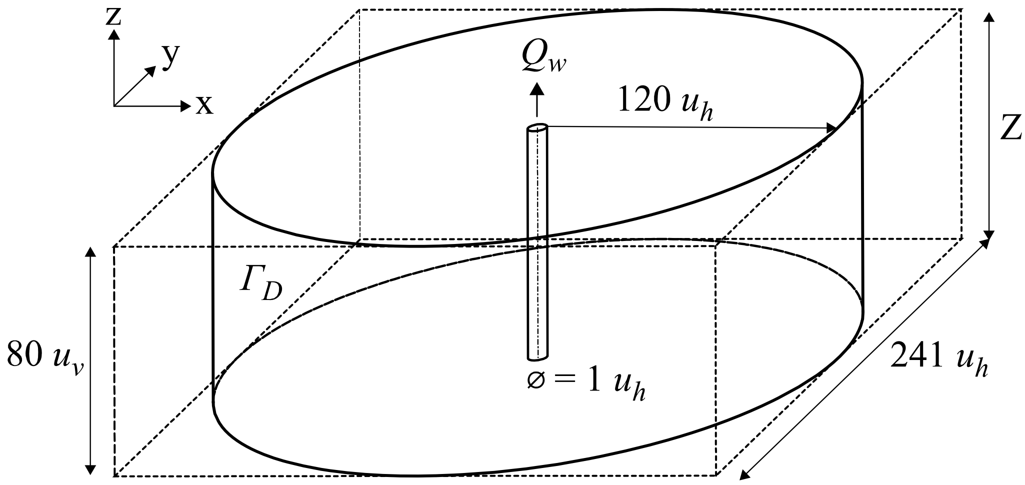

Figure 1. Geostatistical (dashed lines) and numerical (solid lines) domains used for simulations of aquifer architecture and flowmeter tests. Both domains are discretized with a block centered regular grid with cells of volume equal to $1\ u_h \times 1\ u_h \times 1\ u_v$. The well for flowmeter test simulation is located at the centre of the two domains.

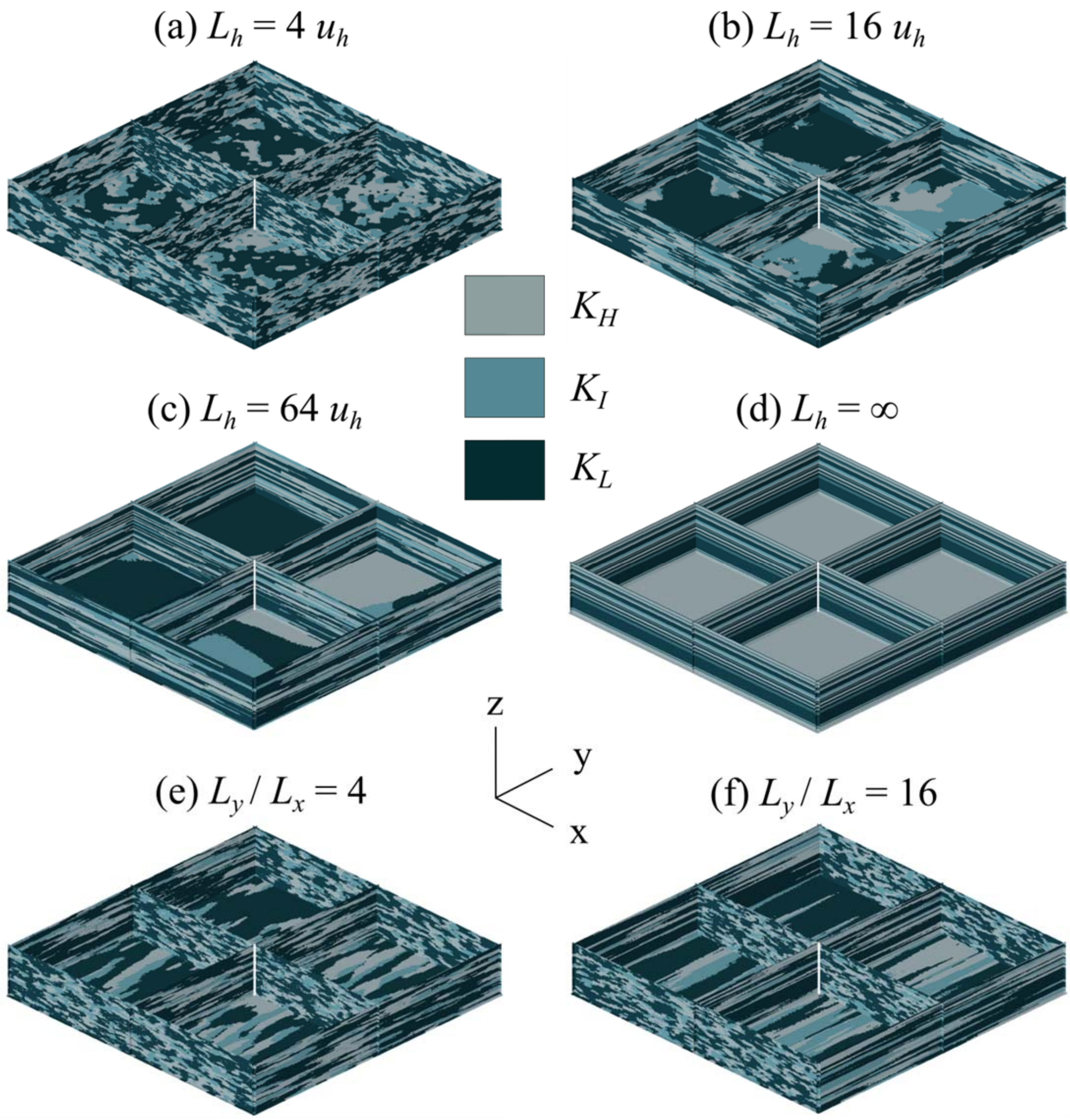

Figure 2. Hydraulic conductivity fields having identical univariate histogram of the K values, identical vertical K profile along the well for flowmeter simulations (shown as a white line), but different structure according to value of the horizontal mean length L_h and anisotropy ratio L_y/L_x.

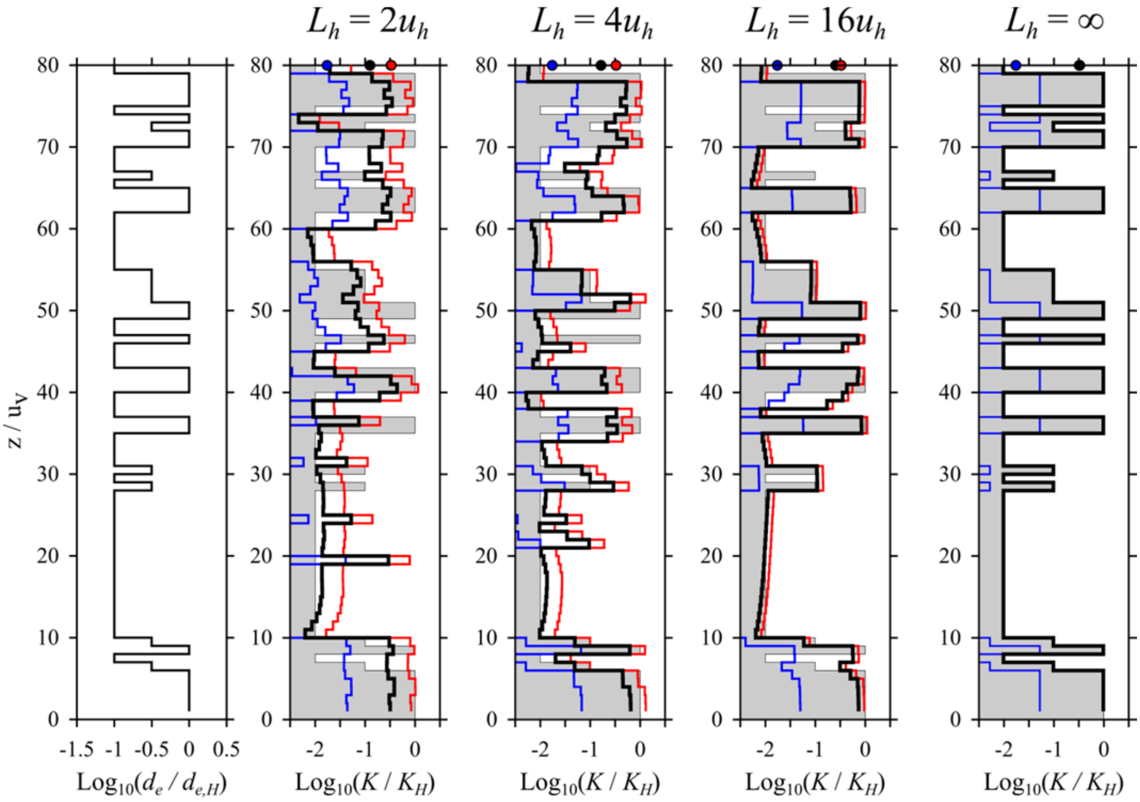

Figure 3. Vertical profile of the effective diameter d_e (normalized by the d_e of hydrofacies "H") and estimated vs. reference K profiles for scenarios considering different L_h values. The reference K profile is indicated by the shaded area. Mean and effective K values are plotted as circles of different colours: blue for harmonic mean; black for $\overline{K}$, and red for arithmetic mean. Simulated flowmeter K profiles considering these mean or effective K values in Equation (1) are plotted as solid lines of corresponding colours.

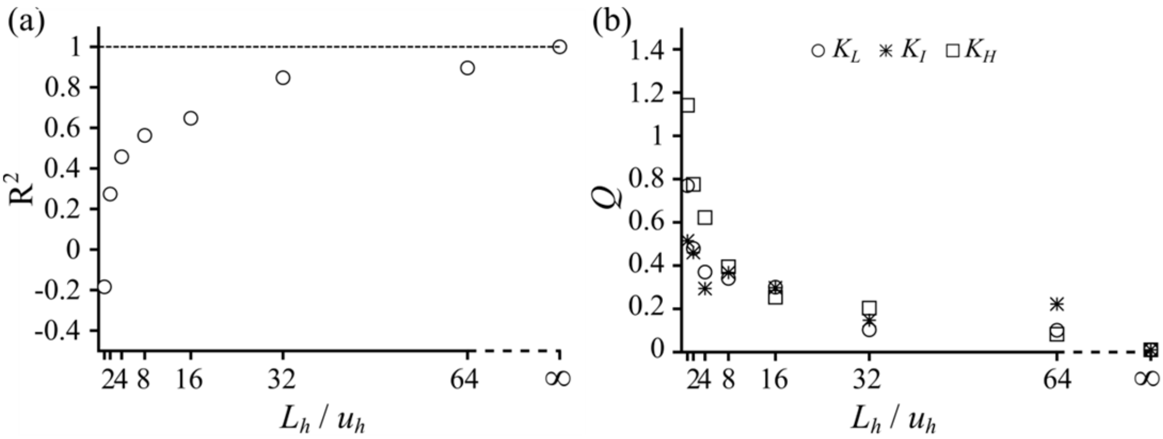

Figure 4. Variations of the accuracy metrics of flowmeter estimates with respect to the L_h of the hydrofacies. (a) Coefficient of determination (R^2). (b) Average accuracy ratio (Q).

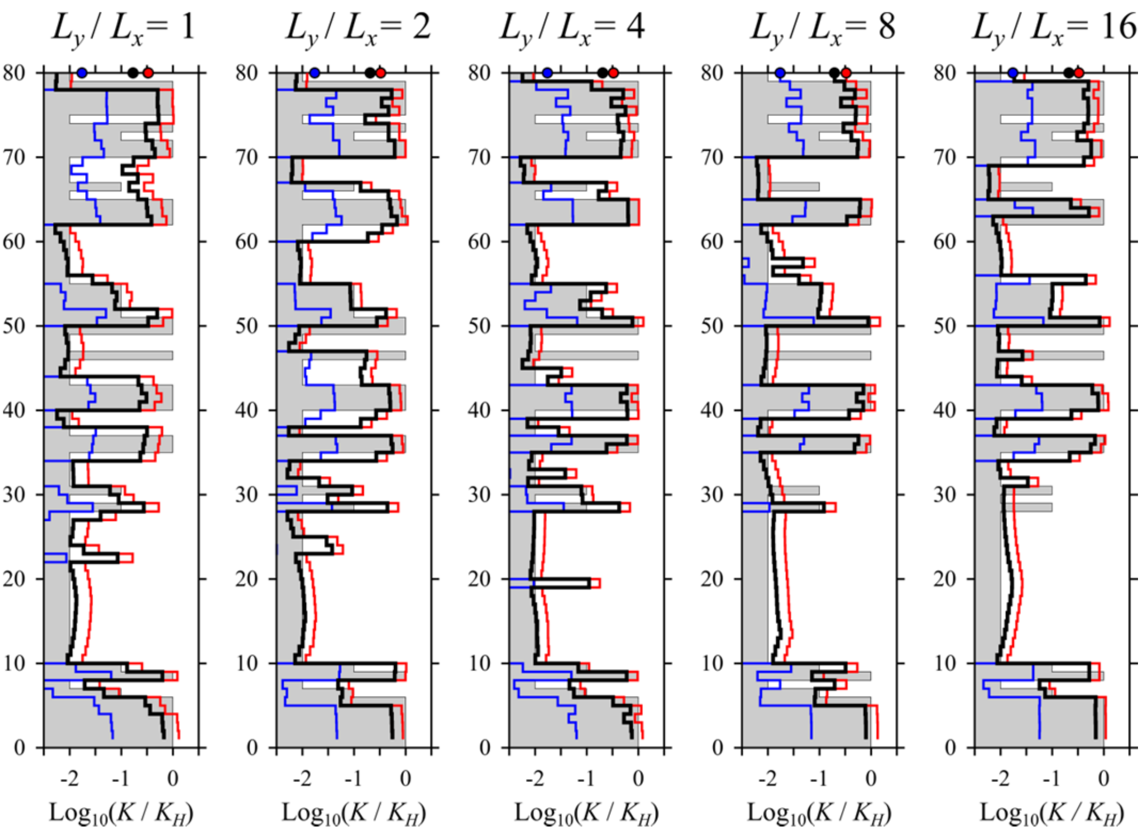

Figure 5. Estimated vs. reference K profiles for scenarios considering different horizontal anisotropy ratios. Refer to the caption of Figure (3) for the meaning of the colours and symbols.

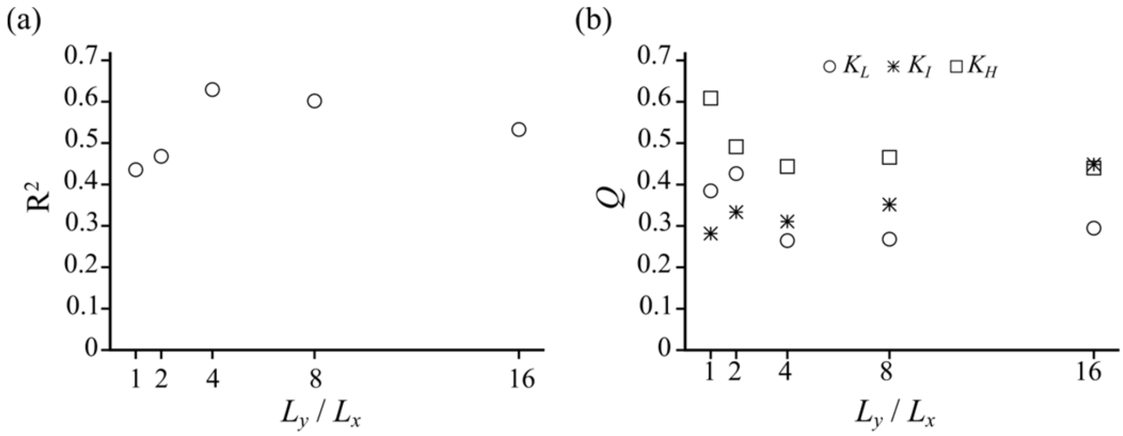

Figure 6. Variations of the accuracy metrics of flowmeter estimates with respect to the anisotropy ratio. (a) Coefficient of determination (R^2). (b) Average accuracy ratio (Q).

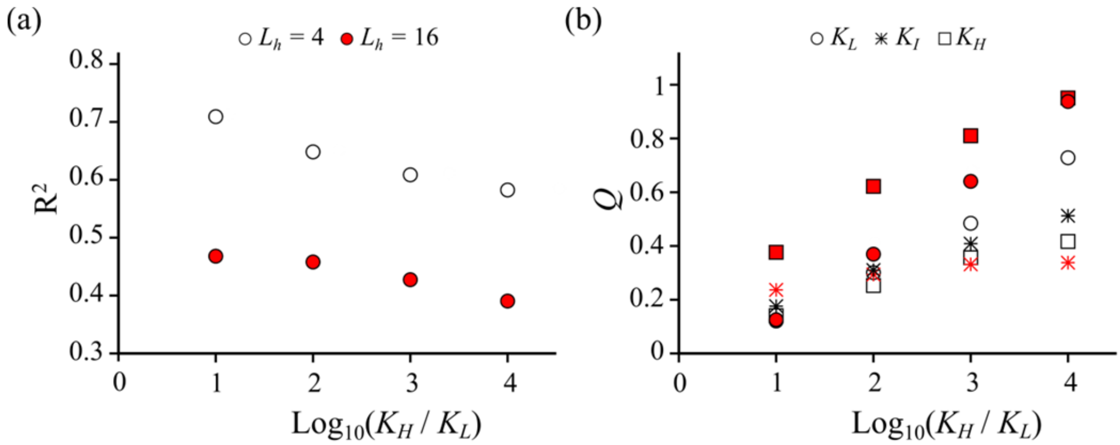

Figure 7. Variation of the accuracy metrics of flowmeter estimates with respect to the K contrasts between hydrofacies "H" and "L" in two architectures. (a) Coefficient of determination (R^2). (b) Average accuracy ratio (Q).The colours of the symbols in (b) correspond the legend of L_h values in (a).

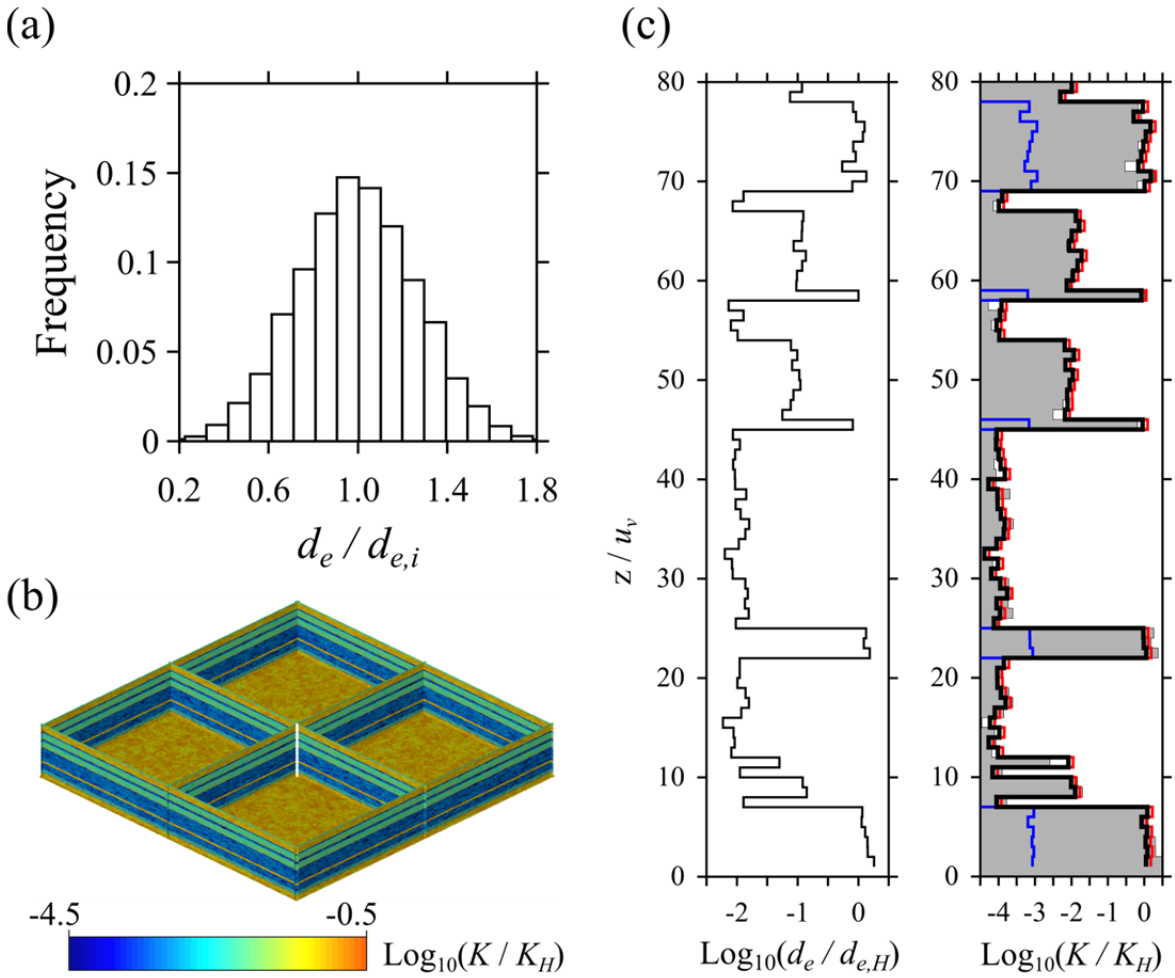

Figure 8. Example of simulation of intrafacies K variability (Scenario Linf_εde08). (a) Histogram of normalized d_e. $d_{e,i}$ is the average effective diameter of a generic hydrofacies i. (b) K field. (c) Vertical profile of normalized d_e (left) and reference vs. flowmeter simulated K profiles (right). Refer to the caption of Figure (3) for the meaning of the colours.

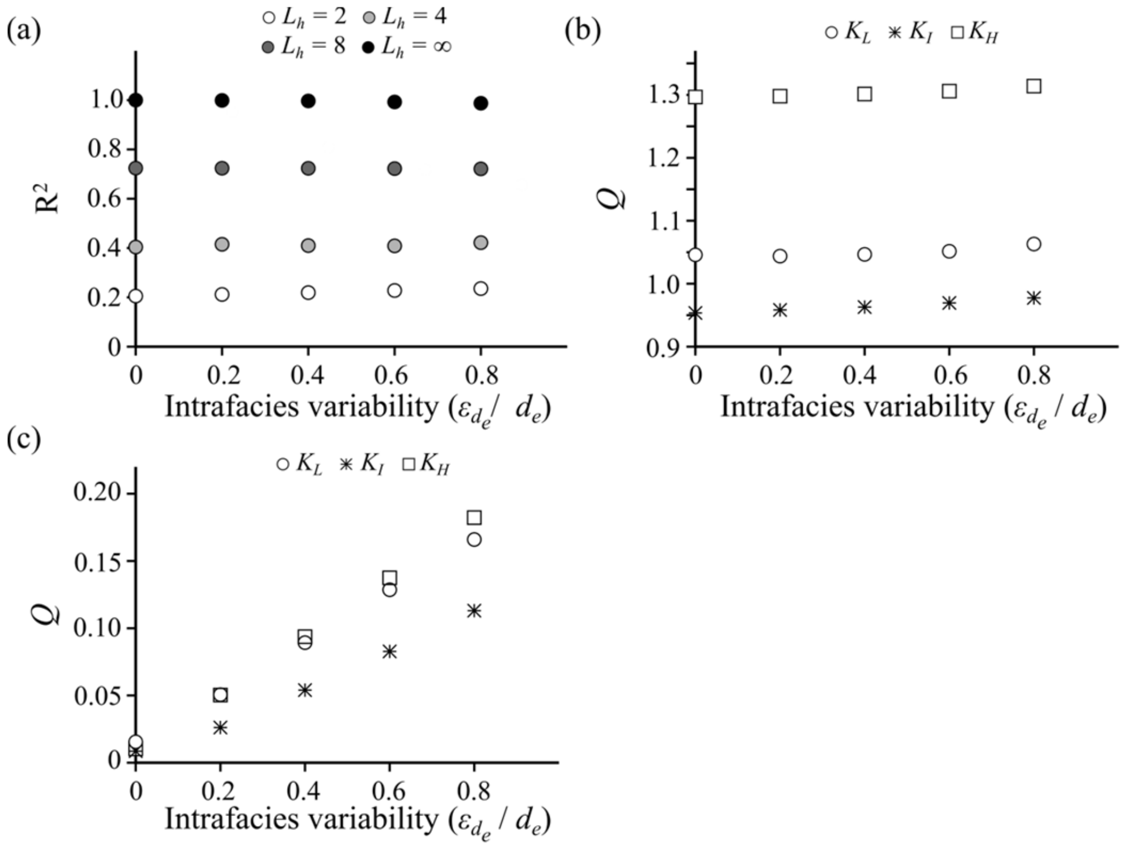

Figure 9. Variation of the accuracy metrics of flowmeter estimates with respect to intrafacies *K* variability. (a) R^2 values for scenarios considering different L_h values. (b) Q values for scenarios considering $L_h = 2$. (c) Q values for scenarios considering $L_h = \infty$.

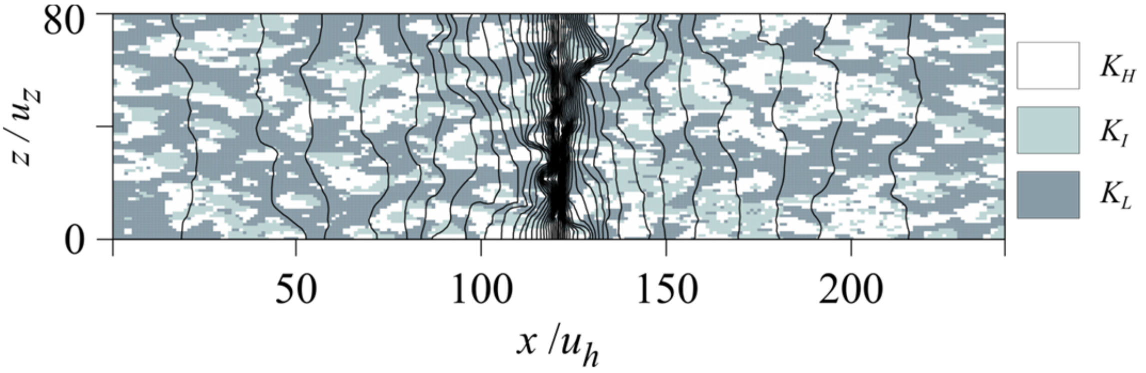

Figure 10. Cross section intersecting the simulated flowmeter test well showing the *K* distribution and simulated head contours for (Scenario L_4).

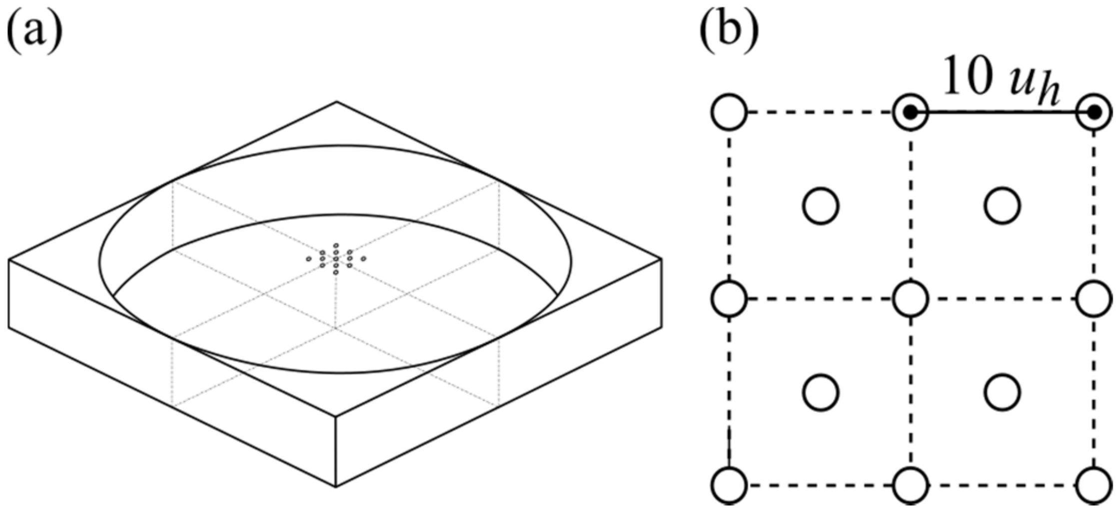

Figure 11. Location of the multiple flowmeter tests simulations.

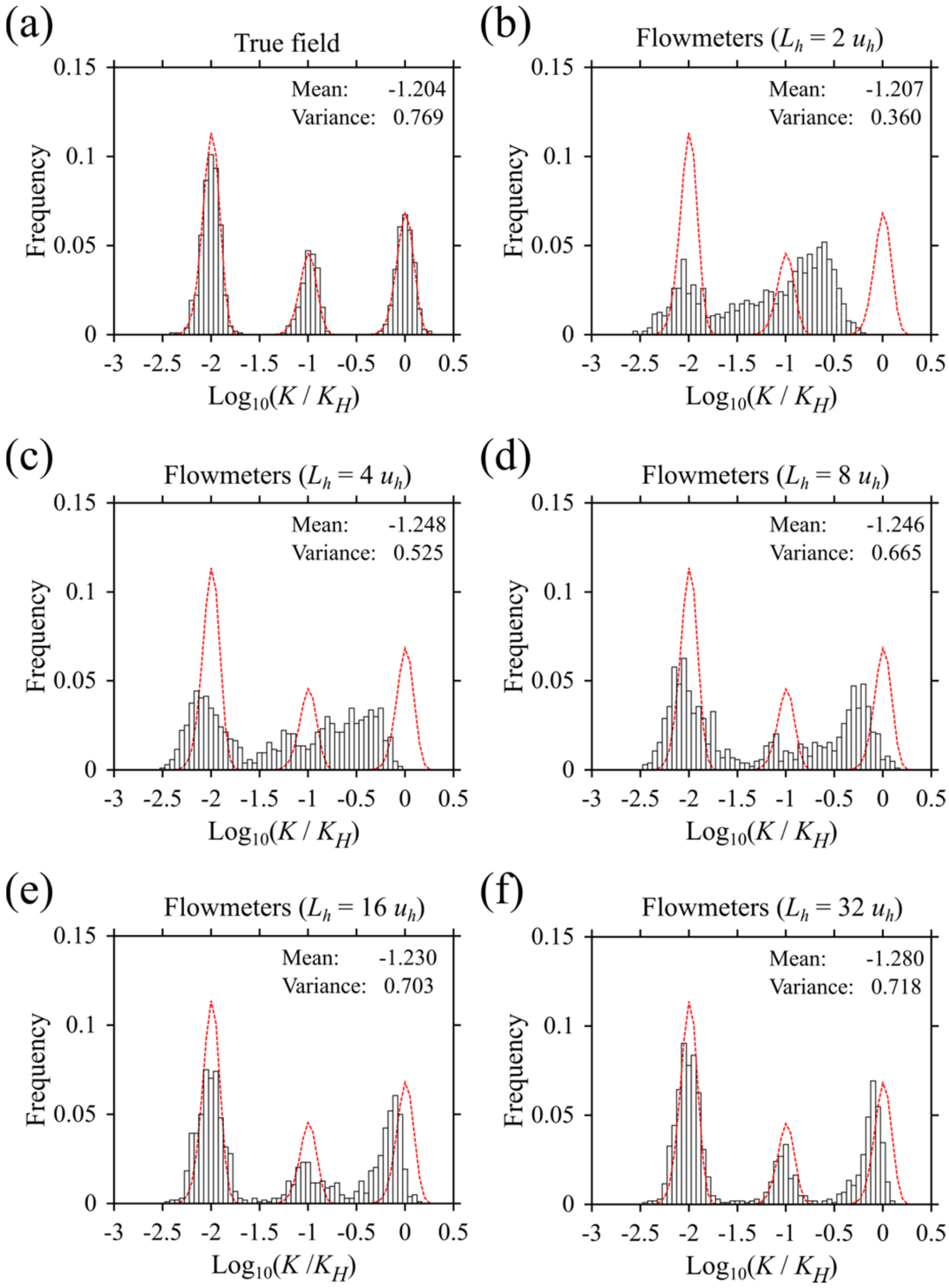

Figure 12. Histograms of the reference and estimated *K* values for different *K* field structures. (a) Reference *K* values sampled from 13 profiles (white bars) vs. reference

values in the entire domain (red dashed line). (b – f). Simulated flowmeter K estimates from 13 locations (white bars) vs. reference values in the entire domain (red dashed line).

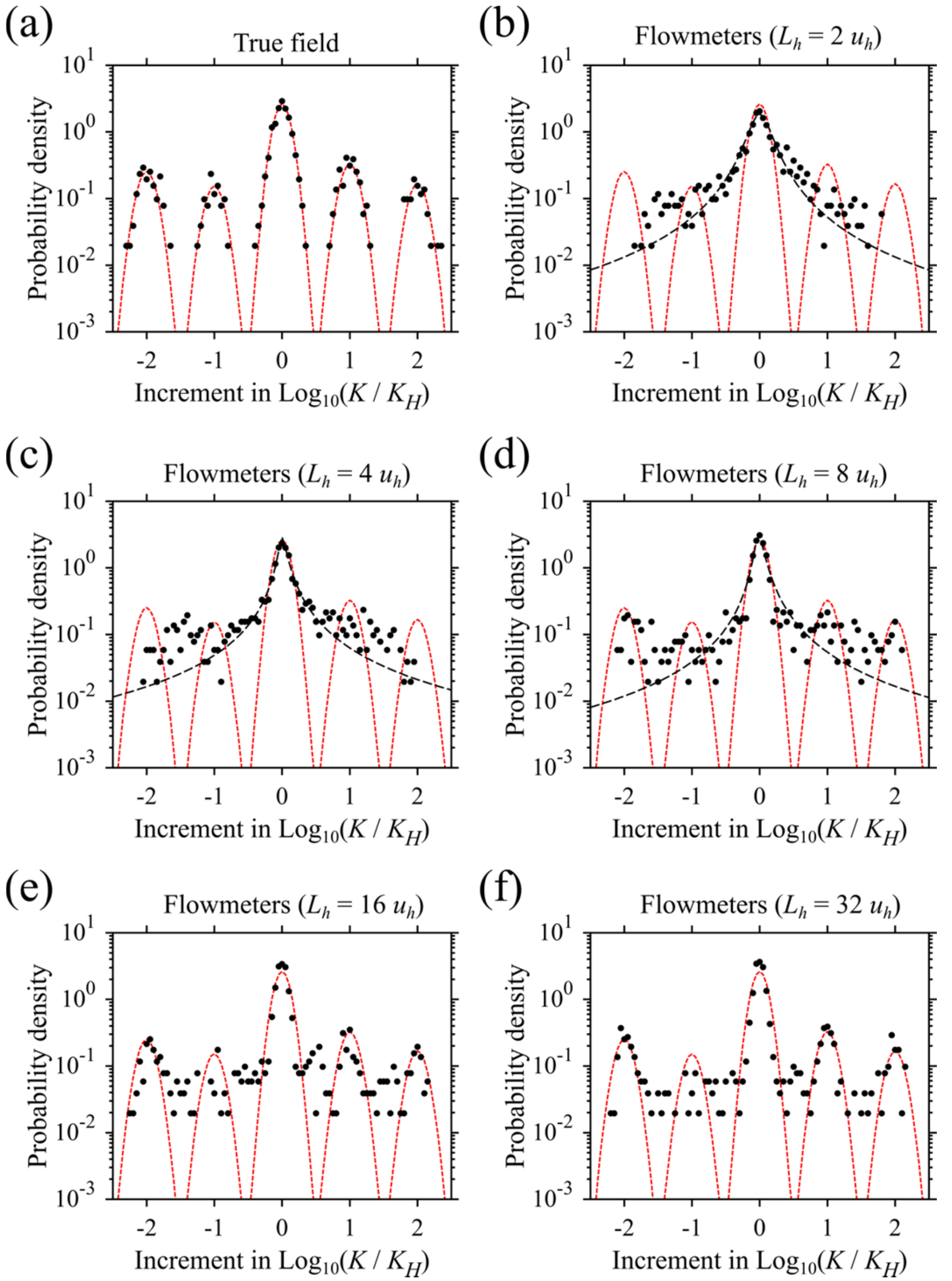

Figure 13. Distribution of the reference and estimated increments in $\text{Log}_{10}(K)$. (a) Increments calculated for reference K values sampled from 13 profiles (black circles) vs. increments calculated for the entire reference K field (red dashed line). (b – f). Increments

calculated for flowmeter *K* estimates from 13 profiles (dots) vs. increments calculated for the entire reference *K* field (red dashed line). Black dashed lines in (b – d) indicate best fitted Levy-stable PDFs. The fitting was performed with an OCTAVE script based on the method of Koutrouvelis (1980, 1981).