AGU PUBLICATIONS

Water Resources Research

RESEARCH ARTICLE
10.1002/2017WR020418

Key Points:
- Use of graph theory to determine connectivity in heterogeneous aquifers
- Relationship between minimum hydraulic resistance and early arrival times of a solute plume
- Definition and computation of a stochastic connectivity measure

Correspondence to:
F. P. J. de Barros,
fbarros@usc.edu

Citation:
Rizzo, C. B., & de Barros, F. P. J. (2017). Minimum hydraulic resistance and least resistance path in heterogeneous porous media. *Water Resources Research*, *53*, 8596–8613. https://doi.org/10.1002/2017WR020418

Received 14 JAN 2017
Accepted 15 SEP 2017
Accepted article online 25 SEP 2017
Published online 30 OCT 2017

Minimum Hydraulic Resistance and Least Resistance Path in Heterogeneous Porous Media

Calogero B. Rizzo[1] and Felipe P. J. de Barros[1]

[1]Sonny Astani Department of Civil and Environmental Engineering, University of Southern California, Los Angeles, CA, USA

Abstract The transport dynamics of a solute plume in porous media are strictly related to the hydrogeological properties. Despite progress in simulation techniques, quantifying transport in strongly heterogeneous geological formations is still a challenge. It is well established that the heterogeneity of the hydraulic conductivity (K) field is one of the main factors controlling solute transport phenomena. Increasing the heterogeneity level of the K-field will enhance the probability of having preferential paths, that are fundamental in predicting the first time arrivals. In this work, we focus on the relationship between the connectivity structure of the K-field to transport quantities. We compute connectivity based on the concept of hydraulic resistance and the corresponding least resistance paths. We present a new efficient algorithm based on graph theory that enables us to extract useful information from the K-field without resorting to the solution of the governing equations for flow and transport. For this reason, an exhaustive and fast analysis can be carried out using a Monte Carlo framework for randomly generated K-fields which allows the computation of the least resistance path and its uncertainty. We examine the minimum hydraulic resistance for both multi-Gaussian (MG) and non-MG log K-fields. The analysis carried out indicates that the expected value of the minimum hydraulic resistance between two points scales exponentially with the standard deviation of the log K-field. Given the strong correlation with plume's first time arrival, our results illustrate how hydraulic resistance and least resistance path can be used as a computationally efficient risk metric.

1. Introduction

Spatial heterogeneity of the hydrogeological properties (e.g., hydraulic conductivity and porosity) has a key role in controlling flow and transport processes. The multi-scale variability of the hydraulic conductivity K causes complex flow patterns which control solute mass fluxes and arrival times. This has been observed in both field (e.g., Bianchi et al., 2011; Zheng et al., 2011) and modeling (e.g., Fiori et al., 2011, 2013; Gómez-Hernández & Wen, 1998; Zinn & Harvey, 2003) studies.

The link between the spatial structure of the flow field and large scale solute transport behavior is well established for aquifers displaying weak to mild heterogeneity in the K-field (e.g., Dagan, 1989; Rubin, 2003). However, establishing this link for highly heterogeneous aquifers is not trivial. The Macrodispersion Experiment (MADE) site located in Columbus, Mississippi (USA) (Boggs, 1990) is one of the examples where high heterogeneity in K can be encountered. Highly heterogeneous K-fields are characterized by the presence of fast flow conduits (i.e., preferential flow paths) and low conductive layers (e.g., Fiori, 2014; Fogg, 1986; Liu et al., 2007; Silliman & Wright, 1988).

Methods for simulating transport in heterogeneous aquifers can be divided into two categories: analytical and numerical methods (see discussion in Chapters 8–10 of Rubin, 2003). In general, analytical solutions are limited to low levels of heterogeneity (e.g., Bellin et al., 1994; de Barros et al., 2015; Fiori, 2001). Many of these analytical solutions are based on perturbation theory. Semi-analytical methods have also been proposed to tackle flow and transport in highly heterogeneous formations (e.g., Cvetkovic et al., 2014; Fiori et al., 2010, 2015). On the other hand, numerical solutions are used to quantify transport in more generic K-fields (e.g., de Dreuzy et al., 2007; Riva et al., 2006). These methods permit to extend many results to high heterogeneity and better understand its impact on anomalous transport (e.g., Le Borgne et al., 2010). The main limitation of numerical methods is the high computational costs associated with simulations of large stochastic systems (e.g., Leube et al., 2013; Moslehi et al., 2015).

The aforementioned challenges and constraints encountered in analytical and numerical methods motivate us to develop alternative approaches able to quantify relevant factors associated with solute transport in aquifers without resorting to the solution of the governing equations for flow and transport. Amongst these factors, we highlight the early time arrivals of a contaminant plume at an environmentally sensitive location. The presence of high K channels is a key factor governing early time arrivals (e.g., Cvetkovic et al., 2012; Deutsch, 1998; Fiori, 2014; Henri et al., 2015; Pardo-Igúzquiza & Dowd, 2003; Trinchero et al., 2008; Tyukhova & Willmann, 2016). Therefore, determining the highly conductive paths between a contaminant source and a receptor is fundamental for risk analysis (e.g., de Barros & Rubin, 2008; de Barros et al., 2016; Maxwell et al., 1999; Tartakovsky, 2007, 2013) due to its correlation with early time arrivals (Gómez-Hernández & Wen, 1998; Tyukhova & Willmann, 2016).

These high K channels lead to the definition of *connectivity*. In general, the connectivity of a field indicates the likelihood that two points are connected by a preferential channel. However, there is not a unique rigorous definition of connectivity and several metrics have been explored (Fernàndez-Garcia et al., 2010; Fiori, 2014; Fiori & Jankovic, 2012; Frippiat et al., 2009; Le Goc et al., 2010; Renard & Allard, 2013; Sánchez-Vila et al., 1996; Tyukhova & Willmann, 2016). Connectivity measures are labeled as static and dynamic (see Knudby & Carrera, 2005). Static connectivity measures are computed using only the K-field and is the subject of interest of the present work. Dynamic connectivity measures use information stemming from the flow field and/or solute transport.

In this paper we develop a computationally efficient framework to identify the connectivity between a source zone and a receptor (which can be represented by distinct combinations of points, lines, areas and volumes). Our goal is to infer dynamic transport quantities (e.g., first arrival times) from static connectivity measures. We aim to find the path that minimizes the hydraulic resistance between the two locations. The *minimum hydraulic resistance* has been successfully used as a static connectivity measure (see Tyukhova & Willmann, 2016) and is equivalent to a spatial distance between two given points. In general, as shown in Tyukhova et al. (2015), the higher the minimum hydraulic resistance is, the more time the solute particle needs to travel between the two points. The curve connecting the two points along which a solute particle encounters less resistance is the *least resistance curve*. The challenge associated to finding the least resistance (shortest or fastest) path is common with other engineering fields, such as traffic or network optimization (e.g., water distribution networks; Rubin, 1987). This class of problems is of key importance in *graph theory* (Bondy & Murty, 1976). Graph theory has also been employed to parametrize spatial hydraulic property characterization (Bhark et al., 2011). By recasting the hydrogeological problem into a graph theory problem, it is possible to find the least resistance path and the corresponding minimum hydraulic resistance using the Dijkstra's algorithm (Dijkstra, 1959). The computational time needed to generate the minimum hydraulic resistance map using this approach is comparable with the time needed to generate a single realization of the K-field using common random field generators. Due to the computational efficiency of the method, we are able to characterize the minimum hydraulic resistance stochastically. Results from our analysis show how the expected value of the minimum hydraulic resistance scales with the heterogeneity of the K-field. We use the methodology to investigate the expected value of the minimum hydraulic resistance in both multi-Gaussian (MG) and non multi-Gaussian (non-MG) log K-fields. As it will be shown, the framework can potentially be utilized to estimate the solute early arrival times at a low computational cost and without resorting to the solution of the flow and transport equations. Therefore, the proposed approach can be used for preliminary screening analysis and adopted in probabilistic risk assessment studies.

This paper is structured as follows. Section 2 presents the procedure to transform the K-field into a graph and illustrate how the Dijkstra's algorithm can be used. In section 3, we explore the relation between the early time arrivals and minimum hydraulic resistance. In section 4, we compute the expected hydraulic resistance of a MG log K-field. Section 5 shows how the method can be employed to compute the minimum hydraulic resistance for connected and disconnected non-MG log K-fields. Finally, section 6 provides a summary and the main conclusions.

2. A Link Between Porous Media and Graph Theory

2.1. Minimum Hydraulic Resistance

Given a source point $\mathbf{s} \in \mathbb{R}^n$ (with $n = 2$ or 3 denoting the spatial dimensionality) and an arrival point $\mathbf{x} \in \mathbb{R}^n$, we define the set of all the possible curves connecting $\mathbf{s}$ to $\mathbf{x}$ as $\mathcal{P}_{\mathbf{s}}^{\mathbf{x}}$. Consider a locally isotropic and

heterogeneous hydraulic conductivity field $K : \mathbb{R}^n \rightarrow \mathbb{R}$. For each $\Gamma \in \mathcal{P}_{\mathbf{s}}^{\mathbf{x}}$ we define the corresponding *hydraulic resistance* as a line integral:

$$\mathcal{R}_\Gamma = \int_\Gamma \frac{1}{K} d\gamma. \tag{1}$$

The hydraulic resistance $\mathcal{R}_\Gamma$ has units of time and it is an indicator of the resistance found by a solute particle traveling along Γ. A curve passing through a high conductivity zone will likely have a low hydraulic resistance, meaning that the time needed to reach the end point will be reasonably low. For this reason, it is of interest to find the curve that minimizes the hydraulic resistance within $\mathcal{P}_{\mathbf{s}}^{\mathbf{x}}$. Formally, we define the *minimum hydraulic resistance* from a point **s** to **x** as:

$$\mathcal{R}_{\mathbf{s}}(\mathbf{x}) = \min_{\Gamma \in \mathcal{P}_{\mathbf{s}}^{\mathbf{x}}} \mathcal{R}_\Gamma \tag{2}$$

Consequentially, the curve $\hat{\Gamma} \in \mathcal{P}_{\mathbf{s}}^{\mathbf{x}}$ that satisfies $\mathcal{R}_{\hat{\Gamma}} = \mathcal{R}_{\mathbf{s}}(\mathbf{x})$ is called the *least resistance curve*. Note that it may be possible to have multiple least resistance curves for a pair of points although in practice, the probability of this event is very low.

Alternatively, we consider a source control volume $A \subset \mathbb{R}^n$ and a target control volume $B \subset \mathbb{R}^n$. Using (2), the *minimum hydraulic resistance* from A to B is defined as:

$$\mathcal{R}_A(B) = \min_{\mathbf{a} \in A} \left[\min_{\mathbf{b} \in B} \mathcal{R}_{\mathbf{a}}(\mathbf{b}) \right]. \tag{3}$$

We point out that the definitions above allow to define the minimum hydraulic resistance between two generic geometries (e.g., a pair of points, an injection plane to a point, a point to a control plane or an injection plane to a control plane).

2.2. Graph Approximation

It is possible to find an analytical solution of equation (2) only for simple scenarios such as homogeneous media (i.e., constant K). In general, finding the minimum resistance (and the corresponding least resistance curve) in a continuum framework is a difficult task since it requires to explore all the possible curves connecting two points. For this reason we propose the following graph-based procedure to develop an approximation for equations (2) and (3).

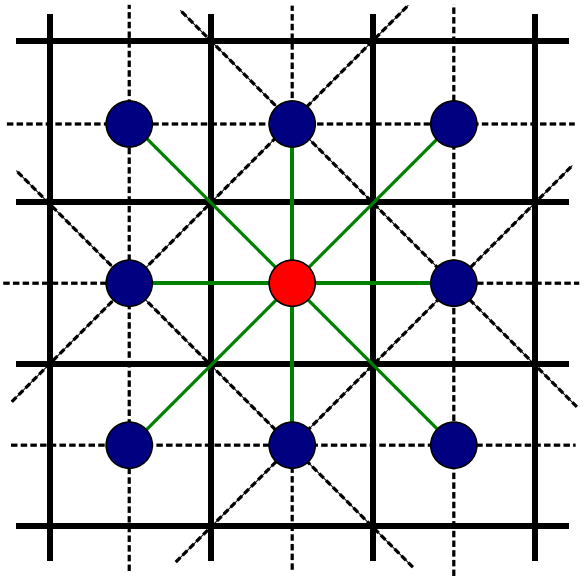

Figure 1. Scheme used to construct the hydraulic resistance graph $R(V, E)$. The circles represent the vertices of the graph, one for each cell. Each vertex is connected to the neighbors through edges. For instance, the red vertex is connected to the neighbors through the green edges.

Using the standard graph theory notation (Bondy & Murty, 1976), we define a graph $G(V, E)$ as a pair of a set V of vertices and a set E of edges. Each edge $e \in E$ connects two vertices of V and it can have a weight w_e (e.g., distance, time or cost). In this paper we will use only *undirected* graphs, meaning that the edges have no orientation (i.e., if a vertex v_i is connected to v_j then v_j is connected to v_i). Finally, a *path* is a sequence of vertices such that adjacent vertices are connected by edges.

In order to have a graph representation of the hydraulic conductivity fields, we partition the domain into cells as shown in Figure 1. The hydraulic conductivity field is discretized such that each cell i is characterized by a hydraulic conductivity value K_i. Note that this domain partition is the same used in classic Finite Volume or Finite Difference methods. The coordinates of the center of each cell represents a vertex of the graph. Two cells are considered neighbors if they share a common face or a common corner. Each vertex is connected to a neighbor vertex through an edge as shown in Figure 2. The weight w_e associated to an edge e connecting two neighboring vertices v_i and v_j is:

$$w_e = \frac{|\mathbf{r}_{ij}|}{K_i} + \frac{|\mathbf{r}_{ji}|}{K_j}, \tag{4}$$

where $|\mathbf{r}_{ij}|$ is the length of the segment connecting the coordinates of v_i with either the center of the face between the cells i and j or the corner coordinates between i and j. This procedure generates the *hydraulic resistance graph* $R(V, E)$

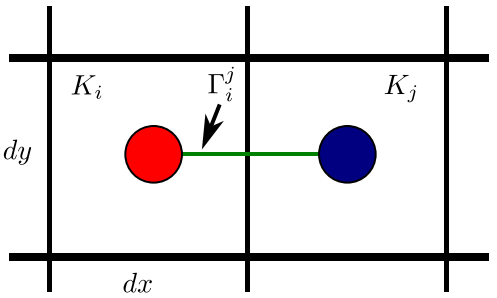

Figure 2. Zoom of two vertices connected by one edge Γ_i^j. The weight of the edge is computed according to the cell conductivities K_i and K_j as in Equation (4).

where V contains as many vertices as the number of cells and E contains the weighted edges as defined in (4). In case of structured grids, the vertex corresponding to an internal cell has exactly 8 neighbors in 2-D and 26 neighbors in 3-D domains. Equation (4) can be extended to the more generic case of anisotropic conductivity tensor using $w_e = |\mathbf{K}_i^{-1} \cdot \mathbf{r}_{ij}| + |\mathbf{K}_j^{-1} \cdot \mathbf{r}_{ji}|$.

We can use the hydraulic resistance graph to find an approximation of the minimum resistance defined in (2). Assuming that the source point **s** and the arrival point **x** are the center coordinates of two cells with corresponding vertices s and x, we can define the set of paths $\overline{\mathcal{P}}_s^x$ from s to x on the graph $R(V, E)$. Since for each path we can build a piecewise linear curve starting from the center of the cell corresponding to s to the center of the cell corresponding to x, the set $\overline{\mathcal{P}}_s^x$ represents a subset of all the possible curves $\mathcal{P}_{\mathbf{s}}^{\mathbf{x}}$. Therefore the approximated minimum hydraulic resistance is given by:

$$\overline{\mathcal{R}}_s(x) = \min_{\bar{\Gamma} \in \overline{\mathcal{P}}_s^x} \mathcal{R}_{\bar{\Gamma}}, \tag{5}$$

where $\mathcal{R}_{\bar{\Gamma}}$ is defined in (1). Using the discrete isotropic hydraulic conductivity field and the weights (4), we obtain a simplified expression of the hydraulic resistance:

$$\mathcal{R}_{\bar{\Gamma}} = \int_{\bar{\Gamma}} \frac{1}{K} \mathrm{d}\gamma = \sum_{e \in \bar{\Gamma}} w_e, \tag{6}$$

where $e \in \bar{\Gamma}$ are all the edges in $\bar{\Gamma} \in \overline{\mathcal{P}}_s^x$. Therefore, the approximated minimum hydraulic resistance becomes:

$$\overline{\mathcal{R}}_s(x) = \min_{\bar{\Gamma} \in \overline{\mathcal{P}}_s^x} \sum_{e \in \bar{\Gamma}} w_e. \tag{7}$$

Note that in this formulation there exists at least one path $\hat{\bar{\Gamma}} \in \overline{\mathcal{P}}_a^b$ such that $\mathcal{R}_{\hat{\bar{\Gamma}}} = \overline{\mathcal{R}}_a^b$ since the number of possible paths is finite. The quality of this approximation can be controlled by the refinement level of the grid. Decreasing the cell size increases the number of possible paths between cells.

Similarly to (3), we can extend (7) for the case of multiple source vertices and multiple target vertices. Given two sets of vertices $\bar{A} \subset V$ and $\bar{B} \subset V$, we add two virtual vertices a and b to the vertices set V. The vertex a is connected to all the vertices in $\bar{A}$ through edges with 0 resistance. By the same token, the vertex b is connected to all the vertices in $\bar{B}$. Therefore, the approximated minimum hydraulic resistance is:

$$\overline{\mathcal{R}}_{\bar{A}}(\bar{B}) = \min_{\bar{\Gamma} \in \overline{\mathcal{P}}_a^b} \mathcal{R}_{\bar{\Gamma}}. \tag{8}$$

2.3. Computation of the Minimum Hydraulic Resistance

The value $\overline{\mathcal{R}}_s(x)$ defined in (7) is the solution of a minimization problem in a graph theory framework. In other words, we need to find a path going from a vertex s to x that minimizes the sum of the weight of its edges. This problem can be solved using the Dijkstra's algorithm (Dijkstra, 1959). The pseudo code to find the minimum hydraulic resistance using the Dijkstra's algorithm is:

```
1: function Dijkstra(R(V, E), A)
2:     W = set of all the vertices V;
3:     for all vertices v ∈ W do
4:         if v ∈ A then
5:             res(v) = 0
6:             pre(v) = undefined
7:         else
8:             res(v) = ∞
9:             pre(v) = undefined
10: while W is not empty do
11:     c = vertex in W with the smallest resistance;
12:     remove c from W;
13:     for all neighbor n of c do
```

14: $r = res(c) + res(c, n)$
15: **if** $r < res(n)$ **then**
16: $res(n) = r$
17: $pre(n) = c$

Details about the implementation can be found in Bondy and Murty (1976) and Thomas et al. (2001). Despite its simplicity, the Dijkstra's algorithm has several features motivating the extensive usage in different fields, such as transport networks and Artificial Intelligence development. The first and probably most important feature of this algorithm is that it is computationally efficient. The optimal worst-case performance is $O(|E|+|V|\log|V|)$, meaning that the algorithm scales well with the number of vertices and edges. Note that the worst-case performance of the original algorithm formulated by Dijkstra is $O(|V|^2)$. In order to achieve a better performance, we need to use a special data structure denoted as min-priority queue so that the operation complexity of selecting the vertex with minimum resistance is $O(1)$ (Fredman & Tarjan, 1987). The second feature is that the algorithm implicitly computes the minimum resistance and the associated least resistance path for all the vertices of the graph. In all our tests, the computational time needed to compute the minimum hydraulic resistance map is less than the time needed to generate a random *K*-field using sequential Gaussian simulators, with a peak of 1.86 seconds for a 2-D field with 1 million cells (using a CPU with 3.40 GHz). Therefore, the time required for the computation of the minimum hydraulic resistance map is negligible compared to the full numerical solution of the flow and/or transport equations. Finally, the algorithm can be applied to any *K*-field using different source and target geometries.

2.4. Algorithm Output and Graph Refinement

Using the algorithm presented in section 2.3, we are able to compute the minimum hydraulic resistance $\overline{\mathcal{R}}_{\bar{A}}(x)$ defined in (8) from a subset of vertices $\bar{A}$ to each vertex x. Since each vertex of the graph is linked to a cell of the grid, we can map the minimum hydraulic resistance to the physical domain obtaining a discrete approximation of $\mathcal{R}_A(\mathbf{x})$ defined in (3) from a subset of the domain A to each point $\mathbf{x}$. Moreover, for each vertex of the graph we are able to retrieve the neighbor vertex where the least resistance path passes through. Thus we can recursively build the least resistance path to each vertex of the graph and therefore map each path to a curve of the physical domain.

The algorithm presented in this paper can compute the minimum hydraulic resistance in two-dimensional (2-D) and three-dimensional (3-D) *K*-fields. With the purpose of illustration, the minimum hydraulic resistance of a 3-D heterogeneous *K*-field is shown in Figure 3. Figure 3 displays the *K*-field (left plot of Figure 3) used to compute the minimum hydraulic resistance (right plot of Figure 3) between the center point of the 3-D formation and all the points of the domain. In the upcoming illustrations, we will focus on 2-D examples.

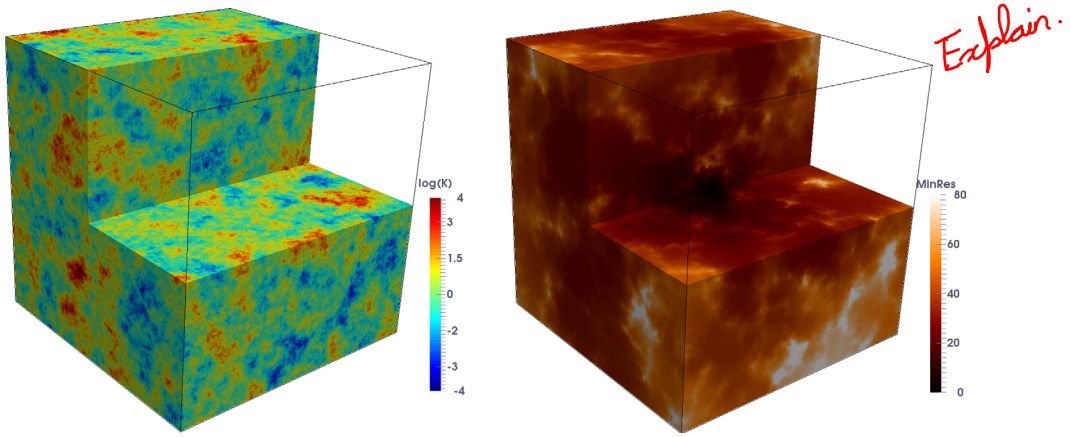

Figure 3. Illustration of the method in a 3-D spatially heterogeneous multi-Gaussian hydraulic conductivity field. The natural logarithm of the hydraulic conductivity (left) and the corresponding minimum hydraulic resistance map (right) obtained with the graph theory algorithm are displayed. The colors on the minimum hydraulic resistance map (right) indicate the resistance between the center of the domain and all remaining points in the 3-D heterogeneous porous medium. The 3-D domain contains 201×201×201 cells.

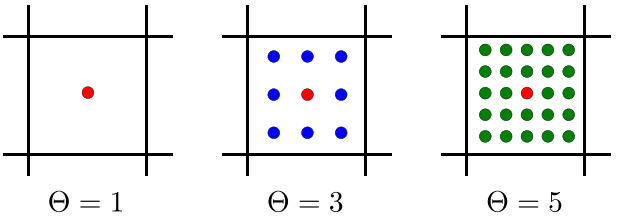

Figure 4. Example of graph refinement with Θ=1, 3 and 5.

The approximation quality of the minimum hydraulic resistance can be controlled by refining the grid. A finer grid provides more discrete paths connecting any two vertices of the graph. However, once the size of the cells is chosen, it is possible to refine the graph while keeping the same grid resolution. An example can be found in Figure 4. The graph refinement level Θ represents the number of vertices per cell in each of the principal directions. It is clear that increasing Θ will increase the approximation quality by adding possible paths to the conductivity graph.

3. Minimum Hydraulic Resistance and Solute Transport

In this section, we explore the relation between the minimum hydraulic resistance and early time arrivals of a contaminant plume. With the purpose of illustration, we consider a 2-D fully saturated porous medium ($n = 2$). We will compare the results obtained using the algorithm described in section 2 with flow and transport simulations.

3.1. Random Field Generation and Physical Formulation

We consider the hydraulic conductivity K of the medium to vary in space $\mathbf{x}=(x,y)$ and, without loss of generality, being statistically stationary and represented by its Random Space Function (RSF). The RSF is characterized by the statistics of the log-conductivity multi-Gaussian field, $Y=\log K$, namely, its mean μ_Y (and geometric mean $K_G=e^{\mu_Y}$), variance σ_Y^2 and isotropic integral scale I_Y. The spatial covariance model of Y is denoted by $\mathcal{C}_Y$. In this work, we will adopt an exponential covariance model such that:

$$\mathcal{C}_Y(r)=\sigma_Y^2 \exp(-r/I_Y), \tag{9}$$

where r is the lag distance between two points. All the simulations presented in this paper use random K-fields generated with HYDROGEN (Bellin & Rubin, 1996). When multiple realizations are needed, we generate fields with $\mu_Y=0$ and $\sigma_Y^2=1$. In order to have fields with a given $\sigma_Y^2 \neq 1$, we multiply each field by the desired σ_Y (e.g., Jankovic et al., 2016).

We consider a steady-state flow in a rectangular domain of dimensions L_x and L_y. The governing equation for the flow field is:

$$\nabla \cdot [K(\mathbf{x})\nabla\varphi(\mathbf{x})]=0, \tag{10}$$

where φ is the hydraulic head. The following boundary conditions are employed:

$$\begin{aligned} \varphi(x=0,y) &= \varphi_L, \quad \varphi(x=L_x,y) = \varphi_R, \\ \left.\frac{\partial\varphi}{\partial y}\right|_{y=0} &= 0, \quad \left.\frac{\partial\varphi}{\partial y}\right|_{y=L_y} = 0. \end{aligned} \tag{11}$$

where φ_L and φ_R are prescribed hydraulic heads. The boundary conditions defined in (11) ensure a longitudinal uniform-in-the-mean flow along the x-direction. The velocity field $\mathbf{v}$ is obtained through Darcy's law:

$$\mathbf{v}(\mathbf{x})=-\frac{K(\mathbf{x})}{n_e}\nabla\varphi(\mathbf{x}) \tag{12}$$

where n_e is the effective porosity.

For the transport problem, we consider an inert tracer released from a line source on the left boundary and aligned with the y-axis. Transport is governed by the advection dispersion equation:

$$\frac{\partial c(\mathbf{x},t)}{\partial t}+\mathbf{v}(\mathbf{x})\cdot\nabla c(\mathbf{x},t)=D\nabla^2 c(\mathbf{x},t), \tag{13}$$

where c is the resident concentration of the tracer and D is the local scale dispersion coefficient (assumed to be constant and isotropic).

The flow equation (10) is solved with the Finite Volume method using the Python library FiPy (Guyer et al., 2009). The transport equation (13) is solved through particle tracking technique following the procedure described in the literature (e.g., Henri & Fernàndez-Garcia, 2014; Salamon et al., 2006). Within this section,

the domain size is $L_x \times L_y = 20I_Y \times 10I_Y$ and 100 hydraulic conductivity fields are generated using $\sigma_Y^2 = 1$ and 100 fields with $\sigma_Y^2 = 4$. A line source (100,000 particles) is placed along all the left boundary (i.e., the size of the vertical line source is equal to L_y) and the particles are evenly distributed along its length. The Péclet number is $Pe = 1{,}000$ where $Pe = UI_Y/D$ and $U = K_G(\varphi_L - \varphi_R)/L_x$ is the mean longitudinal velocity. The parameters are chosen such that the results are in dimensionless units (i.e., $K_G = 1$, $\varphi_L - \varphi_R = L_x$ and $U = 1$).

3.2. Leading Front of the Plume

Although the minimum hydraulic resistance is computed using only the K-field, it can provide useful information related to flow and transport. The front of a plume moving through a porous media is strongly affected by the conductivity distribution. The trajectory that offers less hydraulic resistance will facilitate the passage of the solute. In other words, there is a relation between the least resistance path computed in section 2 and the actual leading front of the plume.

In Figure 5 we compare the numerical solution of the heterogeneous flow and transport model with the least resistance path between left and right boundaries for 9 randomly generated fields. The least resistance paths (red lines) are computed starting from the minimum hydraulic resistance $\mathcal{R}^*(y) = \mathcal{R}_{\Phi_L}(L_x, y)$ between the left boundary of the domain Φ_L and all points within the right boundary of the domain. Subsequently, the coordinate (L_x, y) with minimum resistance is chosen to be the ending point of the least resistance path. Finally, we backtrack the least resistance path using the output of the algorithm described in section 2. Figure 5 shows that the trajectory of the fastest particle computed with the particle tracking code (blue lines) hits the right boundary in the location predicted by the least resistance path in most of the cases. As we can see from the center and center-left panels of Figure 5, for the cases where multiple leading plume edges reach the right boundary of the domain (i.e., the presence of multiple channels with similar low resistance), the least resistance path will point toward one of them, but not necessarily to the final location of fastest particle.

Figure 6 statistically analyzes the normalized difference between the end of the least resistance path y_Γ and the final location of the fastest particle y_t (obtained with the particle tracking code) on the right boundary of the domain for $\sigma_Y^2 = 1$ and 4 over multiple random realizations of the K-fields. In most cases, the difference is approximately zero thus indicating that the least resistance path can be potentially a good estimator of the fastest particle's final location. However, the box plots shown in Figure 6 display the presence of outliers. As previously mentioned, these outliers are an outcome of the presence of multiple low resistance

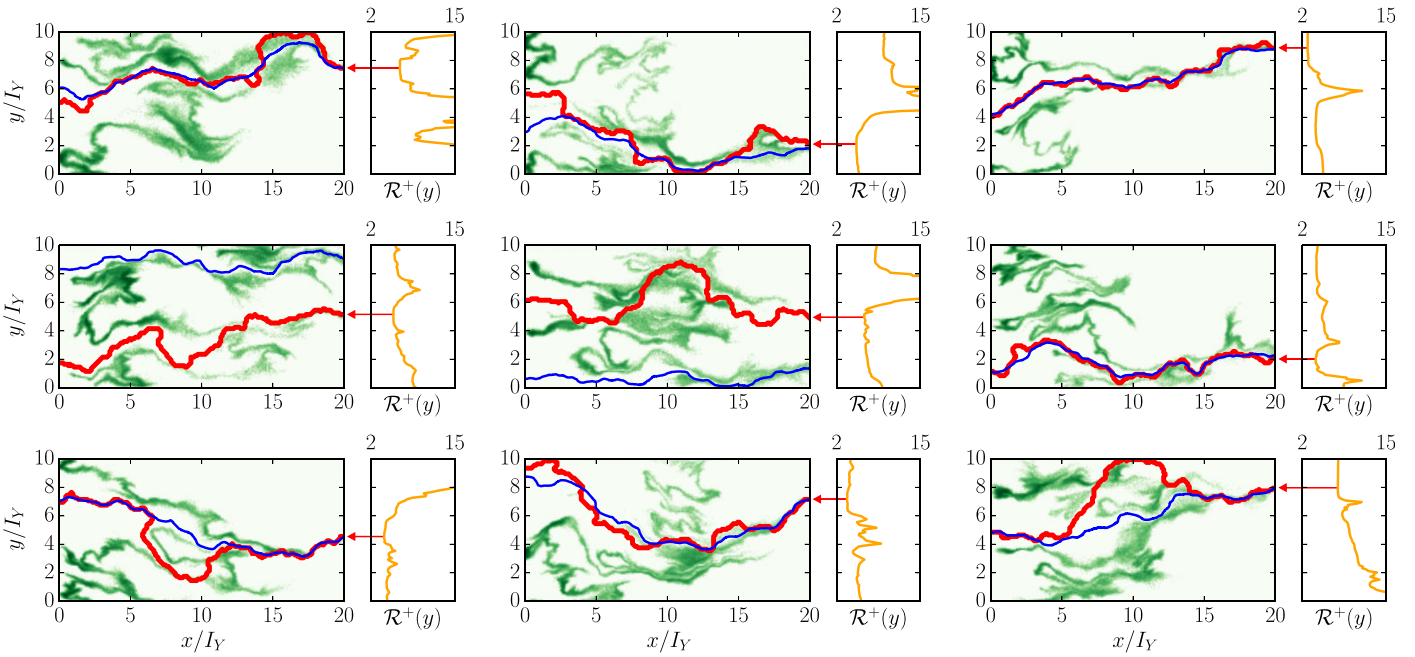

Figure 5. Snapshots of a particle tracking simulation for 9 randomly generated fields ($\sigma_Y^2 = 4$). The least resistance path (red line) is compared with the trajectory of the fastest particle (blue line). On the right side of each panel, the normalized minimum hydraulic resistance from the left boundary to each point of the right boundary is displayed $\mathcal{R}^+(y) = \mathcal{R}^*(y)K_G/I_Y$.

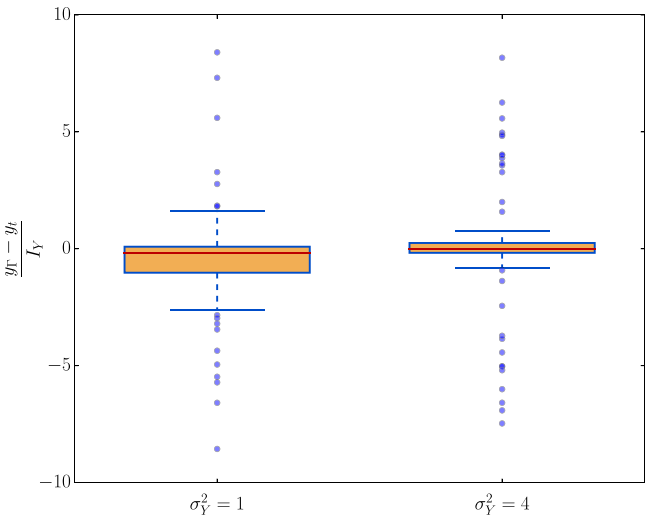

Figure 6. Box plots of the normalized difference between the end of the least resistance path y_Γ and the final location of the fastest particle y_t on the right boundary. The location y_t is computed by simulating flow and transport. In both cases (σ_Y^2=1 and 4), 100 realizations of the hydraulic conductivity field have been used.

channels. It is interesting to note that the box plot for σ_Y^2=4 is narrower than the box plot for σ_Y^2=1 (Figure 6). The reason is that when heterogeneity is higher (i.e., larger values of σ_Y^2), there is an increased probability in the occurrence of a single dominant low resistance channel within the *K*-field. The relationship between the actual trajectory (obtained via flow and transport simulation) and the least resistance path is subject to multiple dynamic quantities that are not embedded in the minimum hydraulic resistance formulation (which is a static quantity by definition) such as dispersivities, boundary conditions and initial condition (e.g., source zone and release history). However, the least resistance path provides a good approximation of the fastest particle trajectory, showing that heterogeneity of the *K*-field is the driving factor governing the leading front of the plume in the cases taken into consideration.

This simple example helps to understand the relation between the leading edge of the plume and the hydraulic resistance, a quantity that can be efficiently computed. Fastest particles tend to flow toward areas with small resistance and motivates the usage for predicting early time arrivals. As we will see in the next subsection, locations with small minimum hydraulic resistance are correlated with first time arrivals.

3.3. Early Time Arrivals

The visual correlation depicted in Figure 5 between the least resistance path and the leading front of the plume indicates that there is a relation between hydraulic resistance and first time arrival. Thus, we compute the time for which 1% of the particles have crossed the right boundary for each of the random field generated. The first arrival time is therefore denoted by $t_{1\%}$.

Figure 7 shows the scatter plot between $t_{1\%}$ and the minimum hydraulic resistance between left and right boundaries of the domain. Even if the model is affected by the presence of boundary conditions, the minimum hydraulic resistance shows a good correlation with early time arrivals for both levels of heterogeneity. We remark that the computation of the minimum hydraulic resistance does not require to solve the flow and transport equations. Similar results have been found by Tyukhova and Willmann (2016) for different kind of fields and using a procedure based on erosion and dilation of a spatial set (Tyukhova et al., 2015). From a risk assessment point of view, the minimum hydraulic resistance has the potential to predict, with some level of uncertainty, the early time arrivals of a given contaminant to a environmentally sensitive area. In Appendix A, we mathematically demonstrate that, under certain conditions (i.e., when the distance between the two control volumes is very large and for a conservative tracer), the minimum hydraulic resistance can provide a lower bound estimate of the first arrival times.

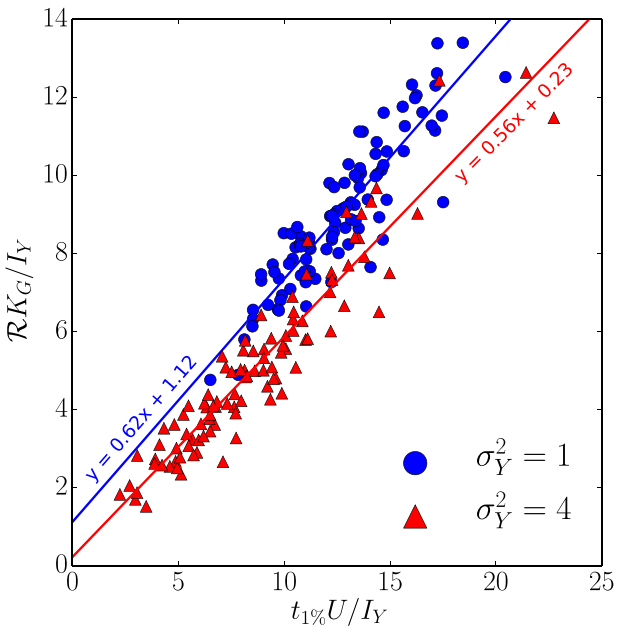

Figure 7. Comparison between early time arrivals ($t_{1\%}$) and minimum hydraulic resistance between left and right boundaries ($\mathcal{R}$). The coefficient of determination of the regression line is R^2=0.815 for σ_Y^2=1 and R^2=0.893 for σ_Y^2=4.

4. Statistical Properties of Minimum Hydraulic Resistance

Given that an extensive site characterization of the subsurface is not feasible, the hydraulic conductivity field is subject to uncertainty. As a consequence, the minimum hydraulic resistance must be treated as a random variable. Using the geostatistical structural parameters defining the spatial variability of the *K*-field and taking advantage of the computational efficiency associated with Dijkstra's algorithm, it is possible to compute the statistical properties of the minimum hydraulic resistance using a Monte Carlo framework. Within this section, the minimum hydraulic resistance is considered to be a stochastic quantity at each location **x**.

For a given starting point **s**, we can compute the hydraulic resistance between **s** and all the other points **x** of the domain using equation (2). If the ensemble

of K is unconditioned on data, the expected minimum hydraulic resistance only depends on the distance $r=|\mathbf{x}-\mathbf{s}|$. Note that radial symmetry only holds for the minimum hydraulic resistance as a stochastic quantity (i.e., $\mathcal{R}_{\mathbf{s}}(\mathbf{x}_1)$ and $\mathcal{R}_{\mathbf{s}}(\mathbf{x}_2)$ have the same pdf if $\mathbf{x}_1$ and $\mathbf{x}_2$ have the same distance from $\mathbf{s}$).

Several hydraulic conductivity fields were generated following the procedure described in the previous section with the covariance model defined in (9) and different values of σ_Y^2. In the upcoming analysis, the domain size is $L_x \times L_y = 30I_Y \times 30I_Y$ with 10 cells per integral scale I_Y in each direction. The source point $\mathbf{s}$ is located at the center of the domain. For each realization of K and for each point of the domain, we compute the minimum hydraulic resistance from the center point $\mathbf{s}$.

4.1. Single Realization Analysis

Before computing the expected value of the minimum hydraulic resistance, we compare K-fields generated with different σ_Y^2. Our goal here is to understand how the connectivity structure (e.g., high connectivity channels) affects the minimum hydraulic resistance. Figure 8 illustrates the minimum hydraulic resistance maps (Figures 8d–8f) generated by different K-field realizations with variance $\sigma_Y^2=0.5$ and 4.5 (Figures 8a–8c). It is clear that the minimum hydraulic resistance map tends to be radially symmetric in low heterogeneous fields (e.g., compare Figures 8d and 8e). The presence of high conductivity areas in strongly heterogeneous fields leads to the formation of preferential channels in which the hydraulic resistance is much lower (see Figures 8e and 8f).

Since the random generated fields are unconditional, the center point of the domain might fall in either a low or high conductivity area as shown in Figures 8b and 8c for $\sigma_Y^2=4.5$. In the first case (Figure 8e), the least resistance path to a given point of the domain must travel through this low conductivity zone until it reaches a high conductivity area. In the second case (Figure 8f), the least resistance path starts from a high

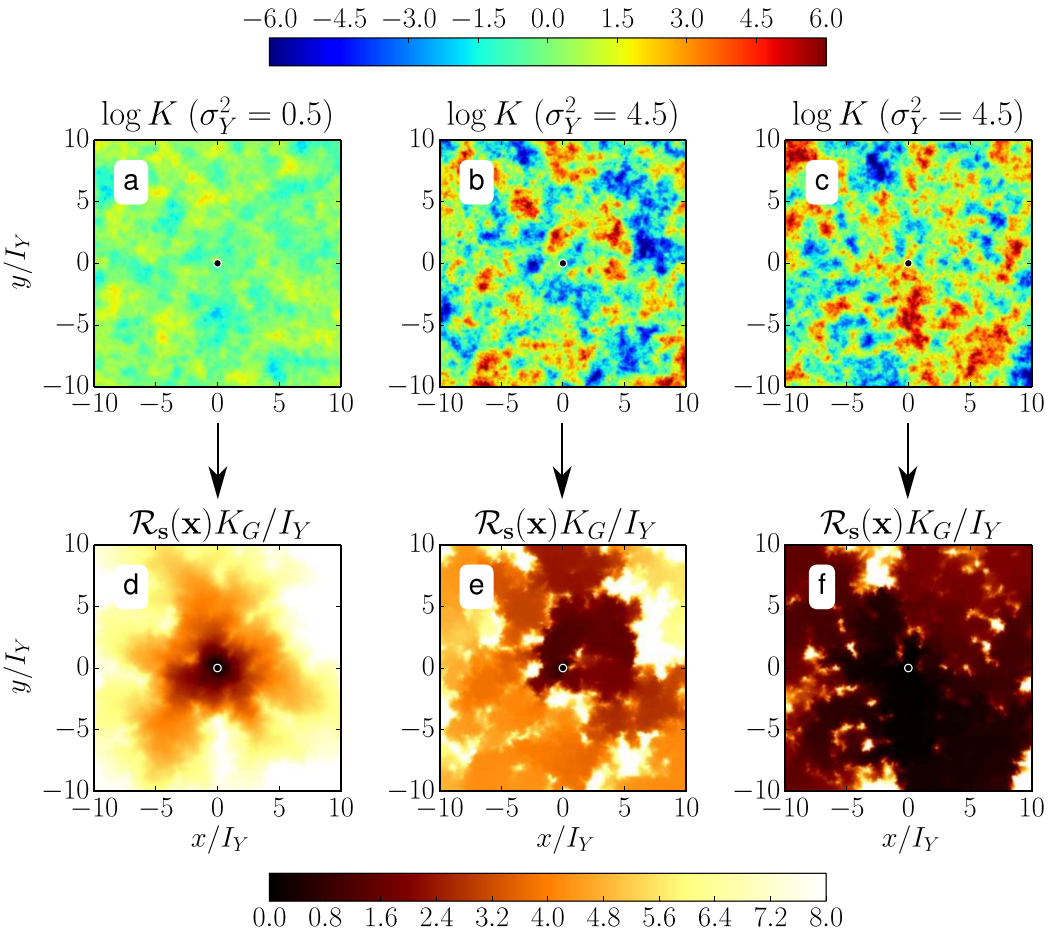

Figure 8. Minimum hydraulic resistance from the center (bottom) obtained using three different K-fields (top).

conductivity zone. In general, if the center point is placed in a high conductivity area, the overall resistance is low since we can find a path from the center to a given point that passes only through high conductivity channels. Therefore, the minimum hydraulic resistance maps shown in 8.e and 8.f are very different even if they are generated using the same *K*-field statistical properties. To better understand this behavior, it is convenient to look at the equivalent graph theory problem defined in equation (7). The shortest path problem can be viewed as a constrained minimization problem; these constraints are the initial and end points. Therefore, the result of the minimization algorithm is highly sensitive to areas surrounding the start and end points of interest.

The results in Figure 8 highlight the important of site characterization of the source zone and are well aligned with theoretical and experimental transport studies that showed the significance of the source zone hydraulics in long distance transport behavior (de Barros & Nowak, 2010; Gueting & Englert, 2013).

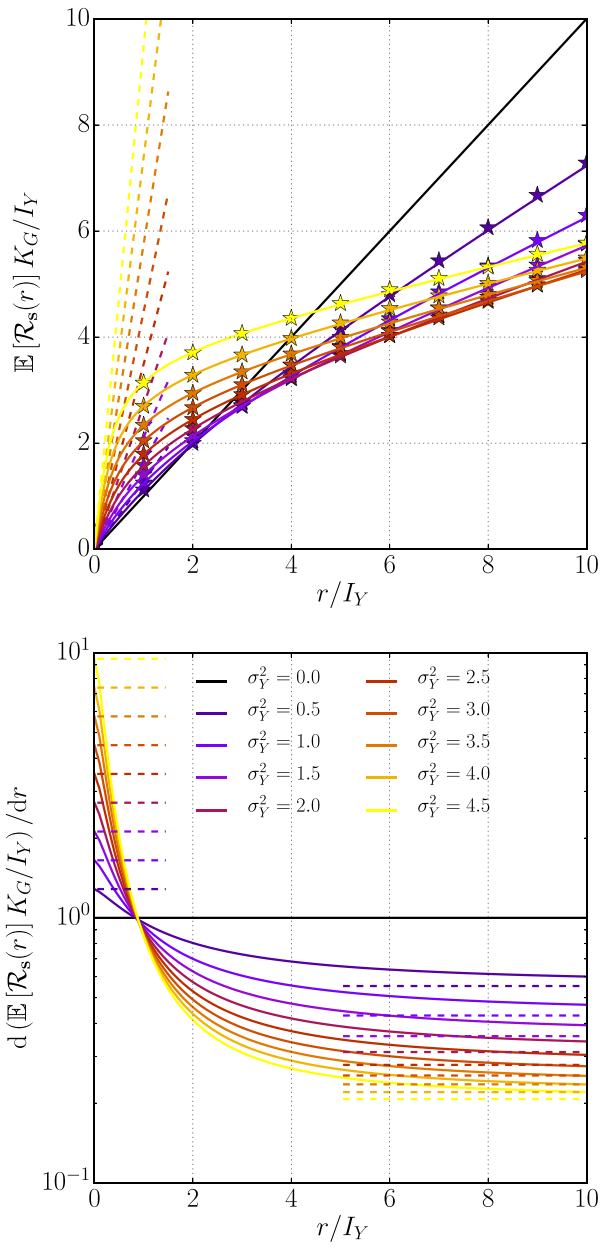

Figure 9. Expected value of the minimum hydraulic resistance between points with distance r (top) and its derivative with respect to *r* (bottom) for different value of σ_Y^2. The numerical results ($\star$) are fitted using Equation (15).

4.2. Average Minimum Hydraulic Resistance

Next, we examine the expected value of the minimum hydraulic resistance. Given a random *K* field, equation (2) can be rewritten as:

$$\mathcal{R}_{\mathbf{s}}(r) = \mathcal{R}_{\mathbf{s}}(\mathbf{x}_{\mathbf{s},r}) = \min_{\Gamma \in \mathcal{P}_{\mathbf{s}}^{\mathbf{x}_{\mathbf{s},r}}} \mathcal{R}_\Gamma = \min_{\Gamma \in \mathcal{P}_{\mathbf{s}}^{\mathbf{x}_{\mathbf{s},r}}} \int_\Gamma e^{-Y} \mathrm{d}\gamma, \tag{14}$$

where $Y = \log K$ and $\mathbf{x}_{\mathbf{s},r}$ is a point distant *r* from **s**. Note that $\mathcal{R}_{\mathbf{s}}(r)$ is a random variable for each value of *r*. The expected value of the minimum hydraulic resistance $\mathbb{E}[\mathcal{R}_{\mathbf{s}}(r)]$ can now be computed numerically using a Monte Carlo approach. The number of realizations needed for the convergence of the mean drastically changes depending on σ_Y^2 We generate several random fields with $\mu_Y = 0$ and different σ_Y^2 using the same procedure described in section 3. We use 50,000 *K*-fields for $\sigma_Y^2 \leq 3.0$, 100,000 otherwise. To improve the convergence, we compute the average using the values of the four cells with same distance from the center (top, bottom, left and right) across all the *K*-fields. Moreover, we use a graph refinement level $\Theta = 3$ as defined in section 2.4.

Figure 9 illustrates the expected value of the minimum hydraulic resistance as a function of *r* for different σ_Y^2. Based on the numerical results in Figure 9 (represented by the symbol $\star$), we observe three distinct regimes corresponding to different behaviors:

Regime 1: if $r \ll I_Y$, the slope of $\mathbb{E}[\mathcal{R}_{\mathbf{s}}(r)]$ increases with σ_Y^2. For an unconditional Gaussian *K*-field, the probability of having high or low *K* in one point is equal. However, since the hydraulic resistance is proportional to $1/K$, low *K* values have a stronger impact on the integral in (2) than high *K* values. The probability of having lower values of *K* increases for fields characterized by larger σ_Y^2. This leads to higher average hydraulic resistance within this regime for large σ_Y^2.

Regime 2: if $r \approx I_Y$, the slope of $\mathbb{E}[\mathcal{R}_{\mathbf{s}}(r)]$ changes with *r*. In this transition regime, the least resistance path of a single conductivity realization is able to escape from a low conductivity area in the vicinity of the source point. Thus, the least resistance path will start to pass through high *K* areas, which is more likely to occur for high values of σ_Y^2. Finally we can observe an inversion of the slope trend.

Regime 3: if $r \gg I_Y$, the slope of $\mathbb{E}[\mathcal{R}_{\mathbf{s}}(r)]$ decreases with σ_Y^2. In other words, if *r* consists of several integral scales, the least resistance path is able to pass through high *K* channels for most of its length. Larger values of σ_Y^2 implies larger likelihood of encountering high *K* zones thus lower hydraulic resistance.

The numerical values of $\mathbb{E}[\mathcal{R}_{\mathbf{s}}(r)]$ displayed in Figure 9 (denoted by the symbol $\star$) have been fitted using a nonlinear least squares method with the following equation:

$$\mathbb{E}[\mathcal{R}_{\mathbf{s}}(r)] = \frac{I_Y}{K_G}\left[c_1\frac{r}{I_Y} + \left(e^{\sigma_Y^2/2} - c_1\right) f^*\left(\frac{r}{I_Y}; c_2, c_3\right)\right], \tag{15}$$

where c_1, c_2 and c_3 are fitting parameters. The function f^*, which represents the transition regime previously mentioned (see Regime 2), is defined as follows:

$$f^*(x; a, n) = x \;\; {}_2\mathcal{F}_1\left(1, \frac{1}{n}; 1 + \frac{1}{n}; -ax^n\right), \tag{16}$$

where ${}_2\mathcal{F}_1$ is the hypergeometric function. Using the hypergeometric function properties, the derivative with respect to x of the function (16) is:

$$\frac{\mathrm{d}}{\mathrm{d}x} f^*(x; a, n) = \frac{1}{1 + ax^n}. \tag{17}$$

Thus, for every pair of positive parameters a and n, the derivative of f^* is 1 for $x = 0$ and tends to zero for $x \to \infty$. The above properties are the motivations for the choice of (16) to model the transition regime.

As shown in Figure 9, equation (15) provides a good approximation for the expected value of the minimum hydraulic resistance. The choice of (15) reflects the numerical results and satisfies the *heuristic* analytical approximations described in the following paragraphs.

First, we derive an approximate solution for the expected minimum hydraulic resistance within the first regime (i.e., $r \ll I_Y$). If the distance between two points is less than an integral scale, the conductivity values in these points are very similar. Therefore, the least resistance path between the two points will likely find similar hydraulic conductivity values along its length (i.e., conduit of similar K values). Thus, we can approximate the least resistance path with a straight line connecting the two points. If $r \ll I_Y$ we obtain:

$$\mathcal{R}_{\mathbf{s}}(r) = \min_{\Gamma \in \mathcal{P}_{\mathbf{s}}^{\mathbf{x}_{\mathbf{s}},r}} \int_\Gamma e^{-Y}\mathrm{d}\boldsymbol{\gamma} \simeq \int_{\Gamma_{\mathbf{s}}^r} e^{-Y}\mathrm{d}\boldsymbol{\gamma} \tag{18}$$

where $\Gamma_{\mathbf{s}}^r$ is a straight line starting from $\mathbf{s}$ with length r. Taking the expected value of (18) we get:

$$\mathbb{E}[\mathcal{R}_{\mathbf{s}}(r)] \simeq \mathbb{E}\left[\int_{\Gamma_{\mathbf{s}}^r} e^{-Y}\mathrm{d}\boldsymbol{\gamma}\right] = \int_{\Gamma_{\mathbf{s}}^r} \mathbb{E}[e^{-Y}]\mathrm{d}\boldsymbol{\gamma} = \int_0^r \frac{e^{\sigma_Y^2/2}}{K_G}\mathrm{d}x = r\frac{e^{\sigma_Y^2/2}}{K_G}. \tag{19}$$

The result is a line with slope $e^{\sigma_Y^2/2}/K_G$. The slope increases exponentially with σ_Y^2. In Figure 9, we compare the solution obtained with the algorithm shown in section 2 with the analytical solution (19) represented by dashed lines for different values of σ_Y^2. As expected, the approximation is valid only within the first regime. Moreover, it does not depend on the spatial covariance model of the conductivity field. The approximation (19) suggests that the solution is linear if r tends to zero. This leads to the choice of (15) and (16) to model $\mathbb{E}[\mathcal{R}_{\mathbf{s}}(r)]$ given their mathematical functional form and properties.

Next, we analyze two points separated by a distance $r \gg I_Y$ (i.e., the third regime). Given two points $\mathbf{s}$ and $\mathbf{x}$ separated by a distance r, we can divide the least resistance path into three sections as shown in Figure 10. First, the path needs to cross the possible low conductivity area around the first point. We assume this area to have a radius δ ($B_{\mathbf{s}}$ in Figure 10). Once the path reaches a high conductivity area, it will follow preferential channels to reach the second point (W in Figure 10). Finally, the path may need to enter a low conductivity area surrounding the second point ($B_{\mathbf{x}}$ in Figure 10). The first and last sections do not depend on the distance r whereas the second section does depend on r. Therefore, we split the expected value of the minimum hydraulic resistance as follows:

$$\mathbb{E}[\mathcal{R}_{\mathbf{s}}(r)] = \mathbb{E}\left[\int_{\hat{\Gamma}\cap B_{\mathbf{s}}} e^{-Y}\mathrm{d}\boldsymbol{\gamma}\right] + \mathbb{E}\left[\int_{\hat{\Gamma}\cap W} e^{-Y}\mathrm{d}\boldsymbol{\gamma}\right] + \mathbb{E}\left[\int_{\hat{\Gamma}\cap B_{\mathbf{x}}} e^{-Y}\mathrm{d}\boldsymbol{\gamma}\right], \tag{20}$$

where $\hat{\Gamma}$ is the least resistance path between $\mathbf{s}$ and $\mathbf{x}$. Note that $\hat{\Gamma}$ is a stochastic path and it is the same for all the integrals in (20). Since $B_{\mathbf{s}}$ and $B_{\mathbf{x}}$

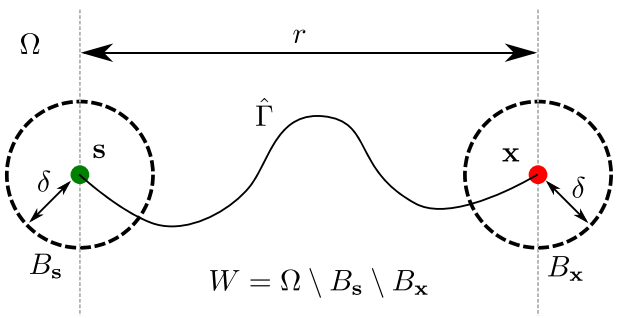

Figure 10. Illustration of the method used to split the least resistance path to derive the large distance approximation.

are independent of r, the first and third terms of (20) are not a function of r. As for the second term, the least resistance path passes mainly through high conductivity areas as a result of the minimization problem. Thus, it is natural to expect that the Y values along this path are larger than the mean value. Therefore, we approximate the log-conductivity in this section as $Y \approx \mu_Y + \alpha\sigma_Y$, where $\alpha > 0$ is an *enhancement* factor due to the high conductivity channels. Inserting this approximation into the second term of (20) yields:

$$\mathbb{E}[\mathcal{R}_{\mathbf{s}}(r)] \approx \beta(Y) + \frac{e^{-\alpha\sigma_Y}(r-2\delta)}{K_G}, \tag{21}$$

where:

$$\beta(Y) = \mathbb{E}\left[\int_{\hat{\Gamma}\cap B_{\mathbf{s}}} e^{-Y} d\gamma\right] + \mathbb{E}\left[\int_{\hat{\Gamma}\cap B_{\mathbf{x}}} e^{-Y} d\gamma\right], \tag{22}$$

$$\mathbb{E}\left[\int_{\hat{\Gamma}\cap W} d\gamma\right] = r - 2\delta. \tag{23}$$

The parameter $\beta(Y)$ contains the contribution, in the average sense, originating from the regions close to the source and target points. We point out that $\beta(Y)$ is a function of the statistical properties of Y (such as σ_Y^2) and does not depend on the distance r.

We observe that the slope of $\mathbb{E}[\mathcal{R}_{\mathbf{s}}(r)]$ in (21) decreases exponentially with σ_Y, explaining the behavior in the third regime observed in Figure 9. By comparing the linear terms in equations (15) and (21), we fit the slope of (21) with the parameter c_1 of (15) using different values of σ_Y^2 ($\sigma_Y^2 = 0.5, 1.0, \cdots, 4.5$) in order to compute the optimal enhancement factor (α). For the Gaussian field characterized with a spatial exponential covariance model for Y, we obtain an enhancement factor $\alpha_{MG} = 0.804 \approx 0.8 = 4/5$ where the subscript MG is used to denote a multi-Gaussian K-field. Based on our results, we point out that the fitted parameter α does not depend on σ_Y^2. Furthermore, equation (21) suggests that $\mathbb{E}[\mathcal{R}_{\mathbf{s}}(r)]$ tends to be linear with increasing r, thus the fitting function (15) has been chosen to reflect this behavior (i.e., the derivative of the function (15) tends to be constant for high values of r). As a result, by comparing the analytical expression (21) with the parameters in the function (15), we find that the expected minimum hydraulic resistance between two points at a distance $r/I_Y \gg 1$ scales linearly with the distance r and exponentially with σ_Y, thus:

$$\mathbb{E}[\mathcal{R}_{\mathbf{s}}(r)] \propto r\, e^{-\frac{4}{5}\sigma_Y}. \tag{24}$$

Equation (24) is valid for the multi-Gaussian field taken into consideration. As shown in section 5, similar results are obtained for fields different than multi-Gaussian, extending the validity of the function (15) and the scaling behavior defined in (24). Based on our analysis, the enhancement factor α is an intrinsic (or static) property of the random field. High values of α imply that the expected resistance decreases rapidly with the standard deviation of the random field σ_Y (see Figure 9).

4.3. Scaling Behavior of First Time Arrivals

The results presented in section 3 showed a clear relationship between the minimum hydraulic resistance and first time arrivals (see Figure 7). Here we examine if the scaling behavior of the expected minimum hydraulic resistance (see equation (24)) also holds for expected first time arrivals.

For this analysis, we will estimate the first time arrivals based on the numerical results reported in Gotovac et al. (2009). Gotovac et al. (2009) presented a detailed numerical study to compute the travel time probability density function (PDF) $f_\tau(t)$ from purely advective transport simulations in a 2-D multi-Gaussian Y-field for different levels of heterogeneity, namely $\sigma_Y^2 =$ 0.25, 1, 2, 4, 6 and 8. The numerical results used for the travel times are extracted from Figure 4 of Gotovac et al. (2009). As noted in Gotovac et al. (2009), there are substantial deviations in early and late arrival times from the log-normal approximation for the travel time PDF which holds for $\sigma_Y^2 < 1$ (e.g., Bellin et al., 1992). The numerical results of Gotovac et al. (2009) show that these deviations increase with σ_Y^2 thus indicating a departure from Fickianity.

In order to verify the scaling behavior (24), we need to define a first time arrival indicator based on the travel time PDF $f_\tau(t)$ obtained from numerical simulations (see circle markers in Figure 4 of Gotovac et al., 2009). Let t_M be the time where $f_\tau(t_M) = \max_t f_\tau(t)$ and let us define t_θ to indicate the time corresponding to $f_\tau(t_\theta) = \theta f_\tau(t_M)$ where $t_\theta < t_M$ (i.e., t_θ is on the left side of the travel time PDF $f_\tau(t)$) and $0 < \theta < 1$. Finally, we

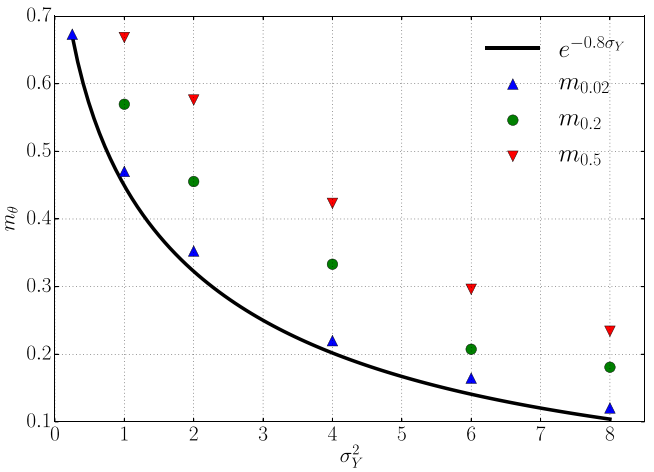

Figure 11. Comparison between the time arrival increase rate m_θ from Gotovac et al. (2009) and the slope of expected minimum hydraulic resistance $e^{-0.8\sigma_Y}$ defined in (24).

define the dimensionless quantity denoted here as the *rate of increase of the first time arrivals* between two control planes situated at distance $x_1 = 10I_Y$ and $x_2 = 20I_Y$ as:

$$m_\theta = U\frac{t_\theta^{x_2} - t_\theta^{x_1}}{x_2 - x_1}, \tag{25}$$

where U is the longitudinal mean velocity. The choice of the control plane locations (i.e., $x_1 = 10I_Y$ and $x_2 = 20I_Y$) and normalization of (25) are based on the numerical results displayed in Figure 4 of Gotovac et al. (2009). We point out that the quantity m_θ is an estimated one since it was extracted from the empirical PDF $f_\tau(t)$ represented by the circle markers in Figure 4 of Gotovac et al. (2009).

Figure 11 shows a comparison between m_θ and the derivative of the expected minimum hydraulic resistance given by (24) obtained by using the parameters described in Gotovac et al. (2009). As we decrease θ (thus t_θ approaches the first time arrivals), the values of m_θ converge to $e^{-0.8\sigma_Y}$. This also occurs for higher levels of heterogeneity as shown for $\sigma_Y^2 = 6$ and 8. Note that these values of σ_Y^2 are higher than the values used to determine expression (24). The results depicted in Figure 11 provide an indication that this scaling behavior may be extended to other values of σ_Y^2. Furthermore, the results in Figure 11 show that the first time arrivals scales according to (24) and that the minimum hydraulic resistance can be used as a computationally efficient metric to estimate first time arrival stochastically.

5. Expected Minimum Hydraulic Resistance of Non-MG Fields

Up to now, the minimum hydraulic resistance was computed only for multi-Gaussian (MG) log-conductivity fields. Here we show how the proposed graph theory algorithm presented in section 2 can be utilized to evaluate the minimum hydraulic resistance in log K-fields displaying non-MG features. Subsequently, we test the performance of the semi-empirical relationship (15) for the expected value of the minimum hydraulic resistance in non-MG fields. As shown in Gómez-Hernández and Wen (1998), the statistical properties of the geological formation influence its connectivity spatial structure as well as the number of extreme values of the permeability field. The spatial statistical properties of the log K-field can potentially have an impact on the magnitude of risk due to groundwater contamination. However, we note that the relative importance of higher-order moments of log K associated with non-MG fields can also depend on the prediction of interest and dimensionality of the computational model. For example, Jankovic et al. (2017) recently performed a comparative analysis between Gaussian and non-Gaussian log-conductivity fields on transport simulations. These authors showed that higher-order moments of log K did not have a significant impact on the bulk of the breakthrough curve at a given control plane for ergodic plumes in 3-D aquifers.

For the purpose of illustration, we consider the non-MG model presented in Zinn and Harvey (2003) for 2-D fields. This model enables us to construct K-fields that are characterized by connected and disconnected structures. We will adopt this model to examine the expected minimum hydraulic resistance for a different ranges of connectivity patterns. The Zinn and Harvey (2003) model has also been adopted by Tyukhova and Willmann (2016) to study the path of smallest resistance computed with a different methodology. In this section, both connected and disconnected fields were generated by the procedure described in Zinn and Harvey (2003). These fields are characterized by the same univariate log-conductivity distributions detailed in the previous section where multi-Gaussian fields were analyzed. The main difference from the MG fields lies in the connectivity between low- and high-K regions (i.e., they differ in higher order moments). For connected fields, high K regions are well-connected whereas low K regions are poorly connected. The opposite occurs for disconnected fields. Additional details regarding the procedure used to generate the K field can be found in Zinn and Harvey (2003).

Figure 12 displays the expected value of the minimum hydraulic resistance as a function of dimensionless distance. Results are reported for different σ_Y^2 and are shown for both connected (Figure 12a) and disconnected (Figure 12b) fields. Similar to Figure 9 (for MG fields), the results originating from the numerical

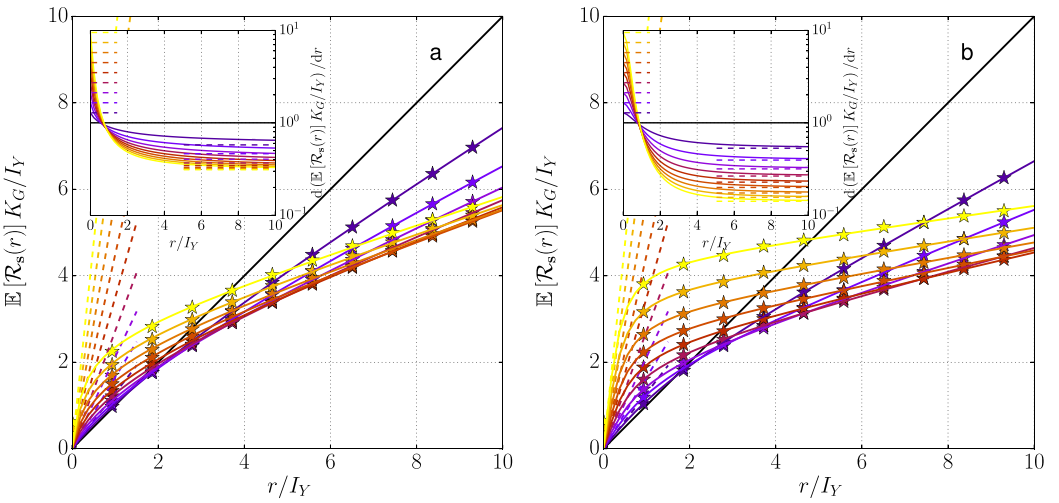

Figure 12. Expected value of the normalized minimum hydraulic resistance versus normalized distance for non-Gaussian log-conductivity fields. Results depicted for (a) well-connected fields and (b) disconnected fields. Insets correspond to the spatial derivative of the expected minimum hydraulic resistance.

simulations are represented by the star symbol while the smooth curves correspond to the fitted semi-empirical relationship (15). The results shown in Figure 12 are remarkable in the sense that the semi-empirical relationship (15) also holds for the non-MG fields investigated. When compared to the results displayed in Figure 9, the only difference lies in the values of the fitted parameters. Similarly to the MG case, we can compute the enhancement factors α. We obtain α_C=0.928 for connected fields and α_D=0.664 for disconnected fields. Here we use the subscript *C* and *D* for connected and disconnected fields respectively. As expected, we find the following relation for the enhancement factor α:

$$\alpha_D < \alpha_{MG} < \alpha_C. \tag{26}$$

Thus, for the same level of heterogeneity σ_Y^2, the expected resistance between two points separated by a distance $r \gg I_Y$ will change more rapidly in disconnected fields than connected fields (see Figure 12). At the same time, by comparing the enhancement factors, the multi-Gaussian case will display an intermediate behavior.

It is also interesting to note the robustness of small distance approximation for this specific log *K* non-MG model (see equation (19)). The small distance approximation holds since the Zinn and Harvey models (both connected and disconnected) are univariate Gaussian (at a point). Thus, the derivative of the expected resistance in zero is the same in all the three cases (compare Figures 9 and 12). As the distance increases, the effect of the higher-order moments becomes more evident. Juxtaposition between Figures 12a with 12b indicates that magnitude of the expected least resistance is higher for the disconnected field, but after the transition regime (i.e., $r \gg I_Y$) there is an inversion of the slopes trend.

Despite the different connectivity patterns analyzed (i.e., MG field in Figure 9 and non-MG fields in Figure 12), the results depicted in this work indicate that the functional form of the semi-empirical expression (15) is quite robust. Nevertheless, additional analysis needs to be carried out to test the range of validity of equation (15) under different conditions such as 3-D structures and other statistical models for the log-conductivity field.

6. Concluding Remarks

Delineating the connected structures of an aquifer is challenging since hydrogeological properties vary in space over multiple scales. Identifying the least resistance path between source and receptor within heterogeneous porous formation can provide crucial information for contaminant site managers. In this paper, we present a computationally efficient method based on graph theory to compute the minimum hydraulic resistance and the corresponding least resistance path. The proposed algorithm converts the hydraulic

conductivity field into a graph and estimates the least resistance path connecting two points without solving the governing equations for flow and transport. Due to its computational efficiency, the algorithm can be cast within a Monte Carlo framework and the statistics of the minimum hydraulic resistance can be obtained.

We compared the resulting minimum hydraulic resistance with solute transport simulations in a 2-D aquifer, confirming the strong correlation between minimum hydraulic resistance and first time arrival reported in Tyukhova and Willmann (2016). Moreover, we showed that our methodology can be used to predict the region where the leading edge of a plume will hit a control plane for the first time. Even if the plume is influenced by boundary effects, a good correlation between the two quantities was computed. Note that the minimum hydraulic resistance is a static quantity and therefore it is not influenced by hydraulic head gradients. For such reasons, further research should be carried out to investigate the relationship between the minimum hydraulic resistance and early arrival times under different boundary and initial conditions.

Our results also indicate the importance of low conductivity zones in the calculation of the minimum hydraulic resistance path. If the starting point is located on the low conductive layer, these low conductivity values will have a strong impact on the expected value of the minimum hydraulic resistance. By making use of the Monte Carlo framework, we were able to provide an equation that describes the expected value of the minimum hydraulic resistance between two points. Analytical approximations considering the two extreme cases of small and large distances support the choice of the semi-empirical equation (15) and was used to verify the results from the graph theory based algorithm.

It is important to note that the connectivity structure will depend on the statistical properties of the hydraulic conductivity field. Nevertheless, we have shown that the proposed approach can be used to investigate the statistical properties of the minimum resistance path in multi-Gaussian (MG) and non-MG log-conductivity fields. Our numerical computations also show that the semi-analytical expression (15) holds for the type of fields investigated, i.e., 2-D statistically isotropic MG log-conductivity fields and connected/disconnected non-MG log-conductivity fields proposed by Zinn and Harvey (2003). The algorithm proposed can be utilized for different types of geological conceptualizations (e.g., Bianchi et al., 2011; Fogg, 1986; Soltanian et al., 2015) and computes the minimum hydraulic resistance in 2-D and 3-D aquifers (e.g., see Figure 3). We also illustrate how the expected minimum hydraulic resistance scales with the distance r and the variance σ_Y^2 following equation (24). Moreover, we have seen that the same scaling equation can be applied to first time arrival, by comparing (24) with numerical simulations presented in the literature. Finally, the expected minimum hydraulic resistance has the potential to efficiently estimate first time arrival in contaminant transport applications which is of importance in risk assessment.

Appendix A: Minimum Hydraulic Resistance and Lower Bound Estimation of the Early Arrival Times

To better illustrate how the minimum hydraulic resistance can be used to estimate early arrivals, consider the case of pure advective flows (i.e., $Pe \rightarrow \infty$). In this case, the minimum hydraulic resistance is a lower bound of the early time arrivals. We demonstrate this by considering two control planes A and B situated at a distance r apart. These planes are orthogonal to the average flow $\langle \mathbf{v}(\mathbf{x}) \rangle$ where the angled brackets denote the average operator. Let Γ^* be the trajectory of the fastest particle (conservative) connecting A and B. This implies that a particle starting from $\mathbf{a} \in A$ traveling along Γ^* will arrive to $\mathbf{b} \in B$ before any other particle released from A at the same time. Therefore, Γ^* can be parametrized using the Lagrangian formulation $\dot{\mathbf{x}} = \mathbf{v}(\mathbf{x}(t))$ with $\mathbf{x}(0) = \mathbf{a}$. Let t^* be the time that a particle needs to travel from A to B through Γ^* (i.e., $\mathbf{x}(t^*) = \mathbf{b}$). Since Γ^* is one of the possible paths connecting A and B, the minimum hydraulic resistance between A and B, as defined in equation (3), is smaller than the hydraulic resistance on Γ^*. Mathematically, this can be expressed as:

$$\mathcal{R}_A(B) \leq \int_{\Gamma^*} \frac{1}{K} \mathrm{d}\gamma. \tag{A1}$$

By rewriting the line integral in (A1) in terms of the Lagrangian trajectory $\mathbf{x}(t)$ we obtain:

$$\int_{\Gamma^*} \frac{1}{K} \mathrm{d}\gamma = \int_0^{t^*} \frac{1}{K(\mathbf{x}(t))} \mathrm{d}(\mathbf{x}(t)). \tag{A2}$$

Using Darcy's law (12) (and setting $n_e = 1$ for simplicity), the differential of the velocity along the trajectory becomes:

$$\mathrm{d}(\mathbf{x}(t)) = K(\mathbf{x}(t)) |\nabla \varphi(\mathbf{x}(t))| \mathrm{d}t. \tag{A3}$$

Substituting (A3) into (A2) and multiplying and dividing the result by t^* we obtain:

$$\int_0^{t^*} \frac{1}{K(\mathbf{x}(t))} \mathrm{d}(\mathbf{x}(t)) = t^* \left[\frac{1}{t^*} \int_0^{t^*} |\nabla \varphi(\mathbf{x}(t))| \mathrm{d}t \right]. \tag{A4}$$

Finally, using (A1) and (A4) we obtain the following inequality:

$$\mathcal{R}_A(B) \leq \eta t^*, \tag{A5}$$

where:

$$\eta \equiv \frac{1}{t^*} \int_0^{t^*} |\nabla \varphi(\mathbf{x}(t))| \mathrm{d}t. \tag{A6}$$

The value η represents the temporal average of the head gradient that a particle traveling along the streamline Γ^* experiences. From (A6), we know that $\eta \leq \max_{\mathbf{x}} |\nabla \varphi(\mathbf{x})|$. Thus, the minimum hydraulic resistance between the two control planes is effectively a lower bound of the early time arrival t^* according to (A5). Understanding the behavior of η over different scenarios would potentially allow to use the minimum hydraulic resistance as a computationally efficient lower bound estimation of the early time arrivals without the need of transport simulations.

Acknowledgments
The first author acknowledges the financial support from the USC Provost's Ph.D. Fellowship. The authors acknowledge the constructive comments of three anonymous Reviewers, the Editor and the Associate Editor. There are no data sharing issues given that all numerical data are provided in the figures and obtained by solving the equations in the paper. All other data for this paper are properly cited and referred to in the reference list. The numerical data are available upon request to the authors.

References

Bellin, A., & Rubin, Y. (1996). Hydro_gen: A spatially distributed random field generator for correlated properties. *Stochastic Hydrology and Hydraulics*, *10*(4), 253–278.
Bellin, A., Rubin, Y., & Rinaldo, A. (1994). Eulerian-lagrangian approach for modeling of flow and transport in heterogeneous geological formations. *Water Resources Research*, *30*(11), 2913–2924.
Bellin, A., Salandin, P., & Rinaldo, A. (1992). Simulation of dispersion in heterogeneous porous formations: Statistics, first-order theories, convergence of computations. *Water Resources Research*, *28*(9), 2211–2227.
Bhark, E. W., Jafarpour, B., & Datta-Gupta, A. (2011). A generalized grid connectivity–based parameterization for subsurface flow model calibration. *Water Resources Research*, *47*, W06517. https://doi.org/10.1029/2010WR009982
Bianchi, M., Zheng, C., Wilson, C., Tick, G. R., Liu, G., & Gorelick, S. M. (2011). Spatial connectivity in a highly heterogeneous aquifer: From cores to preferential flow paths. *Water Resources Research*, *47*, W05524. https://doi.org/10.1029/2009WR008966
Boggs, J. M. (1990). *Hydrogeologic characterization of the MADE site*. Palo Alto, CA: Electric Power Research Institute.
Bondy, J. A., & Murty, U. S. R. (1976). *Graph theory with applications* (Vol. 290). New York, NY: North-Holland.
Cvetkovic, V., Carstens, C., Selroos, J.-O., & Destouni, G. (2012). Water and solute transport along hydrological pathways. *Water Resources Research*, *48*, W06537. https://doi.org/10.1029/2011WR011367
Cvetkovic, V., Fiori, A., & Dagan, G. (2014). Solute transport in aquifers of arbitrary variability: A time-domain random walk formulation. *Water Resources Research*, *50*, 5759–5773. https://doi.org/10.1002/2014WR015449
Dagan, G. (1989). *Flow and transport in porous formations*. Berlin, Germany: Springer Science & Business Media.
de Barros, F. P. J., Bellin, A., Cvetkovic, V., Dagan, G., & Fiori, A. (2016). Aquifer heterogeneity controls on adverse human health effects and the concept of the hazard attenuation factor. *Water Resources Research*, *52*, 5911–5922. https://doi.org/10.1002/2016WR018933
de Barros, F. P. J., Fiori, A., Boso, F., & Bellin, A. (2015). A theoretical framework for modeling dilution enhancement of non-reactive solutes in heterogeneous porous media. *Journal of Contaminant Hydrology*, *175*, 72–83.
de Barros, F. P. J., & Nowak, W. (2010). On the link between contaminant source release conditions and plume prediction uncertainty. *Journal of Contaminant Hydrology*, *116*(1), 24–34.
de Barros, F. P. J., & Rubin, Y. (2008). A risk-driven approach for subsurface site characterization. *Water Resources Research*, *44*, W01414. https://doi.org/10.1029/2007WR006081
de Dreuzy, J.-R., Beaudoin, A., & Erhel, J. (2007). Asymptotic dispersion in 2d heterogeneous porous media determined by parallel numerical simulations. *Water Resources Research*, *43*, W10439. https://doi.org/10.1029/2006WR005394
Deutsch, C. V. (1998). Fortran programs for calculating connectivity of three-dimensional numerical models and for ranking multiple realizations. *Computers & Geosciences*, *24*(1), 69–76.
Dijkstra, E. W. (1959). A note on two problems in connexion with graphs. *Numerische Mathematik*, *1*(1), 269–271.
Fernàndez-Garcia, D., Trinchero, P., & Sanchez-Vila, X. (2010). Conditional stochastic mapping of transport connectivity. *Water Resources Research*, *46*, W10515. https://doi.org/10.1029/2009WR008533
Fiori, A. (2001). The lagrangian concentration approach for determining dilution in aquifer transport: Theoretical analysis and comparison with field experiments. *Water Resources Research*, *37*(12), 3105–3114.
Fiori, A. (2014). Channeling, channel density and mass recovery in aquifer transport, with application to the made experiment. *Water Resources Research*, *50*, 9148–9161. https://doi.org/10.1002/2014WR015950

Fiori, A., Bellin, A., Cvetkovic, V., de Barros, F. P. J., & Dagan, G. (2015). Stochastic modeling of solute transport in aquifers: From heterogeneity characterization to risk analysis. *Water Resources Research*, *51*, 6622–6648. https://doi.org/10.1002/2015WR017388

Fiori, A., Boso, F., de Barros, F. P. J., De Bartolo, S., Frampton, A., Severino, G., . . . Dagan, G. (2010). An indirect assessment on the impact of connectivity of conductivity classes upon longitudinal asymptotic macrodispersivity. *Water Resources Research*, *46*, W08601. https://doi.org/10.1029/2009WR008590

Fiori, A., Dagan, G., Jankovic, I., & Zarlenga, A. (2013). The plume spreading in the made transport experiment: Could it be predicted by stochastic models? *Water Resources Research*, *49*, 2497–2507. https://doi.org/10.1002/wrcr.20128

Fiori, A., & Jankovic, I. (2012). On preferential flow, channeling and connectivity in heterogeneous porous formations. *Mathematical Geosciences*, *44*(2), 133–145.

Fiori, A., Jankovic, I., & Dagan, G. (2011). The impact of local diffusion upon mass arrival of a passive solute in transport through three-dimensional highly heterogeneous aquifers. *Advances in Water Resources*, *34*(12), 1563–1573.

Fogg, G. E. (1986). Groundwater flow and sand body interconnectedness in a thick, multiple-aquifer system. *Water Resources Research*, *22*(5), 679–694.

Fredman, M. L., & Tarjan, R. E. (1987). Fibonacci heaps and their uses in improved network optimization algorithms. *Journal of the Association for Computing Machinery*, *34*(3), 596–615.

Frippiat, C. C., Illangasekare, T. H., & Zyvoloski, G. A. (2009). Anisotropic effective medium solutions of head and velocity variance to quantify flow connectivity. *Advances in Water Resources*, *32*(2), 239–249.

Gómez-Hernández, J. J., & Wen, X.-H. (1998). To be or not to be multi-Gaussian? A reflection on stochastic hydrogeology. *Advances in Water Resources*, *21*(1), 47–61.

Gotovac, H., Cvetkovic, V., & Andricevic, R. (2009). Flow and travel time statistics in highly heterogeneous porous media. *Water Resources Research*, *45*, W07402. https://doi.org/10.1029/2008WR007168

Gueting, N., & Englert, A. (2013). Hydraulic conditions at the source zone and their impact on plume behavior. *Hydrogeology Journal*, *21*(4), 829–844.

Guyer, J. E., Wheeler, D., & Warren, J. A. (2009). Fipy: Partial differential equations with python. *Computing in Science & Engineering*, *11*(3), 6–15.

Henri, C. V., & Fernàndez-Garcia, D. (2014). Toward efficiency in heterogeneous multispecies reactive transport modeling: A particle-tracking solution for first-order network reactions. *Water Resources Research*, *50*, 7206–7230. https://doi.org/10.1002/2013WR014956

Henri, C. V., Fernàndez-Garcia, D., & de Barros, F. P. J. (2015). Probabilistic human health risk assessment of degradation-related chemical mixtures in heterogeneous aquifers: Risk statistics, hot spots, and preferential channels. *Water Resources Research*, *51*, 4086–4108. https://doi.org/10.1002/2014WR016717

Jankovic, I., Maghrebi, M., Fiori, A., & Dagan, G. (2016). When good statistical models of aquifer heterogeneity go right: The impact of aquifer permeability structures on 3d flow and transport. *Advances in Water Resources*, *100*, 199–211.

Knudby, C., & Carrera, J. (2005). On the relationship between indicators of geostatistical, flow and transport connectivity. *Advances in Water Resources*, *28*(4), 405–421.

Le Borgne, T., Dentz, M., Bolster, D., Carrera, J., de Dreuzy, J.-R., & Davy, P. (2010). Non-fickian mixing: Temporal evolution of the scalar dissipation rate in heterogeneous porous media. *Advances in Water Resources*, *33*(12), 1468–1475.

Le Goc, R., de Dreuzy, J.-R., & Davy, P. (2010). Statistical characteristics of flow as indicators of channeling in heterogeneous porous and fractured media. *Advances in Water Resources*, *33*(3), 257–269.

Leube, P. C., de Barros, F. P. J., Nowak, W., & Rajagopal, R. (2013). Towards optimal allocation of computer resources: Trade-offs between uncertainty quantification, discretization and model reduction. *Environmental Modelling & Software*, *50*, 97–107.

Liu, G., Zheng, C., & Gorelick, S. M. (2007). Evaluation of the applicability of the dual-domain mass transfer model in porous media containing connected high-conductivity channels. *Water Resources Research*, *43*, W12407. https://doi.org/10.1029/2007WR005965

Maxwell, R. M., Kastenberg, W. E., & Rubin, Y. (1999). A methodology to integrate site characterization information into groundwater-driven health risk assessment. *Water Resources Research*, *35*(9), 2841–2855.

Moslehi, M., Rajagopal, R., & de Barros, F. P. J. (2015). Optimal allocation of computational resources in hydrogeological models under uncertainty. *Advances in Water Resources*, *83*, 299–309.

Pardo-Igúzquiza, E., & Dowd, P. A. (2003). Connec3d: A computer program for connectivity analysis of 3d random set models. *Computers & Geosciences*, *29*(6), 775–785.

Renard, P., & Allard, D. (2013). Connectivity metrics for subsurface flow and transport. *Advances in Water Resources*, *51*, 168–196.

Riva, M., Guadagnini, L., Guadagnini, A., Ptak, T., & Martac, E. (2006). Probabilistic study of well capture zones distribution at the lauswiesen field site. *Journal of Contaminant Hydrology*, *88*(1), 92–118.

Rubin, Y. (1987). A hierarchical method for the design of water allocation and water distribution networks based on graph-theory. In *Proceedings of the Vancouver Symposium on irrigation and water allocation* (IAHS Publ. 169, pp. 207–220).

Rubin, Y. (2003). *Applied stochastic hydrogeology*. Oxford, UK: Oxford University Press.

Salamon, P., Fernàndez-Garcia, D., & Gómez-Hernández, J. J. (2006). A review and numerical assessment of the random walk particle tracking method. *Journal of Contaminant Hydrology*, *87*(3), 277–305.

Sánchez-Vila, X., Carrera, J., & Girardi, J. P. (1996). Scale effects in transmissivity. *Journal of Hydrology*, *183*(1), 1–22.

Silliman, S. E., & Wright, A. L. (1988). Stochastic analysis of paths of high hydraulic conductivity in porous media. *Water Resources Research*, *24*(11), 1901–1910.

Soltanian, M. R., Ritzi, R. W., Huang, C. C., & Dai, Z. (2015). Relating reactive solute transport to hierarchical and multiscale sedimentary architecture in a lagrangian-based transport model: 1. time-dependent effective retardation factor. *Water Resources Research*, *51*, 1586–1600. https://doi.org/10.1002/2014WR016353

Tartakovsky, D. M. (2007). Probabilistic risk analysis in subsurface hydrology. *Geophysical Research Letters*, *34*, L05404. https://doi.org/10.1029/2007GL029245

Tartakovsky, D. M. (2013). Assessment and management of risk in subsurface hydrology: A review and perspective. *Advances in Water Resources*, *51*, 247–260.

Thomas, H. C., Leiserson, C. E., Rivest, R. L., & Stein, C. (2001). *Introduction to algorithms* (Vol. 6). Cambridge, MA: MIT Press.

Trinchero, P., Sánchez-Vila, X., & Fernàndez-Garcia, D. (2008). Point-to-point connectivity, an abstract concept or a key issue for risk assessment studies? *Advances in Water Resources*, *31*(12), 1742–1753.

Tyukhova, A. R., Kinzelbach, W., & Willmann, M. (2015). Delineation of connectivity structures in 2-d heterogeneous hydraulic conductivity fields. *Water Resources Research*, *51*, 5846–5854. https://doi.org/10.1002/2014WR015283

Tyukhova, A. R., & Willmann, M. (2016). Connectivity metrics based on the path of smallest resistance. *Advances in Water Resources*, *88*, 14–20.

Zheng, C., Bianchi, M., & Gorelick, S. M. (2011). Lessons learned from 25 years of research at the made site. *Ground Water*, *49*(5), 649–662.

Zinn, B., & Harvey, C. F. (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 Resources Research*, *39*(3), 1051. https://doi.org/10.1029/2001WR001146