Inverse Method to Determine Hydraulic Conductivity from a Velocity Field using Graph Theory

Michael E. Mont-Eton[1], Steffen Borgwardt[2], and David C. Mays[1,3]

InterPore Journal

submitted 29 May 2024
revised 4 September 2024
accepted 7 October 2024

[1]University of Colorado Denver, Department of Civil Engineering.
[2]University of Colorado Denver, Department of Mathematical and Statistical Sciences.
[3]Corresponding author: University of Colorado Denver, Department of Civil Engineering, Campus Box 113, PO Box 173364, Denver, CO 80217-3364, USA, david.mays@ucdenver.edu, telephone +1-303-315-7570.

Abstract
A numerical inverse method called FlowPaths is presented to solve for the hydraulic conductivity field of an isotropic heterogeneous porous medium from a known specific discharge field (and constant-head boundary conditions). This method makes possible a new approach to reactive transport experiments, aimed at understanding the dynamic spatial and temporal evolution of hydraulic conductivity, which simultaneously record the evolving reaction and the evolving flow geometry. This inverse method assumes steady, two-dimensional flow through a square matrix of grid blocks. A graph-theoretical approach is used to find a set of flow paths through the porous medium using the known components of the specific discharge, where every vertex is traversed by at least one path from the upstream high-head boundary to the downstream low-head boundary. Darcy's law is used to create an equation for the unknown head drop across each edge. Summation of these edge equations along each path through the network generates a set of linearly independent head-drop equations that is solved directly for the hydraulic conductivity field. FlowPaths is verified by generating 12,740 hydraulic conductivity fields of varying size and heterogeneity, calculating the corresponding specific discharge field for each, and then using that specific discharge field to estimate the underlying hydraulic conductivity field. When estimates from FlowPaths are compared to the simulated hydraulic conductivity fields, the inverse method is demonstrated to be accurate and numerically stable. Accordingly, within certain limitations, FlowPaths can be used in field or laboratory applications to find hydraulic conductivity from a known velocity field.

Article Highlights:

- The two-dimensional groundwater inverse problem is solved for the heterogeneous hydraulic conductivity given a known velocity field.
- Graph theory is used to find down-gradient continuous paths through the porous media.
- The paths are transformed into a system of equations, based on Darcy's law, that is solved for the hydraulic conductivity field.

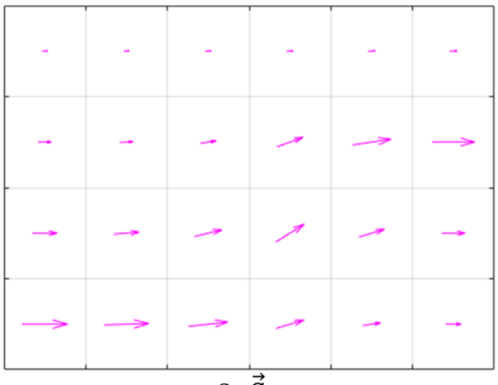
Keywords: Inverse problem, Groundwater, Velocity field, Graph theory

1. Introduction

Modeling flow through porous media supports numerous applications including seepage through dams (*e.g.*, Lyu et al., 2019), subsurface transport of nuclear waste (*e.g.*, Tsang et al., 2015), and groundwater remediation (*e.g.*, Zhao et al., 2024). According to Darcy's law, specific discharge $\vec{q}$ is linearly proportional to the hydraulic head gradient ∇h and to the hydraulic conductivity **K**, a spatially and temporally variable tensor field that reflects the physical properties of water and the geometric intricacies of the porous media. Those geometric intricacies are encapsulated by the permeability **k**, while the specific discharge $\vec{q}$ is the product of the local porosity ε and the local pore velocity $\vec{v}$. Except in idealized cases, such as filter beds packed with monodisperse spherical grains, it is difficult-to-impossible to predict hydraulic conductivity a priori (Baveye et al., 1998); this difficulty is amplified in the context of groundwater remediation, where a host of mechanisms such as clay dispersion, mineral precipitation, biofilm growth, or bubble formation can trigger spatial and temporal changes in hydraulic conductivity (Manga et al., 2012).

Because it cannot be measured a priori, hydraulic conductivity is determined through inverse modeling (*e.g.*, Sun, 1999): One measures the other terms in a known relationship, such as Darcy's law, then solves for the unknown hydraulic conductivity. Indeed, this paradigm is illustrated in Darcy's (1856) original work, in which the discharge through and the head loss across a sand-packed column of known cross-sectional area were measured; those measurements allowed Darcy to calculate the hydraulic conductivity, as an inverse solution, using the equation that now bears his name (Freeze & Cherry, 1979, p. 15). In traditional groundwater engineering, hydraulic conductivity can be determined by pumping tests (Theis, 1935), slug tests (Cooper et al., 1967), or extracting samples for laboratory testing (Fitts, 2002, pp. 67-93). When more than a point measurement of hydraulic conductivity is required, as in most groundwater modeling applications, the hydraulic conductivity field is generally inverted from measured heads, which permits identification of the hydraulic conductivity field (Neuman, 1973). Using measured heads, inverse models have been developed to determine the hydraulic conductivity field (**K**) using stochastic techniques (Chang & Yeh, 2010; Liu et al., 2006; Lu & Zhang, 2002); parametrization (Jiao & Zhang, 2015); or Adomian decomposition (Panawalage et al., 2018). However, to our knowledge no inverse model is available to determine heterogeneous hydraulic conductivity fields from measured velocity fields, which is the gap this work seeks to partially fill.

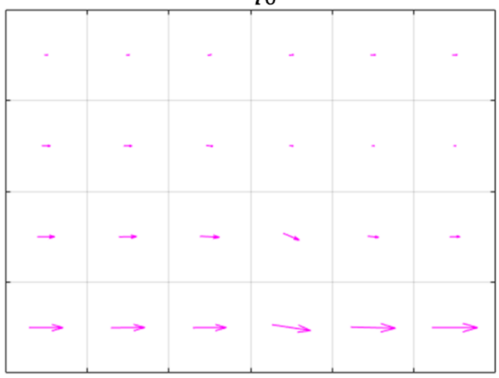
Why determine heterogeneous **K** from the velocity field? Because it makes possible a new approach to reactive transport experiments aimed at understanding the dynamic spatial and temporal evolution of hydraulic conductivity. To envision such experiments, consider the following motivational example: We assume a quasi-2D flow through heterogeneous porous media observed with particle image velocimetry (PIV). Knowing the constant head boundaries and the PIV at time 1 (Figure 1a), we apply FlowPaths to infer the corresponding distribution of hydraulic conductivity (Figure 1b). Then we postulate that some process reduces the hydraulic conductivity in certain cells. That process could include any or all of the mechanisms articulated above (Manga et al., 2012); for concreteness, we assume that mineral precipitation reduces the hydraulic conductivity in the five cells highlighted in Figure 1d, with light shading for a 10-fold decrease and dark shading for a 100-fold decrease. Thus, reaction has altered the hydraulic conductivity, which will in turn alter the flow field. At time 2, we repeat the PIV measurement (Figure 1c) and reapply FlowPaths to find the new distribution of hydraulic conductivity (Figure 1d). The motivation for FlowPaths is that it is essential for such an experiment.



a. $\vec{q}_0$

2	2	5	10	20	20
5	5	10	60	100	100
10	10	20	50	40	40
20	20	30	40	20	20

b. K_0



c. $\vec{q}_1$

2	2	5	10	2	2
5	5	10	6	**1**	1
10	10	20	50	4	4
20	20	30	40	20	20

d. K_1

Figure 1: Hypothetical experiment where (a) the specific discharge field at time 1 determines (b) the hydraulic conductivity field at time 1, and the subsequently measured (c) specific discharge field at time 2 determines (d) the hydraulic conductivity at time 2. The dynamic evolution of hydraulic conductivity, averaged over grid blocks, is shown by light gray shading for 90% clogging and by dark gray shading for 99% clogging.

The dependence of hydraulic conductivity on the complex distribution of porosity continues to be an area of ongoing research (e.g., Ghosh et al., 2020). Here we pose the next question: How do these complex relationships evolve over time? At the heart of these dynamics is a feedback process (Singurindy & Berkowitz, 2003). Reactive processes in groundwater remediation depend on fluid flow; those reactive processes often change the hydraulic conductivity; the changed hydraulic conductivity changes the fluid flow; and the changed fluid flow alters the reactive processes. This feedback process has been recognized for some time but has yet to be studied in detail. Thullner et al. (2002) studied the feedback between biofilm growth and flow; El Mountassir (2014) studied reaction-flow feedback with microbially-mediated mineral precipitation; and Yoon et al. (2019) studied reaction-flow feedback from precipitation of calcium carbonate. Bastidas Olivares et al. (2021) presented a numerical approach to model reaction-flow feedback for a generalized mineral precipitation-dissolution process, and Kelm et al. (2022) extended this approach to include two minerals. In the specific context of groundwater remediation, awareness of reaction-flow feedback has also been reflected in at least one modeling study (Li et al., 2008) and at least one field study (Englert et al., 2009) , and similar observations have been reported at commercial groundwater remediation sites (Kent Sorenson, CDM Smith, personal communication, 2017). In all these cases, hydraulic

conductivity evolves dynamically, and that evolution is both the cause and effect—a feedback process—of reactive transport. Understanding this feedback process calls for a new approach to reactive transport experiments that simultaneously gather evidence for reaction, such as biofilm growth or mineral precipitation, and evidence for the flow geometry, such as the velocity field. Such work, building on prior reactive transport experiments described below, would require an inverse method to determine hydraulic conductivity from velocity.

There are several examples where reaction-flow feedback is relevant. One is clogging near wells during aquifer storage and recovery (Mays, 2013). Another example, relevant to groundwater remediation, is clogging in permeable reactive barriers (Zolla et al., 2009). A particularly illustrative example of reaction-flow feedback is the model reaction where an aqueous plume of calcium chloride is co-injected with an aqueous plume of sodium carbonate. Upon mixing, these plumes precipitate calcium carbonate that has been shown to inhibit dispersive mixing (Tartakovsky et al., 2008) and to separate plumes across interfaces between porous media (Katz et al., 2011). These experiments demonstrate reaction-flow feedback, albeit with a relatively straightforward evolution from a clean-bed system that mixes relatively more to a post-reaction system that mixes relatively less. More sophisticated experiments would mimic reactive transport processes characteristic of groundwater remediation, where reactions determine and alter flow and transport; these experiments would provide both observable reactions and observable velocity fields. In laboratory settings, velocity fields within porous media can be measured by particle image velocimetry (PIV), using micromodels (Yoon et al., 2019), thin two-dimensional flow cells (Larsson et al., 2018), or three-dimensional flow cells filled with refractive index matched porous media (Arthur et al., 2009). The goal of the present study is to show how the velocity field provides a rich set of data that may be used to calculate the hydraulic conductivity field.

In order to formulate a well-posed problem that can be solved by direct matrix inversion, we require an approach to identify unique flow paths, taking advantage of the vector properties of the velocity field. A natural choice is to use graph theory to identify paths by linking the velocity vectors from adjacent grid blocks (henceforth cells), starting at one boundary, and ending at the other. Several previous investigators have applied graph theory to inverse problems in hydrogeology: Eikemo et al. (2009) used graph theory to develop a fast solver for advective transport in fractured media using topological sorting of the discrete fluxes; Rizzo and de Barros (2017) used graph theory to develop a numerical approach to find arrival times of solute plumes via the paths of least resistance through heterogeneous porous media; and Godefroy et al. (2019) used graph theory to analyze the structural interpretation of faults. Complementing these works, there is a body of literature addressing pore-network models in which pore spaces, modeled as vertices, are connected by pore throats, modeled as edges (Berkowitz and Ewing, 1998; Yang et al., 2016). A particular focus of this literature is percolation theory, which describes the fractal paths providing flow and transport in heterogeneous porous media, and which can be used to predict solute transport (Hunt and Ghanbarian, 2016) and hydraulic conductivity (Hunt and Sahimi, 2017). A similar approach has been developed to model transport in fractured rock using discrete fracture networks (Viswanathan et al., 2018). Distinct from these pore- and fracture-network models, but still within the framework of graph theory, are recent works adopting convolutional neural networks (CNNs) to predict macroscopic variables, including hydraulic conductivity, from pore-scale measurements of geometry using machine learning (Santos et al., 2020; Elmorsy et al., 2022; Gärttner et al., 2023). However, to our knowledge, no previous

research has used graph theory—or any other method—to uniquely determine the Darcy-scale hydraulic conductivity field from a measured field of velocity vectors.

This paper is structured as follows. In Section 2 we present an inverse model called FlowPaths that uses graph theory to generate the system of equations needed to determine the hydraulic conductivity field from a known specific discharge field. In Section 3 we present results for one proof-of-concept example and summarize results from model verification using 12,740 simulations, and in Section 4 we discuss these results in light of the literature while suggesting directions for future research. Conclusions are presented in Section 5.

2. Methods

In creating FlowPaths, our intent was to provide a fast solver for the inverse problem that did not rely on iterative methods or additional observations, so that it could be adapted to predict the effects of transport under unsteady conditions in near-real time. Three concepts guide our approach to identifying hydraulic conductivity from specific discharge. First, using Darcy's law to write the head-drop equations on the paths in any tree (the set of paths originating at a single source) from a source to the set of sink vertices (henceforth called target vertices) produces a positive matrix of coefficients. Such a matrix must lead to the solution vector, **K**, being positive, per Farkas' lemma (Matoušek & Gärtner, 2007, pp. 81-105), which limits the solution to the physically plausible condition where each member, K_i, of the solution vector, **K**, is strictly positive. Second, expressing the groundwater flow equation in terms of graph theory produces a convex planar mesh (Tutte, 1960), where every vertex in the flow domain is at least 3-connected, and every line drawn between two vertices must be inside the flow domain. The matrix of head-drop coefficients produces a Hessian matrix which is positive-definite, also implying that the coefficient matrix is a convex set. Strictly convex functions have a global minimum (Rockafellar, 1970), leading to a unique result. And third, the system is solved using a non-iterative method, and its stability is confirmed with a recursive error test.

In Section 2.1 we present a graph-theoretic model for the inverse solution, develop head-drop equations using Darcy's law, find a spanning set of linearly independent equations, and then solve directly for the unknown hydraulic conductivities. We conclude with methods for model verification and error estimation.

2.1 Problem Formulation

In the finite-difference methods used in groundwater modeling software such as MODFLOW (USGS, 2022), cells represent areas where the aquifer properties (*i.e.*, hydraulic conductivity) are constant. As shown in Figure 2, the closed system model is divided into the source region, S, the flow domain, Ω^o, and the target region, T. Together, they form the closure of the model domain, $\overline{\Omega}$ (Trench, 2003). The goal of the analysis is to calculate the hydraulic conductivities $\mathbf{K} = K_{i,j}$ for all cells, where i is the row and j is the column. The cells in regions S and T are known head regions. Ω^o is the region of unknown head, which is calculated after determining **K**. The specific discharge vectors are known everywhere in $\overline{\Omega}$.

The numerical model is reformulated in the framework of graph theory, which conceptualizes the system as a network of nodes and links, where nodes are called *vertices*, links are called *edges*, and together they constitute a *graph* with symbol *G*. Further definitions may be found in the Supporting Information (Section S1). Using block-centered vertices, as shown in Figure 3a, the specific discharge vectors are assigned to the geometric center of each cell. These vectors are broken into components, as shown in Figure 3b, and then these block-centered

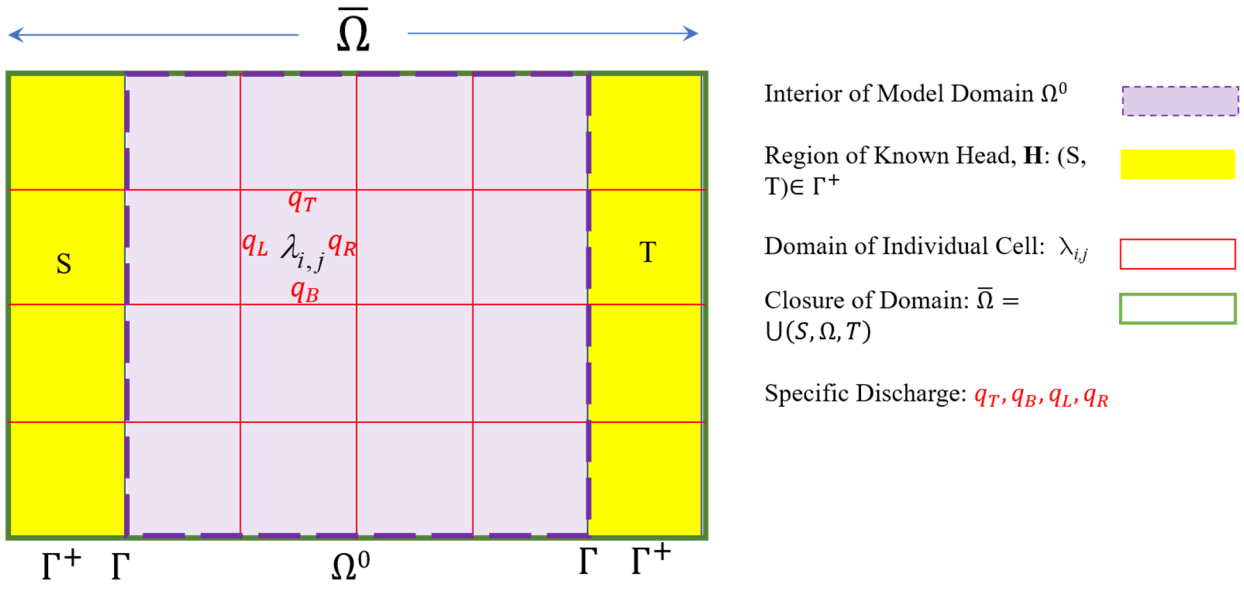


Figure 2. The closed-system model domain is composed of known, constant head grid-blocks (S and T) and a region of unknown head cells Ω^o. No flow occurs across the perimeter of $\overline{\Omega}$ or between the constant-head cells. The specific discharge components q_L, q_R, q_B, and q_T are known for all cell faces, while the hydraulic conductivity K is unknown for all cells.

components are combined into specific discharges passing through the faces of the cells, as shown in Figure 3c. The final step in the reformulation is Figure 3d, which shows a directed graph, where the same flow goes through the edges between vertices as through the faces of the cells.

FlowPaths has been constructed to meet the criteria of a well-posed inverse problem. In the sense of Tikhonov, well-posedness requires that (1) a solution exists, (2) the solution is unique, and (3) the solution is stable (Tikhonov, 1977). Existence is shown by the fact that for every unknown hydraulic conductivity $K(x,y)$, there exists a data source, $q(x,y)$, that occupies a one-to-one relationship. Uniqueness is guaranteed because the minimum number of paths needed for a unique solution is equal to the number of linearly independent paths as detailed in Section 2.3 below. And finally, stability is demonstrated through the recursive error test showing that the cumulative residuals are bounded as detailed in Section 2.4 below. Additional details on the well-posedness of the method are available in the Supporting Information (Section S2).

2.2 Head-Drop Equations

We conceptualize the flow as a set of paths, each of which generates a first-order polynomial equation for hydraulic conductivity using Darcy's law through the faces of the cells (Figure 3c). We define the *head drop* between the predecessor and successor vertices (Figure 3d) a and b as:

$$\Delta h_{a,b} = h_a - h_b. \quad (1)$$

In graph terms, with a,b denoting the edge between vertex a and vertex b, Darcy's law is then:

$$q_{a,b} = -K_{a,b}\frac{h_b - h_a}{L_{a,b}}, \quad (2)$$

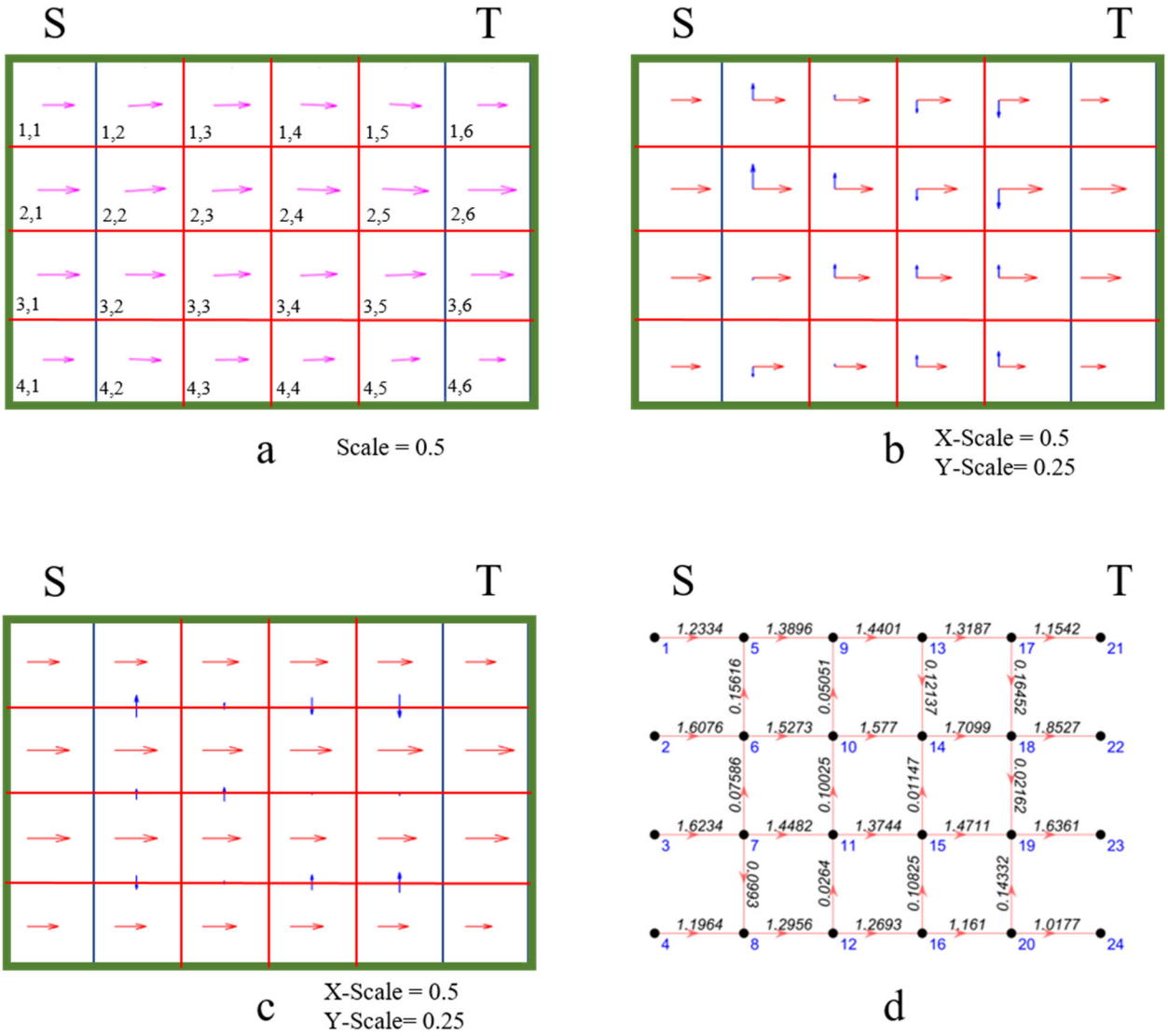


Figure 3. Problem formulation with (a) cell-centered specific discharge vectors, (b) orthogonal components of specific discharge vectors, (c) face-centered specific discharges, and (d) directed graph with edge weights given by the face-centered specific discharges shown in (c). The set S contains the source vertices {1, 2, 3, 4} and the set T contains the target vertices {21, 22, 23, 24}. Note the vertical scale in (b) and (c) has been chosen to exaggerate the vertical component of specific discharge. Results in (d) may not sum to zero at each vertex due to rounding.

where $K_{a,b}$ is the effective hydraulic conductivity over *a,b* and $L_{a,b}$ is the length of *a,b*. The single head-drop equation through one edge is then:

$$\Delta h_{a,b} = \frac{q_{a,b}}{K_{a,b}} L_{a,b}. \tag{3}$$

The sum of a series of individual head drops along *n* vertices in a path can be arranged as

$$\Delta h_{total} = (h_1 - h_2) + (h_2 - h_3) + \cdots + (h_{n-2} - h_{n-1}) + (h_{n-1} - h_n) = \sum_{k=1}^{k=n-1}(h_k - h_{k+1}). \quad (4)$$

In this equation, the total head drop, Δh_{total}, is equal to the sum of the individual head drops. The effective hydraulic conductivity, $K_{a,b}$, is then defined for each head drop. Following Harbaugh (2005, p. 135), since the flow is perpendicular to the cell faces, $K_{a,b}$ is defined as the harmonic mean of the hydraulic conductivity of the adjacent vertices:

$$K_{a,b} = \frac{2}{\frac{1}{K_a}+\frac{1}{K_b}} \quad , \quad (5)$$

which can be derived by analogy to electrical resistors in series. The specific discharge through an edge connecting adjacent vertices *a* and *b* can be expressed using Darcy's law as

$$\frac{h_a - h_b}{L_{a,b}} = \frac{q_{a,b}}{K_{a,b}}. \quad (6)$$

We assume square cells, so by construction, $l = L_{a,b}$ for all adjacent vertices *a* and *b*. Then we substitute the individual head drop expression into the total head drop equation for *n* consecutive vertices along any simple path (*i.e.*, without repeated vertices) between the source head h_S and the target head h_T. The result is a series head loss equation:

$$\frac{h_S - h_T}{l} = \frac{q_{1,2}}{K_{1,2}} + \frac{q_{2,3}}{K_{2,3}} + \cdots + \frac{q_{n-2,n-1}}{K_{n-2,n-1}} + \frac{q_{n-1,n}}{K_{n-1,n}} = \sum_{i=1}^{n-1} \frac{q_{i,i+1}}{K_{i,i+1}}. \quad (7)$$

We now express $K_{a,b}$ as the hydraulic resistance *R* of the edge, where $R = K^{-1}$, so that

$$K_{a,b} = \frac{2}{R_a + R_b}. \quad (8)$$

Now Eq. 7 can be expressed in terms of the resistances of each vertex and the specific discharges of each edge as

$$\frac{h_S - h_T}{l} = \frac{q_{1,2}}{\frac{2}{R_1+R_2}} + \frac{q_{2,3}}{\frac{2}{R_2+R_3}} + \cdots + \frac{q_{n-2,n-1}}{\frac{2}{R_{n-2}+R_{n-1}}} + \frac{q_{n-1,n}}{\frac{2}{R_{n-1}+R_n}} = \sum_{i=1}^{n-1} \frac{q_{i,i+1}}{\frac{2}{R_i+R_{i+1}}}. \quad (9)$$

Inverting the denominators, distributing, and gathering like terms yields

$$\frac{h_S - h_T}{l} = R_1\left(\frac{q_{1,2}}{2}\right) + R_2\left(\frac{q_{1,2}+q_{2,3}}{2}\right) + R_3\left(\frac{q_{2,3}+q_{3,4}}{2}\right) + \cdots + R_{n-2}\left(\frac{q_{n-3,n-2}+q_{n-2,n-1}}{2}\right) + R_{n-1}\left(\frac{q_{n-2,n-1}+q_{n-1,n}}{2}\right) + R_n\left(\frac{q_{n-1,n}}{2}\right). \quad (10)$$

We assume that the hydraulic conductivity of the boundary cells matches that of their neighboring active cells, so their resistances will be equal as well, giving $R_1 = R_2$ and $R_{n-1} = R_n$. Using only the active matrix (*i.e.*, non-boundary grid cells) produces the final head drop equation with combinations of the specific discharges on each vertex for any path:

$$\frac{h_S - h_T}{l} = R_2\left(\frac{2q_{1,2}+q_{2,3}}{2}\right) + R_3\left(\frac{q_{2,3}+q_{3,4}}{2}\right) + \cdots + R_{n-2}\left(\frac{q_{n-3,n-2}+q_{n-1,n-1}}{2}\right) + R_{n-1}\left(\frac{q_{n-2,n-1}+2q_{n-1,n}}{2}\right) \quad (11)$$

To simplify the notation, the parenthetical term following each R is re-defined as the variable q_k^* for each vertex of every path starting with the second vertex ($k = 2$) and ending with the penultimate vertex ($k = n$-1), using the value of the specific discharges through the edges on the path that are incident to the vertex:

$$q_k^* = \begin{cases} q_{1,2} + \frac{q_{2,3}}{2} & k = 2 \\ \frac{q_{k-1,k}}{2} + q_{k,k+1} & k = n-1 \\ \frac{q_{k-1,k}+q_{k,k+1}}{2} & 2 < k < n-1 \end{cases}, \quad (12)$$

where n is the number of vertices on each path. Replacing each specific discharge coefficient with the q_k^* term produces the first-order polynomial equation used for each path:

$$\frac{h_S - h_T}{l} = R_2(q_2^*) + R_3(q_3^*) + \cdots + R_{n-2}(q_{n-2}^*) + R_{n-1}(q_{n-1}^*). \quad (13)$$

The cells corresponding to Eq. 13 are shown in Figure 4. Any valid specific discharge field will be constrained by continuity, so assuming steady flow, it follows that the net flow into (or out of) any vertex $v \neq s,t$ is zero; by analogy to electrical circuits, continuity ensures that Kirchhoff's rules are obeyed.

2.3 FlowPaths Inverse Model

The essential function of the FlowPaths inverse model is to identify a set of pathways, consistent with the known specific discharge field, that will generate the system of equations for the hydraulic conductivity as described above. The flowchart in Figure 5 and Figure 6 summarizes the steps needed to identify the paths and therefore to create the system of equations for the unknown hydraulic conductivity field. The flowchart starts with the full grid-cell domain, $\overline{\Omega}$, the known specific discharge, $\mathbf{q}_o$, and the total head drop Δh_{total}. The steps in the flowchart are explained in detail in the Supporting Information (Section S3).

To solve the system of equations for hydraulic conductivity, we require a set of linearly independent (LI) paths, similar to basis path sets described by Zhu et al. (2021). We define this set as $\mathcal{T}_{\text{LI}}$. The LI paths can be identified by finding all possible paths between the vertices in S and T and discarding those that are redundant, but such an approach would be computationally expensive (Gosnell & Broecheler, 2020). Instead, we subdivide the model domain into directed subgraphs H_{st} for each (s,t) pair and construct the LI paths for each directed subgraph by using the method outlined by Zwick (2013). The maximum number of LI paths in each directed subgraph is related to the cyclomatic number (Berge, 1973, p. 15), as described in the Supporting Information (Section S4). Then we combine the LI paths from each directed subgraph to create a set of paths $\mathcal{T}_{\text{LI}}$, which has the property that every edge in H belongs to some path in P, implying that every vertex in H also belongs to some path in P (Sloane, 1972).

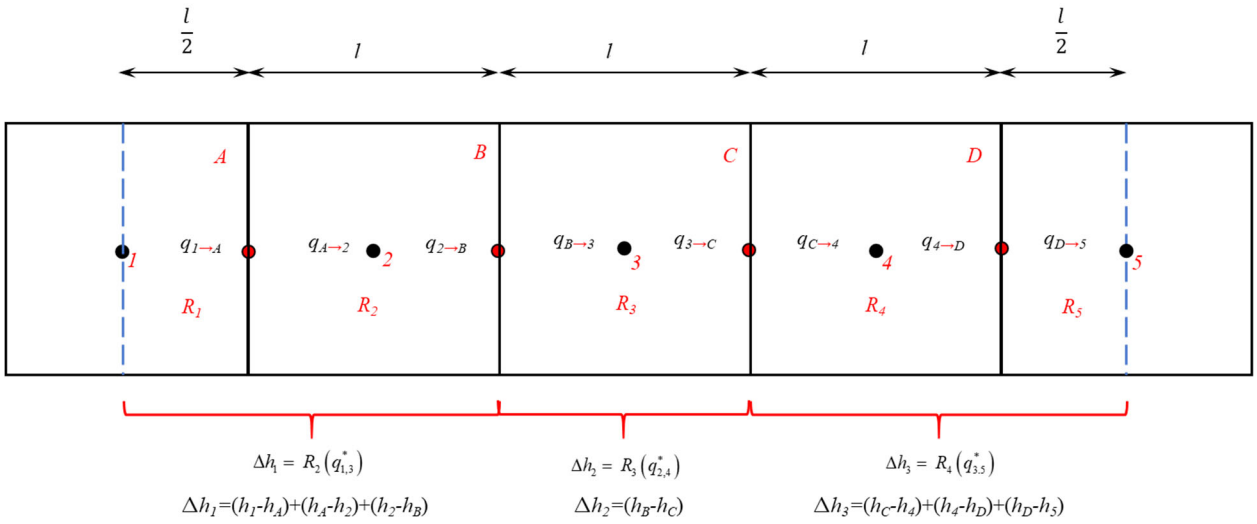


Figure 4. Head drops defined by vertex set {1, 2, 3, 4, 5} or by edge set {*A*, *B*, *C*, *D*}. The first step goes from the first vertex (1) to the center of the first edge after the second vertex (B). The final step goes from the center of the next-to-last edge (in this example, C) to the last vertex (in this example: 5).

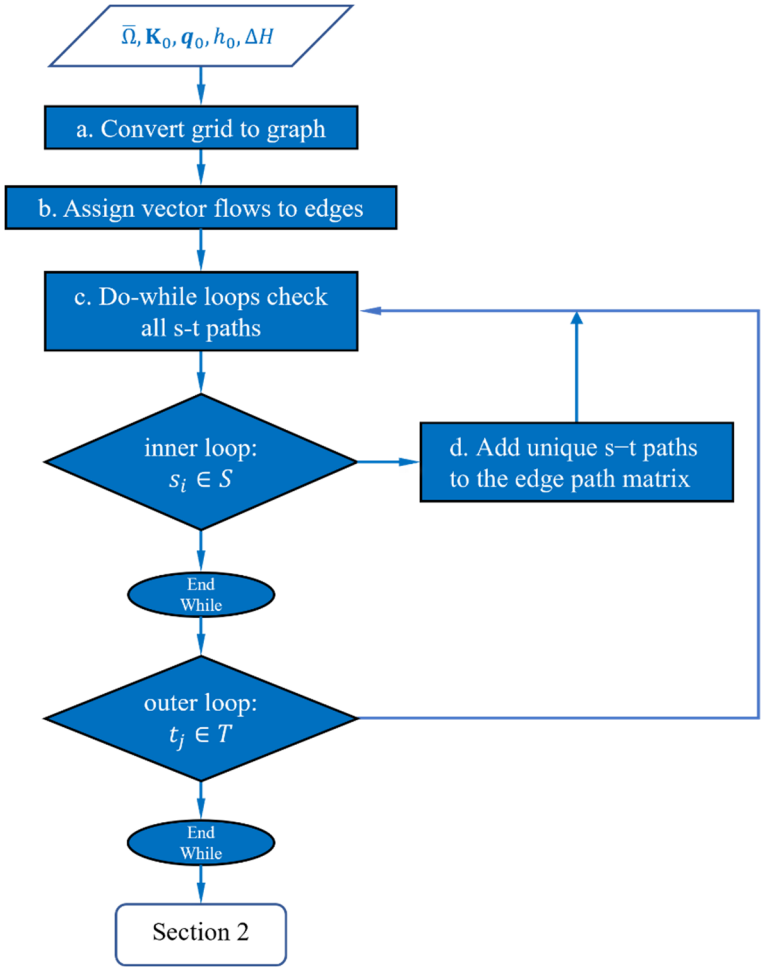


Figure 5. Schematic of the FlowPaths algorithm (Section 1). Inputs include the flow domain dimensions ($\bar{\Omega}$), the given specific discharges ($\mathbf{q}_0$) and the boundary head difference (ΔH). Inputs also include the given hydraulic conductivity ($\mathbf{K}_0$) and head distribution h_0 used for the stability analysis shown in Figure 6. (a) Convert cells in the flow domain $\bar{\Omega}$ to graph G using adjacency matrix $A(G)$. (b) Create directed graph D from the known $\boldsymbol{q}_0$. (c) Start nested loops $s_i \in S$ and $t_j \in T$. (d) Find the intersection of out-trees rooted at source s_i and in-trees rooted at target t_j, which form the envelope of the minimum directed spanning tree (MDST). Find the set of linearly independent (LI) paths for each subgraph. If at least one path exists from s_i to t_j, add the set of paths $P(s,t)$ to the edge path matrix. Continue to Section 2 (Figure 6).

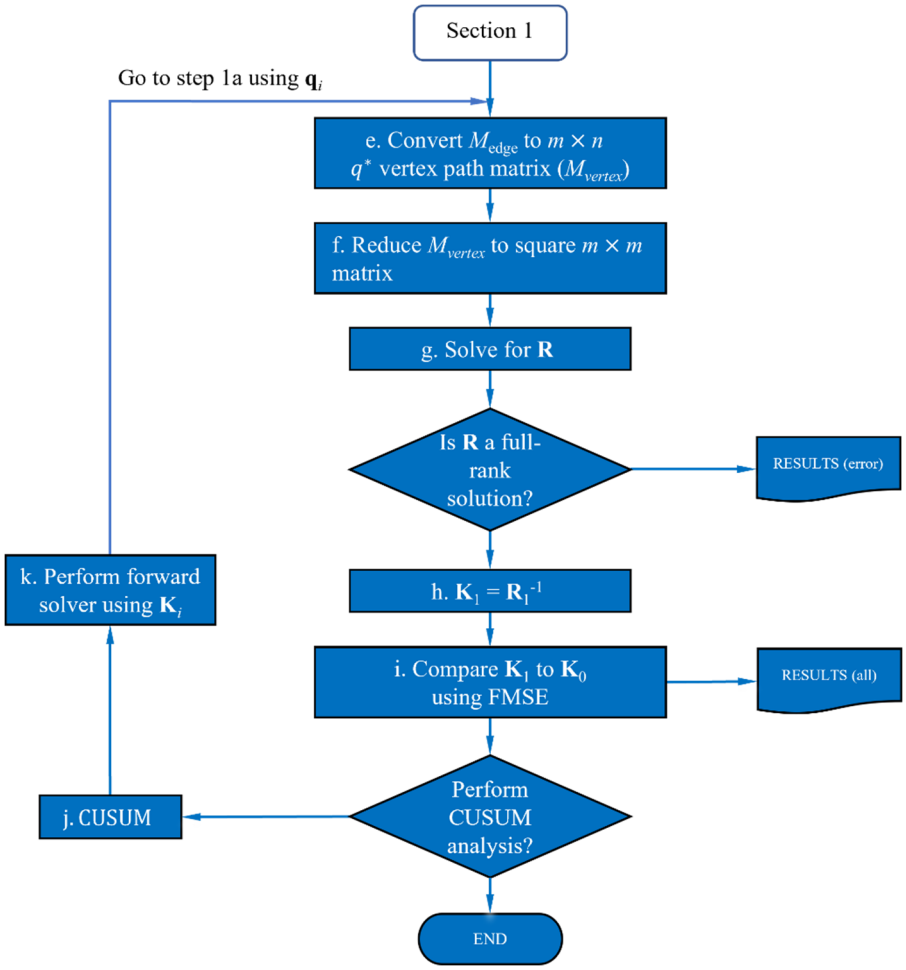


Figure 6. Schematic of the FlowPaths algorithm (Section 2), comprising the following steps: (e) Convert the edge path matrix to the q^* matrix, representing head drop equations for all independent paths. (f) Reduce the overdetermined q^* matrix to a square $m \times m$ matrix. (j) Solve for $\mathbf{R}_1$. (g) Is the result unique and full rank? If so, proceed to (h). If not, return an error. (h) Calculate $\mathbf{K}_1 = \mathbf{R}_1^{-1}$. (i) Calculate the fractional mean symmetric error (FMSE) between $\mathbf{K}_1$ and $\mathbf{K}_o$. Continue to recursive error test? If so, proceed to (j). If not, end FlowPaths. (k) Call the forward solver, using $\mathbf{K}_1$ as input, then return to Section 1 (Figure 5) and use $\mathbf{q}_1$ as input.

Having identified the LI paths for each (s,t) pair, we now combine them into a multi-source multi-sink (MSMS) system. The number of paths is also the number of equations to be solved. Gopalan and Ramasubramanian (2014) show that the maximum number of independent paths, or maximum rank, R_{max}, in an MSMS directed graph is equal to the number of edges, m, minus the number of vertices, n, plus M :

$$R_{max} = m - n + M \,, \tag{14}$$

where M is the total number of measurement vertices (the total number of source vertices and target vertices where the head is known). For example, the directed graph of Figure 3d has $m = 32$ edges, $n = 24$ vertices, and $M = 8$ (4 sources plus 4 targets) giving $R_{max} = 32\text{-}24\text{+}8 = 16$, which is the same as the number of vertices in the interior of the square flow matrix Ω^o (Figure 2). Those 16 paths will generate 16 linearly independent equations to be solved for the 16 unknown values of hydraulic conductivity. If the flow matrix has more columns than rows the inverse problem will be underdetermined.

Knowing the number of required independent paths, our next task is to list the vertices (or equivalently the edges) in each of those independent paths. This task requires us to define the minimum directed spanning tree (MDST). According to Sedgewick & Wayne (2014, p. 604), a minimum spanning tree (MST) is a connected subgraph H of a graph G with no cycles that includes all the vertices in G, with the total weight of its edges no more than any other spanning tree. The minimum *directed* spanning tree is the MST directed from source s to target t, which defines an (s,t) pair. Following the method of Zwick (2013), the MDST is identified by performing a sweep through all the vertices in H to find the smallest-weighted edges directed towards each vertex, which become the branches of the MDST of the subgraph H_{st}.
For example, the directed subgraph for $(s,t) = (3,22)$ is shown in Figure 7, with the MDST outlined in red. The total number of directed paths from vertex 3 to vertex 22 is 14, as enumerated in Table 1 using a breadth–first algorithm, such as the one presented by Korte and Vygen (2012, p. 26).

When all the q_k^* coefficients (Eq. 12) have been sorted into the proper elements for every path, the resulting linear system of equations is solved directly using the Matlab function mldivide (which is informally called backslash) (MathWorks, 2024). The cell-by-cell reciprocal of the resistance field **R** produces the hydraulic conductivity field **K**, as shown in steps *h*-*l* of Figure 6. The FlowPaths inverse model has been implemented in Matlab, and the relevant codes are available through the findable, accessible, interoperable, and reusable (FAIR) online repository HydroShare (see Section 6.5).

We note that the two-dimensional (2D) grid structure of the input is beneficial to obtain a low number of linearly independent paths, as determined by equation (14), which provide a unique solution that can be solved directly. Our path enumeration method is readily generalized to 3D, however, the computational burden would grow by at least an order of magnitude.

2.4 Model Verification

Model verification was performed with 13×98×10 = 12,740 proof-of-concept simulations where the hydraulic conductivity field $\mathbf{K}_o$ and the corresponding specific discharge field $\mathbf{q}_o$ are both known. There are 13 model sizes from 4×4 to 16×16; for each model size, there are 98 levels of heterogeneity quantified by a reservoir heterogeneity index, V_{dp}, from 0.01 to 0.98. The reservoir

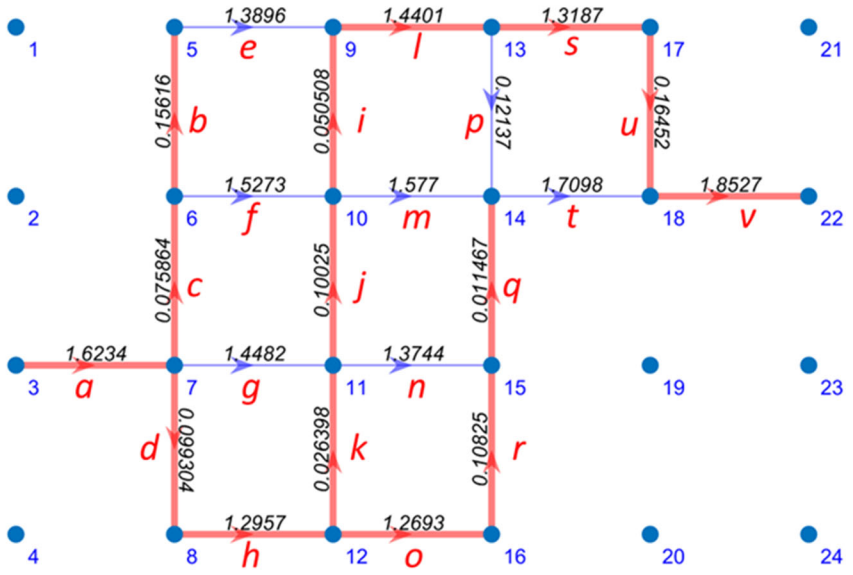


Figure 7. Directed subgraph $H_{3,22}$ connecting source vertex 3 to target vertex 22 in the directed graph shown in Figure 3d. The MDST of $H_{3,22}$ is shown with thick red edges.

Table 1. The 14 possible directed paths from vertex 3 to vertex 22 in subgraph $H_{3,22}$ shown in Figure 7. The eight independent paths, determined using the MDST, are in italics. The trunk line, a part of the MDST, is in bold.

Path	Vertices	Edges
1	[3,7,6,5,9,13,14,18,22]	[a,c,b,e,l,p,t,v]
2	*[3,7,6,5,9,13,17,18,22]*	*[a,c,b,e,l,s,u,v]*
3	*[3,7,6,10,9,13,14,18,22]*	*[a,c,f,i,l,p,t,v]*
4	[3,7,6,10,9,13,17,18,22]	[a,c,f,i,l,s,u,v]
5	[3,7,6,10,14,18,22]	[a,c,f,m,t,v]
6	*[3,7,8,12,11,10,9,13,14,18,22]*	*[a,d,h,k,j,i,l,p,t,v]*
7	***[3,7,8,12,11,10,9,13,17,18,22]***	***[a,d,h,k,j,i,l,s,u,v]***
8	*[3,7,8,12,11,10,14,18,22]*	*[a,d,h,k,j,m,t,v]*
9	*[3,7,8,12,11,15,14,18,22]*	*[a,d,h,k,n,q,t,v]*
10	*[3,7,8,12,16,15,14,18,22]*	*[a,d,h,o,r,q,t,v]*
11	*[3,7,11,10,9,13,14,18,22]*	*[a,g,j,i,l,p,t,v]*
12	[3,7,11,10,9,13,17,18,22]	[a,g,j,i,l,s,u,v]
13	[3,7,11,10,14,18,22]	[a,g,j,m,t,v]
14	[3,7,11,15,14,18,22]	[a,g,n,q,t,v]

heterogeneity index is calculated from the standard deviation of the log-hydraulic conductivity, $\hat{s}$, as follows (Jensen et al., 2020, p. 153):

$$V_{\mathrm{dp}} = 1 - \exp(-\hat{s}). \tag{15}$$

For example, when $V_{\mathrm{dp}} = 0.01$, then $\hat{s} = 0.01$; when $V_{\mathrm{dp}} = 0.98$, then $\hat{s} = 3.9$. For each level of heterogeneity, there are 10 realizations of a spatially correlated stochastic model for hydraulic conductivity, with lognormal univariate statistics and an isotropic Gaussian covariance structure, following the algorithm of Eftekhari (2018), implemented using Matlab code by Bergström (2010). The correlation length scale was chosen to be 1/200 of the cell size, rendering an essentially uncorrelated hydraulic conductivity field. This creates a worst case scenario that tests the ability of FlowPaths to find the hydraulic conductivity for uncorrelated fields with strong cell-to-cell contrast. To quantify the range of order of magnitudes (ROM) of the hydraulic conductivity in each realization, we calculate the ROM as follows:

$$\mathrm{ROM} = \log_{10}\left(\frac{K_{max}}{K_{min}}\right), \tag{16}$$

where K_{max} and K_{min} are the maximum and minimum values of hydraulic conductivity, respectively, for each realization of $\mathbf{K}_{\mathrm{o}}$.

For each realization of $\mathbf{K}_{\mathrm{o}}$ the discrete cellular head field, $\mathbf{H}_{\mathrm{o}}$ and the specific discharge field $\mathbf{q}_{\mathrm{o}}$ are determined by solving the forward problem using an alternating direction implicit (ADI) scheme developed by Esfandiari (2017). The ADI forward solution algorithm is unconditionally stable (Peaceman & Rachford, 1955), and has the advantage of using any stopping criteria for convergence. The ADI scheme, implemented using the Matlab function lsqminnorm (MathWorks, 2020), is executed until the changes in $\mathbf{H}_0$ between iterations are less than 1×10^{-12} cm; an additional stopping criterion limits the scheme to $1{\times}10^9$ iterations. The precision of the simulated data ($\mathbf{H}_{\mathrm{o}}$ and $\mathbf{q}_{\mathrm{o}}$) is on the order of 10 significant digits. Finally we use the simulated $\mathbf{q}_{\mathrm{o}}$ (but not the simulated $\mathbf{H}_{\mathrm{o}}$) as an input to the FlowPaths inverse model to estimate the underlying hydraulic conductivity field $\mathbf{K}_1$. In principle, the estimated field $\mathbf{K}_1$ should match the simulated field $\mathbf{K}_{\mathrm{o}}$.

To quantify the error for each realization, we use an unbiased estimator of error, ζ, which measures the logarithmic differences between all estimated values $y \in \mathbf{K}_1$ and simulated values $x \in \mathbf{K}_{\mathrm{o}}$. This estimator, based on the median symmetric accuracy defined by Morley et al. (2016), is the fractional mean symmetric error (FMSE):

$$\zeta = \exp\left\{\mu\left[\left|\ln\left(\frac{y}{x}\right)\right|\right]\right\} - 1, \tag{17}$$

where $\mu()$ computes the mean. Perfect accuracy would be denoted as $\zeta = 0$, and y and x are interchangeable because FMSE is symmetric. In addition, for each realization, we record the maximum log-difference, ζ_{max}

$$\zeta_{\mathrm{max},k} = \max\left(\left|\ln\left(\frac{y_i}{x_i}\right)\right|\right)_k . \tag{18}$$

We measure the robustness of the inverse solution using a recursive error test (Freebairn, 1978; Galpin & Hawkins, 1984; Pugachev, 1984). The estimated hydraulic conductivity field $\mathbf{K}_1$ is fed into the forward solver, producing $\mathbf{q}_2$, which in turn is fed into the FlowPaths inverse model, producing $\mathbf{K}_2$. The FMSE is calculated for each iteration, and the cumulative FMSE is calculated as follows:

$$C_r = \sum_{i=1}^{i=r} \zeta_i. \tag{19}$$

If the cumulative FMSE remains bounded over *r* recursions, this rules out fundamental error in the model (mis-specification). The recursive error test is implemented with the MATLAB function cusumtest (MathWorks, 2022), and additional information on the forward-inverse problem cycle (Figure 5 and Figure 6) is available in the Supporting Information (Section S5).

3. Results

Results are presented first for a 4×4 proof-of-concept example (Figure 3, Figure 7, Table 1) and then empirical error bounds are reported for the 12,740 model verification simulations. Complete results, with associated metadata, are available through HydroShare (see Section 6.5).

3.1 4×4 Example

The 4×4 example, whose specific discharge field is shown in Figure 4, is based on the particular realization of $\mathbf{K}_o$ shown in Table 2a. For this realization, the average hydraulic conductivity is 10.2 cm/day, the reservoir heterogeneity index is $V_{dp} = 0.5$, and the range of order of magnitudes of the hydraulic conductivities is ROM = 0.82. This average hydraulic conductivity is typical of fractured oil reservoir rock or silty sand (Freeze & Cherry, 1979, p. 29). Using cubic grid blocks of size $(1\text{ cm})^3$, the forward solution produced specific discharge fluxes through the cell faces, with a maximum cell balance error of 1.55×10^{-15} cm^3.

Table 2: Results from the 4×4 example (Figure 2 and Figure 7) in units of [m/d].

a.

0.146	0.146	0.041	0.041
0.213	0.119	0.042	0.075
0.226	0.115	0.034	0.067
0.155	0.142	0.034	0.037

b.

0.146	0.146	0.041	0.041
0.213	0.119	0.042	0.075
0.226	0.115	0.034	0.067
0.155	0.142	0.034	0.037

c.

4.441×10^{-16}	5.274×10^{-16}	-3.469×10^{-17}	-4.857×10^{-17}
5.551×10^{-16}	4.996×10^{-16}	-4.163×10^{-17}	-9.714×10^{-17}
4.441×10^{-16}	7.355×10^{-16}	-4.857×10^{-17}	-6.939×10^{-17}
1.665×10^{-16}	1.277×10^{-16}	-2.776×10^{-17}	-4.857×10^{-17}

d.

3.109×10^{-15}	3.553×10^{-15}	-8.882×10^{-16}	-8.882×10^{-16}
2.665×10^{-15}	3.997×10^{-15}	-8.882×10^{-16}	-1.332×10^{-15}
1.998×10^{-15}	6.217×10^{-15}	-1.776×10^{-15}	-8.882×10^{-16}
1.110×10^{-15}	9.104×10^{-15}	-8.882×10^{-16}	-1.332×10^{-15}

[a] Simulated hydraulic conductivity field $\mathbf{K}_o$.
[b] Estimated hydraulic conductivity field $\mathbf{K}_1$.
[c] Cell-by-cell model error $\mathbf{K}_1 - \mathbf{K}_o$.
[d] Cell-by-cell log error, where the FMSE for the trial is 1.776×10^{-15} m/d and the maximum absolute error is 9.104×10^{-15} m/d.

Using the specific discharge field from the forward model as an input, the FlowPaths inverse model produced $\mathbf{K}_1$ shown in Table 2b. The cell-by-cell error $\mathbf{K}_1 - \mathbf{K}_o$ is shown in Table 2c. These results have FMSE = 2.22×10^{-15} cm/sec and a maximum log-difference error of 9.10×10^{-15} compared to the simulated hydraulic conductivity parameter, $\mathbf{K}_o$. The recursive error test for the 4x4 example is shown in Figure 8. Tables of the specific discharges for the 4×4 example are available in the Supporting Information (Section S6).

3.2 Empirical Error Bounds

For each of the 13 model sizes from 4×4 to 16×16, the one realization (out of 980) with the largest FMSE, calculated with equation (17), is reported on Table 3. For each of these worst-case realizations, we report the reservoir heterogeneity index V_{dp}, the ROM, and FMSE, and the maximum log-difference. Accordingly, Table 3 provides empirical error bounds for the FlowPaths inverse model.

For the 11×11 matrix in Table 3, the FMSE is three orders of magnitude larger than the next largest FMSE. It is the only result out of the 12,740 trials that had a log-difference error for any cell greater than five parts per thousand. The recursive error test for this trial (similar to Figure 8) showed no instability over one million iterations, suggesting that this apparently isolated error resulted from a random grouping of cells with a strong contrast in hydraulic conductivity between adjacent cells that prevented the forward model from producing a correct set of specific discharges.

4. Discussion

FlowPaths opens a new avenue to study the feedback between flow, transport, reaction, and clogging in porous media. Essentially any technology conducted in porous media—including but not limited to groundwater remediation—carries the risk of clogging, which can result from many processes including deposition of suspended colloids, geochemical dispersion of clay minerals, precipitation reactions, biofilm growth, air entrapment, or consolidation (Manga et al., 2012). These processes make clogging a technical challenge in a variety of applications such as managed aquifer recharge, groundwater remediation, and oil and gas development. In all of these applications, the spatial distribution and temporal evolution of hydraulic conductivity $K(x,y,z,t)$ results from a feedback process: It depends on the processes mentioned above, but it also controls flow and therefore the transport and reaction that trigger the processes mentioned above. Simultaneously being a result and a cause is the essence of a feedback process. Following the motivational example presented in the Introduction (Figure 1), for two-dimensional flow domains, FlowPaths opens a new avenue to study this feedback process through laboratory experiments that measure the spatial distribution and temporal evolution of specific discharge $q(x,y,t)$ as an independent variable, for example, using particle image velocimetry (PIV). Specifically, FlowPaths provides a mechanism to interpret the measured $q(x,y,t)$ in terms of the fundamental $K(x,y,t)$, and through this interpretation, to study dynamic hydraulic conductivity, extending the foundational work in this area (Thullner et al. 2002, Li et al. 2008, Englert et al. 2009, El Mountassir et al. 2014, Yoon et al. 2019).

Like any numerical model, computational effort merits a brief discussion. One of the drawbacks of finding a solution for an inverse problem is that the demand on computer resources tends to escalate with the size of the problem (Neuman, 2006, p. 2). Our approach has a comparatively short time complexity due to the use of a fast shortest-path algorithm (Cai et al., 2000, p. 566). Computational times on a standard laptop, with 16GB of RAM running on a

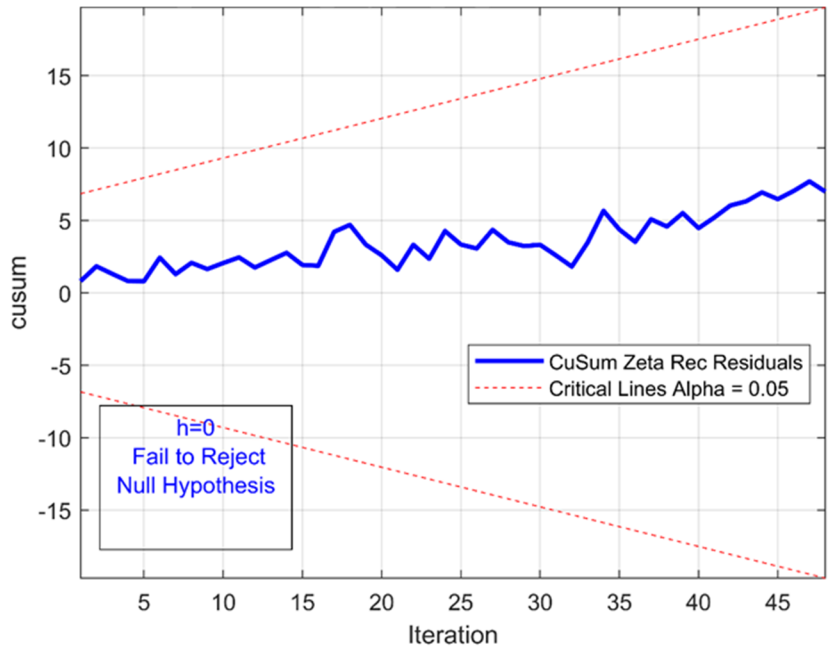


Figure 8. Cumulative sum of the regression residuals, C_r, versus iteration, r, for the 4×4 example. The null hypothesis is that small changes in the data (the result of the inversion) do not produce large changes in the next inversion. The "Fail to Reject Null Hypothesis" is signified by the graphic result that C_r does not exceed the $\alpha = \pm 5\%$ significance levels represented by the dotted critical lines at any step of the recursive error checker.

Table 3. Empirical error bounds for the FlowPaths inverse model. For each model size, the realization with the maximum fractional mean symmetric error (FMSE) is shown.

Model Size	V_{dp}	ROM	FMSE	Maximum Log-Difference
4×4	0.01	0.02	5.6×10^{-11}	8.7×10^{-11}
5×5	0.98	7.33	3.1×10^{-9}	7.4×10^{-8}
6×6	0.93	5.50	2.4×10^{-10}	9.2×10^{-10}
7×7	0.97	7.27	4.7×10^{-8}	3.7×10^{-6}
8×8	0.98	7.57	3.5×10^{-8}	8.8×10^{-7}
9×9	0.98	9.57	4.3×10^{-7}	1.3×10^{-5}
10×10	0.98	8.93	1.2×10^{-7}	5.9×10^{-6}
11×11	0.97	8.08	3.0×10^{-2}	1.8×10^{0}
12×12	0.97	8.15	1.5×10^{-6}	9.6×10^{-5}
13×13	0.98	8.84	2.9×10^{-6}	3.3×10^{-4}
14×14	0.98	8.84	3.0×10^{-6}	5.7×10^{-4}
15×15	0.97	7.46	5.4×10^{-6}	4.2×10^{-4}
16×16	0.92	6.44	3.0×10^{-5}	4.7×10^{-3}

Windows 10 platform scale from about 3 sec for the 4×4 example to about 5600 sec = 1.6 hour for the 16×16 example. The subroutine that finds the MDSTs for the directed subgraphs H_{st} accounts for about 50% of the computation time in smaller flow matrices, but increases to about 90% in larger ones. These computational times stem from the more challenging nature of inverse computations compared to forward computations, specifically, a set of inner loops for each combination of inflow and outflow vertices to generate the sets of unique paths. Although it is beyond the scope of the present work, parallel implementation of FlowPaths could lead to reduced computation times.

FlowPaths has been tested against simulations that provided an assumed ground truth. What might be the effect of (inevitable) experimental errors in specific discharge measurements? Application of FlowPaths requires a specific discharge field that obeys continuity, so the first step would be to allocate experimental errors in a way that preserves continuity (*e.g.*, Deleersnijder, 2001). In our recursive error tests, such as the one displayed in Figure 8, small changes in the inputs (i.e., velocity or equivalently specific discharge) give small changes in the outputs (i.e., hydraulic conductivity). We therefore expect that the small velocity measurement errors from the envisioned future experiments, on the order of epsilon, are likely to generate small errors, some function of epsilon, in the estimated hydraulic conductivity.

We conclude this section by posing a more philosophical question: If one knew the velocity field, why would one want to know the hydraulic conductivity field? This is not a trivial question, and it gets to the heart of the new approach proposed here. If the only goals were to predict water supply and contaminant transport in a static system—that is, a system in which hydraulic conductivity does not change with time—then there would be no need for FlowPaths. Knowing the velocity field, one would know everything there is to know about the flow. But we argue that, in many practical applications, the premise of a static system is too limiting, because hydraulic conductivity evolves dynamically. If our goal is to understand the nature of that evolution, and if we posit that hydraulic conductivity encapsulates the essence of transport in porous media, then we need a way to measure its evolution quantitatively.

5. Conclusions

We show that it is possible to uniquely solve the inverse problem for the hydraulic conductivity from observed flow vectors using the FlowPaths algorithm. The method relies on an analysis of the possible paths from the higher constant head cells to the lower constant head cells (the source vertices to the target vertices) through a square flow matrix. The minimum number of paths needed for a unique solution is equal to the number of vertices (cells) in the flow matrix, with the necessary condition that the paths be linearly independent of each other. Graph theory is used to find the paths between all sources and targets, which are converted to first order linear head-drop equations in the variable q^*, a combination of the specific discharges entering and leaving a vertex following a path. Then matrix algebra is used to solve the problem. The results of the inverse method are verified two ways: (1) by comparing the estimated $\mathbf{K}_1$ to the known $\mathbf{K}_0$, and (2) using a recursive error test, which detects very little variation in the inverse solution when the results are fed back into the forward solution. Future work is required to extend FlowPaths to non-square domains, to three-dimensional domains, and to anisotropic media, but in all these cases, the approach outlined here—based on identifying linearly independent paths using graph theory—is expected to prove itself useful.

6. Statements and Declarations

6.1 Supporting Information

The online Supporting Information provides additional information on graph theory (Section S1), well-posedness (Section S2), details on FlowPaths (Section S3), cyclomatic complexity (Section S4), the forward and inverse problem cycle (Section S5), and 4x4 example data tables (Section S6).

6.2 Acknowledgements

The authors thank Sam Welch and three anonymous referees for their constructive feedback.

6.3 Author Contributions

Conceptualization: MEM and DCM. Formal Analysis: SB. Methodology: MEM. Software: MEM. Supervision: DCM. Writing Original Draft: MEM. Review/Editing: SB and DCM.

6.4 Conflicts of Interest

The authors declare no real or perceived conflicts of interest in this work.

6.5 Data, Code, and Protocol Availability

Matlab codes, results, and associated metadata are available through HydroShare (https://www.hydroshare.org/resource/191c41fcb8294d6ab46484be693999a0/).

6.6 Funding Received

Mont-Eton acknowledges the financial support from the University of Colorado Denver scholarship fund along with conference travel support. Borgwardt acknowledges support through U.S. National Science Foundation award 2006183 and U.S. Air Force Office of Scientific Research award FA9550-21-1-0233.

6.7 ORCiD IDs

Steffen Borgwardt	https://orcid.org/0000-0002-8069-5046
David C. Mays	https://orcid.org/0000-0002-5218-1670
Michael E. Mont-Eton	https://orcid.org/0000-0002-8213-3055

REFERENCES

Arthur, J. K., Ruth, D. W., & Tachie, M. F. (2009). PIV measurements of flow through a model porous medium with varying boundary conditions. *Journal of Fluid Mechanics*, *629*, 343-374. https://doi.org/10.1017/s0022112009006405

Bastidas Olivares, M., Bringedal, C., & Pop, I.S. (2021). A two-scale iterative scheme for a phase-field model for precipitation and dissolution in porous media, *Applied Mathematics and Computation*, 396, 125933. https://doi.org/10.1016/j.amc.2020.125933

Baveye, P., Vandevivere, P., Hoyle, B. L., DeLeo, P. C., & de Lozada, D. S. (1998). Environmental Impact and Mechanisms of the Biological Clogging of Saturated Soils and Aquifer Materials. *Critical Reviews in Environmental Science and Technology*, *28*(2), 123-191. https://doi.org/10.1080/10643389891254197

Berge, C. (1973). *Graphs And Hypergraphs*. North-Holland Publishing Company. http://www.gbv.de/dms/hebis-darmstadt/toc/10727930.pdf

Bergström, D. (2010). *Rough surface generation & analysis*. Retrieved 11/29/2019 from http://www.mysimlabs.com/surface_generation.html

Berkowitz, B., & Ewing, R.P. (1998). Percolation theory and network modeling applications in soil physics, *Surveys in Geophysics*, 19(1), 23-72. https://doi.org/10.1023/A:1006590500229

Cai, M., Yang, X., & Li, Y. (2000). Inverse Problems of Submodular Functions on Digraphs. *Journal of Optimization Theory and Applications*, *104*(5), 559-575. https://doi.org/10.1023/A:1004685508517

Chang, C. M., & Yeh, H. D. (2010). Nonstationary stochastic analysis of flow in a heterogeneous unconfined aquifer subject to spatially-random periodic recharge. *Journal of Hydrology*, *395*(3-4), 163-168. https://doi.org/10.1016/j.jhydrol.2010.10.016

Cooper Jr., H.H., Bredehoeft, J.D., & Papadopulos, I.S. (1967). Response of a finite-diameter well to an instantaneous charge of water, *Water Resources Research*, 3(1), 263-269. https://doi.org/10.1029/WR003i001p00263

Darcy, H. (1856). *Les Fontaines Publiques de la Ville de Dijon*, Victor Dalmont, Paris.

Deleersnijder, E. (2001). Enforcing the continuity equation in numerical models of geophysical fluid flows, *Applied Mathematics Letters*, 14, 867-873. https://doi.org/ 10.1016/S0893-9659(01)00057-X

Eftekhari, A. A. (2018). *How to generate a log-random permeability field*. In https://github.com/simulkade/FVTool/blob/master/FieldGeology/field2d.m

Eikemo, B., Lie, K.-A., Eigestad, G. T., & Dahle, H. K. (2009). Discontinuous Galerkin methods for advective transport in single-continuum models of fractured media. *Advances in Water Resources*, *32*(4), 493-506. https://doi.org/10.1016/j.advwatres.2008.12.010

El Mountassir, G., Lunn, R. J., Moir, H., & MacLachlan, E. (2014). Hydrodynamic coupling in microbially mediated fracture mineralization: Formation of self-organized groundwater flow channels. *Water Resources Research*, *50*(1), 1-16. https://doi.org/ 10.1002/2013WR013578

Elmorsy, M., El-Dakhakhni, W., & Zhao, B.Z. (2022). Generalizable permeability prediction of digital porous media via a novel multi-scale 3D convolutional neural network, *Water Resources Research*, 58(3), e2021WR031454. https://doi.org/10.1029/2021WR031454

Englert, A., Hubbard, S. S., Williams, K. H., Li, L., & Steefel, C. I. (2009). Feedbacks between hydrological heterogeneity and bioremediation induced biogeochemical transformations. *Environ. Sci. Technol.*, *43*(14), 5197-5204. https://doi.org/ 10.1021/es803367n

Esfandiari, R. (2017). *Numerical Methods for Engineers and Scientists Using MATLAB* (2nd ed.). Taylor and Francis. https://doi.org/10.1201/9781315152417
Fitts, C. (2002). *Groundwater Science*. Academic Press.
Freebairn, J. W. (1978). Recursive Coefficient Estimates for the Evaluation of Varying Parameters. *Australian Journal of Statistics*, *20*(3), 219-228. https://doi.org/10.1111/j.1467-842X.1978.tb01104.x
Freeze, R., & Cherry, J. (1979). *Groundwater*. Prentice-Hall. Inc.
Galpin, J. S., & Hawkins, D. M. (1984). The use of recursive residuals in checking model fit in linear-regression. *American Statistician*, *38*(2), 94-105. https://doi.org/10.2307/2683242
Gärttner, S., Alpak, F.O., Meier, A., Ray, N., & Frank, F. (2023). Estimating permeability of 3D micro-CT images by physics-informed CNNs based on DNS, *Computational Geosciences*, *27*(2), 245-262. https://doi.org/10.1007/s10596-022-10184-0
Ghosh, T., Bringedal, C., Helmig, R., & Sekhar, G.P.R. (2020). Upscaled equations for two-phase flow in highly heterogeneous porous media: Varying permeability and porosity, *Advances in Water Resources*, 145, 103716. https://doi.org/10.1016/j.advwatres.2020.103716
Godefroy, G., Caumon, G., Laurent, G., & Bonneau, F. (2019). Structural interpretation of sparse fault data using graph theory and geological rules: Fault data interpretation. *Mathematical Geosciences*, *51*(8). https://doi.org/10.1007/s11004-019-09800-0
Gopalan, A., & Ramasubramanian, S. (2014). On the maximum number of linearly independent cycles and paths in a network. *IEEE/ACM Transactions on Networking*, *22*(5), 1373-1388. https://doi.org/10.1109/tnet.2013.2291208
Gosnell, D. K., & Broecheler, M. (2020). *The Practitioner's Guide To Graph Data: Applying Graph Thinking And Graph Technologies to Solve Complex Problems* (N. Barber, Ed. 1 ed.). O'Reilly Media, Inc. https://www.oreilly.com/library/view/the-practitioners-guide/9781492044062/
Harbaugh, A. W. (2005). *MODFLOW-2005, The U.S. Geological Survey Modular Ground-Water Model*. https://pubs.usgs.gov/tm/2005/tm6A16/PDF.htm
Hunt, A.G., & Ghanbarian, B. (2016). Percolation theory for solute transport in porous media: Geochemistry, geomorphology, and carbon cycling, *Water Resources Research*, 52(9), 7444-7459. https://doi.org/10.1002/2016WR019289
Hunt, A.G., & Sahimi, M. (2017). Flow, transport, and reaction in porous media: Percolation scaling, critical-path analysis, and effective medium approximation, *Reviews of Geophysics*, 55(4), 993-1078. https://doi.org/10.1002/2017RG000558
Jensen, J., Lake, L. W., Corbett, P., & Goggin, D. (2020). *Statistics for Petroleum Engineers and Geoscientists*. Prentice-Hall, Inc.
Jiao, J. Y., & Zhang, Y. (2015). Functional parameterization for hydraulic conductivity inversion with uncertainty quantification. *Hydrogeology Journal*, *23*(3), 597-610. https://doi.org/10.1007/s10040-014-1202-5
Katz, G. E., Berkowitz, B., Guadagnini, A., & Saaltink, M. W. (2011). Experimental and modeling investigation of multicomponent reactive transport in porous media. *120–121*(0), 27. https://doi.org/ 10.1016/j.jconhyd.2009.11.002
Kelm, M., Gärttner, S., Bringedal, C., Flemisch, B., Knabner, P., & Ray, N. (2022). Comparison study of phase-field and level-set method for three-phase systems including two minerals, *Computational Geosciences*, 26(3), 545-570. https://doi.org/10.1007/s10596-022-10142-w
Korte, B., & Vygen, J. (2012). *Combinatorial Optimization: Theory and Algorithms* (5 ed., Vol. 21). Springer Science & Business Media. https://doi.org/10.1007/978-3-642-24488-9

Larsson, I. A. S., Lundström, T. S., & Lycksam, H. (2018). Tomographic PIV of flow through ordered thin porous media. *Experiments in Fluids*, *59*(6), 96-103. https://doi.org/10.1007/s00348-018-2548-6

Li, X., Huang, H., & Meakin, P. (2008). Level set simulation of coupled advection-diffusion and pore structure evolution due to mineral precipitation in porous media. *Water Resour. Res.*, *44*, W12407. https://doi.org/ 10.1029/2007WR006742

Liu, G. S., Zhang, D. X., & Lu, Z. M. (2006). Stochastic uncertainty analysis for unconfined flow systems. *Water Resources Research*, *42*(9), 18, W09412. https://doi.org/10.1029/2005wr004766

Lu, Z. M., & Zhang, D. X. (2002). On stochastic modeling of flow in multimodal heterogeneous formations. *Water Resources Research*, *38*(10), 1190. https://doi.org/10.1029/2001wr001026

Lyu, Z., Chai, J., Xu, Z., Qin, Y., & Cao, J. (2019). A comprehensive review on reasons for tailings dam failures based on case history. *Advances in Civil Engineering*, *2019*, 4159306. https://doi.org/10.1155/2019/4159306

Manga, M., Beresnev, I., Brodsky, E. E., Elkhoury, J. E., Elsworth, D., Ingebritsen, S. E., Mays, D. C., & Wang, C. Y. (2012). Changes in permeability caused by transient stresses: Field observations, experiments, and mechanisms. *Rev. Geophys.*, *50*, RG2004, doi:2010.1029/2011RG000382

MathWorks. (2020). *Minimum Norm Least-Squares Solution to Linear Equation - MATLAB lsqminnorm*. Retrieved 6/26/20 from https://www.mathworks.com/help/matlab/ref/lsqminnorm.html

MathWorks. (2022). *Cusum Test for Structural Change - MATLAB cusumtest*. Retrieved 2/1/23 from https://www.mathworks.com/help/econ/cusumtest.html

MathWorks. (2024). mldivide. Retrieved 8/21/2024 from https://www.mathworks.com/help/matlab/ref/mldivide.html

Matoušek, J., & Gärtner, B. (2007). *Understanding and Using Linear Programming*. Springer Science and Business Media. https://doi.org/10.1007/978-3-540-30717-4

Mays, D.C. (2013). Clogging in managed aquifer recharge: Flow, geochemistry, and clay colloids, in Martin, R., ed., *Clogging Issues Associated with Managed Aquifer Recharge Methods*, International Association of Hydrogeologists (IAH) Commission on Managing Aquifer Recharge, Australia, 14-24.

Morley, S. K., & Los Alamos National Lab. (LANL), L. A., NM (United States). (2016, 2016/07/01). *Alternatives to Accuracy and Bias Metrics Based on Percentage Errors for Radiation Belt Modeling Applications*. Los Alamos National Laboratory, Los Alamos, NM (United States). Retrieved LA-UR-16-24592 from https://www.osti.gov/servlets/purl/1260362

Neuman, S. P. (1973). Calibration of distributed parameter groundwater flow models viewed as a multiple-objective decision process under uncertainty. *Water Resources Research*, *9*(4), 1006-1021. https://doi.org/10.1029/WR009i004p01006

Neuman, S. P. (2006). Blueprint for perturbative solution of flow and transport in strongly heterogeneous composite media using fractal and variational multiscale decomposition. *Water Resources Research*, *42*(6). https://doi.org/10.1029/2005WR004315

Panawalage, S. P., Rahman, M., Biazar, J., & Islam, M. R. (2018). Solution of inverse problem for reservoir permeability. *WIT Transactions on Engineering Sciences*, *45*. https://doi.org/10.2495/AFM040021

Peaceman, D. W., & Rachford, H. H. J. (1955). The numerical solution of parabolic and elliptic differential equations. *Journal of the Society for Industrial and Applied Mathematics*, *3*(1), 28-41. https://doi.org/10.1137/0103003

Pugachev, V. S. (1984). Chapter 7 - Estimator Theory (I. V. Sinitsiya, Trans.). In V. S. Pugachev (Ed.), *Probability Theory and Mathematical Statistics for Engineers* (pp. 242-272). Pergamon. https://doi.org/10.1016/B978-0-08-029148-2.50011-2

Rizzo, C., & de Barros, F. (2017). Minimum hydraulic resistance and least resistance path in heterogeneous porous media. *Water Resources Research*, *53*. https://doi.org/10.1002/2017wr020418

Rockafellar, R. T. (1970). The minimum of a convex function. In *Convex Analysis* (pp. 263-272). Princeton University Press. http://www.jstor.org/stable/j.ctt14bs1ff.31

Santos, J.E., Xu, D., Jo, H., Landry, C.J., Prodanovic, M., & Pyrcz, M.J. (2020). PoreFlow-Net: A 3D convolutional neural network to predict fluid flow through porous media, *Advances in Water Resources*, 138, 103539. https://doi.org/10.1016/j.advwatres.2020.103539

Sedgewick, R., & Wayne, K. (2014). Directed graphs. In *Algorithms, 4th Edition* (pp. 566). Pearson Education, Inc. https://algs4.cs.princeton.edu/42digraph/

Singurindy, O., & Berkowitz, B. (2003). Evolution of hydraulic conductivity by precipitation and dissolution in carbonate rock. *Water Resources Research*, *39*(1), 1016. https://doi.org/10.1029/2001wr001055

Sloane, N. J. A. (1972). On finding the paths through a network. *The Bell System Technical Journal*, *51*(2), 371-390. https://doi.org/10.1002/j.1538-7305.1972.tb01925.x

Sun, N. (1999). *Inverse Problems in Groundwater Modeling*. Springer. https://doi.org/10.1007/978-94-017-1970-4

Tartakovsky, A. M., Redden, G., Lichtner, P. C., Scheibe, T. D., & Meakin, P. (2008). Mixing-induced precipitation: Experimental study and multiscale numerical analysis. *Water Resources Research*, *44*(6), W06s04. https://doi.org/10.1029/2006wr005725

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, *Transactions of the American Geophysical Union*, 2, 519-524.

Thullner, M., Mauclaire, L., Schroth, M. H., Kinzelbach, W., & Zeyer, J. (2002). Interaction between water flow and spatial distribution of microbial growth in a two-dimensional flow field in saturated porous media. *Journal of Contaminant Hydrology*, *58*(3-4), 169-189. https://doi.org/ 10.1016/S0169-7722(02)00033-5

Tikhonov, A. N. A. V. Y. (1977). *Solutions of Ill-Posed Problems*. Halsted Press.

Trench, W. F. (2003). *Introduction to Real Analysis*. Pearson Education.

Tsang, C.-F., Neretnieks, I., & Tsang, Y. (2015). Hydrologic issues associated with nuclear waste repositories. *Water Resources Research*, *51*(9), 6923-6972. https://doi.org/10.1002/2015WR017641

Tutte, W. T. (1960). Convex representations of graphs. *Proceedings of the London Mathematical Society*, *10*(3), 304-320. https://doi.org/10.1112/plms/s3-10.1.304

USGS. (2022). *MODFLOW and Related Programs | U.S. Geological Survey*. USGS. https://www.usgs.gov/mission-areas/water-resources/science/modflow-and-related-programs

Viswanathan, H.S., Hyman, J.D., Karra, S., O'Malley, D., Srinivasan, S., Hagberg, A., & Srinivasan, G. (2018). Advancing graph-based algorithms for predicting flow and transport in fractured rock, *Water Resources Research*, 54(9), 6085-6099. https://doi.org/10.1029/2017WR022368

Yoon, H., Chojnicki, K. N., & Martinez, M. J. (2019). Pore-scale analysis of calcium carbonate precipitation and dissolution kinetics in a microfluidic device. *Environmental Science & Technology*, *53*(24), 14233-14242. https://doi.org/10.1021/acs.est.9b01634

Zhao, Y., Song, J., Cheng, K., Liu, Z., & Yang, F. (2024). Migration and remediation of typical contaminants in soil and groundwater: A state of art review. *Land Degradation & Development*, *35*(8), 2700-2715. https://doi.org/10.1002/ldr.5103

Zhu, J., Meng, Q., Chen, W., & Ma, Z. (2021). Interpreting the basis path set in neural networks. *Journal of Systems Science and Complexity*, *34*. https://doi.org/10.1007/s11424-020-0112-y

Zolla, V., Freyria, F.S., Sethi, R., Di Molfetta, A. (2009). Hydrogeochemical and biological processes affecting the long-term performance of an iron-based permeable reactive barrier, *Journal of Environmental Quality*, 38, 897–908. https://doi.org/10.2134/jeq2007.0622

Zwick, U. (2013). *Lecture Notes on "Analysis of Algorithms": Directed Minimum Spanning Trees*. Retrieved 1/16/2022 from https://www.cs.tau.ac.il/~zwick/grad-algo-13/directed-mst.pdf