Hydrogeological Uncertainty Estimation with the Analytic Element Method

M. Ramgraber[1,2], M. Schirmer[1,2]

[1]Department of Water Resources and Drinking Water, Eawag - Swiss Federal Institute of Aquatic Science and Technology, Switzerland

[2]Centre for Hydrogeology and Geothermics (CHYN), University of Neuchâtel, Switzerland

Corresponding author: *Maximilian Ramgraber (max.ramgraber@eawag.ch).*

Key Points:

- In this study, we explore the use of the Analytic Element Method (AEM) for hydrogeological uncertainty estimation using a MCMC algorithm
- We include a flexible element based on conformal mapping for the influence of uncertain regional flow in a simple Python-based AEM toolbox
- We find that AEM can be a useful tool for direct steady-state uncertainty estimation or act as a support tool during model creation

Abstract

Uncertainty estimation is an important part of practical hydrogeology. With most of the subsurface unobservable, attempts at system characterization will invariably be incomplete. Uncertainty estimation, then, must quantify the influence of unknown parameters, forcings, and structural deficiencies. In this endeavour, numerical modeling frameworks support an unparalleled degree of subsurface complexity and its associated uncertainty. When boundary uncertainty is concerned, however, the numerical framework can be restrictive. The interdependence of grid discretization and the enclosing boundaries make exploring uncertainties in their extent or nature difficult. The Analytic Element Method (AEM) may be an interesting complement, as it is computationally efficient, economic with its parameter count, and does not require enclosure through finite boundaries. These properties make AEM well-suited for comprehensive uncertainty estimation, particularly in data-scarce settings or exploratory studies. In this study, we explore the use of AEM for flow field uncertainty estimation, with a particular focus on boundary uncertainty. To induce versatile, uncertain regional flow more easily, we propose a new element based on conformal mapping. We then include this element in a simple Python-based AEM toolbox and benchmark it against MODFLOW. Coupling AEM with a Markov Chain Monte Carlo (MCMC) routine using adaptive proposals, we explore its use in a synthetic case study. We find that AEM permits efficient uncertainty estimation for groundwater flow fields, and its analytical nature readily permits continuing analyses which can support Lagrangian transport modelling or the placement of numerical model boundaries.

1 Introduction

Groundwater modelling plays an important role in practical hydrogeology. In a discipline in which neither the system nor its properties can be observed in its entirety, it is the task of models to establish spatial and temporal continuity between point-wise information. Where only few observations are available, uncertainty dominates the system characterization. These uncertainties must be quantified to endow any information derived from models with the correct confidence intervals.

Sources of model uncertainty are manifold and somewhat elusive, but arise from three main sources: unknown subsurface parameters (parametric uncertainty: e.g., Linde et al. 2017; Renard 2007), unknown boundary conditions and forcings (forcing uncertainty: e.g., Guillaume et al. 2016; Vrugt et al. 2008), and model structural inadequacies (conceptual uncertainty; e.g., Höge, Guthke, and Nowak 2019). In practice, hydrogeologists tend to devote most resources to parametric uncertainty, focusing on the ambiguity in sediment parameters such as hydraulic conductivity. Where forcing uncertainty is considered, it is rarely explored beyond the addition of white noise to perturb model predictions. This omission can be problematic, since the influence of forcing uncertainty may be indistinguishable from (e.g., Erdal and Cirpka 2016) or even eclipse (e.g., Peeters and Turnadge 2019) the effects of parametric uncertainty.

While all models of unisolated systems require boundary conditions in some form, the specific requirements depend on the chosen modelling framework. Numerical models place particularly stringent requirements, demanding a finite domain along which specified boundary conditions are enforced. Such boundaries are rarely well-defined in reality. Common pragmatic choices are

prescribed hydraulic head conditions inter- and extrapolated from marginal observation wells, or no-flow boundaries along anticipated streamlines. Since the nature of these boundaries intertwines them closely with the extent of the model grid, they do not lend themselves readily to the exploration of uncertainty in their extent or nature without adjusting the grid as well.

Prescribed in- or outflow boundary conditions would be a more versatile choice to represent the uncertain influence of regional flow in a finite domain, but are very difficult to inform and consequently rarely used. Recognizing this limitation, simulation frameworks such as MODFLOW 6 (Langevin et al. 2017) have since implemented multi-level setups which allow the use of simpler, large-scale models to inform the flow boundaries of the main area of interest. This shifts the boundary problem up the hierarchy, but at additional computational expense.

In search of a remedy, an interesting alternative may be found in the analytic element method (AEM: Haitjema 1995; Strack 1989, 2017). Instead of discretizing the model domain into cells or finite elements, AEM constructs a complex-valued, scale-invariant analytic solution to the flow field. This is achieved through superposition (addition) of simpler solutions, the eponymous analytic elements (Figure 1). While initially developed for two-dimensional, steady-state settings, AEM has since been extended to support three-dimensional model domains (Haitjema 1985), smooth inhomogeneities (Craig 2009), and transient dynamics (Furman and Neuman 2003). AEM code has been distributed in modelling frameworks such as TIMML (Bakker 2006; Bakker and Strack 2003), Visual AEM (Craig et al. 2009), or AnAqSim (Fitts et al. 2015).

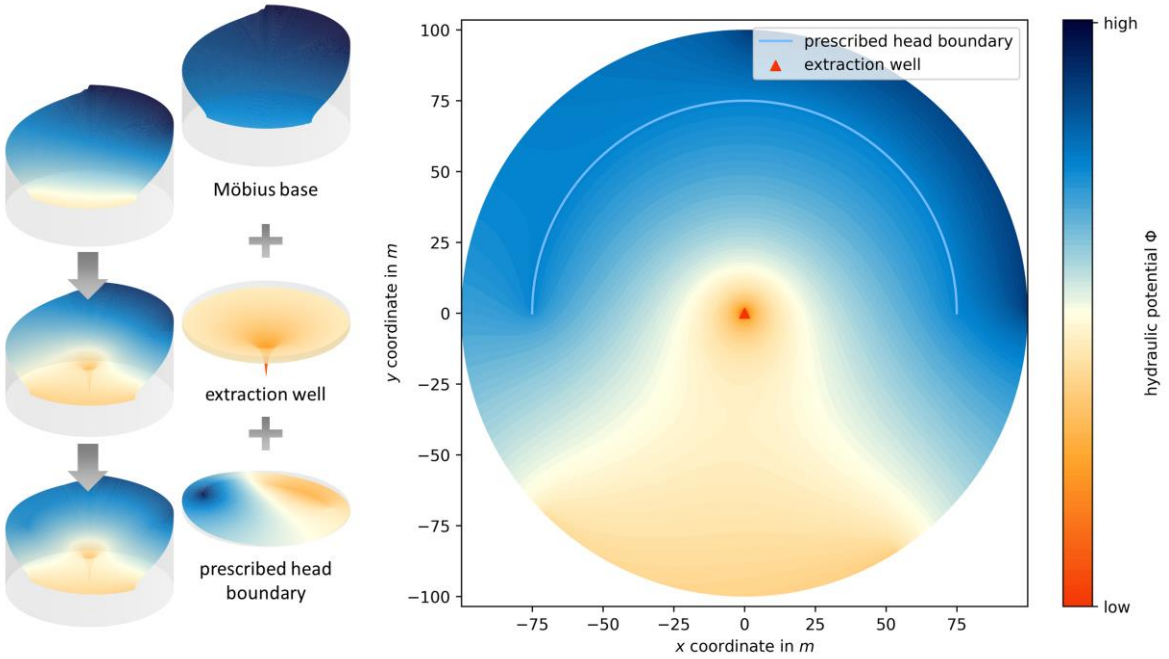

Figure 1. Example of how a AEM model is constructed. The influences of distinct elements (left) can be super-imposed to obtain a single, more complex solution (right, bottom left). Certain elements like prescribed head boundaries require additional preparation to ensure they induce the desired effect (see Section 2.1.8).

In the context of forcing uncertainty, AEM has the desirable property that it does not demand enclosure through finite, specified boundaries. Traditionally, regional flow is implemented as infinite, uniform flow, and subsequently deformed through the placement of farfield elements outside the immediate domain of interest. In principle, this method can induce highly complex flow fields, but is not very well suited for practical uncertainty estimation due to its indirect nature. More substantial changes to regional flow would require changing the position and rotation of the farfield elements.

In this study, our objective is two-fold: we strive to (i) demonstrate the intrinsic suitability of AEM for Bayesian inference, capitalizing on its naturally low parameter count and computational efficiency. Its natural approach to complexity (*start simple, add more complexity as required*) can make it more suitable for exploratory analyses than numerical models, which are often used the

other way around (*start complex, then simplify by aggregating grid parameters*). Towards this end, we (ii) propose a new element based on a Möbius transformation, which can directly induce curving, diverging, or converging regional base flow within a circular model domain of arbitrary size. This improves AEM's suitability for the exploratory analysis of boundary uncertainty.

We subsequently demonstrate the use of AEM and this element for the inference of local, two-dimensional, steady state flow fields. We provide a modular Python code coupling a simple AEM implementation to a MCMC routine, intended for preliminary explorations of plausible flow fields during model conceptualization, or simple Bayesian flow field inference in data-scarce environments.

2 Theory

In this section, we will outline the basic concepts of AEM and some of the most common elements (Section 2.1). The derivations summarized here are mainly based on and explored in much greater detail in the seminal works of Otto Strack (1989, 2017). For our Python implementation, we largely follow the object-oriented procedure suggested by Bakker & Kelson (2009). We present the uncertainty estimation algorithm used in this study in Section 2.2. Variables in **bold** notation denote vectors, matrices, or vector- or matrix-valued functions, while standard notation is reserved for scalar-valued variables and functions. The derivatives of all elements presented in the following are listed in Appendix 1 (supporting information).

2.1 Analytic Element Method

As opposed to conventional numerical models, the Analytic Element Method (AEM) does not seek a solution in terms of hydraulic head, but instead computes a complex valued potential

$$\boldsymbol{\Omega} = \boldsymbol{\Phi} + i\boldsymbol{\Psi} \tag{1}$$

where $\boldsymbol{\Phi}$ is the *discharge* (or *hydraulic*) *potential*, $\boldsymbol{\Psi}$ is the *stream function*, which corresponds to the flow direction, and $i = \sqrt{-1}$ is the imaginary unit. At the heart of the method lies the superposition of simple liear differential equations – the eponymous analytic elements – to yield more intricate solutions:

$$\boldsymbol{\Omega} = \sum_{e=1}^{E} \boldsymbol{f}_e(\boldsymbol{z}; \boldsymbol{\theta}_e) \tag{2}$$

where E is the number of analytic elements, and $\boldsymbol{f}_e$ is the function for a specific analytic element parameterized by $\boldsymbol{\theta}_e$ and evaluated at $\boldsymbol{z}$, a vector of complex-valued coordinates:

$$\boldsymbol{z} = \boldsymbol{x} + i\boldsymbol{y}. \tag{3}$$

where $\mathbf{x}$ and $\mathbf{y}$ are two coordinate components. The conversion of hydraulic potential $\boldsymbol{\Phi}$ into hydraulic heads $\boldsymbol{\phi}$ depends on the aquifer type (Eq. 8.12 and 8.13, Strack 1989):

$$\boldsymbol{\phi} = \frac{\boldsymbol{\Phi} + \frac{1}{2}kH^2}{kH} \qquad \textit{(confined, } \boldsymbol{\Phi} \geq \tfrac{1}{2}kH^2\textit{)} \tag{4}$$

$$\boldsymbol{\phi} = \sqrt{\frac{2\boldsymbol{\Phi}}{k}} \qquad \textit{(unconfined, } \boldsymbol{\Phi} < \tfrac{1}{2}kH^2\textit{)} \tag{5}$$

where k is the hydraulic conductivity and H is the thickness of the aquifer. Its inverse is:

$$\boldsymbol{\Phi} = kH\boldsymbol{\phi} - \frac{1}{2}kH^2 \qquad \textit{(confined, } \boldsymbol{\phi} \geq H\textit{)} \tag{6}$$

$$\boldsymbol{\Phi} = \frac{1}{2}k\boldsymbol{\phi}^2 \qquad \textit{(unconfined, } \boldsymbol{\phi} < H\textit{)} \tag{7}$$

Analytic elements can be broadly classified into two group. Some elements like extraction wells, line sinks, or area sinks only have to induce a *relative* change, for example extract a certain net amount of water. As such, they can simply be added to the stack and induce the desired flow

response. We will refer to such elements as *relative elements*. Other elements, like prescribed head boundary conditions, no-flow boundaries, or inhomogeneities must enforce an *absolute* condition at certain locations. These elements must adapt themselves to the influence of other elements. We will refer to such elements as *absolute elements*. The strength values of absolute elements can be found by setting up a system of linear equations, which we will describe in Section 2.1.8.

In the following, we will present a number of elements which will be used in this study. We will restrict our analysis to circular model domains, but note that all elements except the newly introduced Möbius base can operate on arbitrary – even infinite – domains.

2.1.1 Uniform base flow (relative)

Many analytic elements require some sort of background potential or base flow. A classic choice in AEM is uniform regional flow with specified direction, offset, and gradient:

$$\mathbf{\Omega} = Q\boldsymbol{z}\exp(-i\alpha) + \frac{\Phi_{min} + \Phi_{max}}{2} \tag{8}$$

$$Q = \frac{\Phi_{max} - \Phi_{min}}{2r_d} \tag{9}$$

where α is the flow's rotation in radians relative to the eastern axis, r_d is the radius of the circular domain, and Φ_{min} and Φ_{max} are the minimum and maximum discharge potential. These potentials can alternatively be obtained by defining a minimum and maximum hydraulic head ϕ_{min} and ϕ_{max}, then converting them to hydraulic potentials with Equation (6) or (7). The second right-hand side term in Equation (8) offset the potential to the range between Φ_{min} and Φ_{max}. An example of this flow is illustrated in Figure 2a.

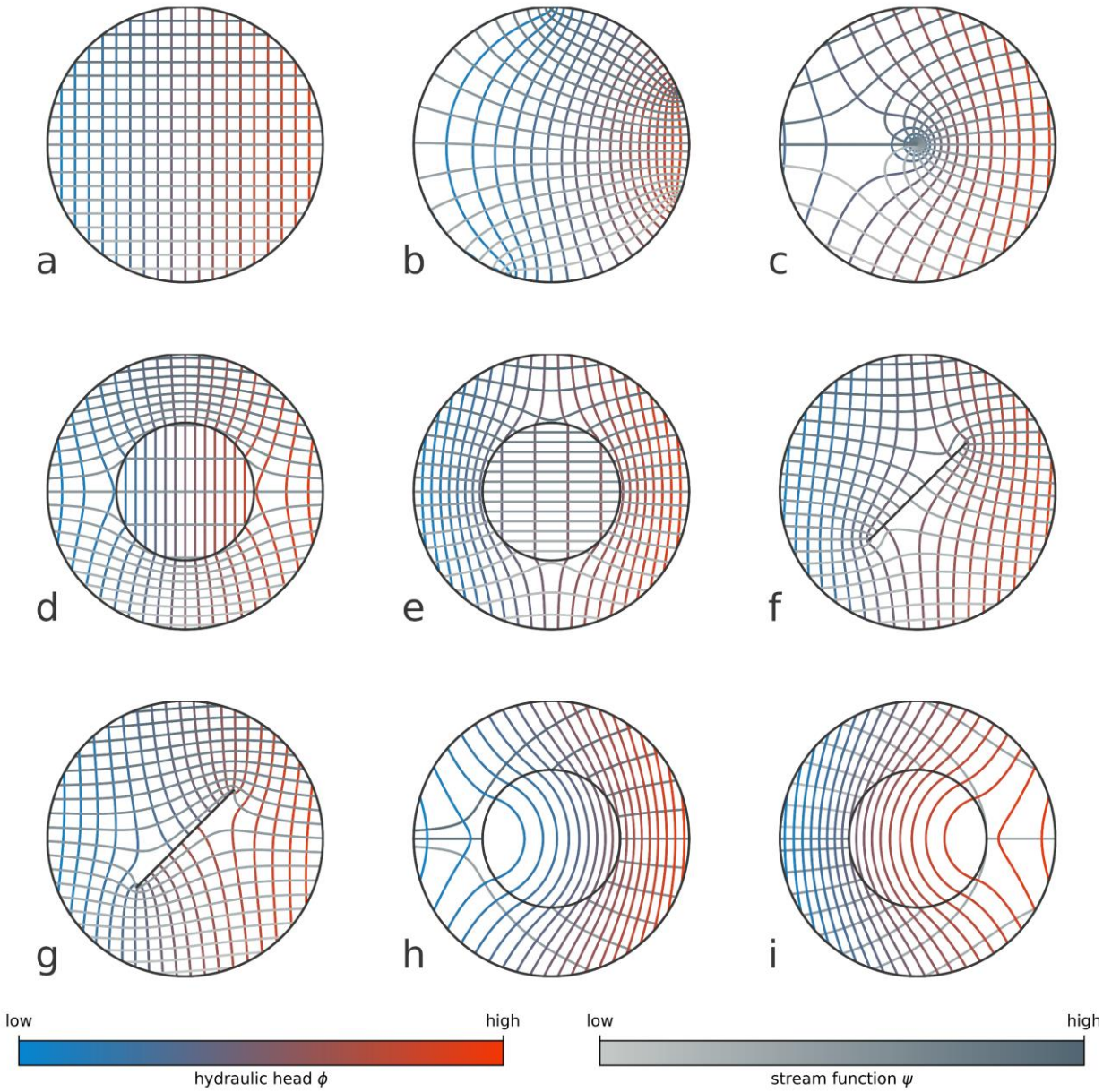

Figure 2. Various analytic elements used in this study. (a) uniform base flow, (b) Möbius base flow, (c) rate-specified extraction well, (d) polygonal inhomoheneity (lower), (e) polygonal inhomogeneity (higher), (f) prescribed head boundary, (g) no-flow boundary, (h) areal sink (negative), and (i) areal sink (positive). All elements are shown for confined conditions, and all elements from (c) onwards use uniform base flow. Solutions for the stream function are not valid inside areal sinks (h, i) and have been masked inside the element. Elements which add or remove water from the system (wells, line sinks, area sinks) induce branch cuts, discontinuities in the stream function from the element westwards.

2.1.2 Möbius base flow (relative)

Analytic elements can be superimposed on fields obtained from *conformal mapping* (e.g., Olver 2018). Conformal mapping is a class of angle-preserving transformations which can convert grids in the complex plane – or, more specifically, simple complex potential fields such as flow on the

unit square from east to west – into more complex shapes, while preserving the validity of the solution. Classic example of such methods are Schwarz-Christoffel transformations (SC: e.g., Driscoll and Trefethen 2009) and Möbius transformations.

To obtain more complex regional flow than the uniform flow defined in Section 2.1.1, we chain a Schwarz-Christoffel transformation with a Möbius transformation. First, the Schwarz-Christoffel transformation maps the unit square onto the unit disk. Then, the Möbius transformation deforms the flow inside the unit disk. This process is illustrated in Figure 3.

For computational purposes, we are mainly interested in the inverse of these maps: from the deformed unit disk (representing points in the model domain $\boldsymbol{z}$, Figure 3c) back onto the unit square (yielding the corresponding complex potential $\boldsymbol{\Omega}$, Figure 3a). The first step consists of casting the model domain back onto the unit disk through translation and scaling:

$$\boldsymbol{z}_{ud} = \frac{\boldsymbol{z} - z_d}{r_d} \tag{10}$$

where $\boldsymbol{z}$ are coordinates in the model domain, z_d is the center and r_d the radius of the circular model domain. The inverse Möbius transformation is defined as:

$$\boldsymbol{\Lambda} = \boldsymbol{M}^{-1}(\boldsymbol{z}_{ud}; a, b, c, d) = \frac{-d\boldsymbol{z}_{ud} + b}{c - a} \tag{11}$$

where $\boldsymbol{\Lambda}$ are the complex coordinates on the standard unit disk (Figure 3b), and a, b, c, and d are the Möbius coefficients. These coefficients can be computed by defining three reference points on the standard unit circle (e.g., points A, B, and C in Figure 3b) and their *images* on the edge of the transformed Möbius unit disk (i.e., A, B, and C in Figure 3c). Instead of working with the complex-valued coordinates on the unit circle, it is easier to express them in terms of polar coordinates of unit length and angle φ:

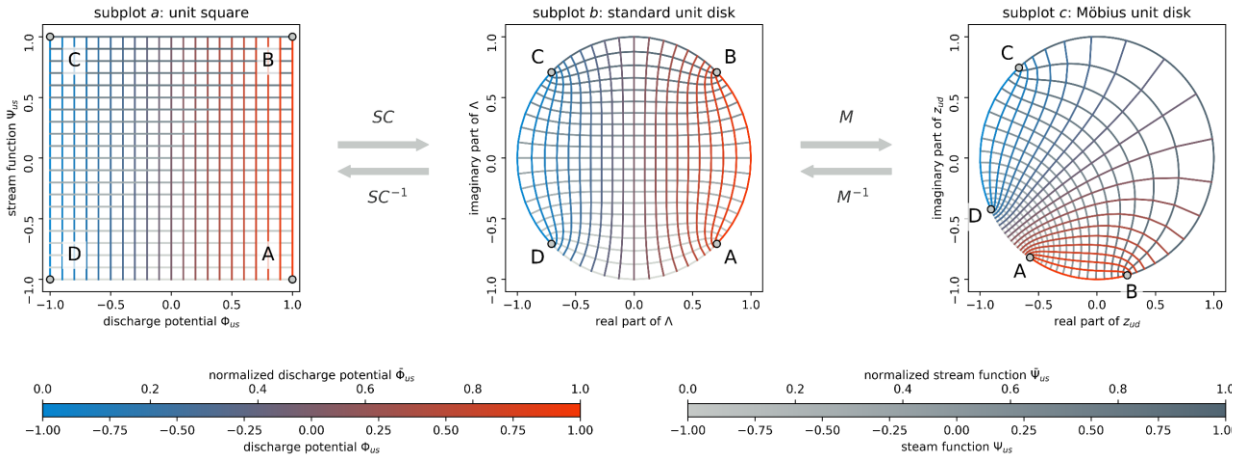

Figure 3. Illustration of the conformal mapping from the unit square (a) through the unit disk (b) to the Möbius-transformed unit disk (c). The control points A, B, and C in the unit disk (b) and the Möbius-transformed unit disk (c) define the coefficients for the Möbius transformation. The forward transformations $\boldsymbol{SC}$ *and* $\boldsymbol{M}$ *are listed in Appendix 2.*

$$Z_\varphi = P(\varphi) = \cos\varphi + i\sin\varphi \tag{12}$$

If we consider the static reference points in Figure 3b

$$\begin{aligned} A := \lambda_A &= \cos(-0.25\pi) + i\sin(-0.25\pi) \\ B := \lambda_B &= \cos(+0.25\pi) + i\sin(+0.25\pi) \\ C := \lambda_C &= \cos(+0.75\pi) + i\sin(+0.75\pi) \end{aligned} \tag{13}$$

we can express their images on the Möbius unit circle of Figure 3c similarly

$$\begin{aligned} z_A &= \cos(\varphi_A) + i\sin(\varphi_A) \\ z_B &= \cos(\varphi_B) + i\sin(\varphi_B) \\ z_C &= \cos(\varphi_C) + i\sin(\varphi_C) \end{aligned} \tag{14}$$

Consequently, we can specify the Möbius transformation through three variables only: φ_A, φ_B, and φ_C. We can then calculate the desired Möbius coefficients:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} z_B - z_C & z_A z_C - z_A z_B \\ z_B - z_A & z_A z_C - z_C z_B \end{bmatrix}^{-1} \cdot \begin{bmatrix} \lambda_B - \lambda_C & \lambda_A\lambda_C - \lambda_A\lambda_B \\ \lambda_B - \lambda_A & \lambda_A\lambda_C - \lambda_C\lambda_B \end{bmatrix} \tag{15}$$

We may then map the standard unit disk onto the unit square with an inverse Schwarz-Christoffel transformation (Fong 2019)

$$\boldsymbol{\Omega}_{us} = \boldsymbol{SC}^{-1}(\boldsymbol{\Lambda}) = \frac{1-i}{-K_e} F\left(\boldsymbol{c} = \cos^{-1}\left(\frac{1+i}{\sqrt{2}}\boldsymbol{\Lambda}\right), m = \frac{1}{\sqrt{2}}\right) + 1 - i \tag{16}$$

where $\boldsymbol{\Omega}_{us}$ is the complex potential on the unit square, $F(\boldsymbol{c}, m)$ is the incomplete Legendre elliptical of the 1st kind with argument $\boldsymbol{c}$ and parameter m:

$$F(\boldsymbol{c}, m) = \int_0^{\boldsymbol{c}} \frac{1}{\sqrt{1 - m\sin^2 t}} dt \tag{17}$$

and $K_e = F(c = 0.5\pi, m = 0.5) \approx 1.854$. This transformation is illustrated in Figure 3a and b. $\boldsymbol{\Omega}_{us}$ may then be transformed into a user-specified range:

$$\boldsymbol{\Omega} = (\Phi_{max} - \Phi_{min})\,(\boldsymbol{\Omega}_{us} + 1)/2 + \Phi_{min} \tag{18}$$

A further examples of a Möbius base flow is shown in Figure 2b.

2.1.3 Extraction or injection wells (relative)

Wells are among the simplest analytic elements. The standard form of this element is (Equation 24.8, Strack 1989):

$$\boldsymbol{\Omega} = -\frac{Q}{2\pi}\ln(\boldsymbol{z} - z_{well}) + C \tag{19}$$

where Q is a positive or negative real-valued discharge, z_{well} is the location of the well in terms of complex coordinates, and C is a real constant. To recognize the need for the constant, consider the effect of the logarithm: If $\boldsymbol{z} = z_{well}$, the hydraulic potential is negative infinity, if $\|\boldsymbol{z} - z_{well}\| \to \infty$, the hydraulic potential approaches positive infinity.

This constant is usually combined with other additive constants in Equation (2) and determined by specifying a reference point of (assumed) known hydraulic potential in the so-called *farfield*, outside the area of interest. Defining this reference point is not a trivial task (Bakker et al. 2016; Haitjema 1995).

For practical purposes, we propose a slight adaptation. We adjust Equation (19) so that its influence on the discharge potential is zero at a distance of $\|z - z_{well}\| = r_i$ by subtracting the induced potential at this distance, which yields:

$$\boldsymbol{\Omega} = -\frac{Q}{2\pi}\left(\ln(\boldsymbol{z} - z_{well}) - \ln(r_i)\right) \tag{20}$$

where r_i is the radius at which we assume the induced drawdown to become zero. Outside the range of r_i, the hydraulic potential increases to positive infinity, so its value should be selected cautiously. By default, we set $r_i = 2r_d$. An example of this element is illustrated in Figure 2c.

2.1.4 Inhomogeneities (absolute)

As the subsurface can be heterogeneous, we may wish to create zones of piece-wise constant hydraulic conductivity. Towards this end, we can define zones of discontinuous hydraulic conductivity enclosed by a polygon:

$$\boldsymbol{\Omega} = \sum_{j=1}^{N_j}\left[\frac{s_j}{4\pi i}\left(\left(\boldsymbol{Z}_j + 1\right)\ln\left(\frac{\boldsymbol{Z}_j - 1}{\boldsymbol{Z}_j + 1}\right) - \left(\boldsymbol{Z}_{j+1} - 1\right)\ln\left(\frac{\boldsymbol{Z}_{j+1} - 1}{\boldsymbol{Z}_{j+1} + 1}\right)\right)\right] \tag{21}$$

Where N_j is the number of vertices spanning up the polygon, s_j are the strengths of each vertex, and

$$\boldsymbol{Z}_j = \frac{2\boldsymbol{z} - \left(z_{j-1} + z_j\right)}{z_j - z_{j-1}} \tag{22}$$

where z_j is the jth vertex of a closed polygon, so that $z_0 = z_{N_j}$ and $z_{N_j+1} = z_1$. The vertices' strengths s_j are determined as part of the linear system (see Section 0) to enforce the desired discontinuity between the conductivity outside (k^-) and inside (k^+) the polygon. Examples for inhomogeneities of lower and higher conductivity are illustrated in Figure 2d and Figure 2e.

2.1.5 Prescribed head boundaries (absolute)

Line sinks inject or remove water along a line. As such, they can be used to simulate infiltration or exfiltration, for example from rivers, or to induce prescribed head boundaries. The extraction rate along a segment of the line element can be spatially varying (e.g., Janković and Barnes 1999; Strack 2018). In this study, we will use the constant rate formulation, and string together multiple *segments* of a line sink to implement spatially-varying extraction rates. Strack (1989) expresses the induced complex potential of line elements in terms of local coordinates:

$$\boldsymbol{Z} = \frac{2\boldsymbol{z} - (z_1 + z_2)}{z_2 - z_1} \tag{23}$$

where $\boldsymbol{Z}$ is the local coordinate, $\boldsymbol{z}$ the corresponding original coordinate, and z_1 and z_2 the original coordinates of the segment's start and end points. This is a small conformal mapping, projecting the line sink onto the real axis between -1 and $+1$, and transforming all evaluation points accordingly. Using these local coordinates, the induced potential is defined as (eq. 8.479, Strack 2017):

$$\boldsymbol{\Omega} = s\frac{L}{4\pi}\left((\boldsymbol{Z}+1)\ln(\boldsymbol{Z}+1) - (\boldsymbol{Z}-1)\ln(\boldsymbol{Z}-1) + 2\right) \tag{24}$$

where s is the strength of the line sink and $L = |z_2 - z_1|$ its length in global coordinates. Similarly to the well element, this element's induced hydraulic potential increases unbounded with distance. As a consequence, we subtract the effect an influence radius similarly to Section 2.1.3:

$$\boldsymbol{\Omega} = s\frac{L}{4\pi}\left((\boldsymbol{Z}+1)\ln(\boldsymbol{Z}+1) - (\boldsymbol{Z}-1)\ln(\boldsymbol{Z}-1) - \left(\frac{2r_i}{L}+2\right)\ln\left(\frac{2r_i}{L}+2\right) + \left(\frac{2r_i}{L}\right)\ln\left(\frac{2r_i}{L}\right)\right) \tag{25}$$

where r_i is the distance of the zero influence point from the segment's end point in terms of global coordinates. This element can be used for prescribed head boundaries (as an absolute element) or as a specified flow line sink (as a relative element). We may further extend

Equation (25) by introducing a connectivity parameter $0 \leq c \leq 1$ to scale each segment's injection or extraction strength s after the linear system has been solved. A connectivity of 0 means the flow field is unaffected by the prescribed head boundary, a connectivity of 1 implies the flow field is fully controlled at the prescribed head boundary. An example of this element is illustrated in Figure 2f.

2.1.6 No-flow boundaries (absolute)

Line doublets create discontinuities in the complex potential while maintaining the mass balance. These elements can be used to create zonal inhomogeneities or no-flow boundaries. Their complex potential influence can be derived as (Strack 1989):

$$\boldsymbol{\Omega} = \frac{s}{4\pi i}\left((\boldsymbol{Z}+1)\ln\left(\frac{\boldsymbol{Z}-1}{\boldsymbol{Z}+1}\right) - (\boldsymbol{Z}-1)\ln\left(\frac{\boldsymbol{Z}-1}{\boldsymbol{Z}+1}\right)\right) \tag{26}$$

where s is the element's real-valued strength, and $\boldsymbol{Z}$ are localized coordinates of Equation (23). No-flow boundaries are absolute elements and can be obtained by chaining together several such segments in sequence. The no-flow condition is enforced by requiring that the gradient of the hydraulic potential across each segment must be zero (see Section 0). An example of such an element is illustrated in Figure 2g.

2.1.7 Area sinks (relative)

Area sinks are among the more complicated elements. Similar to inhomogeneities, they are defined based on closed polygons (Eq. 8.598, Strack 2017):

$$\boldsymbol{\Omega} = -\frac{Q}{32\pi i}\sum_{j=1}^{N_j} {L_j}^2\left(\boldsymbol{Z}_j - \overline{\boldsymbol{Z}}_j\right)\boldsymbol{H}\left(\boldsymbol{Z}_j\right) + \frac{QA}{2\pi}\ln(\boldsymbol{z} - z_1) \tag{27}$$

where Q is the designated flux rate per area, L_j is the length of the side from vertex $j-1$ to j, $\boldsymbol{Z}_j$ is calculated according to Equation (22), $\overline{\boldsymbol{Z}}_j$ is its complex complement, A is the area enclosed by the polygon, and $\boldsymbol{H}(\boldsymbol{Z}_j)$ is calculated according to:

$$\boldsymbol{H}(\boldsymbol{Z}_j) = (\boldsymbol{Z}_j + 1)\ln\left(\frac{\boldsymbol{Z}_j - 1}{\boldsymbol{Z}_j + 1}\right) + 2\sum_{m=j+1}^{N_j} \ln\left(\frac{\boldsymbol{Z}_m - 1}{\boldsymbol{Z}_m + 1}\right) + 2 \tag{28}$$

Examples of this element are illustrated in Figure 2h and Figure 2i.

2.1.8 Solving the system of linear equations

As we established in the beginning of Section 2.1, some elements must enforce absolute conditions at certain locations and thus depend on the influences of all other elements. To simultaneously satisfy the conditions imposed by these elements, their strengths $\boldsymbol{s}$ are estimated jointly as part of a linear system of equations (e.g., Bakker and Kelson 2009; Strack 1989):

$$\boldsymbol{A} \cdot \boldsymbol{s} = \boldsymbol{x} \tag{29}$$

where $\boldsymbol{A}$ is a $N \times N$ matrix, N is the number of the absolute elements' line segments or polygon vertices, $\boldsymbol{s}$ is a vector of length N containing all absolute segments' strength parameters s, and $\boldsymbol{x}$ is a vector of length N specifying the conditions to be met.

To assemble $\boldsymbol{A}$, each entry of the nth row contains the influence of the other absolute elements on the nth absolute control point (the center of a line segment or a polygon vertex) at unit strength ($s = 1$). For segments of no-flow boundaries, each corresponding row is filled with the hydraulic potential gradient $\partial\boldsymbol{\Phi}/\partial\boldsymbol{n}$ along the segment's normal vector $\boldsymbol{n}$. The required derivatives of all elements are reported in Appendix 1. For prescribed head boundaries and inhomogeneities, each row contains influences on the hydraulic potential $\boldsymbol{\Phi}$ (Figure 4, Step 2). A special case applies for the diagonal entries corresponding to inhomogeneities. To these entries,

a term specifying the discontinuity of hydraulic conductivity is added ($k^-/(k^+ - k^-)$, see Section 2.1.4).

The target vector $\boldsymbol{x}$ is assembled so that each absolute element's conditions are fulfilled:

- For *prescribed head boundaries*, the target hydraulic potential at the segment is calculated according to Equation (6) or (7), depending on aquifer type. Then, the $\boldsymbol{\Phi}$ induced by the independent elements is subtracted from the target potential.
- For *no-flow boundaries*, the target is set to $-\partial\boldsymbol{\Phi}/\partial\boldsymbol{n}$, the negative gradient induced by all relative elements on the segment (Figure 4, Step 3).
- For *inhomogeneities*, the target hydraulic potential at the vertex is set to the $\boldsymbol{\Phi}$ induced by the independent elements times -1.

With $\boldsymbol{A}$ and $\boldsymbol{x}$ defined, a standard linear solver can be used to obtain the unknown strength parameters $\boldsymbol{s}$.

2.2 Bayesian inference

Bayesian statistics are a formalized way of assigning probability densities, which can be interpreted as a plausibility metric, to different alternative hypotheses, often defined as a vector of unknown parameters $\boldsymbol{\theta} = [\theta_1, \dots, \theta_D]^\top$, where D is the number of uncertain parameters. When the number of hypotheses is infinite, for example in the case of continuous variables, Bayesian inference works with probability densities instead. These densities are defined through so-called *probability density functions* (*pdf*). Bayes' Theorem formalizes the process of updating one's state of knowledge by combining initial belief – the prior $p(\boldsymbol{\theta})$ – with new information – the likelihood $p(\boldsymbol{y}|\boldsymbol{\theta})$ – to obtain the posterior pdf $p(\boldsymbol{\theta}|\boldsymbol{y})$ after normalization through the model evidence $p(\boldsymbol{y})$:

277 The prior is generally user-specified, and the likelihood is evaluated based on a user-defined

278 function which specifies the probability of obtaining the observations made ($\boldsymbol{y}$) given the current

$$p(\boldsymbol{\theta}|\boldsymbol{y}) = \frac{p(\boldsymbol{\theta})p(\boldsymbol{y}|\boldsymbol{\theta})}{p(\boldsymbol{y})} \tag{30}$$

279

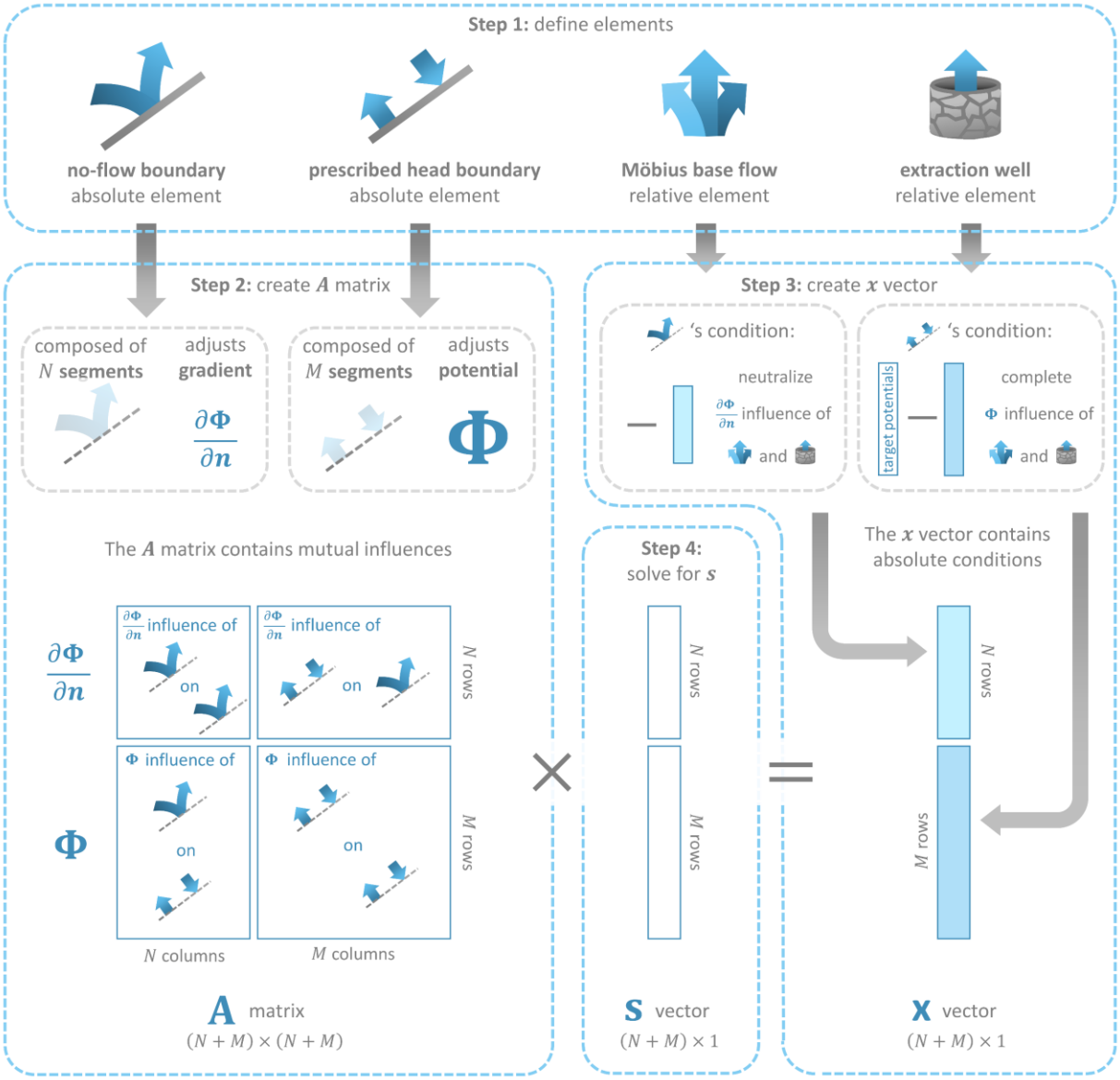

Figure 4. Schema for setting up the system of linear equations, illustrated in an example with two relative elements (Möbius base flow, extraction well) and two absolute elements (no-flow boundary, prescribed head boundary). After the elements are defined (Step 1), we create the $\boldsymbol{A}$ matrix by specifying the mutual influences of the absolute elements' segments onto each other (Step 2). Similarly, we construct the target $\boldsymbol{x}$ vector by evaluating and correcting the influence of the relative elements on the absolute elements' segments (Step 3). Finally, we can solve the linear system $\boldsymbol{A} \cdot \boldsymbol{s} = \boldsymbol{x}$ for $\boldsymbol{s}$ (Step 4). In this figure, $\boldsymbol{n}$ denotes a segment's normal vector, and $\times$ denotes the dot product.

hypothesis $\boldsymbol{\theta}$. Unfortunately, it is generally impossible to solve Equation (30) analytically. The model evidence is rarely known, and it is not always possible to find a tractable analytic formulation for the posterior $p(\boldsymbol{\theta}|\boldsymbol{y})$.

2.2.1 MCMC

However, even if Equation (30) does not have a closed form solution, inference methods such as MCMC can still *sample* from the unknown posterior (e.g., Kruschke 2015). This is useful because a sufficiently large sample set from a random distribution can act as a surrogate for the distribution itself, and consequently be used to infer its properties.

MCMC achieves this by starting from an initial hypothesis – the start point $\boldsymbol{\theta}_0$ – then exploring similar hypotheses nearby by sampling from a proposal distribution, exploring the neighbourhood of the current hypothesis. This proposal distribution can, for example, be Gaussian:

$$\boldsymbol{\theta}_{proposal} \sim \mathcal{N}(\boldsymbol{\theta}_{k-1}, \boldsymbol{\Sigma}) \tag{31}$$

where $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ defines a (possibly multivariate) normal distribution with mean $\boldsymbol{\mu} = \boldsymbol{\theta}_{k-1}$ (the previous hypothesis) and (co)variance $\boldsymbol{\Sigma}$. '$\sim$' represents 'sampled from', and the subscript k denotes the current iteration, or entry in the chain. MCMC then assembles a chain of samples by comparing each new hypothesis (the *proposal*) with the chain's last entry (the *reference*), and accepts or rejects the proposal based on its unnormalized posterior density relative to the reference. If the proposal density is symmetric (it is equally probable to jump from the reference to the proposal than from the proposal to the reference), the acceptance probability p_{accept} can be calculated as:

$$p_{accept} = \min\left(1, \frac{p(\boldsymbol{\theta}_{proposal})p(\boldsymbol{y}|\boldsymbol{\theta}_{proposal})}{p(\boldsymbol{\theta}_{k-1})p(\boldsymbol{y}|\boldsymbol{\theta}_{k-1})}\right) \quad (32)$$

Equation (32) states that if the proposal is more plausible than the reference, the proposal is automatically accepted ($p_{accept} = 1$) and appended to the chain ($\boldsymbol{\theta}_k = \boldsymbol{\theta}_{proposal}$). If it is less plausible, it is either accepted with a probability equal to the posterior density ratio ($p_{accept} < 1$) and appended to the chain ($\boldsymbol{\theta}_k = \boldsymbol{\theta}_{proposal}$), or rejected, in which case the reference is appended to the chain once more ($\boldsymbol{\theta}_k = \boldsymbol{\theta}_{k-1}$).

2.2.2 Adaptive proposals

A practical challenge in MCMC is that if the proposal distribution is sub-optimal, the chain will reject an inordinate amount of proposals, and consequently contain only very few unique samples. To avoid this issue, we implement our MCMC routine with an adaptive proposal distribution, gradually adjusting the proposal's covariance $\boldsymbol{\Sigma}$ so that a desired acceptance rate is achieved. Towards this end, we estimate an uncorrelated covariance matrix $\boldsymbol{\Sigma}_{unique}$ (i.e., all off-diagonal entries are set to zero) from the unique samples in the chain at regular intervals, then scale it with an adjustable factor f. The adjustable factor is increased if the acceptance rate was too high during the last interval, and decreased if it was too low:

$$\boldsymbol{\Sigma} = f\boldsymbol{\Sigma}_{unique} \quad (33)$$

Since it has been shown that adjustable proposals can corrupt the ergodic property, which is critical to the proper functioning of MCMC, it is important that the adjustments vanish asymptotically (e.g., Andrieu and Thoms 2008). For $\boldsymbol{\Sigma}_{unique}$ this should occur automatically as the number of samples increases, and for the adjustable factor f we achieve this by reducing the magnitude of its update exponentially:

$$f_c = f_{c-1}\left(a^{-c}\left(\frac{ratio_{accepted}}{ratio_{target}} - 1\right) + 1\right) \quad (34)$$

where $a > 1$ is a scalar which defines the speed of the decay, c is the current adjustment cycle, $ratio_{accepted}$ the percentage of accepted proposals since the last iteration, and $ratio_{target}$ the desired percentage of accepted proposals.

3 Examples

In this section, we illustrate the performance of the algorithm for two test cases. For the first test case, we benchmark our Python AEM code against MODFLOW, a well-established numerical finite-volume (FVM) framework. In the second test case, we demonstrate the performance of the AEM model and the MCMC inference mechanism in a synthetic test case. The codes for both scenarios are provided under a DOI which does not yet exist, see acknowledgements.

3.1 Benchmarking

For benchmarking, we compare the results from our Python AEM implementation to the results of a steady-state MODFLOW 6 model (Bakker et al. 2016; Langevin et al. 2017) at varying grid resolutions. Towards this end, we design a simple synthetic model employing every element described in Section 2.1 except the base flow elements (uniform or Möbius), as they are redundant in a domain enclosed by absolute boundaries (a prerequisite of the numerical reference). A schematic illustration of the benchmarking model is illustrated in Figure 5a, and the results compared to different resolutions of uniformly-sized hexagonal grids are shown in Figure 5b-d.

The results indicate that our AEM code can faithfully reproduce the FVM predictions, and that conversely the FVM results converge towards the AEM solution at finer grid sizes. This is mainly

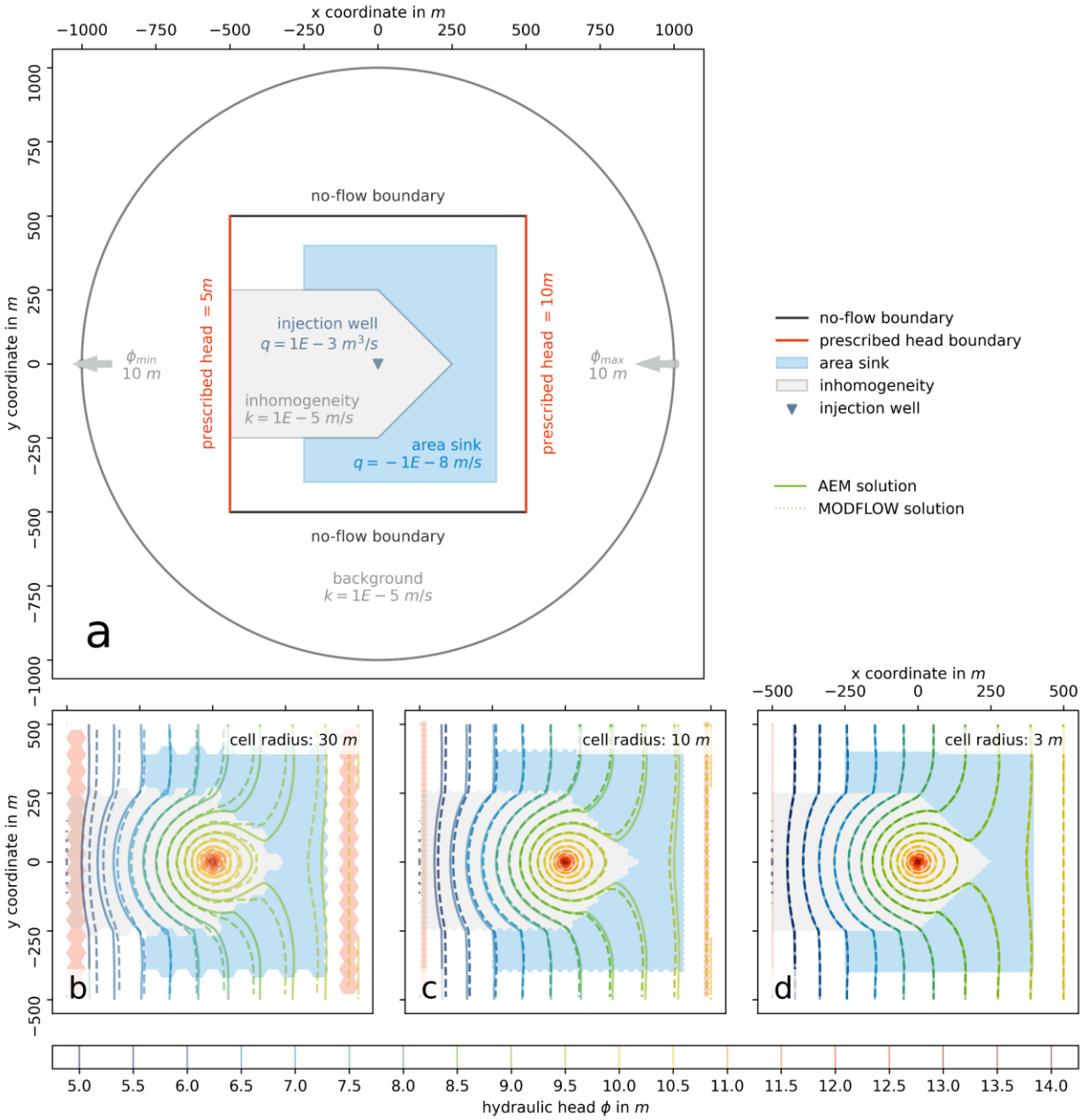

Figure 5. Illustrations of the benchmarking test case. To define a system which can be equivalently evaluated numerically, we embed a square area enclosed by two no-flow and two prescribed head boundaries within (functionally irrelevant) uniform flow. The model domain includes an area sink, an imhomogeneity, and an injection well (a). We compare the AEM results (solid line; b, c, d) with three FVM grids of different cell sizes (dashed line): 30 m (b), 10 m (c), and 3 m (d). The finer the cell resolution, the more the FVM results converge towards the AEM solution. The discrepancies arise because numerical grid sizes determine how well flow-relevant features can be resolved.

owed to the inability of rougher grid sizes to reproduce the boundaries and features in Figure 5a as faithfully as finer resolutions. We do however note that unstructured grids – which may adopt

the cell size and shape locally to reproduce features more prescisely – can yield solutions closer to AEM more efficiently than the uniform regular grids shown here.

3.2 Synthetic test case

To illustrate the potential of AEM for practical groundwater field inference, we apply the algorithm in a synthetic test case. Revisiting the motivating issue of uncertain boundary conditions, we design our synthetic site as local part of a larger catchment, with limited head information and poorly defined boundary conditions. Such scenarios abound in hydrogeological practice.

The water table is assumed to be observed at an extraction well and three surrounding observation wells (Figure 6a). Towards the north-west, a river of unknown connectivity intersects the domain. This river is implemented as a prescribed head boundary with four support nodes for the interpolation of the river's connectivity. The river is embedded in an inhomogeneity of unknown hydraulic conductivity. In the south, we prescribe a no-flow boundary representing an impermeable geological formation.

We define priors centered around the true solution (Table S1, supporting information), independent Gaussian observation errors ($\mu = 0\mathrm{m}$, $\sigma = 0.15\mathrm{m}$), and a MCMC chain length of 10,000. The original proposal distribution is listed in Table S2 (supporting information).

The results are illustrated in (Figure 6b-d). The RMSE ($0.021\ m$) and bias ($-0.004\ m$) reflect the model's lack of structural error, but a glance at its uncertainty (Figure 6c) reveals the ambiguity in the system's states. Uncertainty in the groundwater field is lowest between the observation wells, then swiftly increases towards the domain's edges.

Although this synthetic scenario is somewhat simplistic, we can illustrate some of the potential of AEM beyond simple groundwater field inference. Figure 6d illustrates a selection of uncertain

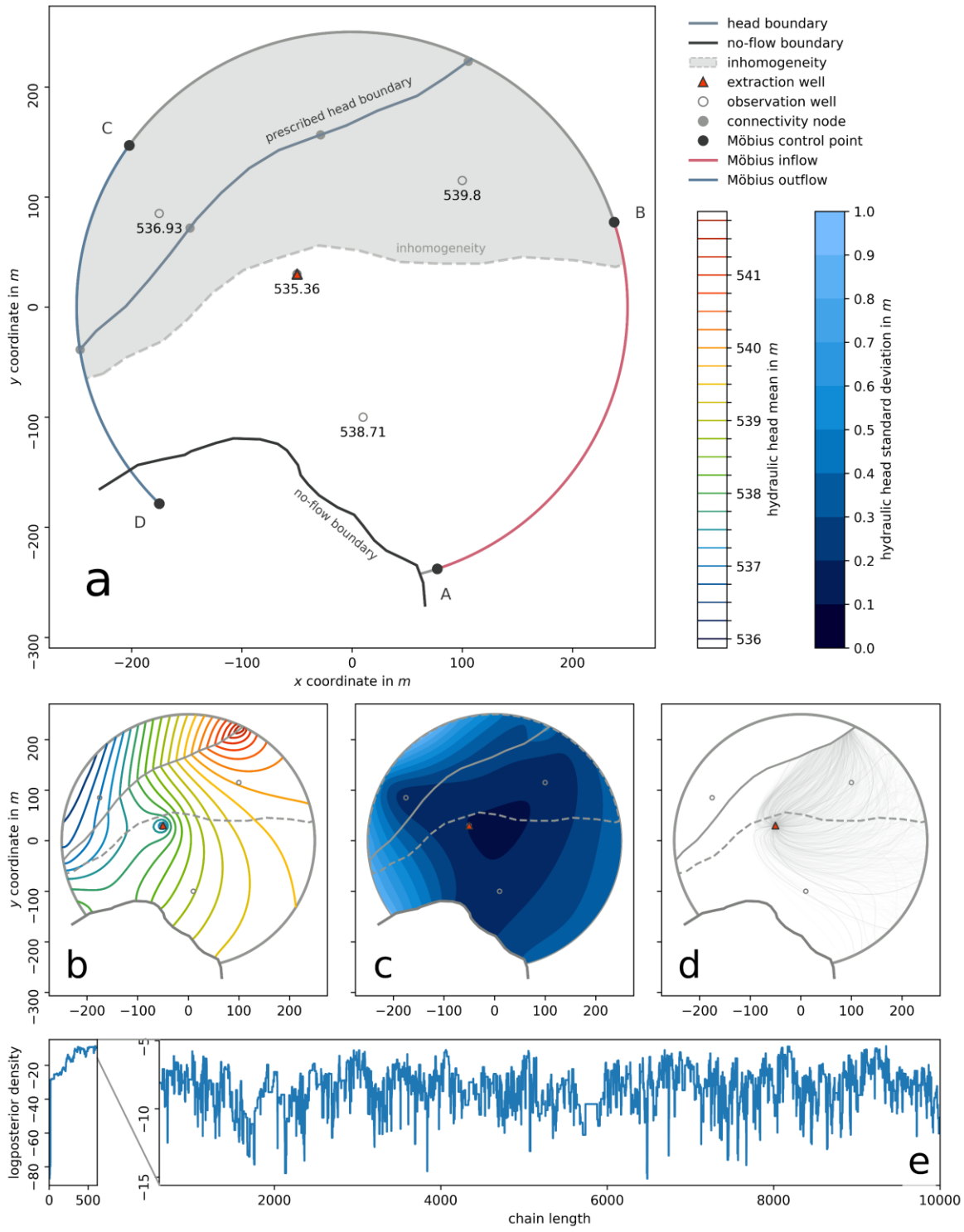

Figure 6. Schematic illustration (a) and results (b, c, d, e) for the synthetic test case. The model is implemented on a Möbius base with an areal inhomogeneity, an extraction well, a prescribed head boundary with spatially interpolated connectivity, and a no-flow boundary. Hydraulic heads are observed at three observation wells and the pumping well. Subplots (b) and (c) show the posterior mean and standard deviation of hydraulic head. Subplot (d) shows a selection of pathlines towards the extraction well. Subplot (e) shows the proportional logposterior density (right) and its cutoff (left) of the MCMC.

flow paths towards the extraction well, which may serve as the basis for Lagrangian transport modelling. Similarly, we could investigate the sign and magnitude of the river segments' strength values to obtain probabilistic estimates about which parts of the river are losing or gaining.

Using the algorithm as a support tool for model conceptualization, we can explore its use for the assignment of numerical model boundaries. This transition to numerical models might be motivated by Eulerian transport simulations or the investigation of transient dynamics.

Naturally, it would be possible to calculate regional fluxes based on the AEM predictions and assign these as inflow or outflow boundaries. For transient dynamics, however, the user might wish to return to no-flow boundaries and time-variable prescribed head boundaries, which are more easily informed through marginal observation wells. If the observations and priors for the steady state AEM simulation were sufficiently representative of the average dynamics, either boundary type is best assigned in regions where the posterior AEM flow direction is relatively certain (deep blue regions in Figure 7).

The map in Figure 7 can be obtained by capitalizing on AEM's analytical nature to directly evaluate the hydraulic potential gradient $\frac{\partial \boldsymbol{\Phi}}{\partial \boldsymbol{z}}$ for each entry in the MCMC chain. The resulting set of gradients at each individual location $z \in \boldsymbol{z}$ can be converted to a set of flow directions $\boldsymbol{\alpha} = \operatorname{atan2}\left(x = \Re\left(\frac{\partial \boldsymbol{\Phi}}{\partial z}\right), y = \Im\left(\frac{\partial \boldsymbol{\Phi}}{\partial z}\right)\right)$, where $\Re(\cdot)$ and $\Im(\cdot)$ are the real and imaginary components of their respective arguments. Since the direction of flow it is irrelevant for the assignment of boundaries (i.e., $\alpha = -0.25\pi \triangleq 0.75\pi$), we can furthermore offset all angles $\boldsymbol{\alpha} < 0$ by π to obtain a more concise estimate, assuming $-\pi \geq \boldsymbol{\alpha} > \pi$. If we then calculate the circular standard deviation for each $\boldsymbol{\alpha}$, we obtain the contours in Figure 7. Prescribed head boundaries with

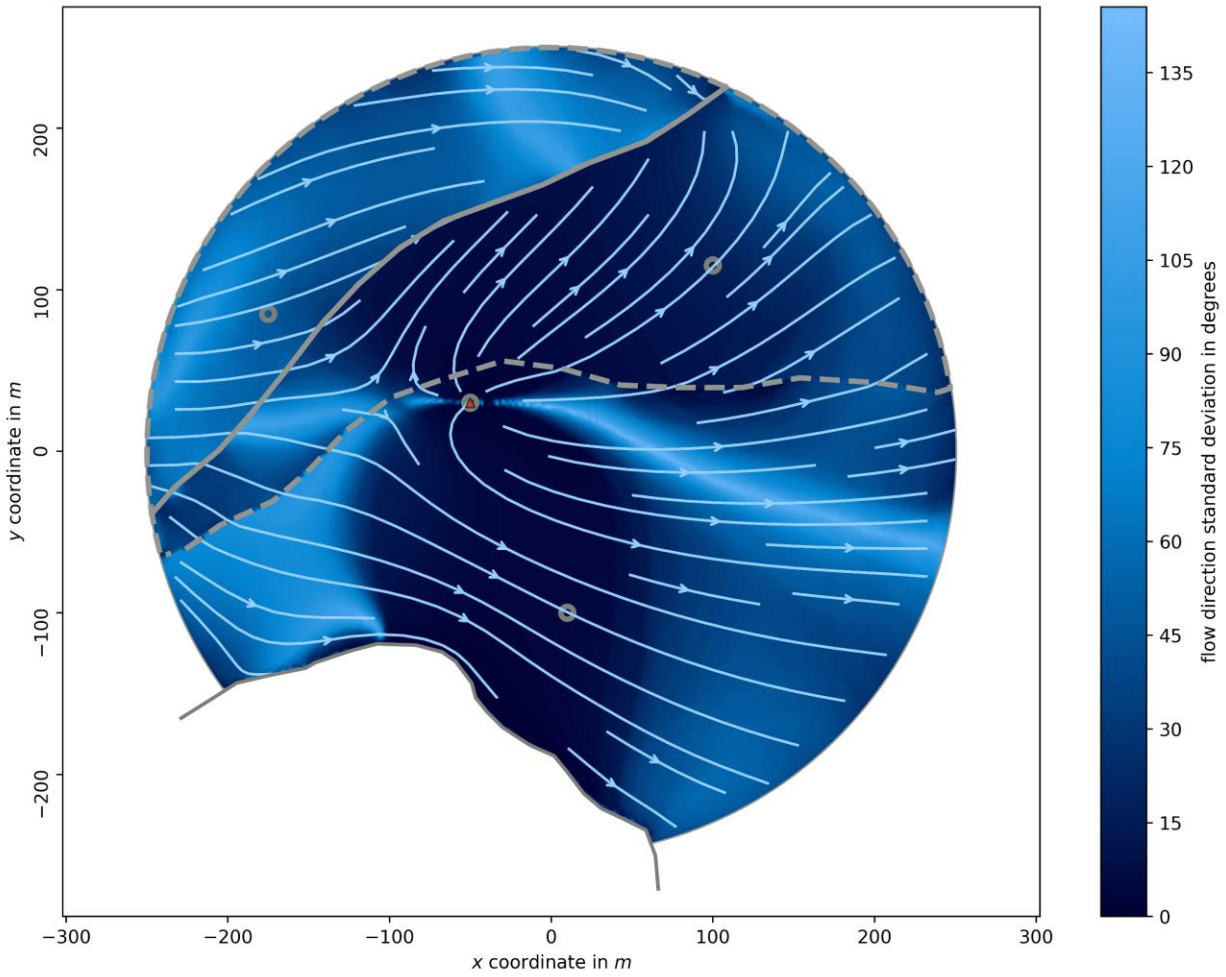

Figure 7. Posterior circular mean (streamlines) and circular standard deviation (filled contours) of the hydraulic potential gradient direction. For the purpose of assigning no-flow boundaries, each direction and its opposite are identical (i.e., $0.5\pi \triangleq -0.5\pi$). Consequently, we estimated the circular standard deviation of the flow direction angles, with all negative angles offset by 180 degrees.

constant head are best assigned perpendicular to the flow lines (light blue). No-flow boundaries are best assigned parallel to the flow lines.

4 Discussion and Conclusions

In this study, we explored the use of the AEM for the inference of uncertain groundwater tables, particularly under the lens of uncertain regional flow. Towards this end, we expanded the standard toolbox of analytical elements with an element based on conformal mapping. This new element flexibly induces curving, converging, or diverging regional flow in a circular model

domain of arbitrary size. We subsequently benchmarked our toolbox against a numerical model (MODFLOW 6) for varying grid resolutions.

To examine the performance of the toolbox in practice, we explored its application in a simple synthetic scenario. Coupling the model to an MCMC routine, we sampled from its parameter posterior and simulated the corresponding uncertain groundwater flow fields. Finally, we illustrated a few possible analyses of the inferred water tables: water table uncertainty, simple pathline tracing, or decision support for the placement of numerical model boundaries.

We find that AEM can be a computationally efficient tool for the exploration of uncertain flow fields in data-scarce environments. Its comparatively simple structure and consequently straightforward uncertainty estimation can make it attractive for the estimation of probabilistic flow maps, particularly in studies without a primary focus on subsurface characterization. Alternatively, it can prove valuable as a support tool in preparation for more complex numerical models, particularly the assignment of boundaries.

In summary, we believe that AEM constitutes a highly attractive compromise between simplistic, often one-dimensional analytical groundwater flow solutions (e.g., method of fragments: Harr 2006), and the sometimes debilitating complexity of full numerical models. Its usually low parameter count and high computational efficiency renders AEM naturally well-suited for most Bayesian uncertainty estimation methods. In light of the push towards more comprehensive uncertainty analyses over the past decades, we remain confident that this property in particular warrants greater attention to AEM in the future. We have provided the AEM and MCMC toolboxes used in this study in the supporting information and on GitHub under https://maxramgraber.github.io/Simple-AEM-Toolbox/.

5 Acknowledgements

We express our gratitude to Prof. Otto Strack, University of Minnesota, and Prof. James Craig, University of Waterloo, who answered our questions about AEM at different points during the preparation of the manuscript. We are also indebted to Prof. Mark Bakker, Delft University of Technology, for clarifying our understanding of the farfield reference point during the preparation of the manuscript. The research leading to these results has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 675120. The data and algorithms to reproduce the results reported in this study have been uploaded under a DOI which does not yet exist. Instead, we have temporarily uploaded all files as a zipped folder under https://drive.switch.ch/index.php/s/nnkqbNHqnroWwvd.

References

Andrieu, Christophe, and Johannes Thoms. 2008. "A Tutorial on Adaptive MCMC." *Statistics and Computing*. doi: 10.1007/s11222-008-9110-y.

Bakker, M., V. Post, C. D. Langevin, J. D. Hughes, J. T. White, J. J. Starn, and M. N. Fienen. 2016. "Scripting MODFLOW Model Development Using Python and FloPy." *Groundwater* 54(5):733–39. doi: 10.1111/gwat.12413.

Bakker, Mark. 2006. "An Analytic Element Approach for Modeling Polygonal Inhomogeneities in Multi-Aquifer Systems." *Advances in Water Resources*. doi: 10.1016/j.advwatres.2005.11.005.

Bakker, Mark, and Victor A. Kelson. 2009. "Writing Analytic Element Programs in Python." *Ground Water*. doi: 10.1111/j.1745-6584.2009.00583.x.

Bakker, Mark, and Otto D. L. Strack. 2003. "Analytic Elements for Multiaquifer Flow." *Journal of Hydrology*. doi: 10.1016/S0022-1694(02)00319-0.

Craig, J., I. Janković, M. Bakker, and S. Matott. 2009. "Visual AEM."

Craig, J. R. 2009. "Analytic Elements for Flow in Harmonically Heterogeneous Aquifers." *Water Resources Research* 45(6). doi: 10.1029/2009WR007800.

Driscoll, Tobin A., and Lloyd N. Trefethen. 2009. *Schwarz-Christoffel Mapping*. Cambridge University Press.

Erdal, D., and O. A. Cirpka. 2016. "Joint Inference of Groundwater–Recharge and Hydraulic–Conductivity Fields from Head Data Using the Ensemble Kalman Filter."

Hydrology and Earth System Sciences. doi: 10.5194/hess-20-555-2016.

Fitts, Charles R., Joshua Godwin, Kathleen Feiner, Charles Mclane, and Seth Mullendore. 2015. "Analytic Element Modeling of Steady Interface Flow in Multilayer Aquifers Using AnAqSim." *Groundwater*. doi: 10.1111/gwat.12225.

Fong, Chamberlain. 2019. "Analytical Methods for Squaring the Disc." *ArXiv Preprint ArXiv:1509.06344*.

Furman, Alex, and Shlomo P. Neuman. 2003. "Laplace-Transform Analytic Element Solution of Transient Flow in Porous Media." *Advances in Water Resources*. doi: 10.1016/j.advwatres.2003.09.003.

Guillaume, Joseph H. A., Randall J. Hunt, Alessandro Comunian, Rachel S. Blakers, and Baihua Fu. 2016. "Methods for Exploring Uncertainty in Groundwater Management Predictions." in *Integrated Groundwater Management: Concepts, Approaches and Challenges*.

Haitjema, H. M. 1985. "Modeling Three-Dimensional Flow in Confined Aquifers by Superposition of Both Two- and Three-Dimensional Analytic Functions." *Water Resources Research*. doi: 10.1029/WR021i010p01557.

Haitjema, H. M. 1995. "Analytic Element Modeling." in *Analytic Element Modeling of Groundwater Flow*.

Harr, Milton. 2006. "Groundwater and Seepage." in *The Handbook of Groundwater Engineering, Second Edition*.

Höge, M., A. Guthke, and W. Nowak. 2019. "The Hydrologist's Guide to Bayesian Model

Selection, Averaging and Combination." *Journal of Hydrology*.

Janković, I., and R. Barnes. 1999. "High-Order Line Elements in Modeling Two-Dimensional Groundwater Flow." *Journal of Hydrology*. doi: 10.1016/S0022-1694(99)00140-7.

Kruschke, John. 2015. *Doing Bayesian Data Analysis: A Tutorial Introduction with R JAGS, and Stan*.

Langevin, Christian D., Joseph D. Hughes, E. R. Banta, Richard G. Niswonger, Sorab Panday, and A. M. Provost. 2017. "Documentation for the MODFLOW 6 Groundwater Flow Model." *U.S. Geological Survey*. doi: 10.3133/tm6A55.

Linde, Niklas, David Ginsbourger, James Irving, Fabio Nobile, and Arnaud Doucet. 2017. "On Uncertainty Quantification in Hydrogeology and Hydrogeophysics." *Advances in Water Resources*. doi: 10.1016/j.advwatres.2017.10.014.

Olver, Peter J. 2018. "Complex Analysis and Conformal Mapping." 84.

Peeters, L. J. M., and C. Turnadge. 2019. "When to Account for Boundary Conditions in Estimating Hydraulic Properties from Head Observations?" *Groundwater*.

Renard, Philippe. 2007. "Stochastic Hydrogeology: What Professionals Really Need?" *Ground Water*. doi: 10.1111/j.1745-6584.2007.00340.x.

Strack, O. D. L. 2018. "Limitless Analytic Elements." *Water Resources Research*. doi: 10.1002/2017WR022117.

Strack, Otto D. L. 1989. *Groundwater Mechanics*. New Jersey: Prentice-Hall, Inc.

Strack, Otto D. L. 2017. *Analytical Groundwater Mechanics*. 1st ed. Cambridge University Press.

Vrugt, Jasper A., Cajo J. F. ter Braak, Martyn P. Clark, James M. Hyman, and Bruce A. Robinson. 2008. “Treatment of Input Uncertainty in Hydrologic Modeling: Doing Hydrology Backward with Markov Chain Monte Carlo Simulation.” *Water Resources Research*. doi: 10.1029/2007wr006720.