



## RESEARCH ARTICLE

WILEY

# Numerical modelling of stream–aquifer interaction: Quantifying the impact of transient streambed permeability and aquifer heterogeneity

Dan Zhou<sup>1,2</sup> | Ye Zhang<sup>3</sup> | Guillaume Gianni<sup>4</sup> | Peter Lichtner<sup>5</sup> | Irina Engelhardt<sup>2</sup>

<sup>1</sup>Institute of Environment, Chengdu University of Technology, Chengdu, China

<sup>2</sup>Department of Hydrogeology, Technische Universität Berlin, Berlin, Germany

<sup>3</sup>Department of Geology and Geophysics, University of Wyoming, Laramie, Wyoming

<sup>4</sup>Center for Hydrogeology and Geothermics, University of Neuchâtel, Switzerland

<sup>5</sup>OFM Research-Southwest, Santa Fe, New Mexico

**Correspondence**

Dan Zhou, Institute of Environment, Chengdu University of Technology, Chengdu, China.  
Email: danzhou0617@gmail.com

**Funding information**

China Scholarship Council

**Abstract**

Stream–aquifer interaction plays a vital role in the water cycle, and a proper study of this interaction is needed for understanding groundwater recharge, contaminants migration, and for managing surface water and groundwater resources. A model-based investigation of a field experiment in a riparian zone of the Schwarzbach river, a tributary of the Rhine River in Germany, was conducted to understand stream–aquifer interaction under alternative gaining and losing streamflow conditions. An equivalent streambed permeability, estimated by inverting aquifer responses to flood waves, shows that streambed permeability increased during infiltration of stream water to aquifer and decreased during exfiltration. Aquifer permeability realizations generated by multiple-point geostatistics exhibit a high degree of heterogeneity and anisotropy. A coupled surface water groundwater flow model was developed incorporating the time-varying streambed permeability and heterogeneous aquifer permeability realizations. The model was able to reproduce varying pressure heads at two observation wells near the stream over a period of 55 days. A Monte Carlo analysis was also carried out to simulate groundwater flow, its age distribution, and the release of a hypothetical wastewater plume into the aquifer from the stream. Results of this uncertainty analysis suggest (a) stream–aquifer exchange flux during the infiltration periods was constrained by aquifer permeability; (b) during exfiltration, this flux was constrained by the reduced streambed permeability; (c) the effect of temporally variable streambed permeability and aquifer heterogeneity were found important to improve the accurate capture of the uncertainty; and (d) probabilistic infiltration paths in the aquifer reveal that such pathways and the associated prediction of the extent of the contaminant plume are highly dependent on aquifer heterogeneity.

**KEYWORDS**

aquifer heterogeneity, groundwater age modelling, multiple-point geostatistics, stream–aquifer interaction, transient streambed permeability

## 1 | INTRODUCTION

Surface water–groundwater interaction has received much attention in recent decades due to its significant impact on the migration of

contaminants from streams to aquifers and vice versa, thus impacting drinking water resources. Examples of studies that focus on surface water–groundwater interaction include bioclogging effects on river infiltration (Newcomer et al., 2016), temporal variation of streambed

permeability in flood seasons (Wu et al., 2015), response of surface water–groundwater exchange flux to varying streamflow (Dudley-Southern & Binley, 2015; Fox, Boano, & Arnon, 2014), and migration of solutes including waste water related contaminants (Engelhardt et al., 2013; Hammond & Lichtner, 2010; Lasagna, De, & Franchino, 2016; Xie, Cook, & Simmons, 2016). Besides topography and climate, a key factor that controls surface water–groundwater interaction is the permeability of both streambed sediments and the aquifer beneath the stream, which together control the flow pattern and commonly displays a high degree of spatial heterogeneity, spanning several orders of magnitude (Fleckenstein, Niswonger, & Fogg, 2006; Sophocleous, 2002; Winter, Harvey, Franke, & Alley, 1998). The permeability of streambed sediments can be affected by continuous sediment deposition, remobilization, clogging, and erosion processes, which can lead to variations of permeability in space and time (Datry, Lamouroux, Thivin, Descoux, & Baudoin, 2015; Geist & Auerswald, 2007; Pholkern et al., 2015). With the exception of the variability introduced from chemical reactions, crustal movements, and human activities, aquifer permeability is commonly considered to be constant in time and can also display spatial heterogeneity due to spatial variability of the hydrofacies of different geological media (Comunian et al., 2016; Fleckenstein et al., 2006; Michael et al., 2010). The variation of permeability in both space and time may have a profound influence on surface water–groundwater interaction.

Many field studies have demonstrated the existence of spatial and temporal variations of streambed permeability (Min, Yu, Liu, Zhu, & Wang, 2013; Wang et al., 2016; Wu et al., 2015). Many approaches are available to estimate streambed permeability, such as by direct measurements (e.g., pumping, slug, and permeameter tests) and indirect estimation (e.g., grain size analysis and seepage meter; Cheong et al., 2008; Kalbus, Reinstorf, & Schirmer, 2006; Pozdniakov, Wang, & Lekhov, 2016). Numerous field studies have demonstrated the importance of variable streambed permeability on surface water–groundwater interaction (Newcomer et al., 2016; Pozdniakov et al., 2016; Simpson & Meixner, 2012). In most numerical investigations of surface water–groundwater interaction, streambed permeability was assumed to be temporally constant due to difficulties in detecting and measuring transient variations in permeability (e.g., Engelhardt, Prommer, Schulz, et al., 2013; Sun et al., 2016; Tian et al., 2015). Gianni, Richon, Perrochet, Vogel, and Brunner (2016) presented an analytical model to identify transient streambed permeability by inverting flood wave responses recorded in the time series of the stream stage and near-stream hydraulic head of the aquifer. The aquifer hydraulic head was simulated using a flood wave function and the streambed permeability, as a parameter of this function, is calculated via an inverse procedure. By testing using synthetic and field data, this method has been shown to be robust and reasonably accurate (Gianni et al., 2016). It further reveals the controlling effect of temporal variation of streambed permeability on surface water–groundwater interaction. However, whether the permeability of the streambed is the single controlling factor, and under what condition the streambed controls surface water–groundwater interaction, is still unclear.

To model aquifer heterogeneities including facies, porosity, and permeability, various geostatistical tools have been developed, such as sequential indicator simulation, hierarchical sequential indicator

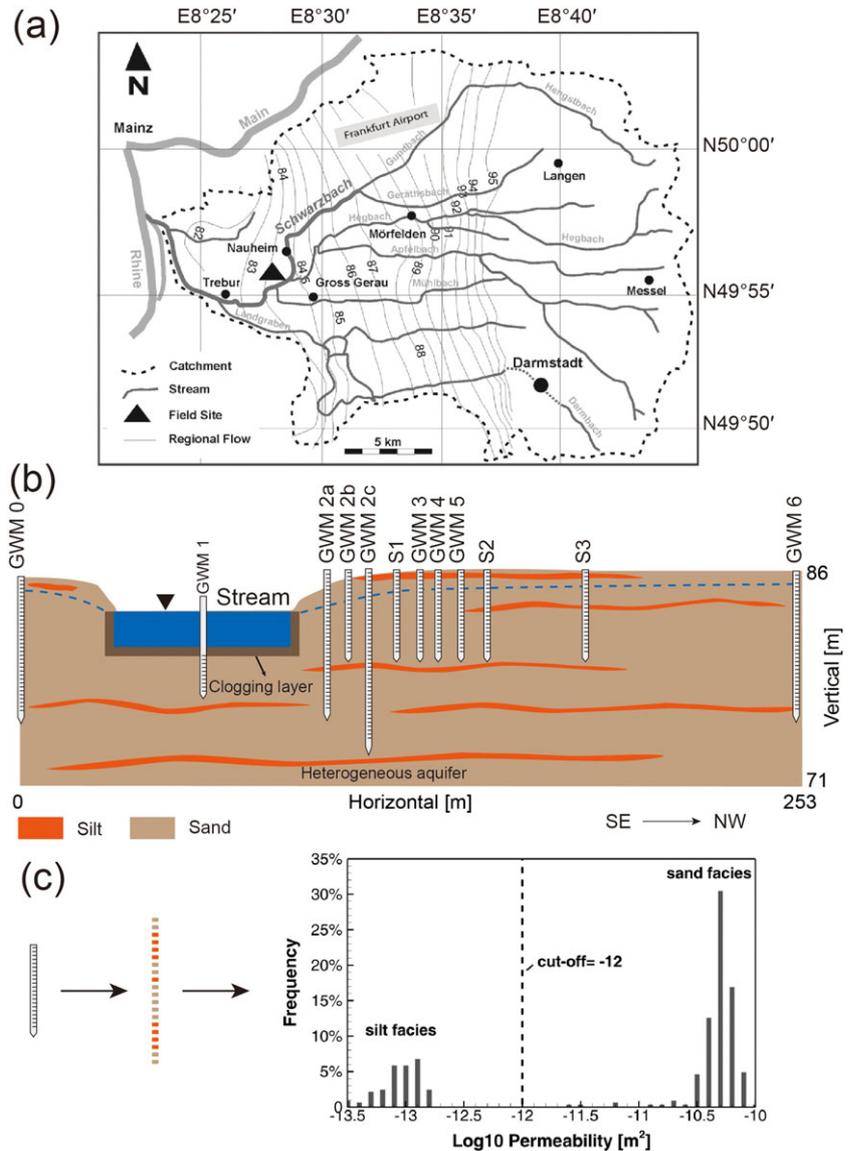
simulation, sequential Gaussian simulation, and multiple-point geostatistics (MPS; Guardiano & Srivastava, 1993; Pyrcz & Deutsch, 2014; Strebelle, 2000; Zappa, Bersezio, Felletti, & Giudici, 2006; Zhou, Gómez-Hernández, & Li, 2012). These geostatistical tools characterize and quantify spatial variability using probabilistic models to generate multiple heterogeneous realizations. Multiple realizations of permeability, for example, can be used to quantify the uncertainty of this parameter and how such uncertainty can propagate into the prediction of flow and transport in aquifers. In studying surface water–groundwater interaction, geostatistical methods have been adopted to investigate the impact of aquifer facies on the spatial variability of stream seepage (Fleckenstein et al., 2006) and the spatial and temporal dynamics of stream–aquifer exchange (Frei, Fleckenstein, Kollet, & Maxwell, 2009). Among these methods, MPS overcome the known limitations of variograms by utilizing the spatial correlation between variables beyond two-point statistics. A training image (TI) is used by MPS to characterize the pattern of subsurface heterogeneity. With a TI, data scarcity due to limited well information from the subsurface can often be addressed satisfactorily. MPS has been developed and applied in many subsurface modelling investigations (e.g., Huysmans & Dassargues, 2009; Malone, Jha, Minasny, & McBratney, 2016; Milliken, Levy, Strebelle, & Zhang, 2008; Zovi, Camporese, Franssen, & Huisman, 2017); however, few studies have adopted MPS to understand the effect of aquifer heterogeneity on surface water–groundwater interaction.

This paper analyses both streamflow and aquifer monitoring data collected at a riparian zone of the Schwarzbach river using computer models to (a) investigate the role of temporal variation of streambed permeability on surface water–groundwater interaction; (b) determine the impact of aquifer permeability heterogeneity, as modelled by MPS, on the prediction of groundwater flow and its uncertainty; (c) identify the most important time-variable factors controlling stream–groundwater interaction under losing and gaining streamflow conditions utilizing uncertainty analysis of the simulated hydraulic head; and (d) determine the most likely stream water infiltration paths as driven by stream–groundwater interaction using Monte Carlo-based groundwater age modelling, which takes model uncertainty into account.

## 2 | METHODS

### 2.1 | Study site

A riparian groundwater monitoring (GWM) campaign was conducted at a field site along the Schwarzbach river, a tributary of the Rhine River, which is located southwest of Frankfurt, Germany (Figure 1a). During the installation of a set of monitoring wells, 44 soil samples were taken from three observation wells (GWM1, GWM2c, and S2) with a vertical sampling density of 0.2 m. Based on the grain size analysis of these samples, intrinsic permeability for each of the 0.2 m intervals was calculated using the Hazen method (Hazen, 1911). Hydraulic head, stream stage, and precipitation were measured from August 15, 2010, to October 9, 2010 (Engelhardt et al., 2013). Using site data, a two-dimensional vertical transect was constructed (Figure 1a,b), which lies perpendicular to the main stream channel. Based on hydraulic



**FIGURE 1** (a) Our study site at a riparian zone of the Schwarzbach river in Germany (Engelhardt et al., 2011). (b) A monitoring well network along a constructed two-dimensional vertical transect; location of the transect is shown in (a) by a triangle symbol. (c) Histogram of aquifer permeability of the collected soil samples. GWM: groundwater monitoring; NW: northwest; SE: southeast

head data collected from all the monitoring wells at this site, the general direction of the groundwater flow is from southeast (SE) to northwest (NW). The transect thus captures the general groundwater flow direction and is used for this modelling study to evaluate spatial and temporal permeability heterogeneity. A detailed description of the hydrology and hydrogeology of the field site and the installation of monitoring wells is presented in Engelhardt, Prommer, Moore, et al. (2013).

## 2.2 | Characterizing aquifer heterogeneity

Currently, four stochastic modelling techniques are commonly used to capture geologic heterogeneity and fluvial-dominated deltaic reservoirs: (a) sequential indicator simulation; (b) object-based modelling; (c) multiple-point statistics (MPS); and (d) spectral component geologic modelling. As tested by Deveugle et al. (2014), unlike MPS, sequential indicator simulation often fails to capture complex facies architecture with curvilinear features. For the fluvial system they examined, object-based modelling can partially reproduce aspects of facies architecture whereas spectral component geologic modelling can introduce

unrealistic artefacts to the modelled facies architecture. MPS was considered the best technique among the four algorithms to reproduce facies architecture and associated sand-body connectivity.

To characterize aquifer heterogeneity, a TI was first generated to capture the facies distribution identified at the study site based on permeability histogram. Multiple realizations of facies distributions were then modelled using MPS, based on multiple point statistics derived from the TI. Permeability distribution was then simulated for each facies using sequential Gaussian simulation (SGSIM). Finally, based on the multiple permeability fields, a set of groundwater flow simulations was conducted. Details of these procedures are described in the following.

### 2.2.1 | Training image construction

The TI is a numerical representation of the perceived geological heterogeneity at a study site, which is required by most MPS algorithms. In this study, a facies categorization was first conducted according to the permeability histogram (Figure 1c), and the proportion of each facies type was calculated. A two-dimensional TI of sedimentary facies was generated by the Stanford Geostatistical Modeling Software

(SGeMS) (Remy, Boucher, & Wu, 2009), which reflects a representative vertical distribution of site sedimentary deposits by conditioning to the information observed in sediment core samples and a geological conceptual model (Figure 1b). The geological conceptual model was created based on a high-resolution seismic survey by Haimberger, Hoppe, and Schäfer (2005) in the northern part of the Upper Rhine Graben, an area that contains our study site. That survey was carried out over a length of 80 km along the Rhine River and a length of 25 km along its tributary, the Neckar River. The borehole logs from the central Upper Rhine Graben show that these Quaternary sediments consist of fluvial and lacustrine unconsolidated calcareous sands and gravels with interbedded clay and silt layers with thickness ranging from 2 to 8 m (Haimberger et al., 2005). The seismically imaged facies of the Pleistocene age in the upper layers shows thick and uniform stratification with low amplitude. Combining this geometrical information with observations from boreholes, the strata at the site are interpreted to lie within a braided to low-sinuosity meandering river system. The TI domain, which represents the geological conceptual model conditioned to borehole facies data, is 150 m × 15 m. It is uniformly discretized with cell dimensions of 0.05 m × 0.05 m, so that the TI consists of 900,000 grid cells. It is twice the size of the subsequent groundwater model (75 m × 15 m) in the horizontal direction in order to reproduce larger scale patterns (Caers & Zhang, 2005). Two types of facies ("sand" and "silt") and their proportions were interpreted from the permeability histogram. Given the knowledge of the geological conceptual model, a sinusoid geobody type was selected in SGeMS to generate the TI.

### 2.2.2 | Facies realizations

Among the available MPS algorithms for facies modelling (Mariethoz & Caers, 2014), the widely used pixel-based single normal equation simulation (SNESIM; Strebelle, 2000) was used to generate facies realizations by conditioning to site borehole facies data while sampling the multiple point statistics from the TI. Compared with other MPS facies modelling algorithms, SNESIM was chosen due to its capacity, both in data conditioning and geological shape reproduction, and its efficient CPU performance in SGeMS. To generate each realization, SNESIM simulates the facies patterns by scanning the TI and applies the sampled statistics to simulate a facies type at each cell of the geostatistical grid. The observed facies type and location of the soil samples served as the conditioning data during the SNESIM simulation. The MPS simulation grid is 75 m × 15 m, with cell dimensions of 0.05 m × 0.05 m, and consists of 450,000 grid cells. Three hundred facies realizations were created by SGeMS, which implements SNESIM (Remy et al., 2009).

### 2.2.3 | Permeability fields

To quantify permeability spatial correlation at the study site, experimental permeability variograms in both the horizontal and vertical directions were calculated and modelled for each facies. The intrafacies permeability field within the silt and sand facies, as modelled in each SNESIM realization, was simulated using SGSIM, a conventional geostatistical method based on permeability histogram and variogram (Deutsch & Journel, 1998). A set of 300 aquifer permeability fields were modelled and then used as the input for subsequent groundwater flow and age modelling.

## 2.3 | Estimating streambed permeability by analytical model inversion

Streambed permeability is one of the key factors controlling surface water–groundwater exchange (Genereux, Leahy, Mitasova, Kennedy, & Corbett, 2008) and is often strongly affected by flood events in natural streams that can erode and redistribute stream sediments (Simpson & Meixner, 2012). In this study, three flood events with different peak magnitudes were recorded between Days 9–18, 29–33, and 45–49 since August 2010. At a given location, the streambed permeability can be expected to be highly transient between alternative losing and gaining conditions when there are significant deposition and remobilization of sediment particles (Simpson & Meixner, 2012). To estimate the temporal variation of streambed permeability and reduce model uncertainty, a technique developed by Gianni et al. (2016) was adopted to rapidly calculate a transient streambed permeability and to identify its impact on the location and timing of water fluxes at the sediment water interface. This technique represents a heterogeneous streambed with homogeneous properties (i.e., an equivalent permeability), which was proven feasible if both the calibration and prediction are made for a connected flow regime (Irvine, Brunner, Franssen, & Simmons, 2012). Because the average stream width and thickness at our study site are on the scale of a few metres and groundwater flow and age distribution are modelled at the full field scale, the extent of the streambed in the coupled surface water–groundwater modelling is represented by only 20 grid cells. Clearly, a homogeneous representation of the streambed is considered suitable for the scale of investigation, and subgrid streambed heterogeneity is assumed to be accounted for by an estimated equivalent permeability of the streambed, as explained below.

The transient streambed permeability is estimated from recorded pressure heads within the aquifer and the stream stage through the inversion of a numerical convolution between discretized stream stage and aquifer response using the analytical model of Gianni et al. (2016). This model numerically computes the aquifer response to stream stage variations by assuming a linear variation of the function between two successive time steps:

$$h(x, n\Delta t) = \sum_{k=1}^{n-1} \frac{\Delta t}{6} \left( F'_k P_{n-k} + F'_{k+1} P_{n-k+1} + 2F'_k P_{n-k+1} + 2F'_{k+1} P_{n-k} \right), \quad (1)$$

where  $n$  is the number of discrete time steps of constant length  $\Delta t$ ,  $h(x, n\Delta t)$  is the hydraulic head at distance  $x$  perpendicular to the stream edge at time  $n\Delta t$ .  $k$  is the index of summation.  $F'_k$  is the stream stage variation at  $k$ -th time step.  $P_{n-k}$  is the hydraulic head at distance  $x$  perpendicular to the stream edge at time  $(n-k)\Delta t$ , given by (Hall & Moench, 1972)

$$P(x, t) = \operatorname{erfc}\left(\frac{x}{2\sqrt{\alpha t}}\right) - \exp\left(\frac{x}{r} + \frac{\alpha t}{r^2}\right) \operatorname{erfc}\left(\frac{x}{2\sqrt{\alpha t}} + \frac{\sqrt{\alpha t}}{r}\right), \quad (2)$$

where  $\operatorname{erfc}$  and  $\exp$  are the complementary error function and the exponential function, respectively.  $\alpha$  ( $\text{m}^2 \text{s}^{-1}$ ) is aquifer hydraulic diffusivity,  $\alpha = K_a \rho g d_a / \mu S$ , where  $K_a$  ( $\text{m}^2$ ) is aquifer permeability,  $\rho$  ( $\text{kg m}^{-3}$ ) is fluid density,  $g$  ( $\text{m s}^{-2}$ ) is gravity acceleration,  $\mu$  ( $\text{kg m}^{-1} \text{s}^{-1}$ ) is fluid dynamic viscosity,  $d_a$  (m) is thickness of the aquifer, and  $S$  is storativity

(-).  $S$  is estimated to be  $6.75 \times 10^{-4}$  from a previous field investigation (Engelhardt, Prommer, Moore, et al., 2013). The retardation coefficient  $r$  (m) represents the required thickness of aquifer to cause the same head loss as the streambed.  $r = dK_a/K$ ,  $d$  (m) is streambed thickness, and  $K$  ( $m^2$ ) is streambed permeability.

When both  $d$  and  $K_a$  are known, then  $r$  can be used as a surrogate for the estimation of the streambed permeability. By the convolution of stream stage variations with Equation (1), water table variation was computed at the observation well GWM2a over time. Parameter  $r$  is then estimated by minimizing the sum of squared error between the observed and modelled pressure heads at GWM2a. The approach is implemented using the mathematical software MATLAB. Due to the high heterogeneity of the aquifer, it is not suitable to calculate the average aquifer diffusivity for the whole transect. Thus, three aquifer diffusivities (0.1, 1, and  $10 \text{ m}^2 \text{ s}^{-1}$ ) were tested in the analytical inversion. The time series of observed groundwater table at GWM2a are segmented in partially overlapping time intervals, referred to as parameter optimization window (POW), which is composed of two parameters, the size (time interval width) and the shift (POW moving length). Based on the segmentation of the time series, a set of varying streambed permeability over time was calculated.

## 2.4 | Stream-groundwater flow model

A previous study by Engelhardt, Prommer, Moore, et al. (2013) used the same monitoring campaign data to set up a groundwater flow model with MODFLOW. In that model, the aquifer permeability was assigned layer by layer in a deterministic sense, and the streambed permeability was assumed temporally and spatially homogenous. Uncertainty in flow and transport simulations due to aquifer spatial variation and streambed temporal variation of permeability was not considered. This research builds upon the previous work and the site streambed and aquifer permeability variation were modelled using analytical inversion of aquifer response and geostatistical MPS simulations conditioned to lithofacies observations at well locations, respectively. Because our lithofacies observations exhibit high vertical resolution at the scale of 0.05 m, the 300 permeability models created by MPS simulations exhibit high spatial resolution at the scale of 0.05 by 0.05 m. This approach thus avoids upscaling the lithofacies model, which can lead to lost resolution in representing the site geology. As a result, a large suite of high resolution permeability MPS realizations were created as input for Monte Carlo flow simulation, thus, computational requirement is much larger and a high performance groundwater flow code is needed. This research conducts stream-groundwater flow simulation using PFLORAN (Hammond & Lichtner, 2010; Hammond, Lichtner, & Mills, 2014), an open source, massively parallel, subsurface flow and reactive transport simulator with parallelization through the PETSc library (Balay, Gropp, & McInnes, 1997). In solving subsurface flow and transport problems ranging from single to two-phase flow, as well as reactive species transport, it has exhibited excellent scaling performance (Zhang, Zhang, & Lichtner, 2017).

### 2.4.1 | Numerical model set-up

A two-dimensional variably saturated flow model was set up for a transect length of 75 m and an aquifer depth of 15 m (refer to the

transect in Figure 1b). The model is discretized into 450,000 grid cells with a cell resolution of  $0.05 \text{ m} \times 0.05 \text{ m}$ , which corresponds to the scale of the MPS simulation model. Given the observed highly transient stream stage and groundwater flow due to summer precipitation and flood events, this high resolution horizontal and vertical cell size was chosen across the model domain to guarantee numerical stability under variably saturated conditions. The stream is surrounded by a 1-m thick streambed which is assumed to be highly controlled by the surface water-groundwater exchange. The simulation period continues over 55 days, that is, from August 15, 2010, to October 9, 2010.

### 2.4.2 | Flow equation

The governing mass conservation equation for the variably saturated flow is given by (Bear, 1979)

$$\frac{\partial}{\partial t}(\rho\theta s) = \nabla \cdot \left[ \frac{\rho k k_r}{\mu} (\nabla P - \rho g \nabla z) \right] + q, \quad (3)$$

where  $\rho$  ( $\text{kg L}^{-1}$ ) is water density,  $\theta$  (-) is porosity,  $s$  (-) is saturation,  $k$  ( $m^2$ ) is permeability,  $k_r$  (-) is relative permeability,  $\mu$  (Pa s) is viscosity,  $P$  (Pa) is pressure,  $g$  ( $\text{m s}^{-2}$ ) is gravitational acceleration,  $z$  (m) is elevation above the sea level (head datum),  $q$  ( $\text{kg L}^{-1} \text{ s}^{-1}$ ) are sources and sinks. The van Genuchten-Mualem model was used to relate fluid pressure, saturation, and  $k_r$  (Mualem, 1976; van Genuchten, 1980):

$$s = \left[ 1 + |bh|^d \right]^{-c}, \quad (4)$$

$$k_r = \sqrt{s} \left[ 1 - \left( 1 - s^{1/c} \right)^c \right]^2, \quad (5)$$

where  $h$  (m) is matric potential.  $b = 14.5 \text{ (m}^{-1}\text{)}$ ,  $c = 0.63$ , and  $d = 2.68$ , which are empirical parameters,  $c = 1 - 1/d$ . As the aquifer site of this study has a sand-rich texture, these parameters were obtained from Carsel and Parrish (1988) as representative average values for such deposits. Due to the thin unsaturated zone compared to the model vertical dimension, the model uncertainty from these empirical parameters would be minor.

### 2.4.3 | Initial and boundary conditions

The initial and boundary conditions were set in a similar manner to the MODFLOW/MT3DMS model by Engelhardt, Prommer, Moore, et al. (2013). The initial hydraulic heads across the model domain were defined from the head gradient measured between the observation well GWM0 and GWM6. The SE boundary was defined by the prescribed hydraulic head using the observed hydraulic head at GWM0. The NW boundary was defined by the hydraulic head that was derived from the interpolation between S3 and GWM6. The groundwater recharge was assigned to the model top layer with a recharge rate of 15% of the precipitation that infiltrated into the groundwater as measured by Berthold and Hergesell (2012). The model bottom layer was defined as a no flow boundary, corresponding to low permeable silty clay layer. The observed time variable stream stage was assigned to the model cells that represent the stream bottom with a time variable head boundary. The porosity for the streambed was 0.12 with homogeneous distribution (Engelhardt, Prommer, Moore, et al., 2013). The

porosity for silt and sand facies were 0.35 and 0.39 throughout the aquifer domain, respectively, which were derived from the mean porosity of each facies. The porosity of streambed and aquifer were assumed to be homogeneous and constant. The time-varying streambed permeability was assumed to be spatially homogeneous and assigned to the cells of the streambed. The permeability fields of the aquifer were assigned to the aquifer cells according to the geostatistical realizations.

#### 2.4.4 | Uncertainty analysis

The multiple realization mode of PFLOTRAN was used to conduct the Monte Carlo simulations of variably saturated flow in streambed and in the connected aquifer using each of the 300 aquifer permeability realizations as input. Simultaneously, 300 Monte Carlo simulations ran on three processor groups with each group consisting of 100 simulations that ran one after another. Ninety-six processor cores were utilized. The computed pressure heads, a total of 300 sets, were obtained to define the influence of permeability uncertainty on the simulated heads. Based on the simulated and measured hydraulic heads at observation wells GWM2a and S3, a mean absolute error (MAE) of each realization was calculated:

$$\text{MAE} = \frac{1}{j} \sum_{i=1}^j |h_{mi} - h_{si}|, \quad (6)$$

$j$  (-) is the number of the measured hydraulic head;  $h_{mi}$  (m) and  $h_{si}$  (m) are the measured and simulated hydraulic head, respectively. It provides a criterion that a realization is acceptable to represent the site geology if its MAE is below a user-defined value (Anderson, Woessner, & Hunt, 2015). A cumulative mean absolute error (CMAE) was further calculated versus the number of realizations to estimate how many realizations are sufficient for convergence:

$$\text{CMAE} = \frac{1}{f} \sum_{i=1}^f \text{MAE}_i. \quad (7)$$

$f$  is the number of realizations, and  $\text{MAE}_i$  is the MAE of the  $i$ -th realization. The maximum, minimum, and arithmetic mean simulated heads were calculated based on the acceptable realizations. Standard deviations of the simulated heads were calculated for the acceptable realizations to represent the uncertainty.

#### 2.5 | Groundwater age simulation

Groundwater age modelling provides a useful technique to facilitate the estimation of groundwater recharge rate and velocity, the calibration of groundwater flow models, and the assessment of the renewability of groundwater reservoirs (Kazemi, Lehr, & Perrochet, 2006). Particle tracking and direct simulation are the two groundwater age modelling techniques most commonly used when a groundwater flow field is known (Suckow, 2014). Compared to particle tracking, which typically considers advection and ignores the mass exchange between flow paths, direct simulation of groundwater age is more favoured in that it also considers dispersion and diffusion (Suckow, 2014). Moreover, because of the uncertainty in aquifer parameters, groundwater flow fields simulated for any field-based studies will suffer uncertainty. However, most studies that applied groundwater age modelling to

investigate groundwater recharge and chemistry evolution do not take such uncertainty into account (Attard, Rossier, & Eisenlohr, 2016; Lemieux & Sudicky, 2010; Yu, Yao, Cao, & Zheng, 2015). To avoid the introduction of bias into modelled groundwater age, model uncertainty cannot be neglected in groundwater age modelling. The groundwater age is calculated by the advection–dispersion equation for nonreactive tracer transport that is implemented in PFLOTRAN following the method described in Goode (1996):

$$\frac{\partial A}{\partial t} = \nabla \cdot \mathbf{D} \cdot \nabla A - \nabla \cdot \mathbf{A} \frac{\mathbf{q}}{\theta} + 1, \quad (8)$$

$$A = \frac{\int_0^{\infty} tC \, dt}{\int_0^{\infty} C \, dt}, \quad (9)$$

where  $A$  (d) is the tracer mean age with concentration  $C$  (mol L<sup>-1</sup>),  $\theta$  (L L<sup>-1</sup>) is the porosity,  $\mathbf{q}$  is the specific-discharge vector, and  $\mathbf{D}$  is the dispersion tensor. The right side term “+1” denotes each molecule getting 1-day older every day. Building upon the flow model, the initial age of  $1 \times 10^{-8}$  days and concentration of  $1 \times 10^{-16}$  mol L<sup>-1</sup> of a dummy tracer was defined across the model domain. A zero gradient boundary condition was defined at the SE, NW, top, and bottom boundary, respectively.

Because the site permeability was explicitly modelled at high resolution, the longitudinal and transverse dispersivity were all set to zero (assuming subgrid dispersion is negligible), thus, the computed age reflects a purely advective travel time distribution along the groundwater streamlines. The resultant groundwater age only refers to the time from the start of the simulation. It reveals the infiltration paths from surface water into groundwater and can help to visualize zones that received surface water fluxes with younger ages. The groundwater age simulation was conducted for each acceptable flow model. A probabilistic infiltration path was then delineated based on the ensemble of simulated infiltration paths, thus taking into account the geological uncertainty. The area with high or close to 100% probability indicates the highest confidence in the predicted infiltration paths.

### 3 | RESULTS AND DISCUSSIONS

#### 3.1 | Facies category

Grain size analysis of the soil samples shows that the aquifer beneath and in close proximity to the stream is highly heterogeneous and is composed of interbedded silt, clay, and sand layers (including fine, median, and coarse sand layers). Permeability is calculated with the Hazen method (Hazen, 1911), using the  $d_{50}$  values from the grain size curve, that vary from  $2.61 \times 10^{-14}$  to  $9.74 \times 10^{-11}$  m<sup>2</sup> (Figure 1c). The distribution of permeability indicates the existence of two hydrofacies populations, one representing a silt-rich facies and the other a sand-rich facies (referred to herein as the silt and sand facies). A cut-off value separating the two populations was identified at  $1 \times 10^{-12}$  m<sup>2</sup>. The bimodal permeability distribution in the histogram shows the primary control of facies on permeability, thus, facies modelling is necessary to capture the permeability variation and to honour the observed sharp spatial transition. Variogram analysis of the permeability data

shows that the horizontal/vertical anisotropy ratio is 5.8 for silt facies and 1.35 for sand facies, respectively. The uncertainty of the Hazen method and insufficient horizontal data are the main reason for lack of agreement between the experimental and the modelled variograms in the horizontal direction, both in silt and sand facies. Moreover, to capture within-facies variability, a subsequent permeability simulation was conducted within each facies using SGSIM, which used the histogram and variogram models of permeability developed for each facies (Figures 1c and 2, Table 1). The approach adopted thus captures both large-scale facies transitions and small-scale within-facies permeability variation. Most of the existing literature (usually modelling  $\ln K$  rather than  $K$ ) is using an exponential function, which creates more “noisy” random  $K$  fields during simulation. We used a spherical model to fit the variogram curves, which can create a smoother random  $K$  fields given the same conditioning data (Koltermann & Gorelick, 1996). A smoother random  $K$  field will lead to smoother streamlines as compared to the noisy  $K$  fields. However, the distinction here is not significant because the facies modelled by MPS dominates the distribution of flowpaths and velocities, rather than the subfacies-scale  $K$  variability (Zhang & Gable, 2008).

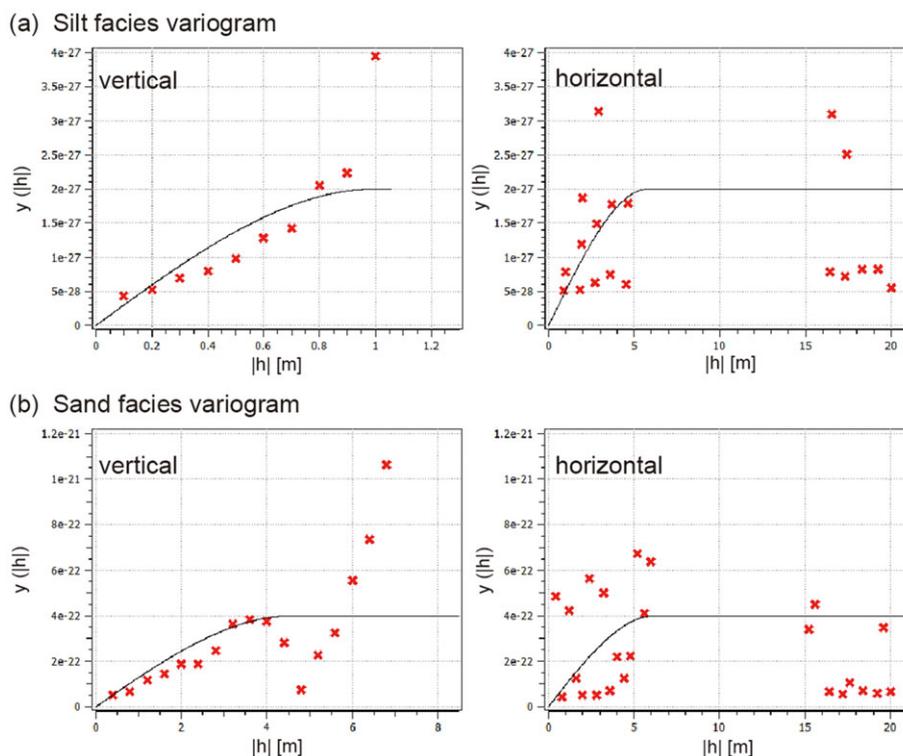
### 3.2 | Aquifer heterogeneity

A stationary TI (Figure 3a) was constructed based on geological information from well data (Figure 3b) and the geological transect (Figure 1 b). In the TI, the red region corresponds to the silt facies with low permeability materials such as silt and clay, whereas the background, or blue region, corresponds to the sand facies with high permeability materials such as sand and gravel. Three hundred facies realizations (Figure 3c) were generated with SNESIM by (a) sampling the local conditional probabilities from the TI using a MPS template, and (b)

constraining the realizations by observed facies identified in soil samples. The optimal reproduction was determined by visual inspection using a 60-node search template with a dimension of  $100\text{ m} \times 100\text{ m}$ , 6 multigrid, and a 0.5 servosystem factor to approach the target facies proportion (0.3:0.7). The template dimension and the number of multigrids are the most influential parameters for the pattern reproduction. The SNESIM realizations preserve the heterogeneous facies structure that is consistent with the geological conceptual model. They show a mean proportion and standard deviation of 0.347 and 0.011 for silt facies and 0.653 and 0.011 for sand facies, respectively. Next, the permeability field was simulated within silt and sand facies using SGSIM in each facies realization. The assembled 300 permeability fields then served as input for the subsequent groundwater flow model (Figure 3d).

### 3.3 | Streambed permeability

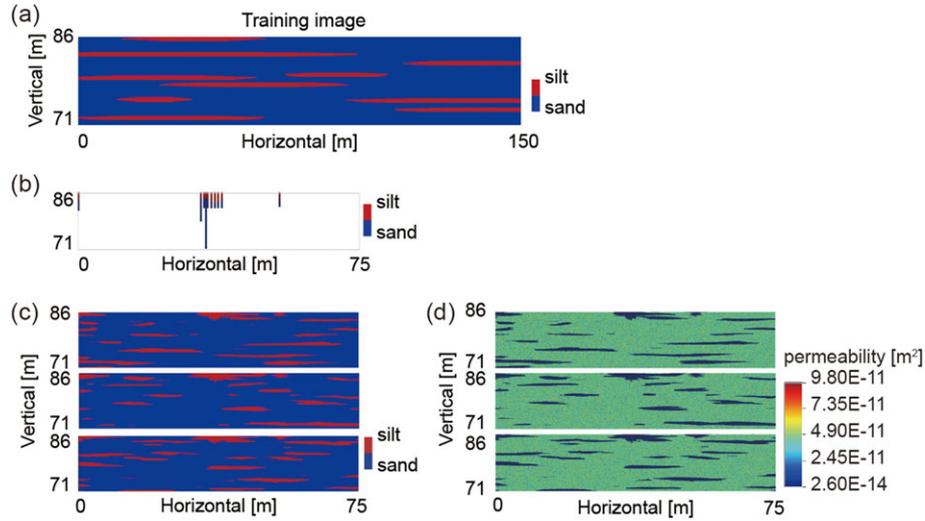
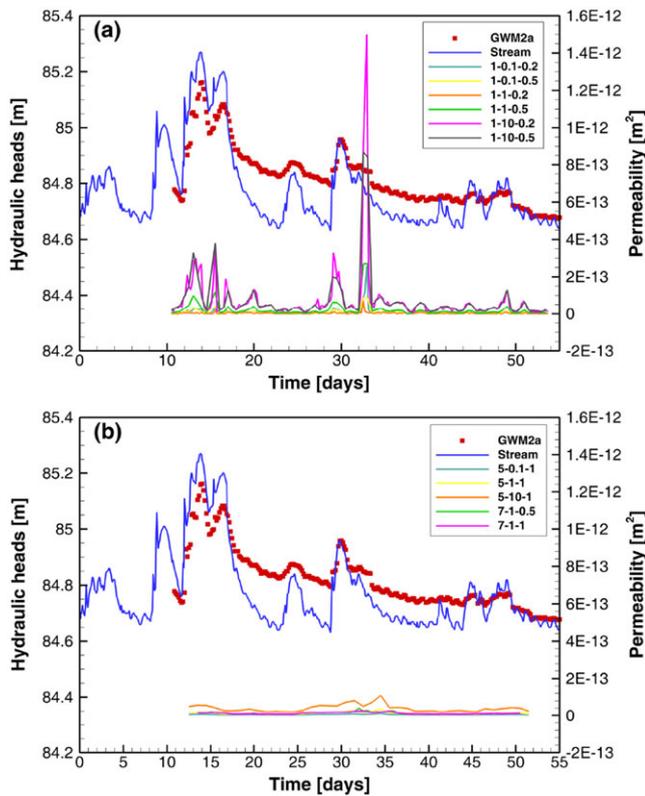
The streambed permeability were calculated for different POW size, shift, and aquifer diffusivities to evaluate the input parameter sensitivity (Figure 4). The streambed permeability were calculated from the 10th day until the end of the simulation period. The aquifer response to the stream stage variation is represented by the measured hydraulic head at GWM2a. Three sets of aquifer diffusivity ( $0.1$ ,  $1$ , and  $10\text{ m}^2\text{ s}^{-1}$ ) were tested to represent the whole transect from the range of minimum values of  $0.01\text{ m}^2\text{ s}^{-1}$  to maximum values of  $19.28\text{ m}^2\text{ s}^{-1}$  calculated from the measured soil sample permeability (Figure 1) and storativity. Under the same POW size (5 days) and shift (1 day), an increase of aquifer diffusivity from  $0.1$  to  $1$  to  $10\text{ m}^2\text{ s}^{-1}$  resulted in a threefold increase in the retardation coefficient. By comparing different POW sizes (1, 5, and 7 days), it can be shown that the model using a larger POW size was able to capture more head observations



**FIGURE 2** Horizontal and vertical experimental (dots) and modelled (curves) permeability variograms for silt and sand facies, respectively

**TABLE 1** Variogram parameters of permeability for the silt and sand facies

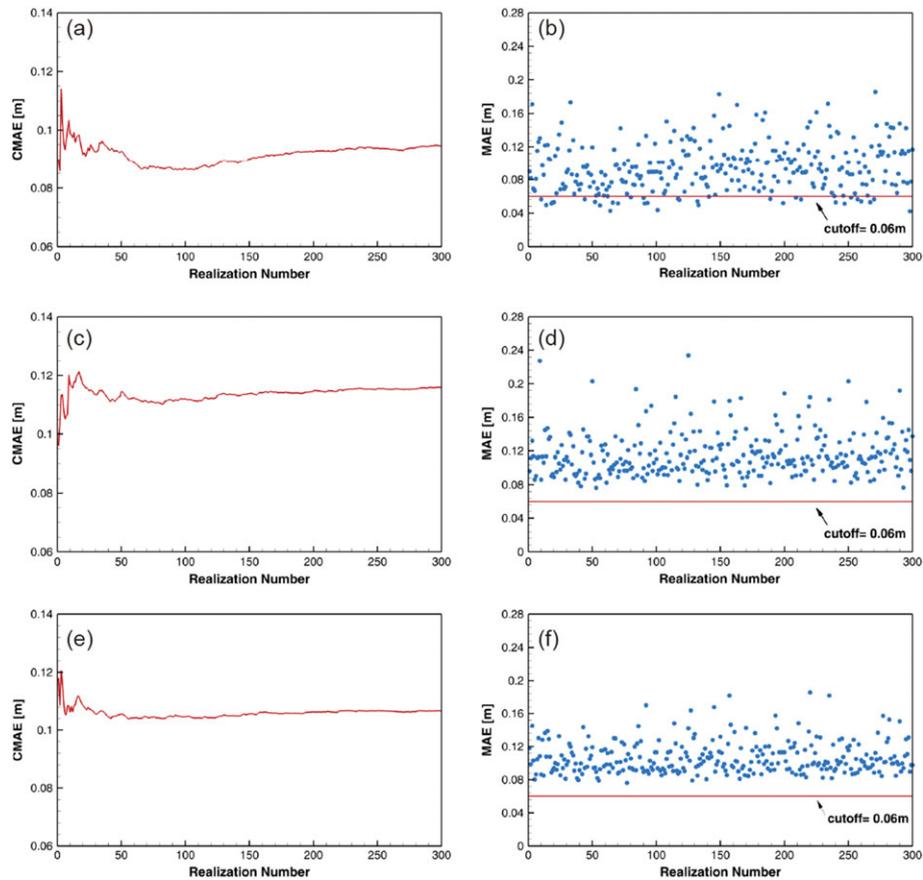
Facies	Variogram type	Nugget ( $m^2$ ) <sup>2</sup>	Ranges (m)			Angles (°)			Sill ( $m^2$ ) <sup>2</sup>
			Max	Med	Min	Azimuth	Dip	Rake	
Silt	Spherical	0	6.2	4.6	0	90	0	0	$4 \times 10^{-22}$
Sand	Spherical	0	5.8	1	0	90	0	0	$2 \times 10^{-27}$

**FIGURE 3** (a) The training image (TI) and (b) well identified facies that were used for constraining the MPS simulations of facies; (c) examples of single normal equation simulation facies realizations; (d) assembled permeability fields**FIGURE 4** The estimated streambed permeability, stream stage, and observed hydraulic heads at GWM2a over time. The 11 sets of streambed permeability indicated with different colours correspond to different parameter combinations of parameter optimization window size, aquifer diffusivity, and parameter optimization window shift, respectively. GWM: groundwater monitoring

and reproduce the general trend, but the transient head fluctuation at small time interval was not well captured. Models with a lower POW size with 0.5 day and aquifer diffusivities of 0.1, 1, and  $10 \text{ m}^2 \text{ s}^{-1}$ , and shifts of 0.2 and 0.5 day per time step were also tested, but no signals were observed due to the limited observations that this POW can capture. Although smaller POW size and shift can capture transient head fluctuation, the decrease of the POW size and shift can produce a number of outliers due to the decrease of the signal (observations) to noise ratio (Figure 4b). Thus, the estimation of the streambed permeability depends highly on the choice of the POW size and shift, and the accurate estimation of the effective aquifer diffusivity. Each set of streambed permeability were transferred into the groundwater flow model. We found that 33 realizations of the 300 simulated hydraulic heads using datasets shown as purple and grey lines in Figure 4a are able to reproduce the recorded hydraulic heads, whereas all of the realizations from the rest of the datasets in Figure 4a,b failed due to their comparable values with one another and less variation with time. Given the quick response of the pressure head at GWM2a to the stream stage, the small window size and shift is reasonably expected. The 300 realizations, using dataset (POW 1 day size and 0.2 day shift, and  $10 \text{ m}^2 \text{ s}^{-1}$  aquifer diffusivity) shown as purple in Figure 4a, were chosen for the subsequential uncertainty analysis.

### 3.4 | Model uncertainty

First, the CMAE in hydraulic heads for the stochastic forward runs using 300 aquifer permeability fields was calculated. Figure 5a shows that the CMAE stabilized after approximately 50 realizations at 0.09 m, indicating that 50 runs would have been sufficient for model



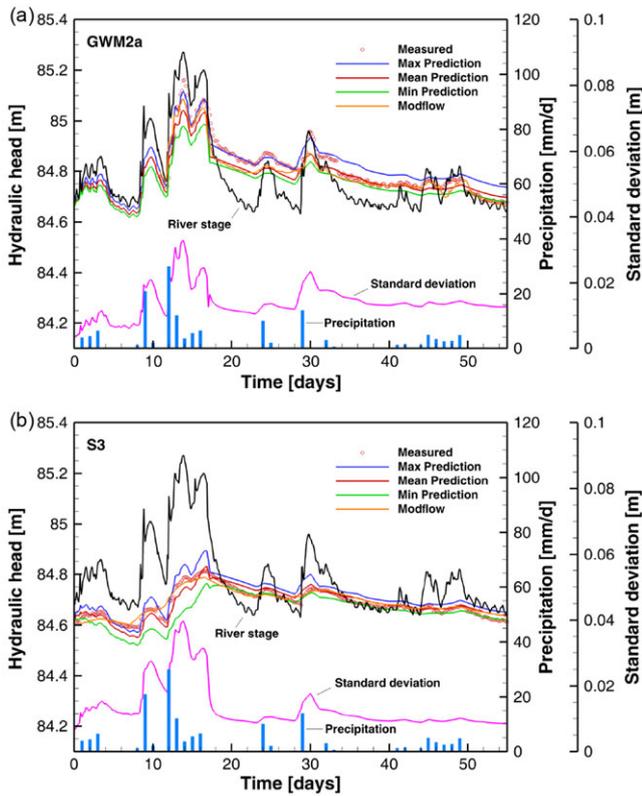
**FIGURE 5** Cumulative mean absolute error (CMAE) and mean absolute error (MAE) of the 300 Monte Carlo simulations with 300 aquifer permeability fields and with the optimal transient streambed permeability (a) and (b), without the streambed layer (c) and (d), and with constant-in-time streambed permeability from Engelhardt, Prommer, Moore, et al. (2013) (e) and (f). (a) The relatively stable CMAE after 50 realizations indicates convergence of the Monte Carlo process. (b) The acceptable simulations were conditioned by removing the runs with MAE greater than 0.06 m. (c) and (e) A higher set of CMAE was observed compared with (a) due to the large deviation of simulated hydraulic head from measured data when removing streambed data. (d) and (f) All the MAEs were above the cut-off line, indicating that the outliers do not yield an acceptable model

convergence (Anderson et al., 2015). If the MAE in hydraulic head is then calculated for each stochastic forward run, the question of whether or not the realization still stays within the acceptable calibration range can be determined (Figure 5b). An optimal cut-off value of 0.06 m was chosen by testing the MAE value from 0.04 to 0.1 m to judge the obtained results. Those realizations with a MAE of more than 0.06 m did not yield an acceptable model and were removed, leaving 33 conditioned realizations for the uncertainty analysis. Results from the 33 forward runs were summarized by plotting the maximum, mean, minimum, and standard deviations of the simulated hydraulic heads, versus the measured hydraulic heads (Figure 6). Figure 6 shows that most of the observed hydraulic heads at GWM2a and S3 lie in the range between the maximum and minimum simulated hydraulic heads, except that the hydraulic heads around 30th day at GWM2a were underestimated, which might be due to the underestimation of the streambed permeability during the short flood event.

The stochastic forward modelling was conducted twice again using the suite of aquifer permeability realizations, but removing the streambed layer and using a constant-in-time homogeneous streambed permeability ( $3.78 \times 10^{-11} \text{ m}^2$ ) from Engelhardt, Prommer, Moore, et al. (2013), respectively. Figure 5c,e shows that CMAE was stabilized at a higher value of 0.115 and 0.105 m than that of 0.09 m in Figure 5a,

respectively. All of the MAEs were higher than the cut-off value, which means none of these simulations yielded an acceptable model (Figure 5d,f). By comparing the simulations of Figure 5c,d to those of Figure 5a,b, which include the streambed layer with temporally transient permeability, the essential role of streambed in the stream-groundwater interaction processes is elucidated. While by comparing Figure 5e,f to Figure 5a,b, the implementation of constant-in-time streambed permeability significantly increases the model uncertainty over those based on a temporally variable streambed permeability. The uncertainty from streambed overwhelms the variation due to aquifer heterogeneity. Thus, temporally variable streambed permeability should be considered for the improved uncertainty analysis of stream-aquifer interaction.

The simulated hydraulic heads from the conditioned stochastic runs match well with the observed values at the observation well GWM2a and S3 (Figure 6a,b). For GWM2a, the maximum uncertainty of 0.033 m occurred at 14th day during the first flood period (9–18th days), followed by a smaller uncertainty of 0.024 m at the 30th day during the second flood period (Days 29–33). The greater distance between the stream and the observation well S3 decreased the response of hydraulic head at S3 to the variation of stream stage, even during the flood periods and heavy precipitation events. However, the



**FIGURE 6** Results of conditioned stochastic forward runs at the observation wells (a) GWM2a and (b) S3, respectively. The maximum, mean, and minimum values of the simulated hydraulic heads are compared with the measured data and the simulated data from the previous base study using MODFLOW (Engelhardt, Prommer, Moore, et al., 2013). Standard deviation shows the uncertainty of simulated hydraulic heads of this work varies with precipitation and flood periods

model uncertainty reached a maximum value of 0.04 m at the 14th day during the first flood period. Another peak value of 0.017 m is observed for the second flood period. In summary, high uncertainties were following the two heavy precipitation events (Days 9–18 and 29–33). Given that the transient streambed permeability was identical in each model run, the high uncertainties are attributed to the aquifer permeability fields that vary between each stochastic simulation. In contrast to infiltration periods, during the exfiltration periods, the uncertainties were distinctively lower. The higher uncertainties during the infiltration periods can be explained by the decreased control of streambed on the exchange fluxes during infiltration periods.

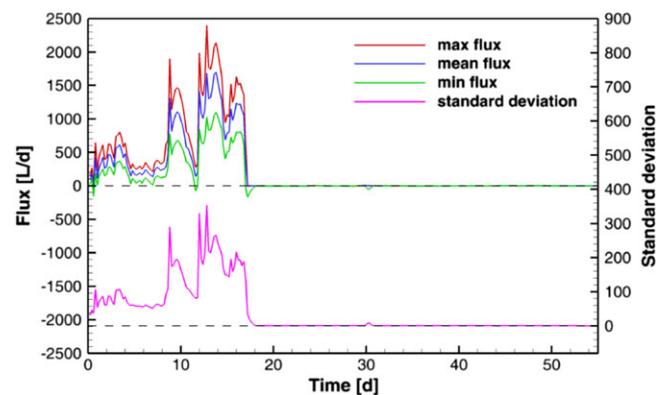
The streambed permeability has been proven by several authors (e.g., Hatch, Fisher, Ruehl, & Stemler, 2010; Kurtz, Hendricks Franssen, & Vereecken, 2012; Taylor, Lamontagne, & Crosbie, 2013; Zhang, Hubbard, & Finsterle, 2011) to be a highly dynamic parameter due to clogging, sedimentation, and remobilization of the sediments (Emmett & Leopold, 1963). Smaller particles have a greater tendency to be mobilized, which in turn will increase the entire streambed permeability during the floods due to the increased porosity caused by larger particles, whereas the clogging and deposition of transported sediments during dry seasons will decrease streambed permeability. Thus, during dry seasons, the streambed limits the surface water–groundwater exchange, while it promotes the exchange during floods.

This is coincident with the simulated water flux at the stream–aquifer interface (Figure 7). During the infiltration periods, the reduced streambed permeability due to sediment remobilization resulted in predicted water fluxes with a large standard deviation, whereas the increased streambed permeability by sediment clogging and deposition limited the exchange flux during the exfiltration periods. The high flux standard deviation during infiltration periods originated from the variation of aquifer permeability fields, which then decreased during exfiltration periods due to the increased control of the streambed on the exchange flux.

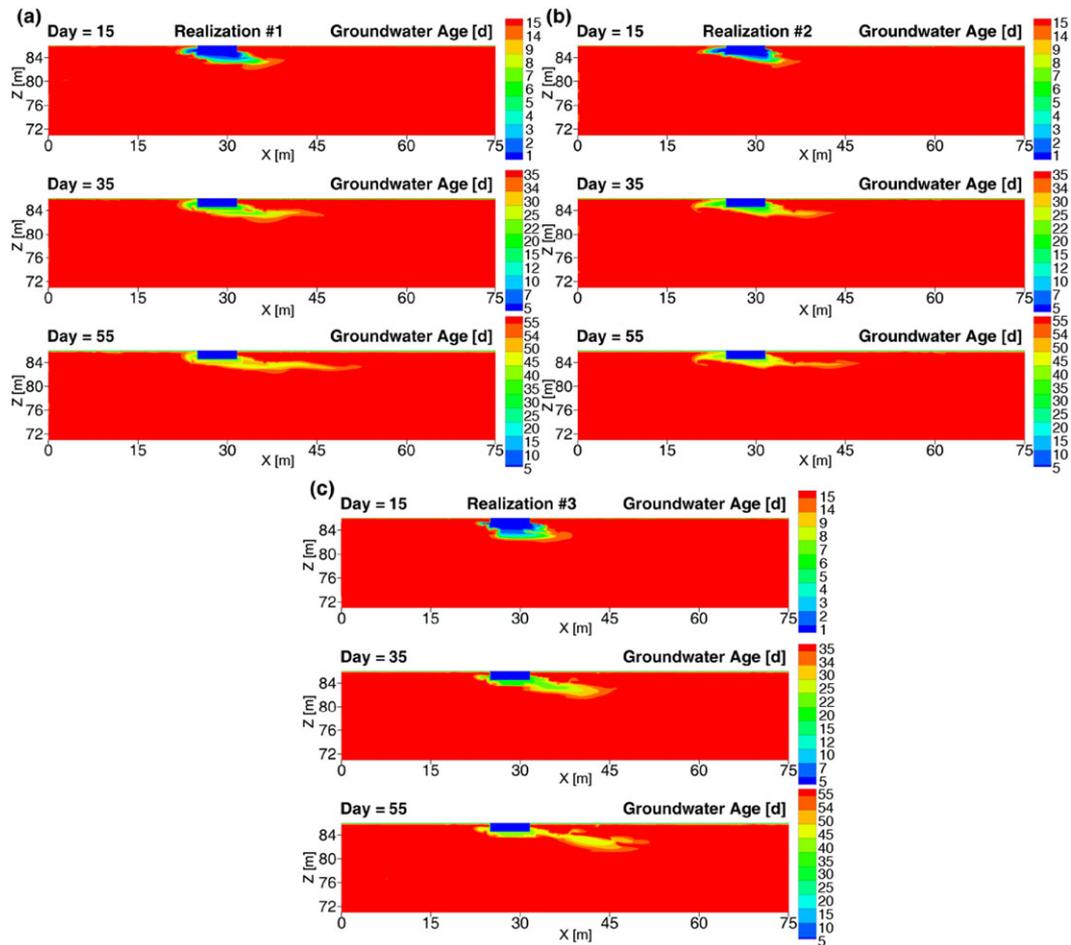
As shown in Figure 6, the simulated hydraulic heads using MODFLOW by the previous study (Engelhardt, Prommer, Moore, et al., 2013) also fitted the measured hydraulic heads well at GWM2a and S3. The modelling results were obtained through an inverse parameter estimation using PEST based on the Gauss–Marquardt–Levenberg method. Given that MODFLOW calibrated the aquifer head falls in the middle of the heads predicted by the stochastic models, the previous calibration using a homogeneous model obtained best-fit equivalent parameters which are not amenable to uncertainty analysis. In this work, spatial heterogeneity in the aquifer, as modelled by a stochastic technique (e.g., MPS), allows quantification of uncertainty in the modelled flow paths, thus capturing not only the mean prediction (as done by MODFLOW) but also uncertainty and spread from the mean. Given that the field site has only limited characterization data, the uncertainty analysis is a necessary and important component of this work, which significantly improves upon the previous study. Moreover, the calibration of temporally homogeneous streambed permeability is highly dependent on the choice of its initial value and the correlation with aquifer permeability, this study calculates temporally heterogeneous streambed permeability only based on the stream stage and measured hydraulic heads which eliminates the impact of parameter correlation.

### 3.5 | Surface water infiltration path

The groundwater age distribution was calculated for each of 33 realizations (Figure 8). Younger groundwater was observed during the infiltration period (Day 15) than the exfiltration periods (Days 35 and 55).



**FIGURE 7** The simulated water flux at the stream–aquifer interface. The maximum, mean, and minimum values of the water flux are summarized from 33 realizations. Positive and negative values indicate water infiltration from the stream into the aquifer and exfiltration from the aquifer to the stream, respectively

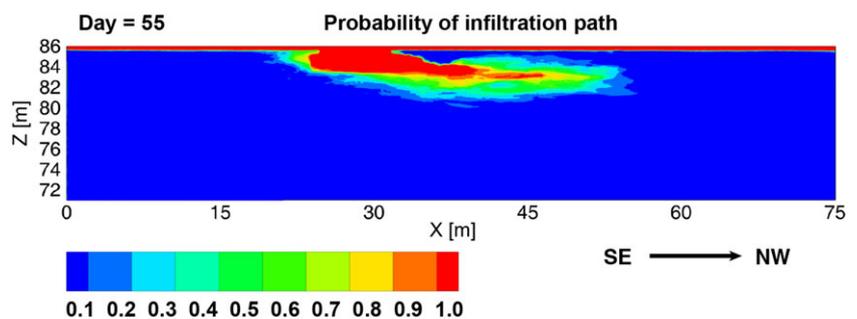


**FIGURE 8** Simulated groundwater age of (a) realization #1, (b) realization #2, and (c) realization #3 at Days 15, 35, and 55. Day 15 corresponds to the infiltration period, whereas Days 35 and 55 correspond to the exfiltration periods

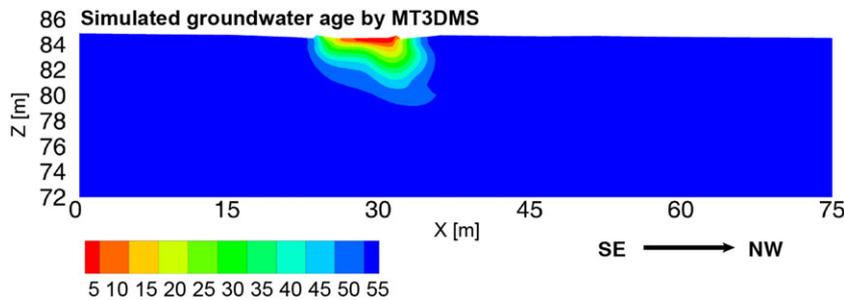
The different aquifer permeability field used in each realization resulted in a highly different age distribution and thus the implicit differences in the flow path. Due to the heterogeneous and anisotropic permeability distribution, longer preferential flow paths were observed in the model domain down hydraulic gradient. The distinctive flow paths shown by each realization indicate the dominant control of aquifer heterogeneity on stream water infiltration.

Figure 9 provides the probabilistic map of the surface water infiltration path after 55 days simulation time. The map summarizes how many times a cell contributed to the path divided by the total number of realizations. The area with a probability of 1.0 represents the case where a cell contributed to the path in all 33 runs. The surface water derived from rainfall penetrated to a depth of 0.5 m from the top layer in 100% probability. The stream water infiltrated into the aquifer with

a horizontal length ranging from 8.4 (100% probability) to 24.2 m (10% probability) in a NW direction from the right stream bank and with a depth ranging from 1.6 (100% probability) to 4.6 m (10% probability) from the stream bottom. The water infiltrated from the stream flowed primarily towards the NW direction as it integrated with the regional groundwater flow system. This method provides a probabilistic estimation of the surface water infiltration path based on the Monte Carlo concept. Although it only considers advective flow without dispersion, sorption, or geochemical reactions, tracking water particle paths in terms of probability can still be used to evaluate a hypothetical wastewater plume in a qualitative sense. For example, areas with extreme probability (0% or 100%) are expected to be less sensitive to particle numbers and discretization levels than those with uncertain probability (around 50%; Juckem, Fioren, & Hunt, 2014). The probability



**FIGURE 9** Probabilistic area of contribution with low (blue) to high (red) for surface water infiltration path based on the 33 groundwater age realizations



**FIGURE 10** Simulated groundwater age by MT3DMS in the previous study with young (red) to old (blue) after 55 days (Engelhardt, Prommer, Moore, et al., 2013)

map also provides a reference for the optimization of monitoring wells and solute sampling depths.

Figure 10 shows the surface water infiltration path simulated by MT3DMS in the previous study (Engelhardt, Prommer, Moore, et al., 2013), which assumes five homogenous aquifer layers and one temporally constant homogenous streambed permeability. By comparing to Figure 9, the significant difference of surface water infiltration shape and scope in Figure 9 is due to the aquifer heterogeneity, which leads to a rougher shape and a larger span in the horizontal direction. Meanwhile, the simulated groundwater age by MT3DMS was based on the calibration of the previous groundwater flow model using Gauss–Marquardt–Levenberg method, the potential parameter non-uniqueness may limit the prediction of all possible scenarios. Thus, aquifer heterogeneity should be considered for the improved uncertainty analysis of groundwater age distribution.

## 4 | CONCLUSIONS

The estimated streambed permeability from the inversion of flood wave responses shows a transient pattern of the streambed permeability, which decreased during infiltration periods and increased during exfiltration periods. The estimation of streambed permeability depends strongly on the selection of the POW size, shift, and aquifer diffusivity. It is worth mentioning that the aquifer anisotropy, especially in the vertical direction, might impose a significant impact on the estimation of the streambed permeability, as layers with lower or higher permeability could be activated alternatively by the varying height of the groundwater table. The aquifer permeability simulated using MPS with multiple realizations are highly heterogeneous and anisotropic. Each realization captures both facies and subfacies scale variability. By incorporating the transient streambed permeability and heterogeneous aquifer permeability, the newly developed groundwater flow model was able to reproduce the recorded fluctuation of hydraulic head at two observation wells. Our numerical results demonstrate that the exchange flux of the stream–groundwater interaction during infiltration periods was mainly reduced by the aquifer, whereas the streambed shows dominant limitation on the exchange flux during exfiltration periods with its reduced permeability. Moreover, by comparing the simulations with constant-in-time homogenous streambed permeability and without streambed layer, respectively, temporally variable streambed permeability can significantly decrease the model uncertainty. Therefore, it is essential to consider the influence of

transient streambed permeability when studying stream–groundwater interaction under variable stream stage and discharge situations.

Groundwater age simulation using multiple permeability realizations quantifies the uncertainty in the stream water infiltration path with a probabilistic map. Stream water infiltrated into the aquifer to a depth of 1.6 (100% probability) to 4.6 m (10% probability) below the stream. Laterally, stream water infiltrated towards the NW direction up to 8.4 (100% probability) to 24.2 m (10% probability) from the stream. Based on the computed probabilities, the region from the stream bottom to a depth of 1.6 m and from the stream right bank to a lateral distance of 8.4 m has experienced stream–aquifer interaction with the highest confidence. The large variation in path probability, however, shows the significant control of aquifer heterogeneity on groundwater flow. The probability map also provides a forecast envelope for a hypothetical wastewater plume that is released from the streambed.

## ACKNOWLEDGMENTS

The authors gratefully acknowledge the computing time granted by the supercomputer HLRN of North German Supercomputing Alliance and the HPC-Cluster of TU Freiberg. The first author also acknowledges generous financial support from the China Scholarship Council (CSC).

## ORCID

Dan Zhou  <http://orcid.org/0000-0001-5841-700X>

## REFERENCES

- Anderson, M. P., Woessner, W. W., & Hunt, R. J. (2015). *Applied groundwater modeling: Simulation of flow and advective transport*. San Diego: Academic press.
- Attard, G., Rossier, Y., & Eisenlohr, L. (2016). Urban groundwater age modeling under unconfined condition—Impact of underground structures on groundwater age: Evidence of a piston effect. *Journal of Hydrology*, 535, 652–661.
- Balay, S., Gropp, W. D., McInnes, L. C., & Smith, B. F. (1997). Efficient management of parallelism in object-oriented numerical software libraries. In: Arge E., Bruaset A.M., Langtangen H.P. (eds) *Modern software tools for scientific computing*. Birkhäuser, Boston, MA.
- Bear, J. (1979). *Hydraulics of groundwater*. New York: McGraw-Hill.
- Berthold, G., & Hergesell, M. (2012). Flächendifferenzierte Untersuchungen zu möglichen Auswirkungen einer Klimaänderung auf die Grundwasserneubildung in Hessen. *Klimafolgen in der Wasserwirtschaft (Grundwasser)*, INKLIM, 16-21.
- Caers, J., & Zhang, T. (2005). Multiple-point geostatistics: A quantitative vehicle for integrating geologic analogs into multiple reservoir models. *AAPG Memoir*, 383–394.

- Carsel, R. F., & Parrish, R. S. (1988). Developing joint probability distributions of soil water retention characteristics. *Water Resources Research*, 24(5), 755–769.
- Cheong, J. Y., Hamm, S. Y., Kim, H. S., Ko, E. J., Yang, K., & Lee, J. H. (2008). Estimating hydraulic conductivity using grain-size analyses, aquifer tests, and numerical modeling in a riverside alluvial system in South Korea. *Hydrogeology Journal*, 16(6), 1129–1143.
- Comunian, A., De Micheli, L., Lazzati, C., Felletti, F., Giacobbo, F., Giudici, M., & Bersezio, R. (2016). Hierarchical simulation of aquifer heterogeneity: Implications of different simulation settings on solute-transport modeling. *Hydrogeology Journal*, 24(2), 319–334.
- Datry, T., Lamouroux, N., Thivin, G., Descloux, S., & Baudoin, J. M. (2015). Estimation of sediment hydraulic conductivity in river reaches and its potential use to evaluate streambed clogging. *River Research and Applications*, 31(7), 880–891.
- Deutsch, C. V., & Journel, A. G. (1998). *Geostatistical software library and user's guide*. New York: Oxford University Press.
- Deveugle, P. E., Jackson, M. D., Hampson, G. J., Stewart, J., Clough, M. D., Ehighebo, T., ... Miller, J. K. (2014). A comparative study of reservoir modeling techniques and their impact on predicted performance of fluvial-dominated deltaic reservoirs: Comparison of reservoir modeling techniques. *AAPG Bulletin*, 98(4), 729–763.
- Dudley-Southern, M., & Binley, A. (2015). Temporal responses of groundwater-surface water exchange to successive storm events. *Water Resources Research*, 51(2), 1112–1126.
- Emmett, W. W., & Leopold, L. B. (1963). Downstream pattern of riverbed scour and fill (pp. 399–409). In Proceedings of the Federal Interagency Sedimentation Conference.
- Engelhardt, I., Piepenbrink, M., Trauth, N., Stadler, S., Kludt, C., Schulz, M., ... Ternes, T. A. (2011). Comparison of tracer methods to quantify hydrodynamic exchange within the hyporheic zone. *Journal of Hydrology*, 400(1), 255–266.
- Engelhardt, I., Prommer, H., Moore, C., Schulz, M., Schüth, C., & Ternes, T. A. (2013). Suitability of temperature, hydraulic heads, and acesulfame to quantify wastewater-related fluxes in the hyporheic and riparian zone. *Water Resources Research*, 49(1), 426–440.
- Engelhardt, I., Prommer, H., Schulz, M., Vanderborght, J., Schüth, C., & Ternes, T. A. (2013). Reactive transport of iomeprol during stream-groundwater interactions. *Environmental Science & Technology*, 48(1), 199–207.
- Fleckenstein, J. H., Niswonger, R. G., & Fogg, G. E. (2006). River-aquifer interactions, geologic heterogeneity, and low-flow management. *Ground Water*, 44(6), 837–852.
- Fox, A., Boano, F., & Arnon, S. (2014). Impact of losing and gaining streamflow conditions on hyporheic exchange fluxes induced by dune-shaped bed forms. *Water Resources Research*, 50(3), 1895–1907.
- Frei, S., Fleckenstein, J. H., Kollet, S. J., & Maxwell, R. M. (2009). Patterns and dynamics of river-aquifer exchange with variably-saturated flow using a fully-coupled model. *Journal of Hydrology*, 375(3), 383–393.
- Geist, J., & Auerswald, K. (2007). Physicochemical stream bed characteristics and recruitment of the freshwater pearl mussel (*Margaritifera margaritifera*). *Freshwater Biology*, 52(12), 2299–2316.
- Genereux, D. P., Leahy, S., Mitasova, H., Kennedy, C. D., & Corbett, D. R. (2008). Spatial and temporal variability of streambed hydraulic conductivity in West Bear Creek, North Carolina, USA. *Journal of Hydrology*, 358(3), 332–353.
- Gianni, G., Richon, J., Perrochet, P., Vogel, A., & Brunner, P. (2016). Rapid identification of transience in streambed conductance by inversion of a floodwave response. *Water Resources Research*, 52, 2647–2658.
- Goode, D. J. (1996). Direct simulation of groundwater age. *Water Resources Research*, 32(2), 289–296.
- Guardiano, F. B., & Srivastava, R. M. (1993). Multivariate geostatistics: Beyond bivariate moments. In: Soares A. (eds) *Geostatistics Tróia'92. Quantitative geology and geostatistics*, vol 5. Dordrecht: Springer.
- Haimberger, R., Hoppe, A., & Schäfer, A. (2005). High-resolution seismic survey on the Rhine River in the northern Upper Rhine Graben. *International Journal of Earth Sciences*, 94(4), 657–668.
- Hall, F. R., & Moench, A. F. (1972). Application of the convolution equation to stream-aquifer relationships. *Water Resources Research*, 8(2), 487–493.
- Hammond, G. E., & Lichtner, P. C. (2010). Field-scale model for the natural attenuation of uranium at the Hanford 300 Area using high-performance computing. *Water Resources Research*, 46(9).
- Hammond, G. E., Lichtner, P. C., & Mills, R. T. (2014). Evaluating the performance of parallel subsurface simulators: An illustrative example with PFLOTRAN. *Water Resources Research*, 50(1), 208–228.
- Hatch, C. E., Fisher, A. T., Ruehl, C. R., & Stemler, G. (2010). Spatial and temporal variations in streambed hydraulic conductivity quantified with time-series thermal methods. *Journal of Hydrology*, 389(3), 276–288.
- Hazen, A. (1911). Discussion: Dams on sand foundations. *Transactions of the American Society of Civil Engineers*, 73, 199.
- Huysmans, M., & Dassargues, A. (2009). Application of multiple-point geostatistics on modelling groundwater flow and transport in a cross-bedded aquifer (Belgium). *Hydrogeology Journal*, 17(8), 1901–1911.
- Irvine, D. J., Brunner, P., Franssen, H. J. H., & Simmons, C. T. (2012). Heterogeneous or homogeneous? Implications of simplifying heterogeneous streambeds in models of losing streams. *Journal of Hydrology*, 424, 16–23.
- Juckem, P. F., Fienen, M. N., & Hunt, R. J. (2014). *Simulation of groundwater flow and interaction of groundwater and surface water on the Lac du Flambeau reservation, Wisconsin (no. 2014-5020)*. Reston, VA, USA: US Geological Survey.
- Kalbus, E., Reinstorf, F., & Schirmer, M. (2006). Measuring methods for groundwater-surface water interactions: A review. *Hydrology and Earth System Sciences Discussions*, 10(6), 873–887.
- Kazemi, G. A., Lehr, J. H., & Perrochet, P. (2006). *Groundwater age*. New Jersey: John Wiley & Sons.
- Koltermann, C. E., & Gorelick, S. M. (1996). Heterogeneity in sedimentary deposits: A review of structure-imitating, process-imitating, and descriptive approaches. *Water Resources Research*, 32(9), 2617–2658.
- Kurtz, W., Hendricks Franssen, H. J., & Vereecken, H. (2012). Identification of time-variant river bed properties with the ensemble Kalman filter. *Water Resources Research*, 48(10).
- Lasagna, M., De Luca, D. A., & Franchino, E. (2016). Nitrate contamination of groundwater in the western Po Plain (Italy): The effects of groundwater and surface water interactions. *Environmental Earth Sciences*, 75(3), 1.
- Lemieux, J. M., & Sudicky, E. A. (2010). Simulation of groundwater age evolution during the Wisconsinian glaciation over the Canadian landscape. *Environmental Fluid Mechanics*, 10(1–2), 91–102.
- Malone, B. P., Jha, S. K., Minasny, B., & McBratney, A. B. (2016). Comparing regression-based digital soil mapping and multiple-point geostatistics for the spatial extrapolation of soil data. *Geoderma*, 262, 243–253.
- Mariethoz, G., & Caers, J. (2014). *Multiple-point geostatistics: Stochastic modeling with training images*. Hoboken, New Jersey, USA: John Wiley & Sons.
- Michael, H. A., Li, H., Boucher, A., Sun, T., Caers, J., & Gorelick, S. M. (2010). Combining geologic-process models and geostatistics for conditional simulation of 3-D subsurface heterogeneity. *Water Resources Research*, 46(5).
- Milliken, W. J., Levy, M., Strebelle, S. B., & Zhang, Y. (2008). The effect of geologic parameters and uncertainties on subsurface flow: Deepwater depositional systems. *Society of Petroleum Engineers, SPE Paper 114099*, 1–16.
- Min, L., Yu, J., Liu, C., Zhu, J., & Wang, P. (2013). The spatial variability of streambed vertical hydraulic conductivity in an intermittent river, northwestern China. *Environmental Earth Sciences*, 69(3), 873–883.

- Mualem, Y. (1976). A new model for predicting the hydraulic conductivity of unsaturated porous media. *Water Resources Research*, 12(3), 513–522.
- Newcomer, M. E., Hubbard, S. S., Fleckenstein, J. H., Maier, U., Schmidt, C., Thullner, M., ... Rubin, Y. (2016). Simulating bioclogging effects on dynamic riverbed permeability and infiltration. *Water Resources Research*, 52(4), 2883–2900.
- Pholkern, K., Srisuk, K., Grischek, T., Soares, M., Schäfer, S., Archwichai, L., ... Wirojanagud, W. (2015). Riverbed clogging experiments at potential river bank filtration sites along the Ping River, Chiang Mai, Thailand. *Environmental Earth Sciences*, 73(12), 7699–7709.
- Pozdniakov, S. P., Wang, P., & Lekhov, M. V. (2016). A semi-analytical generalized Hvorslev formula for estimating riverbed hydraulic conductivity with an open-ended standpipe permeameter. *Journal of Hydrology*, 540, 736–743.
- Pyrz, M. J., & Deutsch, C. V. (2014). *Geostatistical reservoir modeling*. New York: Oxford University Press.
- Remy, N., Boucher, A., & Wu, J. (2009). *Applied geostatistics with SGeMS: a user's guide*. New York: Cambridge University Press.
- Simpson, S. C., & Meixner, T. (2012). Modeling effects of floods on streambed hydraulic conductivity and groundwater-surface water interactions. *Water Resources Research*, 48(2).
- Sophocleous, M. (2002). Interactions between groundwater and surface water: The state of the science. *Hydrogeology Journal*, 10(1), 52–67.
- Strebelle, S. (2000). Sequential simulation drawing structure from training images. PhD thesis, Stanford University, Stanford, CA.
- Suckow, A. (2014). The age of groundwater—definitions, models and why we do not need this term. *Applied Geochemistry*, 50, 222–230.
- Sun, X., Bernard-Jannin, L., Garneau, C., Volk, M., Arnold, J. G., Srinivasan, R., ... Sánchez-Pérez, J. M. (2016). Improved simulation of river water and groundwater exchange in an alluvial plain using the SWAT model. *Hydrological Processes*, 30(2), 187–202.
- Taylor, A. R., Lamontagne, S., & Crosbie, R. S. (2013). Measurements of riverbed hydraulic conductivity in a semi-arid lowland river system (Murray–Darling Basin, Australia). *Soil Research*, 51(5), 363–371.
- Tian, Y., Zheng, Y., Wu, B., Wu, X., Liu, J., & Zheng, C. (2015). Modeling surface water-groundwater interaction in arid and semi-arid regions with intensive agriculture. *Environmental Modelling & Software*, 63, 170–184.
- van Genuchten, M. T. (1980). A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Science Society of America Journal*, 44(5), 892–898.
- Wang, L., Song, J., Zhang, B., Guo, H., Jiang, W., Wen, M., & Zhang, G. (2016). Spatial and temporal variations of streambed vertical hydraulic conductivity in the Weihe River, China. *Water*, 8(3), 70.
- Winter, T. C., Harvey, J. W., Franke, O. L., & Alley, W. M. (1998). Groundwater and surface water: A single resource. US Geological Survey Circular 1139.
- Wu, G., Shu, L., Lu, C., Chen, X., Zhang, X., Appiah-Adjei, E. K., & Zhu, J. (2015). Variations of streambed vertical hydraulic conductivity before and after a flood season. *Hydrogeology Journal*, 23(7), 1603–1615.
- Xie, Y., Cook, P. G., & Simmons, C. T. (2016). Solute transport processes in flow-event-driven stream-aquifer interaction. *Journal of Hydrology*, 538, 363–373.
- Yu, C., Yao, Y., Cao, G., & Zheng, C. (2015). A field demonstration of groundwater vulnerability assessment using transport modeling and groundwater age modeling, Beijing Plain, China. *Environmental Earth Sciences*, 73(9), 5245–5253.
- Zappa, G., Bersezio, R., Felletti, F., & Giudici, M. (2006). Modeling heterogeneity of gravel-sand, braided stream, alluvial aquifers at the facies scale. *Journal of Hydrology*, 325(1), 134–153.
- Zhang, M., Zhang, Y., & Lichtner, P. (2017). Evaluating model complexity in simulating supercritical CO<sub>2</sub> dissolution, leakage, footprint, and reservoir pressure for three-dimensional hierarchical aquifer. *International Journal of Greenhouse Gas Control*, 64, 284–299.
- Zhang, Y., & Gable, C. W. (2008). Two-scale modeling of solute transport in an experimental stratigraphy. *Journal of Hydrology*, 348(3–4), 395–411.
- Zhang, Y., Hubbard, S., & Finsterle, S. (2011). Factors governing sustainable groundwater pumping near a river. *Groundwater*, 49(3), 432–444.
- Zhou, H., Gómez-Hernández, J. J., & Li, L. (2012). A pattern-search-based inverse method. *Water Resources Research*, 48(3).
- Zovi, F., Camporese, M., Franssen, H. J. H., Huisman, J. A., & Salandin, P. (2017). Identification of high-permeability subsurface structures with multiple point geostatistics and normal score ensemble Kalman filter. *Journal of Hydrology*, 548, 208–224.

**How to cite this article:** Zhou D, Zhang Y, Gianni G, Lichtner P, Engelhardt I. Numerical modelling of stream-aquifer interaction: Quantifying the impact of transient streambed permeability and aquifer heterogeneity. *Hydrological Processes*. 2018;1–14. <https://doi.org/10.1002/hyp.13169>