Stochastic Assessment of Nonpoint Source Contamination: Joint Impact of Aquifer Heterogeneity and Well Characteristics on Management Metrics

Christopher Vincent Henri1,2 and Thomas Harter1,2

1Center for Watershed Sciences, University of California, Davis, CA, USA, 2Department of Land, Air and Water Resources, University of California, Davis, CA, USA

Abstract  Nonpoint source (NPS) groundwater contamination in sedimentary basin aquifers with overlying agricultural activities increasingly threatens groundwater supplies. The role of aquifer heterogeneity has not been well understood in the assessment of NPS and in linking pollution sources to impacts in water supply wells. A typical well taps into and mixes groundwater varying in age by decades or even centuries. This study investigates the joint impact of aquifer heterogeneity and pumping well characteristics (well depth, pumping rate, and screen length) on expected values of and uncertainty about key management metrics: (1) travel time of a NPS contaminant to a production well, (2) spatiotemporal characteristics of the source area, and (3) contaminant compliance at production wells. A stochastic approach is employed using Monte Carlo simulation of flow and nonreactive transport in 3-D highly heterogeneous alluvial aquifers systems. Well design is shown to dominate the distribution of travel time mixing and the overall location and spread of the source area. Larger extraction intensity and closer proximity to the land surface are shown to significantly suppress effects of large aquifer heterogeneity on uncertainty on all metrics. Deep wells and wells with lower pumping rates have more uncertain source areas. Long-term NPS contaminants with high source concentration (>10 times compliance level) will be exceeded in most wells within decades, while low-intensity source concentrations (2–4 times compliance level, typical for nitrate and salinity) have large uncertainty about time to exceedance suggesting wide variability among a set of wells with similar well design subject to the same NPS pollution.

1. Introduction

Nonpoint source (NPS) pollution is one of the most common causes of groundwater quality degradation, often from agricultural activities that cause widespread leaching of nitrate, salt, and agrochemicals (Burow et al., 2010; Harter et al., 2017; Howden et al., 2011). A large portion of agricultural activities, and therefore NPS contamination, occurs on plains overlying unconsolidated alluvial deposits offering soil composition and topography favorable to productive farming, particularly where irrigation is employed (Faunt, 2009). These unconsolidated sedimentary basins also represent important groundwater reservoirs (Zektser & Everett, 2004). As the world population continuously grows and disruptive climatic events such as droughts become more frequent, groundwater represents a resource asset to maintain or expand food production and urban water supplies while providing resilience to climate variability (Döll et al., 2014; Wada et al., 2012). Overexploitation of aquifers and synchronous emergence of NPS pollution has led to a rapid water quality deterioration in many groundwater basins that economies around the world depend on for safe and clean water and for global food security (Gurdak & Qi, 2012; Sutton et al., 2011).

NPS groundwater pollution differs from industrial point source contamination in the scale and often in the intensity of the pollution loading. Sources are widespread and continuous across aquifer systems. The most common NPS pollutants—salt and nitrate—lead to pollution of groundwater resources at magnitudes exceeding regulatory limits often by less than a factor of 2 or 3, sometimes as much as an order of magnitude (Harter et al., 2002; Ransom et al., 2017). In comparison, industrial point source contamination often occurs at levels exceeding regulatory targets by several orders of magnitude. But even low-intensity pollution becomes a serious threat to water users when it occurs over broad areas such that it constitutes a significant fraction of groundwater recharge, affecting a large number of water users.
Many countries have or are beginning to address groundwater NPS pollution (e.g., Aarts, 2003; Canada et al., 2012; Health Canada, 2013; Ministry of Health, 2008). Approaches are often two pronged: addressing the immediate needs of well water users through treatment and alternative supplies and reducing future pollution by improving current land use management practices, especially in agriculture.

Effective management of NPSs and affected groundwater users requires, at a minimum, (1) an understanding of time required for pollutants—and, in turn, pollutant reductions—to affect groundwater users; (2) an assessment of the potential source area (also known as the capture zone or zone of contribution; Barlow et al., 2018) and associated sources affecting individual wells; and (3) a prediction of pollutant levels potentially observed at a sensitive location (such as public supply wells).

Addressing NPS pollution is challenged by the inherent complexity of aquifer systems and their associated complex spatial distribution of hydraulic properties (Dagan, 1989; de Marsily et al., 2005; Kourakos et al., 2012). Since exhaustive characterization of this heterogeneity is unfeasible due to technical and financial constraints, significant uncertainty exists about solute transport. Geostatistical methods have been developed to quantitatively describe heterogeneous hydraulic aquifer properties (e.g., Kitanidis, 1997). Stochastic solutions to governing equations of groundwater flow and transport have been obtained analytically or numerically for geostatistical parameter representations (e.g., Rubin, 2003; Weissmann et al., 1999; Wen & Gomez-Hernandez, 1997).

Stochastic analysis has successfully been used to assess the fate and transport of pollutants from point sources (e.g., Fogg & Zhang, 2016; Graham & McLaughlin, 1991; Salamon et al., 2007). The approach is especially well suited for risk analysis (Andricevic & Cvetkovic, 1996; Maxwell & Kastenberg, 1999; de Barros & Rubin, 2008; setting up theoretical foundations) and has recently been applied in diverse simulation settings (e.g., Henri et al., 2015; Siirila-Woodburn et al., 2012, for organic compound contamination and CO₂ sequestration, respectively).

While source behavior has been highlighted as a controlling factor of uncertainty in point source plume predictions (Henri et al., 2016; de Barros & Nowak, 2010; Troldborg et al., 2010), few studies have assessed the role of aquifer heterogeneity and parameter uncertainty on transport from NPS. Zhang et al. (2006) provided stochastic solutions of travel times and concentrations related to a realistic, subregional-scale NPS pollutants in the production horizon of a heterogeneous aquifer system. Some efforts have been made on better understanding the importance of aquifer heterogeneity on source area delineation, focusing either on the role of heterogeneity characteristics (Feyen et al., 2001; van Leeuwen et al., 2000; Varljen & Shafer, 1991) or uncertainty propagation (Franzetti & Guadagnini, 1996; Stauffer et al., 2002). Few studies consider the propagation of uncertainty from aquifer heterogeneity on fate and transport of NPS pollutants in a conceptual setting that explicitly accounts for operating pumping wells. Zhang (2006) illustrates the large variability of salinity travel time to deep production wells in a heterogeneous alluvial system. Riva et al. (1999) investigated the propagation of uncertainty to time-related capture zones under radial flow due to extraction, while Libera et al. (2017) focused on the impact of pumping schedule on well concentration statistics. Both studies considered two-dimensional aquifers, but results indicate that well characteristics may have critical control on NPS management metrics and prediction uncertainty.

In this paper, we focus on better understanding the joint impact of aquifer heterogeneity and large water supply well characteristics (pumping rate, screen length, and depth of extraction) on the transport and fate of NPS pollutants to water supply wells and on the spatial extent and temporally evolving patterns of the contributing source area. Specifically, we focus on thick alluvial aquifer systems with large heterogeneity and significant downward gradients due to recharge and regional pumping, superimposed onto a regional lateral gradient. To do so, we perform high-resolution simulation of three-dimensional groundwater flow and nonreactive solute transport within a stochastic framework. We investigate and contrast the effects of mixing in supply wells, here considered the compliance surface, against the propagation of the uncertainty about the heterogeneous hydraulic conductivity field on and into the NPS management system. Specifically, we assess the uncertainty about three NPS management metrics: (1) travel time distribution in well water, (2) source area location of the well water, and (3) and future pollutant levels in well water.
Figure 1. Illustration of the synthetic aquifer section simulated, showing one realization of the heterogeneous hydraulic conductivity field and a snapshot of a single year's nonpoint source pollutants (particles) 150 years after they entered the water table with recharge. Colored particles are those that will eventually reach one of the three extraction wells (two wells shown; third well, on the right, located within the opaque portion of the aquifer). White particles exit the aquifer section either through the bottom boundary (representing other pumping wells) or laterally at $x = L_x$. The box dimensions in each direction are given in terms of number of cells ($n_i$) times the cell dimension ($d_i$). Yellow, blue, and red particles reached the left, middle, and right wells, respectively. The lower left figure inset zooms onto an extraction well. The lower part of each well (in red) represents its screened sections from where water is extracted at a rate $Q_{out}$ and where pollutant mass arrival is recorded. The top of the well screen is fixed, at a depth $d_s$. The well screen length $l_s$ depends on the location of highly conductive material encountered (gravel and sand, only shown on the well scheme) and the pumping rate (see text).

2. Methodology

2.1. Study Area

In this study, we consider a Central Valley (California, USA)-inspired unconsolidated sedimentary aquifer system with overlying predominantly irrigated agricultural land uses. The climate of the 20,000-km$^2$ basin is semiarid Mediterranean with surrounding mountain ranges capturing significant winter precipitation, while summers are dry. The basin supports an intensive irrigated agricultural industry that depends on both surface water inflows from surrounding mountain ranges, managed through reservoirs, and on groundwater. Historically, groundwater flow dynamics in the Central Valley groundwater system were dominated by mountain front recharge, lateral groundwater flow along weak gradients, and groundwater discharge along the thalweg of the basin (Tulare Lake, San Joaquin River, and Sacramento River). Since the mid-1900s, irrigation reduced or locally seized both, mountain front recharge and discharge along the thalweg. Instead, discharge of groundwater occurs predominantly by groundwater extraction in wells. As a result, recharge from irrigation return flows and groundwater pumping superimpose an anthropogenic downward gradient onto the weak lateral regional groundwater flow system.

The recharge from the irrigated landscape carries often elevated levels of nitrate and salinity across broad portions of the landscape (Central Valley Water Board, 2016; Harter et al., 2017). Here, we account exclusively for these diffuse sources of polluted recharge. We do not consider dilution observed due to recharge from major rivers due to their localized nature (Ransom et al., 2017).

We construct a representative subregional scenario with three pumping wells located along a transverse downstream transect of the simulated alluvial aquifer subregion. For the sake of simplicity and clarity in the analysis, no extraction wells are explicitly represented further upstream. Instead, the bottom of the simulation domain is considered an outflow boundary conceptually representing the general (simplified) impact of other groundwater pumping wells within or below the simulation domain on the subregional flow (Figure 1).

Aquifer hydraulic properties are representative of the highly heterogeneous unconsolidated alluvial sequence of fine- to coarse-grained sediments reflecting sequences of floodplain deposits, stream deposition,
Table 1

| Parameter                        | Value                        |
|----------------------------------|------------------------------|
| Domain discretization            |                              |
| Number of cells, \( n_x \times n_y \times n_z \) | 120 \times 60 \times 625    |
| Cell dimension, \( \Delta_x \times \Delta_y \times \Delta_z \) (m \times m \times m) | 160.0 \times 100.0 \times 0.4 |
| Domain length, \( L_x \times L_y \times L_z \) (m \times m \times m) | 19,200.0 \times 6,000.0 \times 250.0 |
| Flow and transport problem       |                              |
| Porosity, \( \phi \) (–)         | 0.3                          |
| Average longitudinal hydraulic gradient, \( i_z \) (–) | \( 1 \times 10^{-3} \)         |
| Recharge rate, \( R \) (m/day)   | \( 6 \times 10^{-4} \)       |

Table 1: Domain Discretization and Physical Parameters Used in All Simulations

and overbank deposits (e.g., Fleckenstein et al., 2006; Weissmann & Fogg, 1999; Weissmann et al., 1999). Uncertainty about the spatial distribution of aquifer hydraulic properties is accounted for by applying a stochastic approach using Monte Carlo simulations: equiprobable realizations of hydraulic properties are generated and solutions for groundwater flow and solute transport are obtained for each realization. The following sections describe the flow and transport model, the geostatistical model, and the stochastic metrics used to characterize uncertainty in the assessment and management of NPS pollution in groundwater.

2.2. Groundwater Flow

We consider a representative three-dimensional unconfined alluvial aquifer subregion within a rectangular prism domain (Figure 1). Domain dimensions are length \( L_x = 19,200.0 \) m, width \( L_y = 6,000.0 \) m, and thickness \( L_z = 250.0 \) m. The domain is discretized into 120 columns \( \times \) 60 rows \( \times \) 625 layers (total of 4.5 million cells) of dimensions 160.0 \( \times \) 100.0 \( \times \) 0.4 m, respectively (Table 1). We consider long-term average groundwater flow conditions, governed by (Rushton & Redshaw, 1979)

\[
\frac{\partial}{\partial x} \left( K_{xx} \frac{\partial h}{\partial x} \right) + \frac{\partial}{\partial y} \left( K_{yy} \frac{\partial h}{\partial y} \right) + \frac{\partial}{\partial z} \left( K_{zz} \frac{\partial h}{\partial z} \right) + Q'_s = 0,
\]

and groundwater flow is described by Darcy's law

\[
q(x) = -K(x) \Delta h(x),
\]

where \( q \) (m/day) is the specific discharge, \( h \) (m) is the hydraulic head, \( K \) (m/day) is the hydraulic conductivity tensor, and \( Q'_s \) is a volumetric flux per unit volume representing sources and sinks of water. The regional flow is driven by a recharge rate \( R \) of \( 6.0 \times 10^{-4} \) m/day and by a longitudinal mean gradient \( i_z \) of \( 1.0 \times 10^{-3} \), which are typical of southern Central Valley conditions (Faunt, 2009). These two hydraulic forces generate, on average, a top upstream to bottom downstream “diagonal” groundwater flow along the mean regional gradient. The simulated aquifer thickness is 250.0 m, sufficient to simulate the travel paths of contaminant mass reaching the simulated extraction wells (see below). While typical of most production depths in the Central Valley aquifer system, the system itself extends deeper and some well screens are located at larger depth. The overall effect of production depth groundwater pumping and pumping below the main production horizons is captured by assigning prescribed regional downward flow conditions across the lower boundary.

2.3. Groundwater Transport

Solute transport in the saturated aquifer is governed by the advection-dispersion equation:

\[
\phi \frac{\partial c}{\partial t} = \Delta \cdot (\phi D \frac{\partial c}{\partial x}) - \Delta \cdot (\nabla q),
\]

where \( c \) is the concentration of mobile solute, \( \phi \) is the porosity, and \( D \) is the dispersion tensor, usually defined as \( D = (\alpha_L |q| + D_m I + (\alpha_T - \alpha_L) |q|^2) / |q| \), where \( \alpha_L \) and \( \alpha_T \) are the longitudinal and transverse dispersivities, \( D_m \) is the molecular diffusion coefficient, and \( |q| \) is the magnitude of the velocity vector.

We adopt a stochastic framework in order to account for the uncertainty in the hydraulic conductivity (Rubin, 2003). The K-field is considered as a pseudorandom space function. Stochastic flow and transport process outcomes are estimated through Monte Carlo simulations. The Monte Carlo framework consists of generating a realization of the spatial distribution of the hydraulic conductivity from a geostatistical model,
Table 2

| Property                     | g   | s   | ms  | m   |
|------------------------------|-----|-----|-----|-----|
| Proportion (%)               | 0.10| 0.35| 0.25| 0.3 |
| Hydraulic conductivity (m/day)| 200.0| 50.0| 0.5 | 0.01|

solving the flow field driven by the generated heterogeneous K-field, solving the transport problem from the resulting flow field, and analyzing the travel time and the source area location \(x_p(t = 0)\) of the pollutant at each extraction well.

2.4. Geostatistical Model

The spatially correlated distribution of the three-dimensional hydraulic conductivity is simulated using the FORTRAN program T-PROGS (Transition Probability Geostatistical Software). T-PROGS uses the transition probability/Markov chain approach to generate realizations of the hydrofacies field (Carle, 1999; Carle & Fogg, 1996, 1997).

In this work, we describe the variability in the sediments forming the aquifer by considering four hydrofacies (or categories): gravel, sand, muddy sand, and mud. A similar approach has been previously adopted to model alluvial aquifers in the Central Valley (Kings River Valley and San Joaquin Valley) with a satisfactorily representation of the heterogeneity in the hydraulic properties (Weissmann et al., 1999; Zhang, 2006). A single value of hydraulic conductivity \(K\) is allocated to each category. We chose these values to be of the same order of magnitude as values calibrated in the formerly cited studies (see Table 2).

The transition probability/Markov chain approach requires the user to define the proportion of each facies \(j\) of their mean length in each direction \(L_{ij}\) and of the facies \(j\)-to-facies \(k\) transition probability rates in each direction \(r_{j\rightarrow k}\) (Carle & Fogg, 1996, 1997). The last two parameters are not defined for the background category (here mud), which can be interpreted as the material filling the space not occupied by other categories and which is characterized by its volume proportion only. All values are representative of the observed spatial distribution of Central Valley sedimentary deposit (Weissmann & Fogg, 1999; Zhang et al., 2006) and are shown in Tables 2 (proportions) and 3 \(L_{ij}, r_{j\rightarrow k}\). The grid cell dimensions (Table 1) were chosen to be a fifth of the minimum mean length in each direction, which provides for appropriate numerical resolution of flow and transport effects due to aquifer heterogeneity.

Using these parameters, 50 realizations of the K-field were sufficient to converge the average and the variance of the hydraulic conductivity mean and variance. The converged mean and variance of the natural logarithm of the K-fields \(\ln K(x)\) were 0.28 and 14.8, respectively.

2.5. Extraction Well Design

While the impact of production well density on contamination management has been quantitatively assessed in an extensive manner (e.g., Guo & Brusseau, 2017), the impact of well characteristics on transport behavior and uncertainty under NPS contamination is not well understood. Among various well characteristics, three primarily control the arrival of solute mass: the pumping rate \(Q_{\text{out}}\), the depth to the top of the well screen \(d_s\), and the length of the well screen \(l_s\).

In practice, the heterogeneity of the aquifer is taken into consideration when setting the latter \(l_s\). A well screen must intersect a sufficient cumulative thickness of highly conductive material to sustain a targeted pumping rate. Drillers in the Central Valley often use a rule of thumb stating that 10 cumulative feet (3.048 m) of gravel and sand layers crossing the well screen yields about 100 gpm (545.1 m³/day) of pumping. We employ this rule to determine the depth of the simulated well screen, given the targeted pumping rate, the well location, the depth to the top of the screen, and the sequence of vertical facies in the aquifer realization below the top of the screen. Both sand and gravel facies encountered along the simulated borehole are counted toward “highly conductive material” thickness (Figure 1). In each simulation, the three extraction wells are located on a transect near the downstream edge of the domain, at \(x_w = 18,000.0\) m. Wells are separated from one another by a distance of 1,500.0 m, \(y_w = 1,500.0, 3,000.0,\) and 4,500.0 m (see Figure 1).

Two sets of Monte Carlo simulations are designed to assess the impact of pumping rates and screen depths on management metrics. In the first set, we consider four \(Q_{\text{out}}\) values: 750.0, 1,500.0, 3,000.0, and 6,000.0
Table 3
Mean Length (Diagonal Values) of Each Categories (g: Gravel, s: Sand, ms: Muddy Sand, and m: Mud) and Embedded Transition Probability (Nondiagonal Values) in the Longitudinal (x), the Horizontal Transverse (y), and the Vertical Transverse (z) Directions

|   | x       | g       | s       | ms      | m       |
|---|---------|---------|---------|---------|---------|
| g | \(L_{g, x} = 800.0 \text{ m}\) | 0.7     | 0.15    | b       |         |
| s | 0.7     | \(L_{s, x} = 1.500.0 \text{ m}\) | 0.28    | b       |         |
| ms| 0.15    | 0.28    | \(L_{ms, x} = 1000.0 \text{ m}\) | b       |         |
| m | b       | b       | b       | b       |         |

|   | y       | g       | s       | ms      | m       |
|---|---------|---------|---------|---------|---------|
| g | \(L_{g, y} = 500.0 \text{ m}\) | 0.7     | 0.15    | b       |         |
| s | 0.7     | \(L_{s, y} = 850.0 \text{ m}\) | 0.28    | b       |         |
| ms| 0.15    | 0.28    | \(L_{ms, y} = 900.0 \text{ m}\) | b       |         |
| m | b       | b       | b       | b       |         |

|   | z       | g       | s       | ms      | m       |
|---|---------|---------|---------|---------|---------|
| g | \(L_{g, z} = 2.0 \text{ m}\) | 0.7     | 0.15    | b       |         |
| s | 0.7     | \(L_{s, z} = 3.5 \text{ m}\) | 0.28    | b       |         |
| ms| 0.15    | 0.28    | \(L_{ms, z} = 2.0 \text{ m}\) | b       |         |
| m | b       | b       | b       | b       |         |

Note. Matrices read as transition probability from the row facies to the column facies. The background category is designated by the letter b.

m³/day for a fixed depth to the top of the screen of 100.0 m. In the second set, we consider three \(d_s\) values: 50.0, 100.0, and 150.0 m for a fixed pumping rate of 3,000.0 m³/day. These values are all within the range of \(Q_{out}\) and \(d_s\) observed in Central Valley’s production wells (supporting information Figures S1–S4 display histograms of pumping rates, screen depth, and screen length from all wells recorded by the California Department of Water Resources).

### 2.6. Implementation of Flow and Transport Models

**Flow.** For each realization, the flow equation (2) is solved using the widely used finite difference code MODFLOW-2000 (Harbaugh et al., 2000). The flow is constrained by a no-flow boundary condition at the lateral side edges of the domain. Recharge is applied uniformly on the entire top of the aquifer. The pumping rate is nonuniformly distributed over the well screen using the Multinode Node Well 2 package (Konikow et al., 2009). Transfer of water toward the deepest portion of the aquifer system will occur to mimic both natural flow, distributed production aquifer pumping, and deep aquifer pumping. This vertical flux is accounted for by specifying a K-weighted flow boundary condition applied on the entire bottom of the domain. For proper steady flow conditions, the total vertical flux is set to be the difference between the total amount of recharge at the top of the aquifer and the pumping rate in the three explicitly represented wells. On the other hand, a depth-dependent specified head at the upstream and downstream edge of the domain assures the preservation of the average regional flow direction superimposed over the vertical gradient. This yields a downward inclined, quasi-diagonal mean flow.

Just as the longitudinal gradient (\(i_x\)) is maintained by setting a head differential between the upstream and downstream edges of the domain, the vertical gradient (\(i_z\)) can be preserved by imposing a head differential between the upper and lower layers of the upstream and downstream edges of the aquifer. The average vertical gradient \(i_z\) within the plane of specified heads assigned to the upstream and downstream faces of the simulation domain needs to be estimated beforehand: Applying Darcy’s law and assuming recharge as the main driving force of the vertical fluxes, \(i_z\) can be estimated as \(i_z \approx R/(K_h)\), where \(\langle K_h \rangle\) is the ensemble average of the K-fields harmonic mean over the 50 realizations \(\langle K_h \rangle = 3.2 \times 10^{-2} \text{ m/day}\). This gives a vertical gradient \(i_z\) of 1.9 × 10⁻².

**Solute transport.** The advection-dispersion equation (3) is solved using the random walk particle tracking (RWPT) method implemented in the Fortran90 code RW3D (Fernández-Garcia et al., 2005; Henri & Fernández-Garcia, 2014, 2015). This approach presents the advantage of being exempt of numerical
intrafacies transport (port simulation results in highly heterogeneous media are insensitive to local scale dispersivity relevant to

Previous studies (LaBolle, 1999; LaBolle & Fogg, 2001; Weissmann et al., 2002) highlighted that trans-

A total of 500,000 particles is instantaneously released over the entire top edge of the domain, that is, 

\[ 0.1; 0.01; 1.7 \] m

signal resulting from a continuous injection of an identical mass

\[ \alpha \]

time:

we here resolve only the advection equation, leading to an order of magnitude improvement in computation 

observed at intermediate and late times (supporting information Figure S5). For computational efficiency, 

dispersivity to 15.0 m leads to slightly earlier first arrivals, later late arrivals, and lower contaminant levels 

with a continuous mass release from the NPS are obtained from 

\[ c \]

Wesimulate a pulse injection; that is, all particles are instantaneously released at

\[ t \]

n

minimum number of move a particle is likely to do in a cell. After a sensitivity analysis (Figure S7), we fixed 

n to 5.

We simulate a pulse injection; that is, all particles are instantaneously released at \( t = 0 \), representing a single

year’s NPS pollutant release at the water table. However, the principle of superposition states that cumulative

breakthrough curves obtained from a pulse injection of mass \( m_0 \) (\( \dot{m}_0(t) \)) can be interpreted as the mass flux 
signal resulting from a continuous injection of an identical mass \( m_c \). Hence, well concentrations associated 

with a continuous mass release from the NPS are obtained from

\[ c(t) = \dot{m}_c(t)/Q_{\text{out}}. \]

2.7. NPS Management Metrics and Uncertainty

Travel time. A critical decision variable in NPS pollution management is the issue of legacy pollution: deter-

mining the time delay between initiation of improved pollutant source management and the improvement in 
groundwater quality at the receptor of interest (well, stream reach). Assessing the probability distribution 
of the travel time needed for NPS pollution to reach a well screen is therefore an important aspect of NPS 
management.
Within the stochastic framework, travel times are compiled into a probability density function (pdf). Travel times recorded for the three wells are considered to be statistically independent of each other, creating a total of 150 realizations from which to compute pdfs. At constant annual NPS input, travel time pdfs express the likelihood to observe a specific relative concentration or mass breakthrough after elapsed time \( t \) (Cvetkovic et al., 1992; Dagan & Nguyen, 1989). Travel time pdfs have demonstrated their usefulness in probabilistic risk assessment (Andricevic & Cvetkovic, 1996; de Barros & Rubin, 2008; Henri et al., 2015) by allowing for visualization of the expected travel time (mean travel time for Gaussian-like pdfs), and of the main characteristics of travel times uncertainty (through the spread and the shape of the pdfs). Travel times pdfs can be parameterized using the mean and the coefficient of variation (CV), where CV = \( \sigma_t / \mu_t \). Here, the CV represents an explicit measure of the impact of aquifer heterogeneity on the travel time uncertainty, independent of the spatial distribution and temporal behavior of input concentrations at the surface.

The maximum mass is here defined as total cumulative mass emanating from the initial source pulse and extracted by the well over \( t_{end} = 400 \) years. Four hundred years was determined by trial and error to be a sufficient simulation time to reach maximum mass extraction. In this study, travel times will be labeled as \( t_i \), where \( i \) is the portion of the cumulative mass (e.g., \( t_{50} \) will be the time required for 50\% of the cumulative mass to reach the well).

**Spatiotemporal extension of the source area.** Another key metric in NPS contamination management is the extension of the source area, also known as the capture zone, zone of contribution, or well-head protection area (Barlow et al., 2018), that is, the spatial extent of the zone of pollutant recharge that will eventually reach a specific extraction well. In a stochastic framework, the source area or “stochastic capture zone” (e.g., van Leeuwen et al., 1998) is defined as the spatial distribution of the probability that a contaminant leaving the NPS at a location \( x \) will reach a well \( P_w(x_{nps}) \)

\[
P_w(x_{nps}) = \text{Prob}(x_p(t \in [t_0, t_{end}]) = x_{nps} | c_{nps}(t_0) = c_{nps}),
\]

where \( x_p \) is the three-dimensional location of a portion of the plume (represented by a particle in this study), \( x_{nps} \) is a location shared with a well screen, and \( x_{nps} \) is a given location of the NPS. All locations are expressed in the Cartesian coordinate system given by \( x = (x, y, z) \). The spatial extension of nonzero probabilities corresponds to the stochastic source area. In a RWPT framework, the number of particles leaving a given grid cell with coordinate \( x \) and reaching a well \( (n_{nps}^w(x_{nps})) \) can be normalized by the total number of particles leaving the same grid cell \( (n_p^w(x_p)) \) to estimate \( P_w(x_p) \), the probability for the contaminant to leave a given location of the NPS and to be captured (i.e., \( P_w(x_p) = n_{nps}^w(x_{nps}) / n_p^w(x_p) \)). In our setting, each well is considered to be an independent realization of NPS capture by a well. The location of the grid cell \( (x_p) \) is expressed as the relative distance from the well by which a particle is captured. The 50 aquifer realizations generate 150 realizations of wells affected by an NPS pollutant. \( P_w \) reflects the sample statistics for a sample size of 150. We also map the temporal behavior of the source zone, that is, the time required for a contaminant leaving a given location within the stochastic source area.

**Future NPS pollutant levels in supply wells.** Among the most relevant management metrics is the pollutant level observed at an extraction well, which often represent the foundation of control and legislative action. Here, we use a probabilistic mapping tool to visualize the temporal evolution of pollutant exceedance probability over time, for a compliance level of interest (e.g., drinking water maximum contamination level) that is defined here relative to the NPS source concentration at the water table. Under heterogeneous aquifer conditions, the time between the beginning of NPS pollution and the time at which a well exceeds the compliance level is highly uncertain. To account for this uncertainty, we analyze \( P_{c_j}(t) \), the probability to face a normalized concentration \( c^* \) between the values \( c^*_i \) and \( c^*_j \) at a given time \( t \); that is,

\[
P_{c_j}(t) = \text{Prob} \left( c^*_i(t) < c^*(t) < c^*_j(t) \right).
\]

Here again, a single data set is formed by gathering pollutant signals from three pumping wells in each of 50 aquifer realization for a total sample size of 150 wells.

**NPS metrics for equivalent homogeneous aquifer representation.** To assess and illustrate the relative magnitude of uncertainty and variability in NPS metrics due to aquifer heterogeneity against the effects of age mixing in a supply well—with and without pumping—we compute travel time distribution, source area, and future NPS pollutant levels for two heuristic, simple equivalent homogeneous cases without transport...
dispersion: For the nonpumping case, horizontal and vertical fluxes are uniform and set equal to the average of the regional horizontal flux and recharge flux, respectively, obtained from the stochastic flow representation. Well screen depth is set to the average screen depth obtained from the stochastic model (see supporting information Figure S8 and Tables S1–S6 for details).

The effect of pumping on the source area in an equivalent homogeneous aquifer with nonuniform 3-D flow field is heuristically illustrated by first computing the size of the source area as the ratio of pumping rate \((m^3/day)\) and the average recharge rate \((m/day)\) of the stochastic model and then identifying an equally sized area in the stochastic source area maps at locations with the highest source area probabilities (see supporting information Table S7). These two NPS metrics provide an important benchmark, because they can be easily obtained for specific real-world scenarios from well design data and average regional groundwater flow and recharge properties.

3. Results and Discussion: Travel Times to Wells

For purposes of our analysis, well screens are here considered the critical compliance surfaces. The dependence of well screen lengths, \(l_s\), on pumping rate and local aquifer heterogeneity is illustrated by its probability distribution over all K-field realizations (supporting information Figure S9). Simulated screen lengths follow an exponential distribution for low pumping rates and trend to a Gaussian-like shape for largest extractions. Increasing the pumping rate leads also to an increase of both mean and standard deviation of the screen length. \(l_s\) values range from 5 to 50 m for a \(Q_{out}\) of 750 m\(^3\)/day and from 30 to 130 m for a \(Q_{out}\) of 6,000 m\(^3\)/day, which is consistent with screen lengths observed across alluvial systems in the State of California (see supporting information Figure S2).

3.1. Effect of Pumping Rate on Travel Time Distribution

Impact of \(Q_{out}\) on mean travel time. Average simulated travel time range from several decades to well over a century for early travel times (5th percentile), representing the shortest travel paths to the top of the screen. They range from a century to four centuries for late travel times (95th percentile; Figures 2c and 2d), representing mostly the travel time to the bottom portion of the well screen. In comparison, the youngest water in the simple equivalent benchmark well case is 140 years, while the oldest water, arriving at the bottom of the well screen, is from 10 to 100 years older, depending on screen length, which in turn is dictated by pumping rate.
Q_{out} is shown to have only minor effect on expected early travel time since all screen tops are at the same depth. A slight decrease of 〈t_{5}⟩ is only observed for large Q_{out} due to the strong vertical gradient created near the supply well (Figures 2a and 2d). Median (50th percentile) travel time pdfs are also similar across all pumping rates, with slightly later mode of the pdf for the largest pumping rate (Figures 2b and 2d). Larger pumping rates lead to longer screens and significantly deeper water being pumped. Hence, the age (travel time) of the oldest water pumped significantly increases with pumping rate, which is reflected in the sensitivity of the late travel time distribution to Q_{out} (Figures 2c and 2d).

To the degree that the 5% travel time represents water entering near the top of the screen and the 95% represents water entering the bottom of the screen, the results indicate that the distribution of pollutant travel times and, equivalently, of groundwater age within a single water sample from a supply well typically ranges over more than a century (Figure 2). Such distributions of groundwater age may be difficult to measure, given the lack of natural tracers that bracket the decade to few century groundwater age scale. However, Visser et al. (2016) found that, throughout most of the Central Valley, groundwater in wells with a depth of 60 m and more is likely to be tritium free, that is, characterized with groundwater ages now greater than 60 years, confirming our result. Moreover, these simulation results are also roughly consistent with the estimation of groundwater age distribution made by Weissmann et al. (2002) in the San Joaquin Valley from detailed modeling and from chlorofluorocarbon age data. They determined the mean groundwater age ranges from 10 to 50 years in wells that are 2 to 3 times shallower—therefore, having about 2 to 3 times younger water—than the ones simulated in this study.

**Impact of Q_{out} on travel time uncertainty.** The uncertainty about travel time (CV) varies from 20% to 40% of the mean early, median, or late travel time pdf at the well (Figures 2 and 2e). Prediction uncertainty is highest at early travel times as observed in a series of previous work (e.g., Dagan & Nguyen, 1989; Henri et al., 2015) and at low Q_{out} (Figure 2e). For the late travel time pdf, the CV continuously decreases as pumping rate increases. In contrast, the CV decreases at first and is nearly stable for median travel time while it increases again for the early travel times, after a minimum at Q_{out} = 1,500 m³/day (Figures 2 and 2e).

The nearly Gaussian pdfs of the 5%, 50%, and 95% travel time pdfs (Figure 2) rather than a strongly right-tailed distribution result from the strong attraction of mass in the vicinity of the well. Mass is most likely to travel along connected higher-velocity paths, prohibiting the occurrence of significant late mass arrivals from long transport through lower conductive facies in the heterogeneous media. Thus, high pumping rates decrease the uncertainty about travel time that otherwise arise from the variability of aquifer heterogeneity and resulting flow structures across realizations. Including typical values for local dispersion into the simulator does not affect this finding (Figure S5; see Weissmann et al. (2002) for an example of typical values).

Deep wells (i.e., with high pumping rates) show left-skewed late travel time distributions. This may be the result of the finite distance between the bottom of the screen and the bottom of the simulation domain which is about 30 to 50 m (see supporting information Figure S9). Upward flow from greater depths toward the well may be limited by the simulation setup, possibly truncating later travel times pdfs.

Pumping rates have a more moderate impact on mean travel times than the joint impact of screen lengths and regional flow conditions. The latter dictate overall regional downward travel times. Large pumping rates, by favoring the transport in fast, short flow paths, inhibit the propagation of the K-field uncertainty into travel time uncertainty.

**3.2. Well Depth Effect on Travel Time**

**Impact of d_{ts} on mean travel time.** Depth to the top of the well screen is shown to be a strong controlling factor of travel time. Results depict a quasi-linear increase of travel times with d_{ts} for both the stochastic results and the equivalent homogeneous nonpumping benchmark case (Figures 3a–3d). The mode of the median travel time occurs less than 10% earlier than the travel time to the center of the well in the benchmark case. This shows that the strong dependence of travel time on well depth is driven largely by the larger travel distance required to reach the screen as depth to the top of the screen increases.

Most early travel times are significantly earlier than the nonpumping benchmark due to the increased vertical gradient near the pumping well and the presence of well-connected high-velocity travel paths. In contrast, most late travel times occur later than the benchmark due to upcoming of deeper water into the pumping well. The absolute deviation to the nonpumping benchmark remains nearly constant with d_{ts}. 

HENRI AND HARTER
Figure 3. (left panels) Probability density function of the time required for 5% (a), 50% (b), and 95% (c) of the total recorded mass to reach a well as a function of the depth to the top of the well screen: 50.0 m (blue), 100.0 m (red), and 150.0 m (yellow). Here, the pumping rate is fixed at 3,000.0 m³/day. For reference, vertical dashed lines represent the youngest (top), median (middle), and oldest (bottom) water for the equivalent homogeneous, nondispersive, and nonpumping well case. (right panels) Mean (d) and coefficient of variation (e) of the travel time required for 5% (blue), 50% (red), and 95% (yellow) of the total recorded mass to reach a well as a function of the top of the well screen depth. The pumping rate is fixed to 3,000.0 m³/day.

Mean late travel time is about 70 years later, and the mean early travel time is about 20–40 years earlier than the nonpumping benchmark travel time to the bottom and top of the screen, respectively. The nearly linear increase of the mode of the early, median, or late travel time distribution with well depth indicates that there is only a weak effect of travel distance on apparent velocities.

Impact of $d_s$ on travel time uncertainty. The standard deviation of travel time increases for the median and late travel time pdfs but not as rapidly as the means of the early, median, and late travel time pdfs; Hence, for all pdfs, the CV decreases with increasing $d_s$, from a range of 30–45% for $d_s = 50$ m to a range of 13–25% for $d_s = 150$ m (Figure 3e), reflecting the effect that aquifer heterogeneity imparts on travel time variability between wells. Except for the deepest wells, the travel time pdf shape is typical of Gaussian random fields (Figure 3a). For $d_s = 150$ m the pdf shape of the 95% travel times is negatively skewed, with late mode and quickly decreasing probability of very large travel times (Figures 3b and 3c).

The decrease of relative uncertainty about travel time with screen depth (Figure 3e) shows that longer travel distances yield more ergodic transport conditions along the travel path and, therefore, decrease the importance of transport through preferential paths. For the shorter travel distances to the top of the well, travel time uncertainty is large, as some wells (some realizations) connect with the water table through fast flow paths while others may not.

4. Source Area and Source Area Uncertainty

4.1. Impact of Pumping Rates

Stochastic source area delineation. The potential extent of the source area (here referred to as stochastic source area but also called the stochastic capture zone; van Leeuwen et al., 1998) is given by the area characterized by $P_w(x_c) > 0$. A predominant portion of the stochastic source area is characterized by very low to insignificant probabilities to reach a well, which indicates significant uncertainty about its delineation. The zone of the highest probability, delineated graphically by the benchmark equivalent homogeneous case with pumping, is here referred to as the “critical zone,” which is spatially more limited in extent than the stochastic source area (Figure 4).

Increasing the pumping rate extends the stochastic source area in the transverse and (more significantly) in the longitudinal direction (Figure 4). It also pointedly increases the probabilities that pollutants in the
Figure 4. Probability of a particle leaving a given grid cell to reach a well accounting for a pumping rate $Q_{\text{out}}$ of 6,000.0 m$^3$/day (a), 3,000.0 m$^3$/day (b), 1,500.0 m$^3$/day (c), and 750.0 m$^3$/day (d). The depth of the screen top is fixed to 100 m. The axes are expressed in terms of distance from an extraction well (red cross). Nonzero probabilities (colored) represent the maximum probable extension of the source area. The source area of the benchmark nonpumping case is delineated by the gray line, indicating with gray “x” marks the source of water to the top, center, and bottom of the well screen. The source area of the benchmark pumping case is delineated by the closed gray contour.

critical zone actually reach the well. For $Q_{\text{out}} = 3,000$ m$^3$/day and $Q_{\text{out}} = 6,000$ m$^3$/day, $P_w$ is above 50% within the critical area. Decreasing $Q_{\text{out}}$ leads to lower $P_w$ values over the entire source area: $P_w$ within the critical zone fall well below 50%, in the case of $Q_{\text{out}} = 750$ m$^3$/day even below 20% (Figure 4d).

By design, a higher pumping rate is here associated with an increased length of the screen. As discussed previously, deeper extraction will attract mass coming from larger distances and will therefore tend to lengthen and widen the probable source area, as seen by both the nonpumping and the pumping benchmark case in Figure 4. At the same $d_{\text{ss}}$, the downgradient edge of the stochastic source area, most likely associated with early travel time to the top of the screen, remains roughly identical for all $Q_{\text{out}}$ values.

The size of the stochastic source area increases with pumping rate. This is a reflection of the larger area needed to supply the recharge equal to the larger pumping rate. On the other hand, the aquifer heterogeneity structure (hydraulic conductivity distribution) will place the source area in different places between realizations and impart a depth-dependent uncertainty on the location of those areas (see below). The map of $P_w$ values exhibits both phenomena. Low pumping rates have small source areas relative to the heterogeneity scale of the aquifer. Individual source areas between multiple realizations often will not overlap. Overall, then, any location has a low probability for the contaminant to reach a well, while the stochastic source area is much larger than any individual source area or the critical zone. In the theoretical case of an infinitesimally small pumping rate, the size of the source area for a given realization merges into a point with the stochastic source area infinitely larger than the critical zone represented by the point.

For NPS management, the implications may be significant: Land management-related protection efforts would be most effective within the critical zone, where reductions in pollution most likely improve groundwater quality in the water supply well. But identifying a preferred zone of action in case of low pumping rate would be more challenging given the widely distributed area of low $P_w$ values within the relatively large stochastic source area.
The results suggest that, from a practical perspective, a NPS protection zone covering much of the stochastic source area may be delineated by expanding the size of the source area obtained from the benchmark pumping case by a factor of 3–4 in the transverse direction and by a factor of 2–3 toward the well, with the benchmark nonpumping case delineating the third quarter upstream portion of such a protection zone.

Stochastic source area travel time map. While the assessment of the source area is a spatial issue, managers and decision makers also have significant interest in understanding the temporal dynamics of NPS contamination across the source area of water supply wells, specifically the average travel time of particles from a given location of the NPS for various sized wells (various pumping rates, Figure 5). Higher \( Q_{\text{out}} \) results in larger variability of travel times across the source area. Higher pumping rates come with longer screens and, hence, a larger source area that extends further upgradient (Figure 5a). Mass located at the upstream edge of the source area travels longer distances to eventually reach the lower portion of the screen, explaining the increase in variability of travel times across the source area, mirroring the travel time variability across the corresponding well screen.

At low pumping rates, travel time correlates less to the distance between source location and well than under high pumping rates due to the relatively strong effects of aquifer heterogeneity (e.g., Figure 5d). We also observe that, at the same location, the mean travel time of a solute is slightly larger at low pumping rates than at high pumping rates. The difference is more pronounced at source area locations in the vicinity of the well. The increasing travel time for lower pumping rates is caused by the smaller drawdown near the well, resulting in slower flow conditions immediately upgradient of the well and smaller vertical gradients near the well. Thus, for low \( Q_{\text{out}} \) aquifer heterogeneity yields a stronger impact on the spatial distribution of mean travel times. The increased local vertical and horizontal head gradients produced by a large \( Q_{\text{out}} \) attenuate the effects of heterogeneity, yielding better correlation between travel times and travel distances. These findings, here in a three-dimensional setting, are consistent with the results of Libera et al. (2017) who found that pumping operations might overshadow the impact of aquifer heterogeneity on contaminant transport.
Figure 6. Coefficient of variation of travel times associated to particles leaving a given grid cell and reaching a well for a pumping rate $Q_{\text{out}}$ of 6,000.0 m$^3$/day (a), 3,000.0 m$^3$/day (b), 1,500.0 m$^3$/day (c), and 750.0 m$^3$/day (d). The depth of the screen top is fixed to 100 m. The axis is expressed in terms of distance from an extraction well (red cross). The source area of the benchmark pumping case is delineated by the closed gray contour.

Uncertainty about travel times at a given source area location, relative to the mean travel time, is again measured by their CV (Figure 6). Higher CV is observed for locations within the source area that are near the wellhead (Figure 6). Also, at a given distance from a well, relative uncertainty about pollutant travel time slightly increases with $Q_{\text{out}}$. Increasing the pumping rate also leads to stronger correlation between distance from well and travel time uncertainty (Figure 6a).

Large travel time CV in the downgradient section of the source area is consistent with the observed higher CV of the early travel time pdfs ($t_{5}$ in Figure 2b). The mean travel time and travel time CV maps also illustrate that short travel times may be associated with NPS locations from a range of distances from a well and may not systematically originate from near the well.

The travel time maps also reflect another point observed in the travel time pdfs (compare Figure 2 with Figures 4 and 5): larger well extraction leads to decreased uncertainty about the longest travel times but increases the time between the mean early and mean late travel time, that is, the variability of travel times across the source area (Figure 6). Especially near the well, large pumping rates attract the most easily available pollutant mass, that is, the NPS pollutant traveling through randomly located, uncertain connected paths that most optimally link the source to the pumping well. At the same time large pumping rates increase the range of the source locations of the NPS pollutant due to the mobilization of a larger set of preferential channels. Similarly, complex interaction between heterogeneity, pumping, stochastic source zone, and travel time uncertainty has been previously discussed by de Barros et al. (2013) using two-dimensional sequential Gaussian simulator of $K$-fields.

4.2. Analyzing the Impact of the Well Depth

As the depth to the top of the screen increases, the source area moves further away from the well. However, despite similar screen lengths and identical pumping rates, the source area of wells that are located deeper in the aquifer will be significantly larger—more elongated but not significantly more spread out laterally—than for wells with shallow top of screen depth (Figure 7). The larger source area—at the same...
Figure 7. Probability of a particle leaving a given grid cell to reach a well. Depth to the top of the screen \(d_{ts}\) is 50.0 m (a), 100.0 m (b), and 150.0 m (c). The pumping rate is fixed to 3,000.0 m³/day. The axes indicate distance from the extraction well (red cross). Nonzero probabilities (colored) represent the source area—the maximum probable extension of the source area. The source area of the benchmark nonpumping case is delineated by the gray line, indicating with gray “x” marks the source of water to the top, center, and bottom of the well screen. The source area of the benchmark pumping case is delineated by the closed gray contour.

Pumping rate—is caused by pollutants experiencing more of the aquifer heterogeneity along their longer travel path (Figure 7c). The stochastic source area reflects the macrodispersive behavior of the pollutant movement in the strongly heterogeneous media as first discussed by van Leeuwen et al. (1998).

For shallower wells, the critical zone is located toward the downgradient portion of the source area rather than in the center (Figure 7a). But with increasing depth to the top of the screen, the critical zone moves toward a more central position within the stochastic source area, for example, for \(d_{ts} = 150.0\) m (Figure 7c). Also, when the top of the well screen is located deeper, the magnitude of \(P_w\) decreases; in other words, the maximum likelihood that any specific location contributes pollution to the well is decreasing (Figures 7a to 7c). The lower \(P_w\) in the critical zone reflects the larger size of the stochastic source area (larger area with nonzero \(P_w\)).

The further a pollutant needs to travel to enter a well screen, the larger the role of aquifer heterogeneity and the uncertainty about both whether the pollutant will be captured by the well and—if so—when the pollutant is captured by the well. Ye, CV of travel time decreases with larger travel distance as seen earlier, for larger pumping rates. Shallower depth to top of the screen implies shorter travel distances, shorter travel times, (Figure 8), and less absolute uncertainty (standard deviation, not shown), yet more relative uncertainty (CV).

The zone of lowest travel times within the source area is larger for shallow screens (Figure 8a). The proximity of extraction from the surface allows a quick transfer of mass from the source to the well by increasing the local gradient and exacerbating transport through fast paths. The range of mean travel times is significantly larger for the larger well depths, exceeding 300 years in a significant upstream portion of the source area (Figure 8c).

By comparing the spatial distribution of travel time CVs between different depths to top of well screen (Figure 9), we find that the CV for travel time is smaller with deeper top of screen and the source location
Figure 8. Expected travel time (years) of a particle leaving a given grid cell and reaching a well for a depth of the top of the screen \(d_s\) of 50.0 m (a), 100.0 m (b), and 150.0 m (c). The pumping rate is fixed to 3,000.0 m\(^3\)/day. The axis is expressed in terms of distance from an extraction well (red cross). The source area of the benchmark pumping case is delineated by the closed gray contour.

of the contaminants is further away. Near the well, local connectivity may greatly accelerate contaminant extraction. Over longer distances, direct connectivity to the well is less likely, and relative travel time uncertainty decreases. For deep wells, the critical zone is therefore characterized by low CV about travel time (Figure 9c).

5. Stochastic Analysis of the Temporal Evolution of Well Pollutant Concentration

We here use a probability map as a visualization tool to present the stochastic temporal evolution of contaminant levels in a production well at various compliance concentration levels of interest. We normalize concentration using \(c^*(t) = c(t)/c_0\), \(c_0\) is the pollutant concentration in recharge and is obtained as \(c_0 = m_0/V_{rch}\), with \(m_0\) the amount of injected mass and \(V_{rch}\) the volume of recharged water. Particle travel time distributions are integrated into breakthrough curves, thus simulating continuous, constant loading of NPS pollutants across the landscape, beginning at \(t = 0\). Breakthrough curves from the 150 individual realizations are aggregated into an exceedance probability map of normalized concentrations \(c^*(t)\) such that \(P_{c_{crit}}(t) = \text{Prob}(c^*(t) > c_{crit}^*)\), where \(c_{crit}^*\) is the normalized compliance concentration of interest. \(P_{c_{crit}}(t)\) is the exceedance probability that the normalized compliance concentration will be exceeded after a given time \(t\) since the start of the (continuous) NPS contamination. We evaluate the effect of pumping rate and well depth on \(P_{c_{crit}}(t)\). From a practical perspective relevant to policy and decision makers, assuming ergodicity (e.g., Dagan, 1989), \(P_{c_{crit}}(t)\) can also be interpreted as the fraction of wells exceeding the compliance level across an ensemble of supply wells located in the same aquifer system, with similar construction design, and subject to similar NPS pollution levels. We discuss the results from the latter perspective.

As expected, across all cases investigated, the fraction of wells exceeding a specific compliance level increases with time. At a specific time, the fraction of wells exceeding a compliance level decreases for higher compliance level (Figures 10 and 11). Even in the nonpumping benchmark case with no dispersion, the mixing
Figure 9. Coefficient of variation of travel times associated to particles leaving a given grid cell and reaching a well for a depth of the top of the screen $d_{ts}$ of 50.0 m (a), 100.0 m (b), and 150.0 m (c). The pumping rate is fixed to 3,000.0 m$^3$/day. The axis is expressed in terms of distance from an extraction well (red cross). The source area of the benchmark pumping case is delineated by the closed gray contour.

of pumped water within the well screen leads to low compliance levels being exceeded earlier than high compliance levels: The time difference between exceedance occurring at $c/c_0 = 0.1$ versus $c/c_0 = 0.9$ varies from about one decade at screen lengths typical for $Q_{out} = 750$ m$^3$/day to a century for a much longer well screen, typical for $Q_{out} = 6,000$ m$^3$/day (red dashed line in Figure 10). In the heterogeneous aquifer system, it takes typically two to three centuries until most wells exceed a compliance level that is 80% or more of the recharge concentration, while a compliance level that is 1 order of magnitude smaller than the recharge concentration will be exceeded by more than 10% of wells within a few decades.

For lower pumping rates and shallower depth to the top of screen (e.g., $Q_{out} \leq 1,500$ m$^3$/day at $d_{ts} \leq 100$ m or $Q_{out} \leq 3,000$ m$^3$/day at $d_{ts} = 50$ m), the nonpumping benchmark case provides an important result for the specific case $c_{crit}^* = 0.5$ (compliance level is half of the recharge concentration): In those cases, the time at which about half of all wells exceed the compliance level is equal to that for the benchmark case, despite the large heterogeneity in the aquifer system. In the benchmark case the time to reach $c_{crit}^* = 0.5$ is equal to travel time to the center of the well screen in the equivalent homogeneous medium with no pumping.

Pumping rate has a significant impact on the fraction of wells exceeding the compliance levels, consistent with the findings above. For compliance levels that are an order of magnitude or more below the recharge concentration, high pumping rates lead to an early risk that at least 10% of wells are in violation of the compliance level—after one to two decades or less for $Q_{out} = 6,000$ m$^3$/day as opposed to nearly eight decades for $Q_{out} = 750$ m$^3$/day, at $d_{ts} = 100$ m (Figure 10). In contrast, if the recharge concentration is much less than 50% above the compliance level $c_{crit}^* \gg 0.5$, high pumping rates delay the time at which any fraction of wells exceeds the compliance level, relative to the lower pumping case. This is consistent with high production wells accelerating transport in short, fast travel paths connected to the top of the screen relative to wells with lower pumping rates. But high production wells also tap into deeper groundwater that is reached by pollution at a much later time than in wells with less pumping and, hence, shorter screens. This makes high production wells more vulnerable to pollutants that exceed compliance levels even after...
Figure 10. Probability to exceed a given normalized well concentration of contaminant at a given time for the four pumping rates tested. The depth of the screen top is fixed to 100 m.

significant dilution in the well screen (low $c^*_{crit}$) but more resilient to pollutants that are easily diluted in the well screen to below compliance levels (high $c^*_{crit}$).

Increasing the pumping rate narrows the distribution of possible $c^*$ values observed across a group of wells. Hence, at any given compliance level, higher pumping rates reduce the variability in time until a concentration threshold is exceeded across a network of supply wells. At a given time, the variability in concentration across a network of wells is smaller for high production wells than for low production wells.

Depth to the top of the screen also significantly affects the time when and how many wells become noncompliant (Figure 11). Regardless of compliance level, exceedances are observed earlier in shallow wells than in deeper wells. This applies consistently, whether the interest is in the time when a small fraction, half, or most of wells are out of compliance. When recharge concentrations exceed compliance levels by a factor of 2 ($c^*_{crit} < 0.5$), the time when the first 10% of wells are out of compliance until the time when 90% of wells are out of compliance is much shorter in wells with less distance to top of the screen, even if screen length is identical. At $d_s = 50$ m, that time span is several decades to a century, while it is in excess of a century for $d_s = 150$ m (Figure 11). This reflects the macrodispersive effect of the heterogeneous aquifer (van Leeuwen et al., 1998).

6. Concluding Remarks

This study characterizes the impact of key well design aspects on the transport of nonreactive pollutants (e.g., nitrate and salt) originating from a regionally extensive NPS overlying a highly heterogeneous, alluvial aquifer system that is several hundred meters thick. Results highlight the significant role that extraction intensity, screen length, and depth to the top of well screen play on pollutant transport relative to the propagation of uncertainty about aquifer hydraulic properties onto the prediction of breakthrough curves, source area (capture zone), and future supply well concentration.

The work also encompasses, for the first time, a range of practically relevant case studies for the general case of semiconfined, heterogeneous aquifer systems with both a regional horizontal gradient and a significant regional vertical gradient. The latter is created by significant regional recharge and regional deeper aquifer
pumping, as is commonly encountered in semiarid, irrigated agricultural regions overlying unconsolidated sedimentary aquifer systems.

Simplified benchmark cases that can be easily computed for arbitrary aquifer systems serve to show the effect of pumping rate, well screen length, and depth to the top of the well screen the variability of pollutant travel times within the pumped water due to mixing on the location of the source area (capture zone) and on the lateral and transverse extent of the source area. The benchmark cases are for an equivalent homogeneous aquifer system (identical regional horizontal and vertical groundwater flux to the heterogeneous system) with and without pumping and do not consider the effect of local dispersion. In the benchmark cases, the travel time distribution encountered by a well without significant pumping, for example, a monitoring well or domestic well, spans from about one decade to a century if the screen length were the same as the average well screen length encountered in wells pumping 750 and 6,000 m$^3$/day, respectively. Not accounting for local dispersion or aquifer heterogeneity, breakthrough of pollutants is first observed after about 80 to about 200 years for depth to top of well screens ranging from 50 to 150 m, respectively. This highlights the significant groundwater mixing along the well screen, independent of aquifer heterogeneity. The closest point of the source area, without pumping, would be located from over 2 km to over 7 km upgradient of the benchmark nonpumping well.

Aquifer heterogeneity imparts significant uncertainty to the travel time distribution in wells, the location of the source area, and—from a regional perspective—adds significant variability in concentration across wells of similar construction and pumping rate. The travel time of the youngest water within a well screen varies by several decades, while the oldest water in a well may vary over two centuries, within the same aquifer system for similar construction and source exposure. CVs of travel time pdfs vary between 15% and 50%. The stochastic source area—the area possibly contributing pollutants to a well—is consistently 3 to 4 times wider and 2 to 3 times as long as the source area in the benchmark case with a pumping well. The benchmark case for a given pumping rate appears to provide an important practical tool, once modified for

Figure 11. Probability to exceed a given normalized well concentration of contaminant at a given time for the three depths of the top of the screen tested. The pumping rate is fixed to 3,000.0 m$^3$/day.
the generally larger area of the stochastic source area. Travel times linearly scale with distance to the well screen. Relative uncertainty decreases with travel distance, as would be expected, given the fixed heterogeneity scale of the aquifer system considered. Overall, the effect of aquifer heterogeneity on the CV of travel time, the spread of the stochastic source area, and concentration prediction uncertainty are rather limited, however, considering the exceedingly high variability of the aquifer materials (the variance in \( \ln K \) is 15).

High pumping rates decrease the youngest travel times by exerting a steeper vertical gradient on the flow field and more strongly tapping into well-connected fast flow paths. But high pumping rates have less impact on medium travel times, which are mostly controlled by regional flow conditions. In contrast, in wells with higher pumping rate the oldest travel times are mostly controlled by the larger screen lengths, leading to typically much larger travel times than in wells with less pumping. Overall, intense pumping is shown to suppress the propagation of uncertainty, and thus variability between wells, because of the large amount of mixing in their long well screens and the long travel paths taken by much of the pollution in these wells.

High pumping rates result in only slightly wider stochastic source areas when compared to the lower pumping rate. For high pumping rates, the critical zone in the center of the stochastic source area is therefore characterized by much higher probabilities for pollutants to reach the well without affecting the relative uncertainty on the time required for a pollutant to reach a well from a certain location.

Deeper screens lead to longer travel times and increased uncertainty, especially about the early and medium arrival times. The stochastic source area is further upgradient and more disperse, while any specific location within the source area is less likely to contribute to the well.

Overall, the effects of heterogeneity on NPS management metrics are significant but limited. It is most diminished in wells with relatively shallow but long screens and high pumping rates, where mixing rather than the macrodispersive effects of the heterogeneous medium dominates the pollutant breakthrough in the well that results from aerially extensive, continuous NPS pollution.

The results shown here are for aquifer heterogeneity typical of alluvial systems, where the variance of \( \ln K \) is larger than 10 and mean heterogeneity lengths are on the same scale or at most 1 order of magnitude shorter than the benchmark source area length for wells with screen lengths typical of most drinking and irrigation water supply wells (not small domestic wells). The results appear robust, and overall trends are similar to those found for confined and leaky confined aquifers with much smaller log\( K \) variability (Feyen et al., 2001; van Leeuwen et al., 1998). While precise predictions of variability and uncertainty will depend on local aquifer structure, well design, and NPS pollution levels, the examples provided here provide a useful assessment tool of mean travel time, travel time variability, source area location, and future pollutant level in wells. Results can be employed elsewhere after accounting for differences in well design and regional flow rates by considering the NPS management metric behavior relative to the benchmark cases shown here.

Future research is needed to understand the joint effect of NPS source strength variability and uncertainty and aquifer heterogeneity on the NPS management metrics discussed here. For reactive contaminants, the influence of spatiotemporally variable sorption and degradation on NPS management metrics also remains unexplored under the conditions considered here, for example, for the transport of pesticides.

Acknowledgments
The authors gratefully acknowledge the financial support through the California State Water Resources Control Board (SWB), Agreement 15-062-250. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the SWB. The online resources located at the GitHub (see https://github.com/chrishenri/NPS_stochastic_assessment) include the Matlab scripts necessary to reproduce the results.

References
Aarts, H. F. M. (2003). Strategies to meet requirements of the EU-nitrate directive on intensive dairy farms. In Proceedings No. 518 (pp. 28). York, UK: International Fertiliser Society.
Andricevic, R., & Cvetkovic, V. (1996). Evaluation of risk from contaminants migrating by groundwater. Water Resources Research, 32, 611–621.
Barlow, P. M., Leake, S. A., & Fienen, M. N. (2018). Capture versus capture zones: Clarifying terminology related to sources of water to wells. Groundwater, 56, 694–704.
Benson, D. A., Aquino, T., Bolster, D., Engdahl, N., Henri, C. V., & Fernández-Garcia, D. (2017). A comparison of Eulerian and Lagrangian transport and non-linear reaction algorithms. Advances in Water Resources, 99, 15–37.
Boso, F., Bellin, A., & Dumbser, M. (2013). Numerical simulations of solute transport in highly heterogeneous formations: A comparison of alternative numerical schemes. Advances in Water Resources, 52, 178–189.
Burrow, K. R., Nolan, B. T., Rupert, M. G., & Dubrovsky, N. M. (2010). Nitrate in groundwater of the United States, 1991–2003. Environmental Science & Technology, 44, 4988–4997.
Canada, H. E., Harter, T., Honeycutt, K., Jessoe, K., Jenkins, M. W., & Lund, J. R. (2012). Regulatory and funding options for nitrate groundwater contamination. (Technical Report 8 in Addressing Nitrate in California’s Drinking Water with a Focus on Tulare Lake Basin and Salinas Valley Groundwater. Report for the State Water Resources Control Board Report to the Legislature). Davis: Center for Watershed Sciences, University of California.
Carle, S. F. (1999). TProGS: Transition probability geostatistical software, version 2.1. Dep. of Land. Air and Water Resources, Univ. of Calif., Davis.

Carle, S. F., & Fogg, G. E. (1996). Transition probability-based indicator geostatistics. *Mathematics in Geology*, 28(4), 453–477.

Carle, S. F., & Fogg, G. E. (1997). Modeling spatial variability with one-and multi-dimensional continuous Markov chains. *Mathematics in Geology*, 29(7), 891–917.

Central Valley Water Board (2016). Central Valley Region Salt and Nitrate Management Plan, Final Document for Central Valley Water Board Consideration.

Cvetkovic, V., Shapiro, A. M., & Dagan, G. (1992). A solute flux approach to transport in heterogeneous formations: 2. Uncertainty analysis. *Water Resources Research*, 28, 1377–1388.

Dagan, G. (1989). *Flow and transport in porous formations*. New York: Springer.

Dagan, G., & Nguyen, V (1989). A comparison of travel time and concentration approaches to modeling transport by groundwater. *Journal of Hydrology*, 4, 79–91.

de Barros, F. P. J., Guadagnini, A., Fernández-García, D., Riva, M., & Sanchez-Vila, X. (2013). Controlling scaling metrics for improved characterization of well-head protection regions. *Journal of Hydrology*, 494, 107–115.

de Barros, F. P. J., & Nowak, W. (2010). On the link between contaminant source release conditions and plume prediction uncertainty. *Journal of Contaminant Hydrology*, 116(1–4), 24–34.

de Barros, F. P. J., & Rubin, Y. (2008). A risk-driven approach for subsurface site characterization. *Water Resources Research*, 44, W01414. doi:10.1029/2007WR006081

de Marsily, G., Delay, F., Gonçalves, J., Renard, P., Teles, V., & Violette, S. (2005). Dealing with spatial heterogeneity. *Journal of Hydrology*, 316, 161–183.

Döll, P., Müller Schmied, H., Schuh, C., Portmann, F. T., & Eicker, A. (2014). Global-scale assessment of groundwater depletion and related groundwater abstractions: Combining hydrological modeling with information from well observations and GRACE satellites. *Water Resources Research*, 50, 5698–5720. doi:10.1002/2014WR015595

Faust, C. C. (Ed.). (2009). Groundwater availability of the Central Valley aquifer (U.S. Geological Survey Professional Paper, 1766). California. Reston (VA): U.S. Geological Survey.

Fernández-García, D., Illangasekare, T. H., & Rajaram, H. (2005). Differences in the scale-dependence of dispersivity estimated from temporal and spatial moments in chemically and physically heterogeneous porous media. *Advances in Water Resources*, 28(7), 745–759.

Feyen, L., Beven, K. J., De Smedt, F., & Freer, J. (2001). Stochastic capture zone delineation within the generalized likelihood uncertainty estimation methodology: Conditioning on head observations. *Water Resources Research*, 37(3), 625–638.

Fleckenstein, J. H., Niswonger, R. G., & Fogg, G. E. (2006). River-aquifer interactions, geologic heterogeneity, and low-flow management. *Groundwater*, 44(6), 837–852.

Fogg, G. E., & Zhang, Y. (2016). Debates—Stochastic subsurface hydrology from theory to practice: A geologic perspective. *Water Resources Research*, 52, 9235–9245. https://doi.org/10.1002/2016WR019699

Franzetti, S., & Guadagnini, A. (1996). Probabilistic estimation of well catchments in heterogeneous aquifers. *Journal of Hydrology*, 174(1), 149–171.

Graham, W. D., & McLaughlin, D. B. (1991). A stochastic model of solute transport in groundwater: Application to the Borden, Ontario, tracer test. *Water Resources Research*, 27(6), 1345–1359.

Guo, Z., & Brusseau, M. L. (2017). The impact of well-field configuration and permeability heterogeneity on contaminant mass removal and plume persistence. *Journal of Hazardous Materials*, 333, 109–115.

Gurdak, J. J., & Qi, S. L. (2012). Vulnerability of recently recharged groundwater in principle aquifers of the United States to nitrate contamination. *Environmental Science & Technology*, 46(11), 6004–6012.

Harbaugh, A. W., Banta, E. R., Hill, M. C., & McDonald, M. G. (2000). MODFLOW-2000, the U.S. Geological Survey modular ground-water model user guide to modularization concepts and the ground-water flow process, (Open File Rep., 00–92, 121 pp.). Reston, VA.

Harter, T., Davis, H., Mathews, M., & Meyer, R. (2002). Shallow groundwater quality on dairy farms with irrigated forage crops. *Journal of Contaminant Hydrology*, 2002(55), 287–315.

Harter, T., Duzrell, K., Kourakos, G., Hollandier, A., Bell, A., Santos, N., et al. (2017). Nitrogen fertilizer loading to groundwater in the Central Valley (Final Report to the Fertilizer Research Education Program, Projects 110301 and 150454) (333 pp.). Davis, CA: California Department of Food and Agriculture and University of California Davis.

Health Canada (2013). Guidelines for Canadian drinking water quality: Guideline technical document nitrate and nitrite. Water and air quality bureau, healthy environments and consumer safety branch, health Canada, Ottawa, Ontario. (Catalogue No H144-13-2/2013EIPDF).

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

Henri, C. V., & Fernández-García, D. (2015). A random walk solution for modeling solute transport with network reactions and multi-rate mass transfer in heterogeneous systems: Impact of biofilms. *Advances in Water Resources*, 86, 119–132.

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

Henri, C. V., Fernández-García, D., & de Barros, F. P. J. (2016). Assessing the joint impact of DNAPL source zone behavior and degradation products on the probabilistic characterization of human health risk. *Advances in Water Resources*, 88, 124–138.

Howden, N. J. K., Burt, T. P., Worrall, F., Mathias, S., & Whelan, M. J. (2011). Nitrate pollution in intensively farmed regions: What are the prospects for sustaining high-quality groundwater? *Water Resources Research*, 47, W00L02. https://doi.org/10.1029/2011WR010843

Kitanidis, P. K. (1997). *Introduction to geostatistics: Applications in hydrogeology*. Cambridge, UK: Cambridge University Press.

Konikow, L. F., Hornberger, G. Z., Halford, K. J., & Hanson, R. T. (2009). Revised multi-node well (MNW2) package for MODFLOW ground-water flow model. *US Geological Survey Techniques and Methods, 6-A30, 67.

Kourakos, G., Klein, F., Cortis, A., & Harter, T. (2012). A groundwater nonpoint source pollution modeling framework to evaluate long-term dynamics of pollutant exceedance probabilities in wells and other discharge locations. *Water Resources Research*, 48, W00L13. https://doi.org/10.1029/2011WR010813

LaBolle, E. M. (1999). Theory and simulation of diffusion processes in porous media (Ph.D. dissertation), 202 pp.; Univ. of Calif., Davis, CA.

LaBolle, E. M., & Fogg, G. E. (2001). Role of molecular diffusion in contaminant migration and recovery in an alluvial aquifer system. *Transport in Porous Media*, 42, 155–179.
Libera, A., de Barros, F. P. J., & Guadagnini, A. (2017). Influence of pumping operational schedule on solute concentrations at a well in randomly heterogeneous aquifers. *Journal of Hydrology*, 546, 490–502.

Maxwell, R. M., & Kastenberg, W. E. (1999). Stochastic environmental risk analysis: An integrated methodology for predicting cancer risk from contaminated groundwater. *Stochastic Environmental Research and Risk Assessment*, 13, 1–2.

Ministry of Health (2008). Drinking-water standards for New Zealand 2005. revised 2008. Wellington: Ministry of Health.

Ransom, K. M., Nolan, B. T., Traun, J. A., Faunt, C. C., Bell, A. M., Gromberg, J. M., et al. (2017). A hybrid machine learning model to predict and visualize nitrate concentration throughout the Central Valley aquifer, California, USA. *Science of The Total Environment*, 601-602, 1160–1172.

Riva, M., Guadagnini, A., & Ballio, F. (1999). Time-related capture zones for radial flow in two dimensional randomly heterogeneous media. *Stochastic Environmental Research and Risk Assessment*, 13(3), 217–230.

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

Rushton, K. R., & Redshaw, S. C. (1979). Seepage and groundwater flow—Numerical analysis by analogue and digital methods.

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

Salamon, P., Fernández-Garcia, D., & Gómez-Hernández, J. J. (2007). Modeling tracer transport at the MADE site: The importance of heterogeneity. *Water Resources Research*, 43, W08404. https://doi.org/10.1029/2006WR005522

Siirila-Woodburn, E. R., Navarre-Stitchler, A. K., Maxwell, R. M., & McCray, J. E. (2012). A quantitative methodology to assess the risks to human health from CO2 leakage into groundwater. *Advances in Water Resources*, 36, 146–164.

Stauffer, F., Attinger, S., Zimmermann, S., & Kinzelbach, W. (2002). Uncertainty estimation of well catchments in heterogeneous aquifers. *Water Resources Research*, 38(11), 1238. https://doi.org/10.1029/2001WR000819

Sutton, M. A., Howard, C. M., Erisman, J. W., Billen, G., Bleecker, A., Grennfelt, P., et al. (2011). The European nitrogen assessment: Sources, effects, and policy perspectives. Cambridge, UK: Cambridge University Press.

Trolldborg, M., Nowak, W., Tuxen, N., Bjerg, P. L., Helmig, R., & Binning, P. J. (2010). Uncertainty estimation of mass discharge estimates from a contaminated site using a fully Bayesian framework. *Water Resources Research*, 46, W12552. https://doi.org/10.1029/2010WR009227

van Leeuwen, M., Butler, A., te Stroet, C., & Tompkins, J. (2000). Stochastic determination of well capture zones conditioned on regular grids of transmissivity measurements. *Water Resources Research*, 36(4), 949–957.

van Leeuwen, M., te Stroet, C. B. M., Butler, A. P., & Tompkins, J. A. (1998). Stochastic determination of well capture zones. *Water Resources Research*, 34(9), 2215–2223.

Weissmann, G. S., Zhang, Y., LaBolle, E. M., & Fogg, G. E. (2002). Dispersion of groundwater age in an alluvial aquifer system. *Water Resources Research*, 38(10), 1198. https://doi.org/10.1029/2001WR000907

Wen, X. H., & Gomez-Hernandez, J. J. (1997). Numerical modeling of macrodispersion in heterogeneous media: A comparison of multi-Gaussian and non-multi-Gaussian models. *Journal of Hydrology*, 190, 129–156.

Zektser, I. S., & Everett, L. G. (Eds.). (2004). *Groundwater resources of the world and their use, Ser. Groundwater* (Vol. 6, pp. 342). Paris: Int. Hydrol. Programme, U. N. Educ., Sci., and Cultural Organ.

Zhang, H. (2006). Role of heterogeneity in flow and solute transport - Case study in Western San Joaquin Valley (Master's thesis), University of California, Davis.

Zhang, H., Harter, T., & Sivakumar, B. (2006). Nonpoint source solute transport normal to aquifer bedding in heterogeneous, Markov chain random fields. *Water Resources Research*, 42, W06403. https://doi.org/10.1029/2004WR003808

Zheng, C. (1994). Analysis of particle tracking errors associated with spatial discretization. *Groundwater*, 32(5), 821–828.