



Stochastic assessment of the effect of land-use change on nonpoint source-driven groundwater quality using an efficient scaling approach

Christopher V. Henri¹ · Thomas Harter¹ · Efstathios Diamantopoulos²

Accepted: 28 August 2020 / Published online: 11 September 2020
© The Author(s) 2020

Abstract

Assessing the impact of a land-use change (LUC) or change in land-use management on nonpoint source-driven groundwater quality in heterogeneous aquifers requires complex analysis. Stochastic methods have been used to account for prediction uncertainty but at high computational cost, which significantly limits the application of these approaches. As an efficient alternative, this study evaluates the application of a meta-analytical solution for evaluating the change in contaminant breakthrough curves at extraction wells in response to LUC. The solution uses the concentration percentiles from a reference stochastic simulation of water flow and solute transport in a groundwater system, assuming a reference land-use distribution pattern. Reference land-use controls the spatially variable rates of both, recharge and contaminant mass loading. The effect of a LUC is evaluated by scaling the ratio between the reference and the new (post-LUC) average input concentrations. The validity of the proposed meta-analysis tool is tested by comparing the results of the meta-analytical solution with those from a full stochastic simulation of the post-LUC scenario. Simulation results show that the accuracy of the meta-analytical solution is best when the regional average recharge rates for both pre- and post-LUC remain approximately unchanged, for any change in contaminant mass loading. Results also indicate that changes in spatial variability and pattern of the recharge rate do not significantly impact the flow field, travel times, and resulting concentrations, if the magnitude of local recharge remains about the same. Lastly, the results show large variability among wells of (and—for an individual well—uncertainty about) the time lag between the time of LUC and the time of consequential effective change in concentrations across wells in the affected region, captured here using statistical metrics.

Keywords Land-use change · Stochastic analysis · Non-point source · Heterogeneity

1 Introduction

The degradation of groundwater quality by large scale agricultural non-point source (NPS) contamination is becoming an environmental and economic concern worldwide (Nolan et al. 2002; Zektser and Everett 2019;

Chen et al. 2005; Sinha and Elango 2019). Implementing classic remediation actions at the agricultural basin scale is rather challenging and often impractical (King and Jensen 2012). Therefore, assessing the impact of proactive measures (e.g., source control) is crucial for regulators, although very complex in practice. Potential measures for reducing the environmental impact of food production on water bodies include practices focused on reducing leaching of agrochemicals to the aquifer, e.g., through land-use changes (LUCs) from high to low agrochemicals-demanding crops or through more efficient practices (Molénat and Gascuel-Oudou 2002; Gascuel-Oudou et al. 2009; Yang et al. 2015). However, the effectiveness of such LUCs (including changes in land-use management), relies on the understanding and on the assessment of complex processes which are associated with high predictive uncertainty.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00477-020-01869-y>) contains supplementary material, which is available to authorized users.

✉ Christopher V. Henri
chenri@ucdavis.edu

¹ Department of Land, Air and Water Resources, University of California, Davis, Veihmeyer Hall, Davis, CA 95616, USA

² Department of Plant and Environmental Sciences, University of Copenhagen, Thorvaldsensvej 40, 1871 Copenhagen, Denmark

One of the main challenges to assessing the effectiveness of potential LUCs is the aquifer heterogeneity, the main process controlling the transport/fate/movement of a contaminant in groundwater (Dagan 1984; Cvetkovic et al. 1992; Weissmann et al. 2002) and the most important source of uncertainty in this predictive effort (Rubin 2003). During the last 4 decades, stochastic assessment approaches have been developed to quantify this uncertainty (Maxwell and Kastenbergh 1999; Rubin 2003; Henri et al. 2015). These methods support contaminated aquifer management by enabling stakeholders to take informed decisions based on the statistical characterization of different management metrics such as travel times, contributing area extension and contaminant breakthrough curves (BTCs) at sensitive locations, e.g. wells (Henri and Harter 2019).

While much research has focused on the development of stochastic assessment methods (Rubin 2003), those are only used to a limited degree by decision makers/stakeholders (Rajaram 2016). Available analytical solutions provide general insights, but site-specific applications of the stochastic approach require computationally expensive Monte Carlo simulations (Henri et al. 2015). Despite the potential of such approaches, testing a series of source control strategies through LUC represent a burden that significantly limits the practicability of stochastic methods (Sanchez-Vila and Fernández-García 2016). There is a need for computationally more inexpensive decision-support tools, which account for complex stochastic processes and evaluate the long term changes in groundwater quality after a LUC.

The overall objective of this work is to evaluate the effectiveness of a meta-analytical approach to evaluating LUC impacts on concentration statistics at a series of extraction wells of concern without re-simulating flow and transport. We propose a new, simple and efficient method in a stochastic framework. We first present the governing equations for simulating the transport of nitrate from the soil surface to wells of concern. We then develop a meta-analytical solution and evaluate its performance by comparing results from the meta-analytical solution against an explicit solution. Finally, we illustrate the new approach by assessing the impact of LUC (crop type or practice changes) in a basin within the Central Valley, CA, USA.

2 Methodology

We first review the basic mathematical framework frequently used to stochastically assess groundwater quality. We then develop an efficient meta-analytical approach to quantify the impact of LUC within the stochastic framework.

2.1 Stochastic quantification of well concentration

The groundwater flow condition is governed by:

$$\frac{\partial}{\partial x} \left(K_{xx} \frac{\partial h}{\partial x} \right) + \frac{\partial}{\partial y} \left(K_{yy} \frac{\partial h}{\partial y} \right) + \frac{\partial}{\partial z} \left(K_{zz} \frac{\partial h}{\partial z} \right) + Q'_s = 0, \quad (1)$$

where \mathbf{K} [m d^{-1}] is the hydraulic conductivity tensor, Q'_s is the volumetric flux per unit volume representing sources and sinks of water [$\text{m}^3 \text{d}^{-1}$], and h [m] is the hydraulic head. Groundwater fluxes are described by the Darcy's law:

$$\mathbf{q}(\mathbf{x}) = -\mathbf{K}(\mathbf{x})\nabla h(\mathbf{x}), \quad (2)$$

where \mathbf{q} [m d^{-1}] is the specific discharge.

Given the often highly uncertain spatial distribution of aquifer geological features, the hydraulic conductivity field can be described by a space random function. Subsequently, the spatially variable flow field is described as a random process. Finally, uncertainty propagates from the discharge into the concentration signal at any location of interest fulfilling the advection–dispersion equation:

$$\phi \frac{\partial c}{\partial t} = \nabla \cdot (\phi \mathbf{D} \nabla c) + \nabla \cdot (\mathbf{q}c) + s(x, y, t) \quad (3)$$

where c is the solute concentration [g m^{-3}], ϕ is the porosity [–], \mathbf{D} is the dispersion tensor [$\text{m}^2 \text{d}^{-1}$], and $s(x, y, t)$ [$\text{g m}^{-3} \text{d}^{-1}$] is the source term describing the rate of contaminant mass released from the nonpoint source (located in any x and y at the top of the domain). In this work, we are interested in evaluating concentration signals in extraction wells.

Consequently, describing the aquifer K -field as a random function allows for the description of the temporal evolution of well concentration by probabilistic metrics such as mean, variance, exceedance probabilities, or other statistical moments. In practice, these can be obtained by Monte Carlo simulation of aquifer heterogeneity and corresponding solutions to the flow (2) and transport (3) equations for each aquifer realization. Estimating the implication of a LUC using such classical stochastic framework would require a Monte Carlo simulation of flow and transport (Eqs. 2, 3) for each scenario (Breuer et al. 2006).

2.2 A simple solution for LUC assessment

We assume here that well concentrations resulting from a potential LUC (c_w^{LUC}) can be related to pre-LUC (reference) concentrations at the same well (c_w^{INI}) by the simple expression:

$$c_w^{LUC}(t) = \frac{\sum_{i_c} k_{i_c}^{LUC} \times \bar{c}_{i_c}^{LUC}}{\sum_{i_c} k_{i_c}^{INI} \times \bar{c}_{i_c}^{INI}} \times c_w^{INI}(t), \tag{4}$$

where k_{i_c} is the proportion of the i_c -th crop over the basin (area of crop over area of basin), and \bar{c}_{i_c} is the spatially (i.e., over the basin) and temporally averaged concentration of the NPS chemical of interest associated with the i_c -th crop. In simple terms, we assume that the proportional change in source concentrations (in recharge at the water table) resulting from a LUC scales proportionally with the concentrations due to the initial land use, observed at any well.

Through Eq. 4, a LUC can be defined as a change in the proportion of individual crops over the basin (k_{i_c}) and/or a change in agricultural practices leading to a change in the averaged recharge concentration of the contaminant (\bar{c}_{i_c}).

Equation 4 is only valid for testing a LUC occurring at $t = 0$. To add a subsequent landuse change, i.e., to account for any LUC at a later time t , the temporal evolution of well concentrations can be obtained by superposition of signals from the initial (pre-LUC) and new (post-LUC) conditions:

$$c_w(t) = \int_0^t \dot{c}_w^{INI}(t - \tau) \times \delta^{INI}(\tau) + \int_0^t \dot{c}_w^{LUC}(t - \tau) \times \delta^{LUC}(\tau), \tag{5}$$

after introducing binary step functions specifying which source concentration signal is activated: $\delta^{INI}(t)$ equals 1 when input concentrations at the initial rate is occurring, 0 otherwise; $\delta^{LUC}(t)$ equals 1 when input concentrations after the LUC is occurring, 0 otherwise. The notation \dot{c} indicates that the concentration signals result from an instantaneous injection of mass (pulse) at time equal 0. An illustration of the process and a detailed explanation are shown in Fig. 2.

The solution (5) can be extended to any number of land-use change incidences (n_{LUC}) as:

$$c_w(t) = \int_0^t \dot{c}_{w,INI}(t - \tau) \times \delta_{INI}(\tau) + \sum_{i=1}^{n_{LUC}} \int_0^t \dot{c}_{w,i}^{LUC}(t - \tau) \times \delta_i^{LUC}(\tau), \tag{6}$$

For the sake of simplicity, we develop and illustrate the method for a single LUC. We further assume temporally stepwise source functions, with each value representing the average (source or well) concentration during a time-step (e.g., 1 year), such that the convolution (5) can be expressed in terms of a finite number of discrete time intervals:

$$c_w(t) = \sum_{j=1}^{j=k} \sum_{i=j}^{i=k} \dot{c}_w^{INI}(t_j - t_i) \times \delta^{INI}(t_i) + \sum_{j=1}^{j=k} \sum_{i=j}^{i=k} \dot{c}_w^{LUC}(t_j - t_i) \times \delta^{LUC}(t_i) \tag{7}$$

where i and j are summation indexes.

Within the stochastic framework, well concentrations can be statistically described by their exceedance probability function to provide an estimate of prediction uncertainty. Since a simple linear scaling function operates on concentration signals, a similar scaling operation can be used for the i_p -th percentiles of the concentration exceedance probability function (P_{i_p}), i.e., the signal superposition principle (Eq. 7) can also be used in a stochastic framework by substituting concentration signals by percentiles of concentration exceedance probability functions obtained using:

$$P_{i_p}^{LUC}(t) = \frac{\sum_{i_c} k_{i_c}^{LUC} \times \bar{c}_{i_c}^{LUC}(t)}{\sum_{i_c} k_{i_c}^{INI} \times \bar{c}_{i_c}^{INI}(t)} \times P_{i_p}^{INI}(t). \tag{8}$$

In short, the new method requires (1) prior estimation of well concentration exceedance probability functions due to current (initial) landuse configuration, (\dot{c}_w^{INI}), using explicit Monte Carlo simulations, and (2) prior knowledge of source concentrations (\bar{c}_{i_c}) and crop area proportions within the likely source area (k_{i_c}), both, before (initial) and after LUC. Once this is achieved, the future effect of LUC on the well BTCs can be computed stochastically using (8).

3 Method validation and limitations

The simple approach (8) relies on 2 main assumptions: (1) the spatial variability of leaching concentrations associated with a given crop over a basin does not significantly impact the amount of mass recovered at extraction wells at time t and (2) a LUC does not significantly impact the transport pathways of the contaminant and the related travel times to a well, which allows for scaling of concentrations (Eq. 4) to be sufficient to quantify post-LUC contaminant signals at extraction wells.

The first point is discussed in Henri et al. (2020). Their study shows that the spatial variability of source terms (effective recharge and contaminant leaching rate) has a minor impact on management metrics for NPS contamination such as travel times, capture zones or contaminant BTCs. Yet, we implicitly test this assumption further in the validation approach employed below. The second assumption—that a LUC-induced change in effective recharge does not significantly affect use of (8)—remains to be tested.

We test the validity of (8) by comparing full Monte Carlo simulation results to the analytical solution (8). We chose to test the scenario most likely to deteriorate the solution of the proposed method: we consider a LUC where all crops distributed over the entire probable capture zone (i.e., the source area from where all the solute reaching any extraction well is expected to originate) are changed into a single land-use with an effective recharge rate that differs significantly from the average recharge rate prior to LUC. This quasi homogenization of the land-use and, therefore of the effective recharge rate, has the most influence on the flow field, travel paths and well concentrations (Henri et al. 2020).

The step-by-step validation approach consists of (Fig. 3):

1. the stochastic simulation of flow and transport for a reference case to evaluate probabilistic capture zones and BTCs;
2. changing the land use over the stochastic capture zone (the area where the probability for the contaminant to reach a well is larger than zero);
3. run a new batch of stochastic simulation accounting for the LUC and evaluate the concentration statistics;
4. evaluate the same concentration statistics using the simple expression (Eq. 8) and
5. compare concentration statistics obtained from the explicit stochastic simulations with those from the meta-analytical solution.

3.1 Reference case

As a reference scenario, we used the simulation set-up of Henri et al. (2020), key aspects of which are summarized here. We model a synthetic but realistic NPS contamination of a heterogeneous aquifer from nitrate leaching due to crop fertilization and irrigation return flow. We represent a sub-regional domain of length $L_x = 19,200.0$ m, width $L_y = 6000.0$ m and depth $L_z = 250.0$ m, discretized in 120

$(n_x) \times 60 (n_y) \times 625 (n_z)$ cells (Table 1). Spatial variability of aquifer hydraulic properties is treated stochastically, using a series of equiprobable realizations to evaluate the uncertainty of the capture zone extension and of contaminant BTCs in extraction wells.

3.1.1 Aquifer heterogeneity representation

Heterogeneity in the hydraulic conductivity (K) field is described using a geostatistical approach in order to account for the uncertainty in its representation. We generate a series of realization of the hydrofacies field using the program T-PROGS, which uses the transition probability/Markov chain method (Carle 1999). Details of this approach are described in Carle and Fogg (1996, (1998)). Following previous work on aquifer heterogeneity in the Central Valley (CA, USA) (Weissmann and Fogg 1999a, b), we consider 4 different constitutive materials of the aquifer (or hydrofacies, categories): gravel, sand, muddy sand, and mud. All geostatistical parameters (hydrofacies proportions, mean lengths, K values, facies-to-facies transition probability rates; see aforementioned references for details) are set to be representative of a typical California Central Valley alluvial aquifer system (Weissmann and Fogg 1999a, b) (Tables SM1, SM2). 50 realizations of the K -field are generated (see a subset of realizations in Fig. 1). 50 realizations were shown to be sufficient to converge the lower statistical moments of K and of the apparent velocities (Henri et al. 2020).

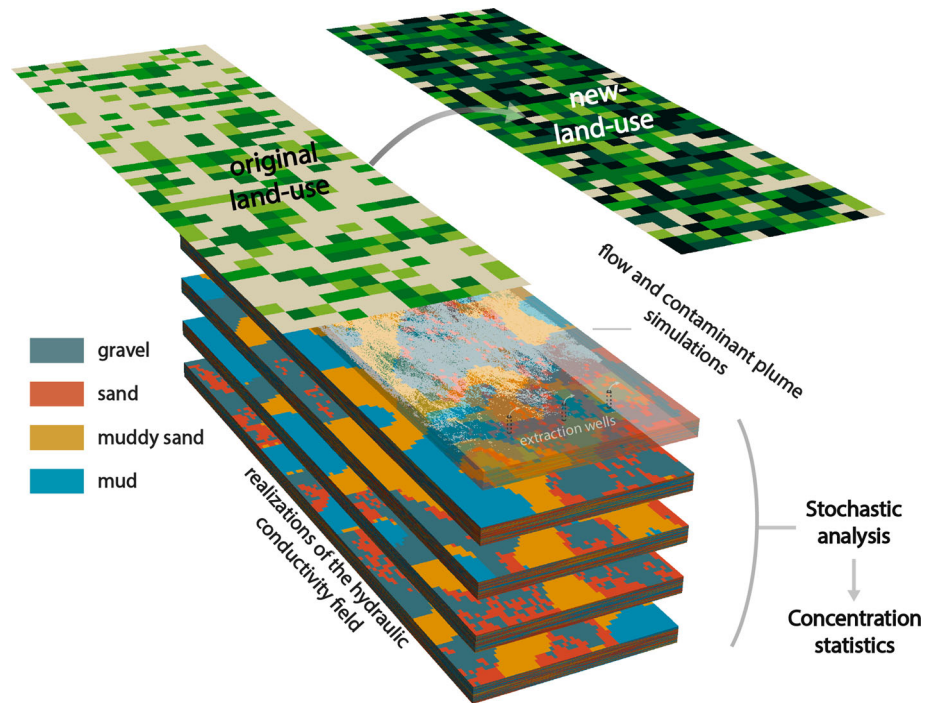
3.1.2 Source term simulation

Crop- and soil-dependent recharge and nitrate leaching rates are estimated through a series of numerical simulations. Water flow in the unsaturated zone is described by 1-D vertical Richards equation (Diamantopoulos and Durner 2012) and contaminant transport by the advection–dispersion equation. Hydrus 1D (Šimunek et al. 2016) is used to numerically solve the two equations. Simulations

Table 1 Domain discretization and physical parameters used in all simulations

Parameter	Value
<i>Domain discretization</i>	
Number of cells, $n_x \times n_y \times n_z$	$120 \times 60 \times 625$
Cell dimension, $A_x \times A_y \times A_z$ [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]	$19200.0 \times 6000.0 \times 250.0$
<i>Flow and transport problem</i>	
Porosity, ϕ [–]	0.3
Average longitudinal hydraulic gradient, i_x [–]	1×10^{-3}
Extraction rate, Q_{out} [m ³ d ⁻¹]	3×10^3

Fig. 1 Illustration of a stochastic analysis of the impact of a land use change. After quantifying the land-use map dependent effective amount of water and contaminant leaching into the aquifer, we solve flow and transport in a series of equiprobable realizations of the hydraulic conductivity field. Statistics of contaminant breakthrough curves at a series of extraction wells can then be evaluated



are run for 6 crop types and 3 typical soil profiles, for a total of 18 simulations. We simulated dominant Central Valley crops: almond, citrus, corn, cotton, grain and grapes. The soil profiles accounted for are: gravel/sand (S1), muddy sand (S2) and mud. Resulting average recharge rates and effective nitrate mass fluxes reaching the water table are listed in Table 2. We refer to Henri et al. (2020) for a full description of the methodology used to quantify these crop/soil dependent leakage rates.

For each realization, the soil map is represented by the top layer of the K-field generated by the TPROGS model. The land-use map, i.e., the spatial distribution of crops over the basin, is generated randomly assuming a uniform distribution. The initial land-use (reference case) is composed of 24% almond (*Prunus dulcis*), 24% citrus, 18% corn (*Zea mays*), 12% cotton (*Gossypium* sp.), 12% grain, 10%

grapes (*Vitis* sp., Fig. 4), which represents a typical Central Valley land-use pattern (Harter et al. 2017). Each cell of the top layer of the realization-dependent domain is characterized by a crop type and a soil type. These two determine the effective recharge rate and contaminant leaching rate (Table 2). The resulting spatial distribution of soils, crops, and of corresponding recharge and contaminant leaching rates characterizes the initial and boundary conditions for a given realization of the groundwater flow domain (pre-LUC).

3.1.3 Flow and transport

Flow Steady state 3-dimensional groundwater fluxes are simulated by numerically solving the groundwater flow equation. Water extraction is simulated by implementing 3

Table 2 Recharge rate and nitrate mass flux applied for each crop—soil type combination in three scenarios. Scenarios S1, S2 and S3 represent a soil characterized as “sand”, “muddy sand”, and “mud”, respectively

	Almond	Citrus	Corn	Cotton	Grain	Grapes
r [m/m ² /d]						
S1	2.3×10^{-3}	1.7×10^{-3}	1.5×10^{-3}	1.7×10^{-3}	4.1×10^{-4}	1.8×10^{-3}
S2	1.7×10^{-3}	1.4×10^{-3}	1.3×10^{-3}	1.4×10^{-3}	5.0×10^{-4}	1.4×10^{-3}
S3	2.9×10^{-5}	2.3×10^{-4}	2.6×10^{-4}	2.0×10^{-4}	2.0×10^{-5}	1.5×10^{-4}
m_f [g/m ² /d]						
S1	2.8×10^{-2}	2.5×10^{-2}	2.0×10^{-2}	2.1×10^{-2}	1.4×10^{-2}	4.6×10^{-3}
S2	2.8×10^{-2}	2.1×10^{-2}	1.5×10^{-2}	1.4×10^{-2}	1.8×10^{-2}	4.3×10^{-3}
S3	0.0	1.9×10^{-10}	3.3×10^{-9}	0.0	0.0	0.0

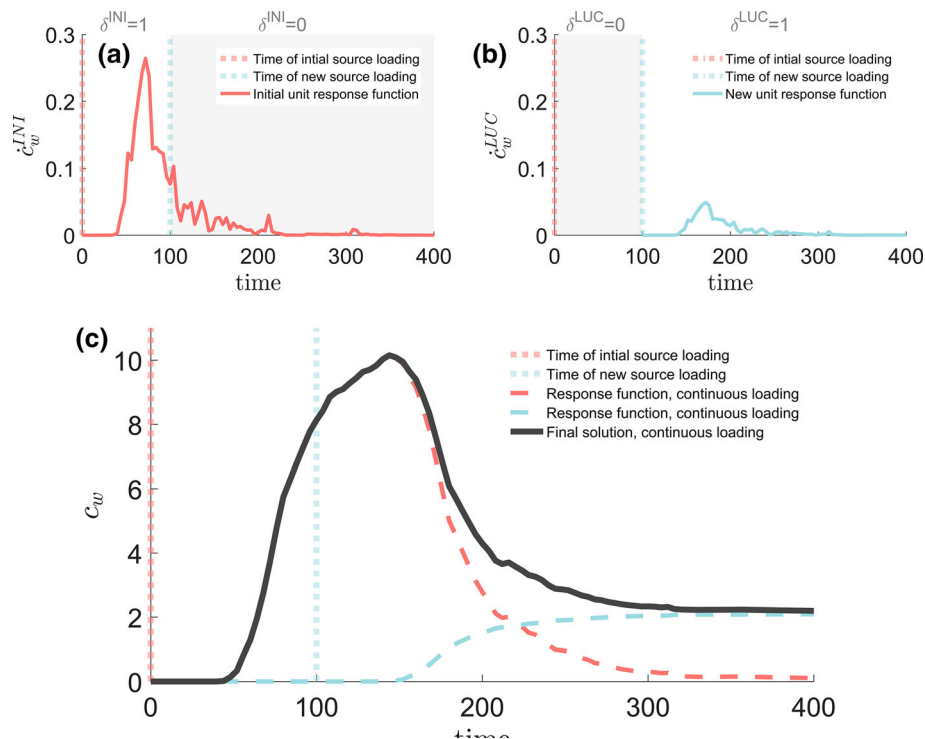


Fig. 2 Illustration of the convolution process used in Eq. 5 for a landuse change occurring at $t = 100$ (grey vertical line). The frame **a** (top left) shows a hypothetical concentration signal from a pulse injection accounting for an initial landuse (c_w^{INI}). The frame **b** (top right) shows a hypothetical concentration signal from a pulse injection accounting for the new landuse (c_w^{LUC}). The convolution of these signals would generate breakthrough curves corresponding to a

constant and continuous injection of mass. In case of a LUC, the convolution is effectively computed only at times when mass injected ($\delta^{INI}=1$ for the initial landuse, $\delta^{LUC}=1$ for the new landuse). The frame **c** shows the resulting 2 signals (dashed lines, red for the initial landuse, blue for the new landuse) and the final well concentration breakthrough curve obtained from their addition (plain black line)

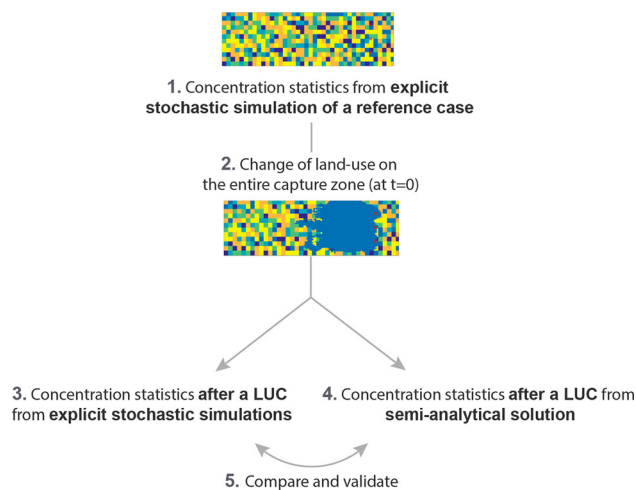


Fig. 3 Flow-chart of the validation approach

extraction wells located on a transverse transect at $x_w = 18,000$ m. The extraction (pumping) rate, Q_{out} is fixed to $3000.0 \text{ m}^3 \text{ d}^{-1}$. The depth of the top of the well screen is set to 100.0 m. In order to sustain the extraction well, a certain amount of conductive material (gravel, sand) needs

to be crossed by the well screen. The length of the well screen is therefore realization dependent. In this study, we consider that 10 cumulative foot (3.05 m) of gravel or sand has to be crossed for every 100 gallon-per-minute ($545.1 \text{ m}^3 \text{ d}^{-1}$) of water extracted.

We impose a longitudinal regional gradient of 1.0×10^{-3} . The vertical flow is controlled by the recharge and by a downward flux from the bottom of the domain simulating water flux toward a deeper part of the aquifer and the potential influence of non-simulated extraction over the sub-basin.

Transport Nitrate transport is simulated accounting for advection only in order to decrease the computational time. As demonstrated by a series of previous studies (LaBolle 1999; LaBolle and Fogg 2001; Weissmann et al. 2002; Henri and Harter 2019), the fine discretization of the velocity field and the explicit simulation of aquifer hydraulic conductivity heterogeneity captures most relevant macro-dispersive characteristics of the transport.

The advective transport is numerically solved using the particle tracking technique implemented in the code RW3D

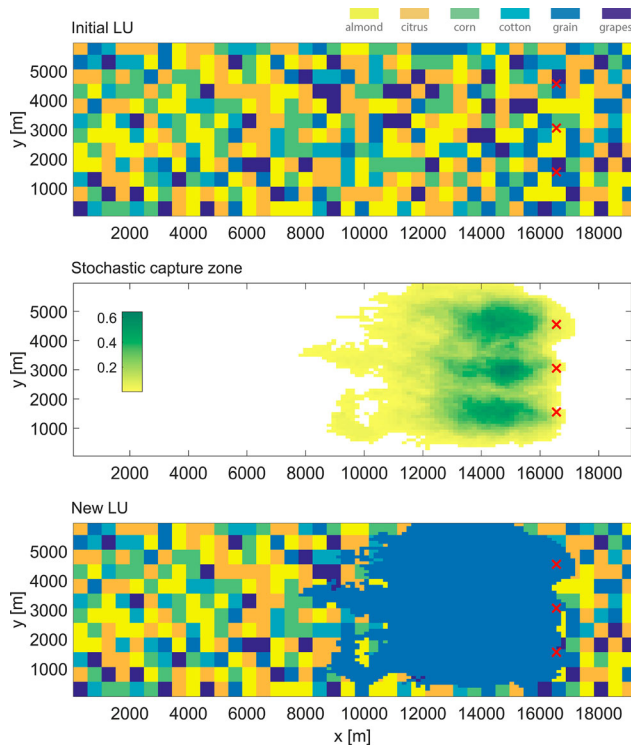


Fig. 4 Initial (top) and post- land-use change (bottom) land-use map. For the latter, all crops located in the stochastic capture zone were changed to grain. The middle frame displays the probability for a particle to leave a given grid cell and to reach an extraction well. The red crosses represent the location of the simulated extraction wells

(Fernàndez-Garcia et al. 2005; Henri and Fernàndez-Garcia 2014, 2015) and expressed as:

$$\mathbf{X}_p(t + \Delta t) = \mathbf{X}_p(t) + \vec{\mathbf{v}}(\mathbf{X}_p(t))\Delta t, \tag{9}$$

where \mathbf{x}_p is the particle position and $\vec{\mathbf{v}}$ is the velocity vector.

For each simulation, 500,000 particles are injected over the entire top of the domain. The local density of particles (i.e., the number injected in each cell) follows the realization dependent spatial variability of the nitrate leaching rate described in Sect. 3.1.2. We simulate a pulse injection of mass (i.e., all mass is injected at $t = 0$). Given the steady state flow conditions, the BTCs from a NPS with continuous release of nitrate at the water table is mathematically obtained by integrating the mass arrival at extraction wells resulting from the pulse input. Well concentration observed at any time in each of the 3 wells for each of the 50 realization (i.e., 150 concentration values) are gathered in order to compute statistics.

3.2 Land use change and models comparison

Simulation outputs from the 50 realizations are also used to determine the stochastic capture zone. The latter is defined

as the surface area from where a contaminant particle has a probability to reach a well that is larger than 0. For the reference scenario, the stochastic capture zone covers an area of about 8 km by 5 km (Fig. 4, middle frame). For the alternative scenario, the land-use over this area is only a single crop (Fig. 4, bottom frame). To test the approximation presented in Eq. 8, the simulation framework described in Sect. 3.1 is repeated with the new land-use map. To better understand the solution performance, we compare the average recharge rate for pre-LUC (\bar{r}^{INI}) and post-LUC (\bar{r}^{LUC}), using:

$$\bar{r}^{INI} = \frac{\sum_{i_c, i_s} r_{i_c, i_s} \times k_{i_c}^{INI} \times p_{i_s}}{n_s \times n_c}; \tag{10}$$

$$\bar{r}^{LUC} = \frac{\sum_{i_c, i_s} r_{i_c, i_s} \times k_{i_c}^{LUC} \times p_{i_s}}{n_s \times n_c} \times S_{cz} + \bar{r}^{INI} \times (1 - S_{cz}), \tag{11}$$

where the index i_c refers to a given crop, p_{i_s} is the proportion of the i_s th soil type (Table SM1), n_s is the number of soil types (3), n_c is the number of crops (6), and S_{cz} is the proportion of the domain covered by the stochastic capture zone (0.33, Fig. 4).

We test two different scenarios. First, the land over the capture zone is converted to grapes only. This LUC significantly lowers the nitrate losses but preserves the mean recharge rate ($\bar{r}^{LUC}/\bar{r}^{INI} = 1.02$; over the capture zone only, the ratio is 1.07). In the second scenario, the area is converted into grains, which leads to a significant reduction of the average effective recharge ($\bar{r}^{LUC}/\bar{r}^{INI} = 0.75$; over the capture zone only, the ratio is 0.28).

Well concentration percentiles obtained from simulations with the original land use (P_i^{INI}) are used to analytically estimate the well concentration percentiles post-LUC (P_i^{LUC}), using the crop proportions of scenario 1 in Eq. 8:

$$k_{i_c}^{INI} = \{0.24; 0.24; 0.28; 0.12; 0.12; 0.10\};$$

$$k_{i_c}^{LUC} = \{0.0; 0.0; 0.0; 0.0; 0.0; 1.0\}$$

and for scenario 2:

$$k_{i_c}^{INI} = \{0.24; 0.24; 0.28; 0.12; 0.12; 0.10\};$$

$$k_{i_c}^{LUC} = \{0.0; 0.0; 0.0; 0.0; 1.0; 0.0\},$$

where $i_c = \{\text{almond; citrus; corn; cotton; grain; grapes}\}$. The average input concentration of a crop i [g m^3] remains unchanged by the LUC, and is obtained as $\bar{c}_{i_c} = \langle m_{f_{i_c, i_s}}/r_{i_c, i_s} \rangle$, where the brackets indicate the average over all i_c soil types, i.e., considering the Table 2:

$$\bar{c}_i^{INI} = \bar{c}_i^{LUC} = \{13.8; 13.9; 11.5; 10.7; 36.4; 2.6\}.$$

4 Results

The performance of the semi-analytical solution is assessed by calculating ϵ_{i_p} , defined as the mean difference over the 400 year period between Monte Carlo simulation concentration and semi-analytical solution concentration representing the i_p -th concentration exceedance probability. If the land covering the stochastic capture zone is changed to grapes, applying the semi-analytical solution produces a mean error of 0.20, 0.087, and 0.067 g/m³ for the 90th (ϵ_{90}), 50th (ϵ_{50}) and 10th (ϵ_{10}) exceedance probability concentrations, respectively (Table 3). In other words, medium and high concentrations associated with least (P_{10}) and median (P_{50}) exceedance probabilities—those typically of most concern—are well reproduced by the semi-analytical approach, with minimal errors (Fig. 5). Low concentrations exceeded by 90% of wells (P_{90}) are overestimated at all times. In a resource management setting, this overestimation may not be significant since wells in the low 10th percentile most likely are below the MCL. An overestimation of the 90th percentile exceedance concentration provides a conservative analysis for decision-making.

Contrary to the first scenario, the “LUC to grain” scenario significantly affects the quality of all semi-analytical solutions, generating errors $\epsilon_{90} = 5.09$, $\epsilon_{50} = 6.63$, and $\epsilon_{10} = 6.83$ g/m³ (Table 3). Relative to simulated values, semi-analytically predicted contaminant mass arrives earlier, and contaminant concentration BTCs associated with the high, median, and low exceedance probabilities are overestimated (Fig. 6). The decay in the accuracy of the semi-analytical approach is due to much lower recharge, which leads to a decrease of the effective vertical velocity. That in turn delays first arrivals of mass and allows for heterogeneity to increase relative macrodispersion—all of which is appropriately captured in the full MC simulation, but not with the semi-analytical approach. If conservative (overly protective) predictions are needed for management purposes, the observed errors may be acceptable as concentrations tend to be over-predicted by the semi-analytical tool, at least in scenarios similar to the one considered here.

The results indicate that the solution (8), as well as the signal superposition method (7), is best applied when the LUC does not significantly modify the average recharge

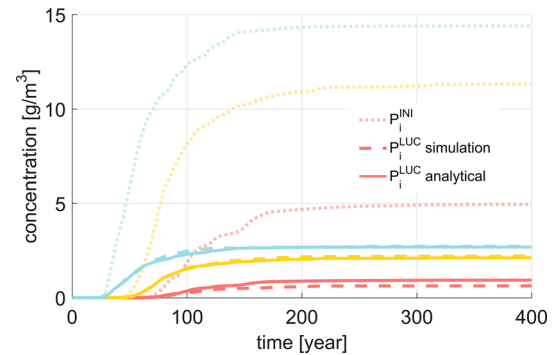


Fig. 5 Concentrations representing the 90th (P_{90} , red), 50th (P_{50} , yellow) and 10th (P_{10} , blue) exceedance probability. Dotted line: initial landuse scenario obtained using Monte Carlo simulation of flow and transport. Dashed lines: “LUC to grapes” scenario (lower nitrate mass loading, similar recharge), with LUC inside the stochastic capture zone only, obtained using Monte Carlo simulation. Solid lines: same scenario as dashed lines, but obtained analytically using Eq. (8)

rate over the LUC area in the basin. Simulations of the effective recharge rate obtained from realistic scenarios of flow and transport in the unsaturated zone indeed indicate that recharge may not vary significantly from one crop to another, especially among permanent crops. Here only r associated with grain significantly differs from the others. In practice, this allows us to employ the analytical approach under a wide range of scenarios of interest. We note that changes in agricultural practices that only lead to a reduction or increase of contaminant mass flux (e.g. a change in the fertilizing practice) but not to a large change in recharge also will allow an application of the semi-analytical solution (8) without particular restriction.

5 Application

This section aims to illustrate the potential of the semi-analytical approach using realistic, but hypothetical conditions. The initial land-use, groundwater flow and transport conditions are similar to the example developed above (Sect. 3.1), representing typical conditions in a Central Valley (CA, USA) agricultural basin. For simplicity, we assume that nitrate leaching began in the middle of the twentieth century, at constant rates representing practices typical for the 1970s and 1980s. Beginning in 2020, we assume a few extreme LUC conditions, changing land-use

Table 3 Mean error produced on concentration statistics by using the meta-analytical solution for different LUC scenarios

Percentile	Mean error; LUC to grapes [g/m ³]	Mean error; LUC to grain [g/m ³]
90th	0.20	5.09
50th	0.087	6.63
10th	0.067	6.83

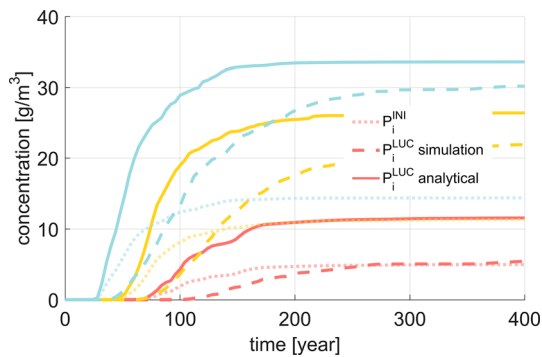


Fig. 6 Concentrations representing the 90th (P_{90} , red), 50th (P_{50} , yellow) and 10th (P_{10} , blue) exceedance probabilities. Dotted line: initial landuse scenario obtained using Monte Carlo simulation of flow and transport. Dashed lines: “LUC to grains” scenario (significant reduction in recharge), with LUC inside the stochastic capture zone only, obtained from the pre-LUC Monte Carlo simulation. Solid lines: same scenario as dashed lines, but obtained semi-analytically using Eq. (8)

uniformly to one of six crops. We apply the scaling of concentration percentiles (8) and the superposition method (7) to obtain results for these conditions.

To test if the semi-analytical solution is applicable, the average recharge pre- and post-LUC are compared (10) and shown on Fig. 7. Only a change of land-use to grain, which has a significantly different average recharge rate ($\bar{r}^{LUC}/\bar{r}^{NI} = 0.76$), cannot be accurately assessed using the proposed solution (Fig. 7). A change of land-use to any of the other 5 crop leads to a minimal change in the average recharge rate (<10%) and is therefore expected to be appropriately assessed by the proposed solution.

Under LUC conditions, nitrate concentrations in recharge from individual crops are similar to or lower than those obtained from the same crop under current conditions. Lower leaching rates are achieved by improvements in nutrient management. Therefore, all scenarios lead to a reduction in well concentrations (Fig. 7) but the reduction is more pronounced for crops with lower input concentrations (grapes, cotton). The reduction in concentration is largest for the highest concentrations, corresponding to low exceedance probabilities (blue lines in Fig. 7).

We observe that the time at which a reduction of concentration associated with a given exceedance probability is observed, i.e., the time to trend reversal, differs for different exceedance probabilities. In our simulation setting, trend reversal occurs very late for wells with the lowest concentrations (exceeded by 90% of wells, P_{90}), after only about 100 years, in the early 2100s. In contrast, wells with the highest concentration, exceeded by only 10% of wells (P_{10}) will see a trend reversal already after about 20 years, around 2040 (Fig. 7). Considering other exceedance probability levels than 90%, 50%, and 10% (red, yellow, and blue lines, respectively), we find that the

time to trend reversal is by far longest for the (low) concentrations with the highest exceedance probabilities, but linearly declines from 70 years for concentrations with exceedance probabilities of 85% to 20 years for (high) concentrations with exceedance probabilities of 5% (Fig. 8).

For a single production well, the results highlight the large uncertainty about the time until future improvements can be seen—improvements in concentrations may occur after already 2 decades or after only 8 decades or more. Alternatively, the stochastic results can be considered representative of the concentration distributions across a large number of wells in a basin with hydrogeologic and land-use conditions similar to those considered here. In that context, our results show that wells of largest concern—those with high concentrations and least exceedance probabilities—are most likely to see improvements at the earliest, after about 2 decades.

This is explained by the fact that the highest well concentrations are observed where source concentrations are not only high but also “connected” to extraction wells by highly conductive channels forming quasi-preferential paths. Under these conditions, a change in source concentration over the effective contributing area is transmitted relatively quickly to a well. This outcome is of significant and promising importance to NPS management: It suggest that the worst nitrate pollution can be most quickly addressed through the right LUC, i.e., the right management practices in agriculture, whether that is a change to crops with lower nitrate loading or a change to improved nutrient management practices. On the other hand, a large percentage of wells—while continuing to increase in concentration—will remain at relatively lower concentrations. The results also indicate that a wide range of outcomes must be expected across a set of regional wells, even under a relatively uniform change in land-use practices.

6 Online tool

The proposed semi-analytical solution allows a computationally inexpensive estimation of basic statistics of stochastic well concentration BTCs resulting from a LUC. We take advantage of this high computational efficiency to develop an online tool for users to test LUC scenarios of their choice. The application, accessible at this link, is coded using R (R Core Team 2019) and Shiny (Chang et al. 2019). In the online tool, users have three tabs “Background”, “Concentration Statistics”, and “Land Use Change”. The “Concentration Statistics Tab” provides infographics on concentration statistics as function of user-defined effective porosity, depth to top of the well screen,

Fig. 7 Concentrations for the 90th (P_{90} , red), 50th (P_{50} , yellow) and 10th (P_{10} , blue) exceedance probabilities. Land-use change (LUC) occurs in year 2020 (vertical gray dashed line). LUC occurs in the entire stochastic capture zone and involves a change to one crop only, as indicated on each panel, resulting in concentration reductions (solid lines), each of which also includes a change in management practices leading to some reduction in nitrate losses when compared to the pre-LUC scenario. A LUC to grain would generate too large a change in the average recharge rate for the semi-analytical solution to be valid (grayed frame). A business as usual scenario, continuing with current landuse conditions is shown in dotted lines

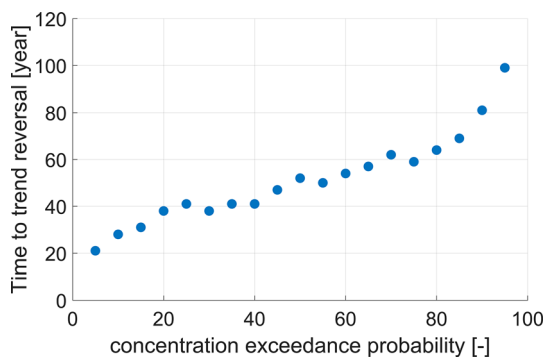
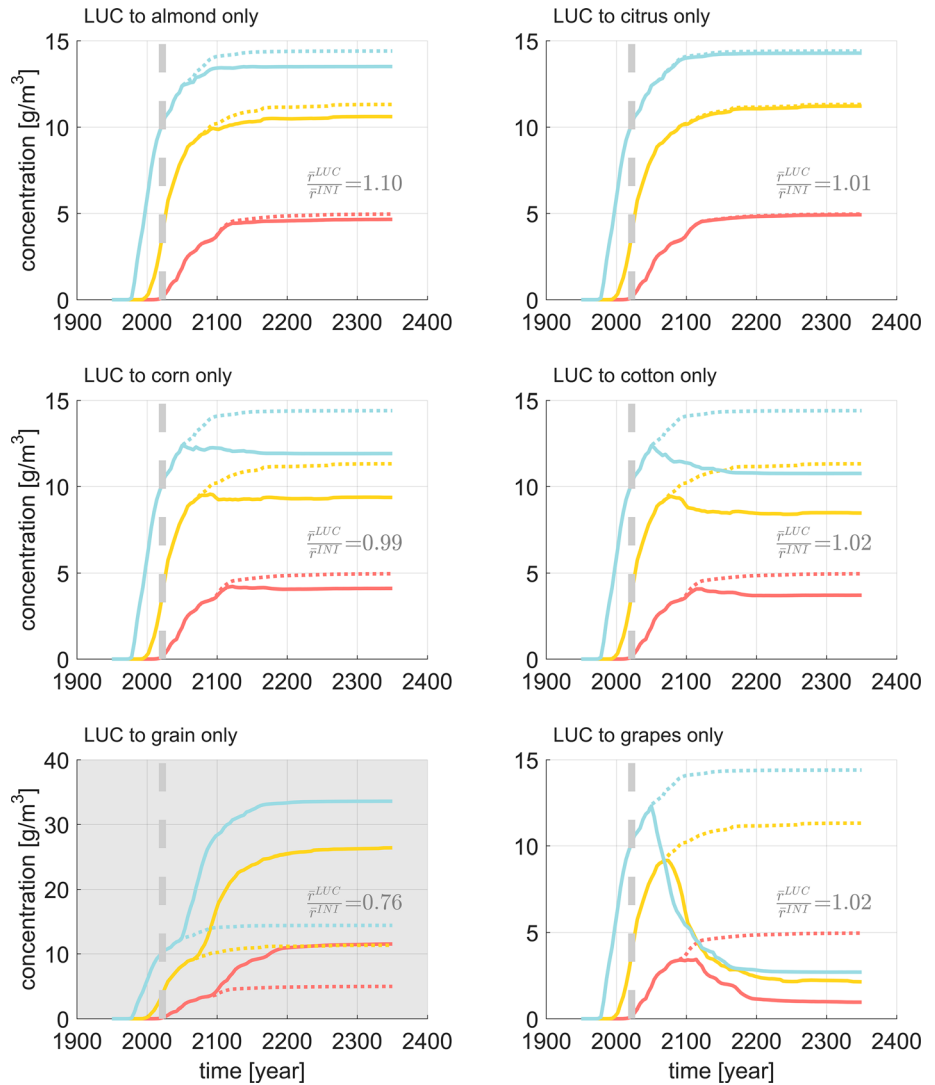


Fig. 8 Time to trend reversal, i.e., the period after a LUC required to observe a reduction of concentration in wells characterized by a given concentration exceedance probability, P_i . Note that higher exceedance probabilities are associated with lower (nitrate) concentrations

and threshold (control) concentration level: Concentration over time is shown for various exceedance probabilities. Concentrations as a function of time for various

exceedance probabilities are also provided in a heat map. And concentrations histograms can be graphed for user-defined future time.

Under the “Land Use Change” tab, the user defines an initial landuse distribution in their region of interest, a recharge concentration associated with each landuse, and the time at which this initial landuse distribution begins. Similarly, the user defines a new landuse distribution in their region of interest, a recharge concentration associated with the new landuses, and the time at which the new landuse occurs. The user also defines the depth to the top of the well screen, the effective porosity, and the exceedance probability of interest. The online application uses the methodology developed in this paper to estimate the concentration evolution resulting from the user-defined LUC, i.e., a change in input concentration of leaching nitrate (c_i^{NI}) and/or proportions (k_i^{NI}) associated to a series of crops. The tool considers the initial well concentration

statistics (c_w^{INT}) as a first LUC by applying the proposed analytical solution (Eq. 8), with the reference scenario described in Sect. 3.1 as the starting point. The user-defined LUC is a second LUC, also applied using our analytical scaling approach. The numerical analysis is instantaneous to the user, with instant graphical results displayed.

In order to provide an indication about the quality of the concentration statistics estimation, the tool systematically estimates the change in the average recharge rate produced by the user-specified initial and new land use and provides feedback on the solution accuracy. When recharge rates change by over 7%, a warning is given to the user, based on the results in this study, where a change in average recharge rate of 7% has allowed for a good performance of the semi-analytical solution (Fig. 5).

Outputs from this tool are, at the moment, only representative of the basin represented in this paper (Central Valley alluvial aquifer system). Codes required to compute flow and transport and to update the online tool for other groundwater basins are provided via an open access repository.

7 Conclusion

Managing non-point source contamination of groundwater requires the assessment of the outcomes from potential remedial actions such as a change in land-use or in land-use practices, while also considering prediction uncertainty due to heterogeneous aquifer conditions. Ideally, this predictive effort would consist of simulating flow and contaminant transport using stable and accurate numerical Monte Carlo solutions in a stochastic framework that can quantify uncertainty. In practice, such an approach is often impossible to apply due to the significant computational burden.

Here, we propose a simple and computationally instantaneous scaling method to estimate the evolution of concentration statistics after a land-use change for situations with limited changes in regional recharge. The approach is implemented into an online tool for illustration and application (in a specific basin) purpose. The method is applicable for many agricultural basins annual as recharge is driven by irrigation and climate, both of which are variable, but within a limited range. Using the semi-analytical tool, we show that the concentrations of most concern (but least exceedance probability) are reversible by LUC much sooner than lower concentrations among an ensemble of wells, suggesting a narrowing of the observed concentration range over time.

Acknowledgements 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 purpose only and does not imply endorsement by the SWB.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Breuer L, Huisman JA, Frede HG (2006) Monte Carlo assessment of uncertainty in the simulated hydrological response to land use change. *Environ Model Assess* 11(3):209. <https://doi.org/10.1007/s10666-006-9051-9>
- Carle SF (1999) TProGS: transition probability geostatistical software. Department of Land, Air and Water Resources, University of California, Davis
- Carle SF, Fogg GE (1996) Transition probability-based indicator geostatistics. *Math Geol* 28(4):453
- Carle SF, Fogg GE (1998) Modeling spatial variability with one-and multi-dimensional continuous Markov chains. *Math Geol* 29(7):891
- Chang W, Cheng J, Allaire J, Xie Y, McPherson J (2019) Shiny: web application framework for R. R package version 1.4.0. <https://CRAN.R-project.org/package=shiny>
- Chen J, Tang C, Sakura Y, Yu J, Fukushima Y (2005) Nitrate pollution from agriculture in different hydrogeological zones of the regional groundwater flow system in the North China Plain. *Hydrogeol J* 13(3):481–492
- Cvetkovic V, Shapiro AM, Dagan G (1992) A solute flux approach to transport in heterogeneous formations: 2. Uncertainty analysis. *Water Resour Res* 28:1377
- Dagan G (1984) Solute transport in heterogeneous porous formations. *J Fluid Mech* 145:151
- Diamantopoulos E, Durner W (2012) Dynamic nonequilibrium of water flow in porous media: a review. *Vadose Zone J*. <https://doi.org/10.2136/vzj2011.0197>
- Fernández-García D, Illangasekare TH, Rajaram H (2005) Differences in the scale-dependence of dispersivity estimated from temporal and spatial moments in chemically and physically heterogeneous porous media. *Adv Water Res* 28(7):745
- Gascuel-Oudou C, Arousseau P, Cordier MO, Durand P, Garcia F, Masson V, Salmon-Monviola J, Tortrat F, Trepos R (2009) A decision-oriented model to evaluate the effect of land use and agricultural management on herbicide contamination in stream water. *Environ Model Softw* 24(12):1433
- Harter T, Dzurella GKK, Hollander A, Bell A, Santos N, Hart Q, King A, Quinn J, Lampinen G, Liptzin D, Rosenstock T, Zhang M, Pettygrove G, Tomich T (2017) Nitrogen fertilizer loading to groundwater in the central valley. Final report to the fertilizer research education program. Technical report. Projects 11-0301 and 15-0454, California Department of Food and Agriculture and University of California Davis

- Henri CV, Fernández-García D (2014) Toward efficiency in heterogeneous multispecies reactive transport modeling: A particle-tracking solution for first-order network reactions. *Water Resour Res* 50(9):7206
- Henri CV, 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. *Adv Water Resour* 86:119
- Henri CV, Harter T (2019) Stochastic assessment of non-point source contamination: joint impact of aquifer heterogeneity and wells characteristics on management metrics. *Water Resour Res* 55(8):6773
- Henri CV, Fernández-García D, de Barros FPJ (2015) Probabilistic human health risk assessment of degradation-related chemical mixtures in heterogeneous aquifers: Risk statistics, hot spots, and preferential channels. *Water Resour Res* 51:4086
- Henri CV, Diamantopoulos E, Harter T (2020) On the conceptual complexity for non-point source management: impact of spatial variability. *Hydrol Earth Syst Sci* 24:1189–1209. <https://doi.org/10.5194/hess-24-1189-2020>
- King GFTHA, Jensen V (2012) Groundwater remediation and management for nitrate. Report for the state water resources control board report to the legislature. Technical report 5 in: addressing nitrate in California's drinking water with a focus on Tulare Lake Basin and Salinas Valley groundwater. Center for Watershed Sciences, University of California, Davis
- LaBolle EM (1999) Theory and simulation of diffusion processes in porous media. Ph.D. dissertation, University of California, Davis, CA
- LaBolle EM, Fogg GE (2001) Role of molecular diffusion in contaminant migration and recovery in an alluvial aquifer system. *Transp Porous Media* 42:155–179
- Maxwell R, Kastenber WW (1999) Stochastic environmental risk analysis: an integrated methodology for predicting cancer risk from contaminated groundwater. *Stoch Environ Res Risk Assess* 13:27
- Molénat J, Gascuel-Oudou C (2002) Modelling flow and nitrate transport in groundwater for the prediction of water travel times and of consequences of land use evolution on water quality. *Hydrol Process* 16(2):479. <https://doi.org/10.1002/hyp.328>
- Nolan BT, Hitt KJ, Ruddy BC (2002) Probability of nitrate contamination of recently recharged groundwaters in the conterminous United States. *Environ Sci Technol* 26(10):2138–2145
- R Core Team (2019) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria
- Rajaram H (2016) Debates-stochastic subsurface hydrology from theory to practice: introduction. *Water Resour Res* 52(12):9215. <https://doi.org/10.1002/2016WR020066>
- Rubin Y (2003) Applied stochastic hydrogeology. Oxford University Press, Oxford
- Sanchez-Vila X, Fernández-García D (2016) Debates-Stochastic subsurface hydrology from theory to practice: why stochastic modeling has not yet permeated into practitioners? *Water Resour Res* 52(12):9246. <https://doi.org/10.1002/2016WR019302>
- Šimunek J, van Genuchten M, Šejna M (2016) Recent developments and applications of the HYDRUS computer software packages. *Vadose Zone J.* <https://doi.org/10.2136/vzj2016.04.0033>
- Sinha SP, Elango L (2019) Deterioration of groundwater quality: implications and management. In: Singh A, Saha D, Tyagi A (eds) *Water governance: challenges and prospects*. Springer water. Springer, Singapore
- Weissmann GS, Fogg GE (1999a) Multi-scale alluvial fan heterogeneity modeled with transition probability geostatistics in a sequence stratigraphic framework. *J Hydrol* 226:48
- Weissmann GS, Fogg GE (1999b) Three-dimensional hydrofacies modeling based on soil surveys and transition probability geostatistics. *Water Resour Res* 35:1761
- Weissmann GS, Zhang Y, LaBolle EM, Fogg GE (2002) Dispersion of groundwater age in an alluvial aquifer system. *Water Resour Res* 38(10):1198
- Yang Z, Zhou Y, Wenninger J, Uhlenbrook S, Wan L (2015) Simulation of groundwater–surface water interactions under different land use scenarios in the Bulang Catchment, Northwest China. *Water* 7(12):5959
- Zektser IS, Everett LG (2019) Groundwater resources of the world and their use. Series on groundwater, vol 6. United Nations Educational, Scientific and Cultural Organization, Paris

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.