Interactive comment on "On the Conceptual Complexity of Non-Point Source Management: Impact of Spatial Variability" by Christopher V. Henri et al.

Editor Decision: Publish subject to minor revisions

Comments to the Author:

The reviews offer detailed points that require consideration, with appropriate clarifications in a revised manuscript.

The authors' responses appear well thought out and convincing, and a revised manuscript reflecting the additional information and clarifications can fully address the reviews. Please proceed with a thorough revision.

I note that some of the author responses (particularly to referee #2) consisted of a detailed commentary, but actual revision to the manuscript was not included. Please reconsider - in most cases, where a detailed response is needed, even to justify/clarify why something was or was not included in the analysis, at least a short additional explanation is helpful in the manuscript itself (if a referee had a question, it's reasonable to assume that other readers may have the same one).

Thank you for your comments and for inviting us to submit a revised manuscript.

Indeed, a few points discussed while answering the referee #2 deserved an additional clarification in the manuscript. We took in consideration your point, and all concerns of both reviewers are now addressed in the new version of the paper. These added sentences are listed in our point-by-point response to the reviews that you can find below. Changes made in the manuscript are listed in green. A marked-up manuscript version is also provided.

Referee #1 Evaluations

Overview

This is an interesting work that brings together many contributions in the field of probabilistic risk assessment (PRA) in aquifers to investigate transport of non-point sources (NPS). The authors explore parameters such as recharge rates and contaminant loadings in the final model output. Furthermore, the authors attempt to reduce the complexity of the model by upscaling a set of spatially/temporally variable quantities, such as the hydraulic conductivity, on the management of NPS. Through the use of numerical simulations, the authors provide an analysis that couples vadose zone, aquifers and land use in a single framework.

I enjoyed reading this paper given that it aims in bringing in elements of stochastic hydrogeology into decision making. The material is well written and organized. The illustrations are clear and well depicted. The referencing is also appropriate although some contributions in the PRA of contaminated aquifers are missing. This is not a big issue. Through the use of scaling arguments (i.e. compliance planes, source sizes etc) the authors make a compelling argument to evoke upscaling for the problem at hand. They claim that due to significant mixing in the compliance plane and the lack of significant variability in NPS solutes, the uncertainty in predictions are reduced thus leading more simplified approaches for modeling such complex systems. Results indicate that the mass arrival time distributions are not that sensitive to the spatial variability of recharge and solute loading whereas some sensitivity is observed for the concentration signal and capture zone estimation. The authors also show that homogenization of the conductivity affects the uncertainty of arrival times.

Thank you for your overall positive evaluation of the manuscript and for your interest. We hope that our answers to your specific comments will clarify your concerns.

Specific comments

• The authors refer to the word ergodicity multiple times. Ergodicity in what? I think they are referring to ergodicity in the transport behavior. If so, provide a quantitative measure of what ergodicity is. For example, the ratio between the source zone dimension and correlation scale needs to be large. If this is the case, then why one would need to quantify uncertainty due to the conductivity field? The spatial statistics is representative of the ensemble statistics. This needs to be better discussed.

See response after the next bullet.

 I am not sure if I missed this in the text but it would be interesting to see if the upscaled dispersion reaches its Fickian limit. Looking at figure SM6, it seems that this is not the case and therefore, transport is still subject to uncertainty. To my understanding, based on the histograms, these upscaled dispersion coefficients reported in figure SM6 are not the ones in the Fickian limit and therefore ergodicity is not attained. So how is it that the authors claim "ergodicity" in this paper?

Thank you for raising your concern about ergodicity in these two comments. The context in which the term was used in the manuscript was, indeed, unclear.

The assumption of ergodicity is here employed NOT in the sense that we assume the NPS plume to be so large that there is no variability in plume moments between realizations. Instead, we here employ the ergodicity hypothesis in the same way as described, e.g., by Rajaram (2002) or Gelhar (1993). There, a single large realization (spatially extending over many correlation scales) of the K field is used to compute sample moments (e.g., of head,

concentration). The ergodic hypothesis is employed to justify comparing these sample moments to the theoretical ensemble moments obtained analytically. But the ensemble moments also characterize the uncertainty about, e.g., head or concentration at a single unmeasured location within the aquifer represented by the simulation.

Analogous, we here use the ergodic hypothesis to postulate that the statistical sampling across the 150 well and 150 source areas simulated is representative of the ensemble moments (the stochastic management metrics) at an individual unmeasured well and its source area in the real-world aquifer represented. The 150 simulated wells and source areas are analogous to the single large realization in Rajaram (2002), since the K field is stationary.

In other words, we must assume ergodicity and stationarity to be able to equate the stochastic NPS management metrics obtained from across 150 samples (wells, source areas) to their ensemble properties. The ensemble properties in turn apply to any (unmeasured) well and its source area in the aquifer system. They characterize the uncertainty about these metrics at each (unmeasured) location (well, source area).

We propose to clarify this point in the manuscript (see below) and also refer to the histogram of the mean and variance of the K fields in supplementary material (Figure SM3, to be updated in the new version of the manuscript and shown below) showing stationarity in the moments of the K field, i.e., a very narrow range of values for the mean and variance, and therefore displaying that structural ergodicity can be assumed.

"Then, stochastic management metrics quantify both, the mean and variability of pollution levels across a large sample of production wells encountered over a basin as well as the expected value and uncertainty about pollution levels at an individual well. This is done by simulating stationary random fields (Figure SM3) and assuming ergodic conditions [e.g., Gelhar, 1993; Rajaram, 2002]." (line 122)

Reference: Rajaram (2002), Perturbation theories for the estimation of macrodispersivities in heterogeneous aquifers, in Rao S. Govindaraju, Stochastic methods in subsurface contamination hydrology, p13-62.

Gelhar, L. W. (1993), Stochastic subsurface hydrology: Englewood Cliffs, New Jersey, Prentice-Hall, 390 p.



Figure SM3: Histogram of the mean (top) and variance (bottom) of the lognormal hydraulic conductivity.

• It would be interesting to see how the conclusions regarding recharge reported in this paper compare with the ones reported in the works of Rubin and Bellin (1994) WRR and Li and Graham (1999) WRR. These authors investigate the impact of recharge and its randomness on travel time pdfs.

Thank you for drawing our attention to these relevant papers. First, it might be interesting to note the differences between our simulation setup and the one used in Rubin and Bellin (1994) and Li and Graham (1999): The two aforementioned papers are (semi-) analytically analyzing 2D transport, with a point source and no pumping. Rubin and Bellin (1994) assumes uniform recharge and found that recharge increases longitudinal plume spreading. Li and Graham (1999) do account for spatial variability in recharge further increases longitudinal spreading, and that uncertainty in the recharge spatial variability increases the uncertainty in solute concentrations.

On the other hand, we here simulate 3D transport, with non-point source transport, and significant pumping. Our results show little effect of recharge spatial variability (correlated to the top of the K-field) on travel times, while homogenization of recharge leads to increased uncertainty in well concentrations.

Results highlight that a potential superficial increase in lateral spreading is less (or non-) relevant in a 3D nonpoint source setting as indicated by our analysis of travel time statistics.

Concerning concentration statistics, the difference in outputs between Li and Graham (1999) and our work could indicate that decorrelated spatial variability of the recharge rate adds uncertainty, while a r-K correlation increases the conditioning of the flow field, which leads to decreased uncertainty. We recall that in our setting, the existence of a r-K correlation is a result of the explicit simulation of water flow in the unsaturated area.

In addition, considering three dimensions, as well as the presence of extraction wells adds some degree of complexity in transport and uncertainty propagation. Fully understanding these processes would require a significant effort of its own.

We propose to discuss these two points in a new version of the manuscript:

"Previously, Li and Graham (1998) investigated the impact of recharge spatial variability in a more theoretical and simplified 2D heterogeneous aquifer contaminated by a point source under non-pumping conditions. The work highlights that spatial variability in recharge increases spreading, especially in the transverse direction. In our 3D NPS setting, transverse spreading is less relevant (Figure 3) and we do not observe the increase in variability." (line 327)

"Li and Graham (1998) stochastically analyze the impact of spatially random recharge rate on transport in a 2D point source setting. Their work concluded that, for those conditions, large variability in – and therefore uncertainty about - recharge increases uncertainty in solute concentration. In our work, we observe the opposite. The difference may be partly due to the 3D non-point source transport, and partly caused by the implicit correlation between the hydraulic conductivity and the recharge rate in our scenarios, which may increase the conditioning of the flow field that leads to the observed decrease of uncertainty relative to the homogenized scenario." (line 401)

• Line 460: "The results here confirm that..., but also put the macro-dispersive process. . .". I could not understand the meaning of this sentence. Please revise its structure. Thanks.

Thank you for pointing it out. This sentence has been changed to:

"The results presented here confirm this observation for the case of non-point source contaminations, but also highlight the generation of a quasi-macro-dispersive process through the (vertical) well mixing process." (line 460)

Referee #2 Evaluations

The present work deals with the feasibility of simplify the conceptual modelling of non- point source (NPS) contamination of an aquifer, in the presence of an active pumping well. The simplification consists of spatial homogenization of otherwise heterogeneous terms (recharge rate and contaminant concentration of the recharging water) and of the aquifer hydraulic properties (conductivity). I liked the general goal/purpose of the study (even though there is no theoretical/technical novelty in the employed method- ology) and its practically oriented nature. Nevertheless, it is my opinion that there are several unclear points and unstained observations that should be addressed before acceptance for HESS. I list my concerns below. I do see a very good paper after addressing these issues/confusions. I also hope that my comments help in making the paper shorter and clearer, it becomes quite hard to follow it from the beginning to the end.

Thank you for your constructive and well-thought-out comments. We appreciate the attention to the details of our work. Please see below our answers (in blue) to your concerns. Proposed additions to, deletions from, or modification of the manuscript are written in green.

Line 5-9: 'On the other hand, concentration levels of some key NPS contaminants (salinity, nitrate) vary within a limited range (<2 orders of magnitude); and significant mixing occurs across the aquifer profile along the most critical compliance surface: drinking water wells with their extended vertical screen length. Here, we investigate, whether these two unique NPS contamination conditions reduce uncertainty such that simplified spatiotemporal representation of recharge and contaminant leakage rates and of hydraulic conductivity are justified when modeling NPS pollution.'

I don't agree with the fact that the Authors tested that the two mentioned conditions imply a reduction in the uncertainty (they do not explore scenarios where these two conditions aren't meet!). I would say that the Authors investigate, under these two peculiar NPS conditions, the possibility of introducing simplifying modelling hypotheses which do or do not affect the ensuing uncertainty about targeted quantity. Please consider revise the sentence, here and throughout the whole paper where needed.

Thank you for making this excellent point. To clarify our approach, we propose to change this sentence to:

"Given these two unique NPS contamination conditions, we here investigate the degree to which NPS travel time to wells and the NPS source area associated with an individual well can be appropriately captured, for practical applications, when spatiotemporally variable recharge, contaminant leakage rates, or hydraulic conductivity are represented through a sub-regionally homogenized parametrization." (line 7)

Line 14-15: 'Surprisingly, regional statistics of well concentration time series are fairly well reproduced by a series of equivalent homogeneous aquifers, highlighting the role of NPS solute mixing along well screens.'

I have two comments here: (i) the term regional statistics is somewhat obscure; (ii) once clarify it in the section 6 (see also my comment 3), it is my understanding that the diverse homogenization here proposed work fine for low (i.e., 0.1) and middle (i.e.,0.5) probability of exceeding a given concentration in the well, while for high probability (i.e., 0.9) the homogenization-based results are unreliable (see e.g., the discrepancy between the curves in Figure 6a and Figure 10a, in Fig. 10a I don't see the 'regional' counterpart (red dashed curve)).

The term regional is not necessary here. We propose to delete that in a revised version. And indeed, the concentrations with a high probability to be exceeded are not well reproduced if K is homogenized. Yet, these lower concentrations are rarely relevant in nonpoint source contamination management. The last sentence of the abstract will be rewritten to clarify this point:

"Surprisingly, the statistics of relevant NPS well concentrations (fast and intermediate travel times) are fairly well reproduced by a series of equivalent homogeneous aquifers, highlighting the role of NPS solute mixing along well screens." (line 14)

Lines 90-92: 'Assuming ergodicity (Dagan, 1990), stochastic management metrics are quantified both, for the pollution variability across an ensemble of production wells encountered over a basin, and for the uncertainty about pollution levels at an individual well.' As a matter of fact the Author verified the validity of the ergodicity principle (here in the sense that the results of one single realization are representative of the results across the whole ensemble of realizations, or, in other words, there is no variability across the ensemble of realizations of the investigated output) only with respect to the soil crop arrangement (see Fig. 1), but ergodicity does not hold with respect to the (either heterogeneous or homogenized) aquifer conductivity (i.e., their results clearly show that there is a variability in the investigated NPS management metrics as the conductivity distribution varies among realizations)! Please clarify this point.

We agree that indeed the paper shows ergodicity w.r.t. to crop coverage, that is, ergodicity in the sense that the crop variability across the source area is sufficiently large that any realization of crop produces the same outcome (for the stochastic management metric).

But we consider ergodicity here also in a different context. The issue has been raised by reviewer #1. So, we here copy our response to reviewer #1 as this should address reviewer #2's concern as well:

Thank you for raising your concern about ergodicity in these two comments. The context in which the term was used in the manuscript was, indeed, unclear.

The assumption of ergodicity is here employed NOT in the sense that we assume the NPS plume to be so large that there is no variability in plume moments between realizations. Instead, we here employ the ergodicity hypothesis in the same way as described, e.g., by Rajaram (2002) or Gelhar (1993). There, a single large realization (spatially extending over many correlation scales) of the K field is used to compute sample moments (e.g., of head, concentration). The ergodic hypothesis is employed to justify comparing these sample moments to the theoretical ensemble moments obtained analytically. But the ensemble moments also characterize the uncertainty about, e.g., head or concentration at a single unmeasured location within the aquifer represented by the simulation.

Analogous, we here use the ergodic hypothesis to postulate that the statistical sampling across the 150 well and 150 source areas simulated is representative of the ensemble moments (the stochastic management metrics) at an individual unmeasured well and its source area in the real-world aquifer represented. The 150 simulated wells and source areas are analogous to the single large realization in Rajaram (2002), since the K field is stationary.

In other words, we must assume ergodicity and stationarity to be able to equate the stochastic NPS management metrics obtained from across 150 samples (wells, source areas) to their ensemble properties. The ensemble properties in turn apply to any

(unmeasured) well and its source area in the aquifer system. They characterize the uncertainty about these metrics at each (unmeasured) location (well, source area).

We also refer to the histogram of the mean and variance of the K fields in supplementary material (Figure SM3, to be updated in the new version of the manuscript and shown below) showing stationarity in the moments of the K field, i.e., a very narrow range of values for the mean and variance, and therefore displaying that structural ergodicity can be assumed.

These points have been clarified in the new version of the manuscript:

"Then, stochastic management metrics quantify both, the mean and variability of pollution levels across a large sample of production wells encountered over a basin as well as the expected value and uncertainty about pollution levels at an individual well. This is done by simulating stationary random fields (Figure SM3) and assuming ergodic conditions [e.g., Gelhar, 1993; Rajaram, 2002]." (line 122)



Figure SM3: Histogram of the mean (top) and variance (bottom) of the lognormal hydraulic conductivity.

Reference: Rajaram (2002), Perturbation theories for the estimation of macrodispersivities in heterogeneous aquifers, in Rao S. Govindaraju, Stochastic methods in subsurface contamination hydrology, p13-62.

Gelhar, L. W. (1993), Stochastic subsurface hydrology: Englewood Cliffs, New Jersey, Prentice-Hall, 390 p.

The Authors named the regional analysis ('pollution variability across an ensemble of production wells encountered over a basin') (see also Section 6) the scenario in which both the conductivity and the land crop usage varies between the Monte Carlo simulations. They referred to the single well analysis ('uncertainty about pollution levels at an individual well') when only the conductivity varies among the MC simulations (with the crop arrangement fixed). To me this distinction is not meaningful, since it implies that in the groundwater basin there are sub-portion (the simulated domain) subjected to the very same boundary conditions (aside from the infiltration rate) and that these sub-portions do not influence each other (they are far away from one another). In my vision the Authors have, given a domain of interest and deterministic initial and boundary conditions (aside from the infiltration rate), conducted the uncertainty analysis (i) for a given (i.e., conditional) to a crop arrangement and (ii) considering the uncertainty in the crop arrangement, proven then that the uncertainty in the latter is not an influential factor (i.e., there is ergodicity w.r.t. to the crop usage). Please consider this aspect, the regional and one-well distinction seems to me confusing and not well supported by the investigated set-up.

Indeed, we assume that there are sub-portions of the aquifer that do not influence each other. The size of the simulation domain was chosen to fully accommodate the source area of three production wells. Since we show that their source areas do not overlap, each well's area of capture can indeed be considered independent. This point is now explicitly mentioned in the manuscript:

"To assess the spatial variability of NPS management metrics across an ensemble of well locations in a groundwater basin, the equiprobable realizations of the aquifer system represent the variety of locations across a basin with geostatistically similar geological features. This is true since the domain is designed to ensure that source areas of the three production wells are fully accommodated and that each well's area of capture can be considered independent. In the case of a regional analysis, land-use is simulated as a random process." (line 120)

The reviewer is also correct that we make a conceptual distinction between the case of assessing the uncertainty about an individual well's source area and contaminant travel time, and the case of assessing the variability of source area and contaminant travel time of many wells in a region with a stationary hydraulic conductivity, recharge, NPS mass loading, and landuse pattern. We argue that this is an important theoretical distinction, for the same reason that the concept of ergodicity is relevant.

From a theoretical standpoint, therefore, generating 50 realizations (each with 3 wells), without changing actual landuse is, in principal, only assessing the uncertainty about recharge, loading, and K under a given spatial configuration of landuse (which would be specific to each of the three wells).

Then also making landuse spatial pattern a stochastic variable is conceptually consistent with the idea of simulating 150 different wells, each with its own spatial landuse arrangement.

But the reviewer is correct in pointing out that there is negligible difference between these two, effectively proving ergodicity w.r.t. landuse. But we do not believe it is useful to say this explicitly (e.g., in a sentence at the end of section 6), because we don't want to confuse the reader with the "other" (valid) interpretation of ergodicity that the reviewer is in fact employing here.

We therefore propose to add the word "pattern" to the line 533, make a small change in line 536

"To adapt the simulation setup to the regional stochastic analysis, the spatial distribution pattern of crops...." (line 533)

"..... interpretation of the stochastic results, the range OF individual travel times, capture zones," (line 536, capitalized word to be inserted)

Line 115-117: 'Assuming ergodicity (Dagan, 1990), stochastic analysis is applied to first quantify uncertainty about pollution outcomes at individual wells and to secondly quantify regional spatial variability in pollution outcomes across an ensemble of wells.' Please revise this sentence according with Comment 2-3. Furthermore, in the way it is written it means that ergodicity is needed in order to quantify uncertainty (this is the case for example in geostatistical approach where spatially distributed measure of conductivity in a field, i.e., in a single realization, are employed to describe the ensemble statistic of conductivity) in this study, i.e., that automatically a single realization is sufficient to describe the behavior of the ensemble, whereas, it is my understanding that the Author did the other way around: prove the validity of ergodicity (for the investigated quantity) w.r.t. to the crop usage. If ergodicity was originally assumed, no need to do many simulations.

See our answer to the previous two comments.

Line 129-130: 'The histograms of the mean and the variance of the logarithm of K are shown in Figure SM3. Fifty realizations were sufficient to converge the lower statistical moments of K and of the resulting mean velocities (Figure SM7)'.

It is not clear at all that the Authors are referring to the spatial mean and spatial variance (this is my impression) evaluated for each field of K, and then doing the histogram of these quantity. Is it so? Why do we care about it? How do the Author prove the convergence of these spatial mean and spatial variance employing 50 realizations? To which values should these quantities converge? I am way more concerned about: are the 50 realizations enough to ensure the convergence of the statistics (e.g., pdf, CDF) of the output quantities of interest (i.e., travel time, breakthrough curves at the well and capture zone)? This is aspect is not investigated at all by the Authors and looking to the high spatial variance of K (from 10 to 18 approximately) I am afraid that 50 realizations are not enough, even if the source is spatially distributed. Please analyze the convergence of the investigated results w.r.t. to the Monte Carlo simulations. Furthermore, regarding the histogram in SM3 and SM4 for the recharge rate and the concentration of pollutant in the recharge, why do we care about them? Please clarify.

As mentioned before, the histogram of the mean and variance of the InK fields were erroneous (they were considering the first layer of the domain only). Please apologize the mistake. Below are the corrected histograms, which provide a better measure of the high degree to which the entire range of K variability is captured in each of the 50 realizations.

The convergence of the apparent velocities and of apparent dispersivities, which are derived from particle plumes, give an indication about convergence of transport metrics. The dataset used to compute statistics is composed of 150 values of well concentrations and of travel times (3 wells per realization). For the sake of clarity, we now also show in Figure SM6 and SM7 (see below) the convergence of the mean and variance of the travel times and of the 50% highest concentrations, respectively. We propose to add the following text to the manuscript:

"Figure SM6 shows satisfactorily the convergence of the mean and variance of t_{50} ." (line 252)

"Figure SM7 shows satisfactorily the convergence of the mean and variance of the concentration exceeded by 50% of wells." (line 259)

The histograms for the recharge rate (SM4) and the input concentrations (SM5) are given in supplementary material in case a reader may have a concern about the range of values used. The following text would be added to the manuscript:

"For indications about the range of values and degree of variability, histograms of the mean and variance of the recharge rates applied in the 50 realizations of heterogeneous cases are shown in Figure SM4." (line 211)

"Histograms of the mean and variance of the initial concentration over the 50 realizations are shown in Figure SM5 if a visualization of the range of values and of the variability is needed." (line 238)



Figure SM6: Convergence of the mean and variance of the time required for 50% of the total recorded mass to reach a well.



Figure SM7: Convergence of the mean and variance of the concentration exceeded by 50% of the wells.

• Line 151-195: I see the detailed description of the estimation of the recharge rate to better fit in an Appendix.

We would be open to moving this section to an appendix and replacing it with a short paragraph describing the core concept, if the other reviewers and associate editor agree.

Line 230-233: 'The detailed discretization of the velocity field described above is capturing the most relevant characteristics affecting the macro-dispersive transport behavior (LaBolle, 1999; LaBolle and Fogg, 2001; Weissmann et al., 2002; Henri and Harter, 2019). Therefore, effects of grid-scale dispersion are assumed to be negligible.' So, does the Author set the tensor D in (10) equal to zero? Please clarify.

Yes, the dispersion tensor is equal to zero. We propose to clarify with the following text:

"Therefore, effects of grid-scale dispersion are assumed to be negligible, i.e., **D**=0 and Eq. 10 is simplified to $x_p(t + \Delta t) = x_p(t) + \Delta t. v(x_p(t))$. This assumption, which potentially impacts NPS management metrics, is further evaluated in Appendix B." (line 233)

Line 244-246: 'Three relevant nonpoint source (NPS) pollution management metrics are considered to measure the stochastic simulation outcomes: the probability distribution of pollutant travel times to wells, the probability distribution of pollutant con- centration in wells, and the probability distribution of source locations.'

These are the 3 quantities of interest, why do you introduce them in the 2.3.1 Pollutant travel times section? Better just before.

Excellent observation. This sentence is now moved where you advised it should be.

Line 250-251: 'Following a stochastic approach, probability density functions (pdfs) of travel times ti are obtained by determining the histogram of ti in 150 simulated wells'. Please note that the histogram is not a pdf, whereas the latter is associated with a continuous variable and the former to a discrete variable. I would limit to say to that the pdf are estimated on top of the 150 simulated wells. Furthermore, at lines 244-246 the probability distribution is mentioned, this is not the pdf. Also, at line 214 the authors say that they analyzed the probability distribution, but then in Figure 3 they depict the pdf. Please check for the consistency of the terminology/results through all the work.

Thank you for point this out. The terminology has been made consistent. For clarification, we do not analyze the *pdf* of concentrations but of travel times only. Also, a histogram can be obtained for both, discrete and continuous variables (e.g., velocity, concentration). It is a binned representation of an (unknown) pdf. Among others, the aforementioned sentence has been changed to:

"Following a stochastic approach, probability density functions (pdfs) of travel times t_i are obtained using time series from 150 simulated wells (50 realizations, each with 3 wells)." (line 250)

• Lines 262-263: 'NPS pollution management may also require the assessment of the effective source area, i.e., the capture zone or contributing area of the pollution observed in a production well.'

Please avoid to use source zone, this is typically used to indicate the area covered by the contaminant at the initial time (regardless if it reach the well or not), the capture zone of the well is way more clear as the wording in my opinion.

This terminology is, indeed, often confusing. Barlow et al. (2018) provide some good thoughts and clarification about the different terms used in the literature under different contexts (water budget, transport, ...). The authors define "capture zone" as "the three-

dimensional, volumetric portion of a groundwater flow field that discharges water to a well". This is not what we show in our work. We analyze the 2D projection of this so-defined capture zone, which would correspond more to a source or contributing area. However, we do understand that both terms (source area and capture zone) has been used in the literature to designate this 2D projection. In order to avoid an eventual confusion, the new manuscript acknowledges this fact by providing the abovementioned reference and by explicitly mentioning that these terms have equivalent meaning throughout the paper:

"Important aspects of NPS pollution are pollutant travel times, the location of well source areas (also known as capture zones; Barlow et al. 2018) to identify specific pollution sources, and the long-term evolution of contaminant levels in and across affected wells and streams." (line 23)

"The stochastic capture zone (or source area) is the area characterized by Pw(x,y)>0." (line 336)

Reference: Barlow, P. M., Leake, S. A. and Fienen, M. N. (2018), Capture Versus Capture Zones: Clarifying Terminology Related to Sources of Water to Wells. Groundwater, 56: 694-704. doi:10.1111/gwat.12661.

Line 275-277: 'The NPS metrics from fully heterogeneous simulations are compared to the NPS metrics obtained from a range of upscaled, homogenized simulations that employ effective homogeneous properties rather than the original heterogeneous distribution of the *K*, *r*, and c0 terms.'

Please consider avoiding the word upscale/upscaling (here and in the whole text), since this inherently implies a change of scale (e.g., from pore to continuum or from continuum to continuum) and it is typically associated with a change of the governing equations used to describe the process (e.g., effective model for the solute transport involving non-local terms) whereas here the Author conducted a simple homogenization (with the arithmetic average as a rule) of the diverse terms.

We, in fact, employ a change of scale here: Instead of characterizing recharge, loading, or K at the local scale, we utilize a sub-regional scale "effective" value. Previous work refers to upscaling of the hydraulic conductivity as the estimation of effective parameters aimed to be used in regular flow equations (e.g., Fleckenstein and Fogg, 2008). In our work, we also estimated effective velocities and dispersivities in our homogenization. The terms upscaling and homogenization are both used regularly along the manuscript, which, we believe, makes our methodology clear.

In order to clarify the meaning of this term, the new manuscript refers to the abovementioned paper:

"The NPS metrics from fully heterogeneous simulations are compared to the NPS metrics obtained from a range of upscaled (e.g., Fleckenstein and Fogg, 2008), homogenized simulations that employ effective homogeneous properties rather than the original heterogenous distribution of the K..." (line 275)

Reference: Fleckenstein, J.H., and Fogg, G.E. (2008). Efficient upscaling of hydraulic conductivity in heterogeneous alluvial aquifers. Hydrogeology Journal, 16:1239. https://doi.org/10.1007/s10040-008-0312-3.

Line 286-292: 'To simulate flow and transport in an equivalent homogeneous, upscaled K conditions, we estimate the effective longitudinal and transverse vertical hydraulic conductivity, K*x and K*z , and dispersion, alpha*L and alpha*T. Effective parameters in the

longitudinal direction (K*x and alpha*L) are determined from the first and second spatial moments of a plume resulting from an injection of mass in a vertical plane of width 3000.0 m and depth 50.0 m. The same approach is adopted to estimate the effective parameters in the transverse vertical direction (K*z and alpha*T) by injecting particles in a horizontal plane covering the entire top of the domain. No extraction is considered in both cases in order to capture the natural behavior of the plume.'

What about the K*y and dispersivity in the y direction? Furthermore, 'No extraction is considered in both cases in order to capture the natural behavior of the plume', but the ensuing macro-dispersion of a plume is influenced by the presence of pumping well (e.g., radial versus uniform flow conditions), please clarify/justify this aspect.

Transverse longitudinal (y-direction) transport is negligible in our simulation setting (nonpoint source and no gradient in y).

We estimate apparent parameters (velocity and dispersion) under "natural" conditions in order to only capture the impact of K heterogeneity on transport. Scenarios accounting for homogenized K simulate explicitly extractions.

"The transverse horizontal (y-direction) component of transport is considered negligible given the size of the NPS plume and given that no gradient in y was applied." (line 288)

Lines 317-316: 'For all simulations, early mass travel times are within a range of 10 to 100 years with an expected value (highest probability) of 50 years (Figure 3a).'
 For a continuous variable is not possible to define the value with the highest probability (it is possible for a discrete variable). Furthermore, the expected value of a pdf does not always coincide with the value where the value of the pdf is the highest. Clarify.

Thank you for the concern and clarification. The sentence has been changed to:

"For all simulations, early mass travel times are within a range of 10 to 100 years with a peak of probability of 50 years (Figure 3a)." (line 317)

Lines 346-347: 'Spatial variability in the recharge is responsible for somewhat more uncertainty in the exact delineation of the capture zone along its margins than what is captured by the homogenization of *r*.'

Is not r the symbol to indicate the recharge? I think r should be replaced with C0 or viceversa. Furthermore, focusing on the impact of the homogenization of the recharge (compare Fig. 4a with Fig. 4c) I would say that the homogenization of the latter is associated with a less spatially extended capture zone w.r.t. to the case in which the recharge is treated as heterogeneous, this does not mean that the there is more uncertainty in the delineation of the capture zone. Furthermore, looking to the Fig. 4 I would say that capture zone is not well-delineated along its border, due to the low number of Monte Carlo realizations. Please check it and revise the sentence.

Yes, it should have been c_0, sorry for the typo. The dark blue area in Figure 4 is associated with very low probability to reach a well (Pw). In our setting, it may seem complex to reach a "smoother" delineation of the area defined by Pw values tending to 0.

Furthermore, despite that Pw>0 in these locations, we do not believe that it should be interpreted as an indicator of the extend of the capture zone in a management effort. On the other hand, we interpret a decrease of probability in the "hot spot", i.e., the source area characterized by the highest Pw values as an increase of uncertainty. It is now made clearer:

"Spatial variability in the recharge is responsible for somewhat more uncertainty (i.e., a decrease of the highest P_w values) in the exact delineation of the capture zone along its margins than what is captured by the homogenization of c_0 ." (line 346)

Lines 379-380: 'For an individual well, the results indicate that there is a 10% chance for nitrate concentrations to start to rise before 30 years, a 50% chance to rise no later than 50 years, and a 90% chance to rise before 70 years.' Should not be after X years?

We do believe that before is right. In Figure 6, P_{90} represents the concentrations exceeded with a 90% probability. These concentrations rise at t=70 years, while the concentrations exceeded with a 10% probability (P_{10}) rise earlier, at t=30 years. Since we analyze probabilities of exceedance, there is a 90% chance for well concentrations to rise before 70 years, and a 10% chance to rise earlier, before 30 years.

Lines 402-404: 'Homogenizing only concentration also leads to an underprediction, by about 20%, of concentrations exceeded by either 90%, 50%, or 10% of well, relative to the fully stochastic land use treatment (green lines in Figure 6).' Is green correct?

Thank for pointing out this typo. It is not green but blue.

"Homogenizing only concentration also leads to an underprediction, by about 20%, of concentrations exceeded by either 90%, 50%, or 10% of well, relative to the fully stochastic land use treatment (compare blue and red lines in Figure 6)." (line 402)

 Lines 471-472: '... only adds a moderate degree of uncertainty to the capture zone delineation...'
 Commont 15

see Comment 15.

Here again, we associate uncertainty to a decrease of the highest probabilities to reach a well (P_w), and not with an increase of the surface covered by very low P_w values. It is now made clearer earlier in the paper (*line 346*).

• Line 480: 'The observed gradient of travel times'

I don't see any gradient (i.e.,â LG) of the travel times evaluated by the Authors. There is a spatial variation of the mean travel time in Fig. 9, but this is different from a proper evaluation of the gradient. Please revise the sentence

The sentence has been modified to:

"The observed spatial variation of the mean travel times, increasing with the distance from a well, is overestimated when K-fields are homogenized." (line 480)

 Lines 487-489: 'Thus, contaminant mass reaching the top of the well has little variability – here only to the degree that the homogenization is done individually for each realization, leading to some minor variability in the homogenized K between realizations.' Why does the fact that the mass reaching the top of the well exhibits a low variability (note that the Authors do not provide a quantification of it) lead to have minor variability in the homogenized K? Please clarify.

This sentence has been reworked in light of your following remark. Please see below.

• Lines 489-491: 'More uncertainty is observed on the upstream side of the capture zone since it represents mass reaching the bottom of the screen, the vertical position of which is

realization dependent'

I suppose this comment is related to Fig. 9, which depicts the expected value of the travel time to the well. The latter gradually increases as we move upstream w.r.t. to the well location. I don't see this as a measure of an increasing level of uncertainty! It could be that the expected value of the travel time increases, but it can also be that spread (e.g., measured through the variance) of the pdf decreases. Please consider evaluating at least the variance of the pdf of travel times as a quantification of the degree of uncertainty.

Thank you for the comment. Indeed, Figure 9 does not display an estimation of travel time uncertainty. This paragraph discussing variability focused more on Figure 8 displaying the probabilities to reach a well. The phrase has been reworked and the appropriate Figure has been referred to in order to avoid confusion.

"Thus, contaminant mass reaching the top of the well has little variability - here only to the degree that the homogenization is done individually for each realization - leading to some minor realization-to-realization variability at the downstream side of the capture zone for the homogenized K (Figure 8). More uncertainty is observed on the upstream side of the capture zone since it represents mass reaching the bottom of the screen, the vertical position of which is realization dependent." (line 486)

Given the already extensive length of the manuscript, we believe that it is not necessary to add a more detailed discussion about the uncertainty of travel times associated to any location of the capture zone. Here, we focus more a more practical aspect of the problem, i.e., the assessment of the performance of a homogenization of K in reproducing expected travel times.

For your information, our previous paper (Henri and Harter, 2019) discusses the uncertainty of travel times associated to any location of the capture zone in similar conditions.

Reference: Henri, C., & Harter, T. (2019). Stochastic assessment of nonpoint source contamination: Joint impact of aquifer heterogeneity and well characteristics on management metrics. Water Resources Research, 55. https://doi.org/10.1029/ 2018WR024230.

Lines 520-521: 'Results show that homogenized K-fields perform more poorly to predict the lowest concentrations (P90) than the highest ones (P50 and P10).' Why P90 is associated with low concentrations? P90 is a probability, looking at Fig. 10a I

see that for a given concentration (either high or low) there is a time necessary to exceed this level of concentration in the well with a probability of 0.9. The way I interpret Fig. 10 is that homogenized solutions are in good agreement with the heterogeneous case, when it is a matter of evaluating low (P10) and middle (P50) probability of exceedance of a given concentration, but the homogenized solutions do not work well for high (P90) probability of exceedance of whatever concentration.

Thank you for raising out this point. This paragraph was more intended to be an interpretation of what we call a "regional" analysis.

 P_{90} represents, indeed, a high (90%) probability of exceedance. In other words, only there is only 10% of chance that a well will present a lower concentration. In what we describe as a regional analysis, this means that 10% of wells present concentration lower than the one shown in P_{90} . We believe that it a fair assessment to qualify these concentrations as the lowest concentrations.

The paragraph has been moved into the appropriate section.

"For instance, results show that homogenized K-fields perform more poorly to predict the lowest concentrations (P90) than the highest ones (P50 and P10). From a NPS pollution management perspective, the accuracy of the higher concentrations exceeded by half of wells or even by just 10% of wells is most critical, since they are more likely to exceed the MCL. The homogenized predictions are least accurate during the transition (breakthrough) period when concentrations in the vertically mixed sample obtained from a well are strongly controlled by travel time pdfs, which in turn are affected by the heterogeneity in the land use and aquifer dynamics." (line 552)

Conclusion: I note a change of style in the conclusion, whereas there is a more consistent and proper use of the terminology with respect to the rest of the work.

Thank you.

•

On the Conceptual Complexity of Non-Point Source Management: Impact of Spatial Variability

Christopher Vincent Henri¹, Thomas Harter¹, and Efstathios Diamantopoulos²

¹University of California, Davis, Center for Watershed Sciences, Veihmeyer Hall, Davis, CA 95616, USA. ²University of Copenhagen, Department of Plant and Environmental Sciences, Thorvaldsensvej 40, Copenhagen, DK-1871, Denmark.

Correspondence: Christopher Vincent Henri (chenri@ucdavis.edu)

Abstract. Non-point source (NPS) pollution has degraded groundwater quality of unconsolidated sedimentary basins over many decades. Properly conceptualizing NPS pollution from the well scale to the regional scale leads to complex and expensive numerical models: Key controlling factors of NPS pollution - recharge rate, leakage of pollutants, and soil and aquifer hydraulic properties - are spatially and, for recharge and pollutant leakage, temporally variable. This leads to high uncertainty in

- 5 predicting well pollution. On the other hand, concentration levels of some key NPS contaminants (salinity, nitrate) vary within a limited range (<2 orders of magnitude); and significant mixing occurs across the aquifer profile along the most critical compliance surface: drinking water wells with their extended vertical screen length. Here, we investigate, whether Given these two unique NPS contamination conditionsreduce uncertainty such that simplified spatiotemporal representation of recharge and, we here investigate the degree to which NPS travel time to wells and the NPS source area associated with an individual well can
- 10 be appropriately captured, for practical applications, when spatiotemporally variable recharge, contaminant leakage ratesand of, or hydraulic conductivity are justified when modeling NPS pollutionrepresented through a sub-regionally homogenized parametrization. We employ a Monte Carlo-based stochastic framework to assess the impact of model homogenization on key management metrics for NPS contamination. Results indicate that travel time distributions are relatively insensitive to the spatial variability of recharge and contaminant loading, while capture zone and contaminant time series exhibit some sensitivity to
- 15 source variability. In contrast, homogenization of aquifer heterogeneity significantly affects the uncertainty assessment of travel times and capture zone delineation. Surprisingly, regional statistics of well concentration time series the statistics of relevant NPS well concentrations (fast and intermediate travel times) are fairly well reproduced by a series of equivalent homogeneous aquifers, highlighting the dominant role of NPS solute mixing along well screens.

1 Introduction

20 The use of agrochemicals to address an ever growing food demand has led to the contamination of many sedimentary ground-water basins underlying intensively farmed regions (Nolan et al., 2002; Zektser and Everett, 2004; Rockstrom et al., 2009). Given the broad, continuous expanse of agricultural pollution sources across affected groundwater basins, this type of large scale pollution is often referred to as non-point source (NPS) pollution (Ritter and Shirmohammadi, 2000). The development of effective protection or remediation strategies in groundwater bodies affected by NPS pollution will require understanding

- 25 of the dynamics of NPS pollution in groundwater systems. Important aspects of NPS pollution are pollutant travel times, the location of well source areas (capture zonesalso known as capture zones; Barlow et al. (2018)) to identify specific pollution sources, and the long-term evolution of contaminant levels in and across affected wells and streams. The predictive modeling of these processes and associated management metrics is challenged by the inherent complexity of NPS pollution in groundwater systems.
- 30 Spatial variability represents a key source of complexity to be considered in understanding pollutant transport in the subsurface. Decades of investigation at contaminated industrial sites have highlighted the critical role that aquifer heterogeneity (e.g., the hydraulic conductivity) has in accurately understanding the solute transport behavior, to identify polluters, and to design effective remediation schemes and assess associated risk (e.g., Dagan and Nguyen, 1989; Cvetkovic et al., 1992; de Barros and Nowak, 2010; Henri et al., 2016). Large aquifer heterogeneity significantly affects the macro-dispersive behavior of contam-
- 35 inant plumes emanating from point sources. Lacking data to characterize subsurface properties in sufficient detail introduces significant uncertainty in the prediction of solute transport, the design of remediation measures, and the prediction of future concentrations at wells of interest (e.g., Dagan, 1984; Rubin, 2003). The prediction of solute transport from the NPS to a compliance area of interest (e.g., extraction or observation wells) has been shown to be critically impacted by aquifer heterogeneity, but also by mixing along the screen of production wells: contaminant mass arrivals in extraction wells can occur over may take

40 decades to centuries and are characterized by significant uncertainty (Hua and Harter, 2006; Henri and Harter, 2019).

Unique to nonpoint sources, the spatial (and temporal) variability of the source itself across a groundwater basin introduces an additional level of system complexity. NPS pollution of groundwater is typically associated with dissolved solutes associated with groundwater recharge across the landscape. Both, recharge rates and contaminant concentrations in nonpoint sources are subject to large spatial and temporal variability. The variability is partly due to spatially variable soil properties (e.g., Nielsen

- 45 et al., 1973). These properties control infiltration, recharge to groundwater, and the fate and transport of contaminants in the unsaturated zone (Hillel, 1980). Landscape management that leads to NPS pollution releases, e.g., irrigation, fertilization, construction and maintenance of urban, domestic, and other distributed waste systems leaking incidentally or intentionally into groundwater, is also subject to large spatial and temporal variability (Jordan et al., 1997). As with aquifer properties, the minutia of such spatial and temporal variability cannot be measured (or estimated) except at larger scales. For example, to
- 50 the degree that differences exist in average recharge and pollutant loading between mappable landscape management systems, these may be explicitly represented in space and time (e.g., Loague and Corwin, 1998; Nolan et al., 2018). This includes NPS differences between different farming systems (Kladivko et al., 2004) or between crops (Logsdon et al., 2002). Similarly, the degree to which mappable soil units affect recharge and pollutant fate and transport to the water table can also be explicitly represented (Biggar and Nielsen, 1976). However, spatial variability at smaller spatial scales or between individual units of the
- 55 same mappable class are subject to stochastic variability (Sisson and Wierenga, 1981; Vereecken et al., 2007). Furthermore, both, the timing and the spatial distribution of mappable and smaller scale unknown landscape processes is a stochastic process from a regional management perspective, which is concerned with pollution dynamics across an ensemble of wells.

The dual complexity of aquifer heterogeneity and spatio-temporal source variability represent a largely unexplored challenge in the assessment and management of NPS pollution in aquifers. Yet, conceptually simplified approaches have been 60 successfully employed to predict general trends and expected (average) contaminant behavior across ensembles of pollutant receptors of interest (wells, stream reaches) (e.g., Conan et al., 2003). Typically, these assessments lack any measures to also assess predictive uncertainties.

Some key characteristics of NPS contamination on the other hand make the NPS pollution system in groundwater well suited for upscaling without loss of information relevant to understanding the range of impacts on receptors: First, the individual

65 compliance surface of interest (the groundwater-well interface, the groundwater-stream reach interface) is subject to complete mixing prior to exposure (extracted well water, stream reach baseflow contribution). For example, production wells for urban water supplies are typically screened over dozens of meters (Henri and Harter, 2019). Even domestic wells are typically screened vertically over several meters of an aquifer system (Horn and Harter, 2009; Perrone and Jasechko, 2019). Similarly, stream reaches mix across an aquifer area of several tens to tens of thousands of square meters. The source area associated with such significantly sized compliance surfaces typically has length scales exceeding 100 m and frequently exceeding 1km (Horn and Harter, 2009; Henri and Harter, 2019). As a result, extracted water will be a mixture of groundwater age and source water

quality (Weissmann et al., 2002; Koh et al., 2018).

Secondly, while the source is wide-spread, compliance levels of key NPS contaminants (e.g., salt, nitrate) are commonly much less than one order of magnitude lower than the concentration in NPS recharge. This is characteristically different from most point-source contamination, where concentrations at the source may exceed compliance levels by many orders of magnitude (e.g., Frind et al., 1999). With the smaller difference between compliance and NPS recharge concentration, mixing at the compliance surface (i.e., in the well screen, at the stream reach scale) acts to homogenize the NPS recharge signature both, in space and time, thus reducing the need to accurately characterize the variability in space or in time to determine the

mixed concentration at an individual compliance surface (Kourakos et al., 2012). In contrast to assessing point sources of

- 80 industrial pollution, significant simplification in the spatiotemporal representation of, both or either, water and contaminant leakage rates and hydraulic conductivity may be possible without loss of accuracy. Bastani and Harter (submitted) explored the homogenization of temporal variability in NPS behavior. Yet, little work has been done to better understand and quantify the degree to which spatio-temporal variability in NPS representation or spatial variability in aquifer representation can be homogenized in NPS simulation tools while still accurately predicting NPS management metrics, including concentrations at
- 85 compliance surfaces.

In this paper, we assess the degree to which detailed spatial representation of both, the aquifer hydraulic conductivity and of contaminant source parameters - recharge rate of water and contaminant loading from the NPS to the groundwater table - can be homogenized in NPS models without reducing model accuracy. We consider three NPS management metrics and use a comparative simulation approach for our assessment.

90 Our starting point is a set of simulations that predict the long term contamination of an aquifer from NPS pollution under highly resolved heterogeneous aquifer and NPS source conditions. Results are compared against various simulation scenarios with homogenized representations of the aquifer and source heterogeneity. We compare results from various homogenization scenarios by focusing on three stochastic management metrics: the travel time distribution to production wells, the stochastic capture zone, and the stochastic contaminant time series in well water. Assuming ergodicity (Dagan, 1990), stochastic

3

Table 1. Proportion and hydraulic conductivity of the four categories (g: gravel; s: sand; ms: muddy sand; m: mud)

	g	S	ms	m
proportion [%]	0.10	0.35	0.25	0.3
hydraulic conductivity [m/d]	200.0	50.0	0.5	0.01

95 Stochastic management metrics are quantified both, for the pollution variability across an ensemble of production wells encountered over a basin, and for the uncertainty about pollution levels at an individual well. The later assumes structural ergodicity (Dagan, 1990), i.e., that the mean and variance of a single realization of the hydraulic conductivity field are close to the same statistics of the ensemble distribution (see histograms in Figure SM1).

2 Methodology

100 2.1 Reference case

We consider an unconsolidated sedimentary aquifer sytem typical of the Central Valley (California, USA), initially uncontaminated (e.g., pre-development state) and subject to nitrate pollution from agricultural NPS sources. The sub-region is characterized by a semi-arid Mediterranean climate, with dry summers and significant winter precipitation occurring mostly via the surrounding mountain ranges. The Central Valley groundwater basin is subject to intensive irrigated agricultural activities

- 105 supported by reservoirs managing surface water inflows from surrounding mountain ranges and by groundwater. Over the past eight decades, irrigation and groundwater pumping added a significant vertical flow component: Lateral groundwater flow, fed by mountain front recharge and discharged along the thalweg used to dominate the groundwater system dynamic. Modern groundwater discharge is mostly due to groundwater extraction. Recharge from intensive irrigation is superimposed on a weak lateral gradient, significantly increasing the importance of downward flows (Faunt, 2009). Water recharged from the irrigated
- 110 landscape to groundwater bodies carries significant loading of agricultural NPS pollutants, such as salt or nitrate (e.g., Baram et al., 2016).

The simulated soil and aquifer contamination setting represents conditions typically encountered in Central Valley's agricultural basins, but are not specific to a particular location. We represent heterogeneity in the hydraulic conductivity as well as the spatial variability in soil types and land use. The latter two are key characteristics that control spatial variability in recharge

115 and contaminant leakage rates. The transfer of water and nitrate from the soil surface to the aquifer is estimated through the modeling of flow and transport in the unsaturated zone for a series of typical soil types and crops found in the Central Valley.



Figure 1. Illustration of the methodology used in the study, considering a representative quadri-lateral subsystem of a highly heterogeneous alluvial aquifer system, $19.2 \times 6 \times 0.25 \text{ km}^3$ (12 miles x 3.75 miles x 820 feet). A land use map is randomly generated (top-left). Each color represents a different crop. A soil map (top-right) is extracted from the top layer of each geostatistical realization of the hydraulic conductivity (*K*). For each combination soil-type/crop, effective leakage rates of water and nitrate are numerically estimated. White particles represent a snapshot of the NPS pollution (particles) eighty years after a single year of contaminant loading. In each simulation, three extraction wells are explicitly represented at downstream location of the domain. The lower part of each well (in red) represent its screen from where water is extracted and from where mass arrival is recorded. Other wells are implicitly represented by the flux into the lower boundary of the domain.

2.1.1 Representation of aquifer and soil heterogeneity

2.1.2 Stochastic analysis

- Uncertainty in the representation of the spatial variability of the aquifer and soil hydraulic conductivity is systematically accounted for through the use of a geostatistical model in a Monte Carlo framework (Rubin, 2003). The propagation of variability and uncertainty into management metrics is assessed across an ensemble of production wells. Assuming ergodicity (Dagan, 1990), stochastic analysis is applied to first quantify uncertainty about pollution outcomes at individual wells and to secondly quantify regional spatial variability in pollution outcomes across an ensemble of wells: To characterize the uncertainty at an individual well, a large number of realizations of individual wells with equiprobable aquifer and soil realizations is generated.
- 125 Flow and transport processes across each are solved using a specified (fixed), mappable land use representation. To assess the spatial variability of NPS management metrics across an ensemble of well locations in a groundwater basin, the equiprobable

realizations of the aquifer system represent the variety of locations across a basin with geostatistically similar geological features. In this case This is true since the domain is designed to ensure that source areas of the three production wells are fully accommodated and that each well's area of capture can be considered independent. In the case of a regional analysis, land-use

130 is simulated as a random process.

Then, stochastic management metrics quantify both, the mean and variability of pollution levels across a large sample of production wells encountered over a basin as well as the expected value and uncertainty about pollution levels at an individual well. This is done by simulating stationary random fields (Figure SM3) and assuming ergodic conditions [e.g., Gelhar, 1993; Rajaram, 2002].

135 2.1.3 Aquifer

140

Spatial variability in the aquifer hydraulic conductivity (K) is represented using the transition probability/Markov chain method (Carle, 1999) for generating random realizations of the hydrofacies field (Carle and Fogg, 1996, 1998). Here, we consider 4 hydrofacies: gravel, sand, muddy sand and mud. The geostatistical model requires the characterization of the proportion, mean length and hydraulic conductivity of each facies, and of the facies-to-facies transition probability rates. We set these parameters to be representative of Central Valley aquifer conditions (Weissmann and Fogg, 1999a, b) (Table 1 and 2). A total of 50 realizations of the *K*-field were generated. An example of *K*-field can be observed in Figure 1. The histograms of the

mean and the variance of the logarithm of K are shown in Figure SM3. Fifty realizations were sufficient to converge the lower statistical moments of K and of the resulting mean velocities (Figure SM7SM9).

2.1.4 Soil map

145 The top layer of each K-field realization is here considered to represent the (unmapped) spatial variability of the soil type. Thus, a soil map, displaying the spatial distribution of the 4 hydrofacies at the land surface, is geostatistically consistent with each realization of the aquifer K-field underlying the soil horizon.

2.1.5 Land-use

The landscape of the simulated sub-basin is considered to be exclusively occupied by agricultural activities. Six different crop 150 types are randomly distributed over a domain of 19200.0 m × 6000.0 m. The crops are: almond, citrus, corn, cotton, grain and grapes. All fields are of the same rectangular dimension, 360 m × 300 m (11 ha, 27 acres). The spatial distribution of crop types is generated randomly and fulfills the following proportions of the 6 crop types: 24% of Almond, 24% of Citrus, 18% of Corn, 12% of Cotton, 12% of Grain, 10% of Grapes (Table 3). Crop types and proportion are representative of what may be encountered in the southeastern Central Valley (Harter et al., 2012).

Table 2. Mean length (diagonal values) of each categories (g: gravel; s: sand; ms: muddy sand; m: mud) and embedded transition probability (non-diagonal values) in the longitudinal (x), the horizontal transverse (y), and the vertical transverse (z) directions. Matrices reads as transition probability from the row facies to the column facies. The background category is designated by the letter *b*.

(x))	g		g s		m
g		$\bar{L}_{g,x}$	= 800.0m	0.7	0.15	b
s			0.7	$\bar{L}_{s,x} = 1500.0m$	0.28	b
ms	5		0.15	0.28	$\bar{L}_{ms,x} = 1000.0n$	ı b
m			b	b	b	b
(y)		g	S	ms	m
1	g	$\bar{L}_{g,g}$	y = 500.0m	0.7	0.15	b
1	s		0.7	$\bar{L}_{s,y} = 850.0m$	0.28	b
n	ns		0.15	0.28	$\bar{L}_{ms,y} = 900.0m$	b
ſ	n		b	b	b	b
		(z)	g	S	ms	m
		g	$\bar{L}_{g,z} = 2.0r$	n 0.7	0.15	b
		s	0.7	$\bar{L}_{s,z} = 3.5m$	0.28	b
		ms	0.15	0.28	$\bar{L}_{ms,z} = 2.0m$	b
		m	b	b	b	b

155 2.1.6 Estimation of recharge and contaminant leakage

Numerical simulations were conducted to simulate the vadose zone flow and transport processes across all possible crop and soil type combinations. Here, the gravel category in a soil layer was assumed to represent the same sandy soil as the sand category. Hence, a total of 18 vadose zone profiles represent all possible combinations of the 6 different land types (crops) and 3 different hydraulic soil profiles (sand, muddy sand, and mud). Simulations provide time-varying recharge and pollutant leakage rates for each of the 18 possible land-use and soil combinations at the water table of the respective underlying aquifer system. The time-series of the 18 simulations are computed *a priori* and then applied to define the water table boundary conditions of the groundwater flow and transport simulations.

Governing equations

160

One dimensional water flow in soils is described by the Richards equation (Diamantopoulos and Durner, 2012):

165
$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z} \left[K'(\psi) \left(\frac{\partial}{\partial z} + 1 \right) \right] - S,$$
 (1)

where θ [-] is the volumetric water content, t [d] is time, z [m] is the vertical spatial coordinate, positive upward, ψ [m] is the pressure head, $K'(\psi)$ [m²d⁻¹] is the saturated/unsaturated hydraulic conductivity as a function of ψ and S is the sink term, representing root water uptake [d⁻¹]. The water retention curve $\theta(\psi)$ and the hydraulic conductivity curve are required. The

two function are described by the van Genuchten-Mualem (van Genuchten, 1980) model:

170
$$\theta(\psi) = \begin{cases} \theta_r + (\theta_s - \theta_r) \times (1 + |\alpha \psi|^n)^{-m} & \psi < 0 \\ \theta_s & \psi \ge 0 \end{cases}$$
(2)

$$S_e = \frac{\theta(\psi) - \theta_r}{\theta_s - \theta_r} \tag{3}$$

$$K'(S_e) = K_s \times S_e^l \times \left[1 - \left(1 - S_e^{1/m}\right)^m\right]^2$$
(4)

where θ_s and θ_r [-] are the saturated and residual water contents, respectively, α [m⁻¹], n [-], m [-], and l [-] are shape parameters, m = 1 - 1/n, n > 1, and S_e [-] is the effective saturation.

175 Solute transport for a conservative tracer is described using standard advection-dispersion equation of the form:

$$\frac{\partial\theta c}{\partial t} = \frac{\partial}{\partial z} \left(\theta D \frac{\partial c}{\partial z}\right) - \frac{\partial qc}{\partial z} - S \times c \tag{5}$$

where $c \text{ [g m}^{-3}\text{]}$ is the concentration of the solute in the liquid phase, $D \text{ [m}^2 d^{-1}\text{]}$ is the dispersion coefficient, q is the volumetric water flux density (m d⁻¹) evaluated with the flow equation and $S \times c \text{ [g m}^{-3} d^{-1}\text{]}$ is the root nutrient uptake for the case of passive uptake. By focusing on hydrodynamic dispersion, D is defined as

$$180 \quad D = \lambda q/\theta \tag{6}$$

where λ is the dispersivity [m].

Parametrization of Hydrus 1D

For the numerical solution of equations 1 and 5, we used the Hydrus 1D software (Šimunek et al., 2016). Root water uptake for the six different crops was simulated by assuming a macroscopic root water uptake approach (Feddes et al., 1978). The
parameters for equations 2 and 5 were estimated by using Rosetta pedotransfer function (Scaap et al., 2004) and are shown in Table 4. For each soil horizon, dispersivity values were calculated by using the pedotransfer function of Perfect et al. (2002).

The simulation time was 21 years, from January 1, 1995 until December 31, 2015. Of the 21 years, the first 11 were used as a warm-up period and the remaining 10 were used to represent temporally variable boundary conditions at the top of the groundwater system. For an initial condition of equations 1 and 5, we assumed a uniform distribution of the pressure head and a solute free profile, respectively. The upper boundary condition for the flow problem accounts for precipitation, irrigation, and crop evapotranspiration. Daily reference (grass) evapotranspiration (ET_0) and precipitation (P) from the Stratford Meteorological station (California Irrigation Management Information System (CIMIS)) are used to represent southeastern Central Valley climate conditions. For each crop, ET_0 values were converted to potential crop evapotranspiration (ET_c) by using the single crop coefficient method (Allen et al., 1998). Daily time series of boundary conditions are used in Hydrus 1D.

Table 3. Area distribution, fertilization application and root zone depth for the 5 crops considered in this study.

Crop	Area ^a (%)	Fertilization application ^b [kg/ha/year]	Root zone [cm]
Almond	24.0	246.0	137.0
Citrus	24.0	157.0	107.0
Corn	18.0	239.0	91.0
Cotton	12.0	195.0	122.0
Grain	12.0	198.0	91.0
Grapes	10.0	39.0	91.0

^a from Harter et al. (2012)

^b from United States Department of Agriculture (1997)

Table 4. Horizon depth, soil hydraulic properties and dispersity of each horizon, for the three different soil profiles assumed in this study.

Soil type	Horizons [cm]	$\theta_r [\mathrm{cm}^3 \mathrm{cm}^{-3}]$	$\theta_s [\mathrm{cm}^3 \mathrm{cm}^{-3}]$	$\alpha [{\rm cm}^{-1}]$	n [-]	$K_s [\mathrm{cm} \mathrm{d}^{-1}]$	l [-]	λ [cm]
	0-20	0.036	0.389	0.033	1.41	20.30	-1.03	2.0
	20-56	0.035	0.389	0.036	1.43	23.44	-1.06	3.6
S 1	56-86	0.038	0.388	0.034	1.42	20.39	-1.07	3.0
	86-147	0.035	0.389	0.038	1.44	25.49	-1.06	6.1
	147-1000	0.033	0.391	0.040	1.47	29.72	-1.03	3.0
	0-30	0.039	0.376	0.057	1.66	50.17	-1.08	3.0
S2	30-107	0.048	0.370	0.055	1.86	46.55	-1.07	7.7
	107-1000	0.032	0.343	0.056	1.66	46.66	-1.07	4.3
	0-23	0.091	0.4988	0.022	1.19	7.86	-3.01	2.3
S 3	23-94	0.084	0.4978	0.014	1.22	0.09	-1.58	7.1
	94-1000	0.094	0.4740	0.018	1.27	4.28	-1.52	5.6

195 Based on calculated ET_c and P values, we created an irrigation schedule for each combination of crop-soil type, using the so-called evapotranspiration method (Allen et al., 1998). Irrigation was assumed to take place only during the crop period and not through the winter period (Figure 2). For all crop-soil combinations, we assume three fertilization events per year with the total amount of fertilizer application given in Table 3.

Preparing Water Table Boundary Conditions from Unsatured Zone Simulation Results

200 Simulations led to an estimation of the temporal evolution of water and nitrate leakage rate at the bottom of the 1D profile (Figure SM1 and SM2) for each crop - soil type combination, at daily time-steps. For the sake of simplicity, our groundwater simulation are conducted for steady state flow (but transient transport) conditions. Following Bastani and Harter (submitted), we homogenize both, recharge and pollutant flux in time and compute average, effective recharge and nitrate leakage rates over



Figure 2. Illustration example of daily values of potential crop evapotranspiration (orange points) used for the description of the upper boundary condition. The blue bars represent rainfall events and the light blue bars irrigation events. The red bar defines fertilization application for an amount equal to 65 Kg/ha (195 Kg/ha/year). Crop is cotton.

r [m/m²/d]	almond	citrus	corn	cotton	grain	grapes
S 1	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}
S 3	2.9×10^{-5}	2.3×10^{-4}	2.6×10^{-4}	2.0×10^{-4}	2.0×10^{-5}	1.5×10^{-4}
$m_f \left[g/m^2/d \right]$	almond	citrus	corn	cotton	grain	grapes
S 1	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}
S 3	0.0	1.9×10^{-10}	3.3×10^{-9}	0.0	0.0	0.0

Table 5. Recharge rate and nitrate mass flux applied for each crop - soil type combination

205

a 10 year time series (illustration in Figure SM1 and SM2 and average values in Table 5). For each of the 50 K-field (and, therefore, soil map) realization, the temporally homogenized results for each crop-soil combination are then used to describe the spatial distribution of the effective recharge (r(x,y)) and nitrate mass flux $(m_0(x,y))$. Histograms of the mean and the variance of recharge rate (r) and initial concentrations (c_0) are shows in Figure SM4 and SM5, respectively.

2.2 Groundwater flow and transport

2.2.1 Flow

210 Groundwater flow and nitrate transport are numerically simulated. We consider a 3-dimensional aquifer with a length (L_x) of 19200.0 m, a width (L_y) of 6000.0 m, and a depth (L_z) of 250.0 m (Table 6). Average steady-state groundwater flow conditions are governed by (Rushton and Redshaw, 1979):

$$\frac{\partial}{\partial x} \left(K_{xx} \frac{\partial h}{\partial x} \right) + \frac{\partial}{\partial y} \left(K_{yy} \frac{\partial h}{\partial y} \right) + \frac{\partial}{\partial z} \left(K_{zz} \frac{\partial h}{\partial z} \right) + Q'_s = 0, \tag{7}$$

where h (m) is the hydraulic head, and Q'_s is a volumetric flux per unit volume representing sources and sinks of water. 215 Groundwater fluxes are simulated by solving numerically the Darcy's law:

$$\mathbf{q}(\mathbf{x}) = -K(\mathbf{x})\nabla h(\mathbf{x}),\tag{8}$$

where q (m d⁻¹) is the specific discharge in the 3 dimensions x={x,y,z}. The longitudinal flow is defined by a regional gradient of 1.0×10^{-3} . The vertical flow is impacted by recharge and by a downward flux from the bottom of the domain. The spatially distributed fixed flux boundary condition across the bottom of the domain represents water flux to pumping wells in the deeper

- 220 part of the aquifer and the effect of implied, non-simulated groundwater extraction by wells distributed throughout the lower part of the simulated aquifer sub-basin. Domestic wells are not considered to significantly affect flow and transport processes and are not simulated. Recharge is considered spatially variable to account for realistic spatial distribution of crop and soil types (see Section 2.1.6). Histograms of For indications about the range of values and degree of variability, histograms of the mean and variance of the recharge rates applied in the 50 realizations of heterogeneous cases are shown in Figure SM4...
- Three extraction wells are located near the downstream edge of the domain. The extraction rate, Q_{out}, is set to 3000 m³d⁻¹ (551 gpm), corresponding to an actual irrigation pumping rate of 6000 m³d⁻¹ (1102 gpm) or 9000 m³d⁻¹ (1653 gpm) during a six month or four month annual irrigation season, respectively. The length of the well screen is location and realization dependent, depending on the vertical distribution of highly conductive material (gravel, sand) (Appendix A). The total outflow (downward flux at the domain bottom plus extraction at the three wells) is set to be equal to the inflow of water by recharge.
 The resulting water flow system representation is solved using MODFLOW (Harbaugh et al., 2000) for each realization of the
- *K*-field and the upper boundary.

2.2.2 Transport

Nitrate transport is modeled using the advection-dispersion equation (ADE) given by:

$$\phi \frac{\partial c}{\partial t} = \nabla \cdot (\phi \mathbf{D} \nabla c) + \nabla \cdot (\mathbf{q}c) \tag{9}$$

where c is the solute concentration, D is the 3-dimensional dispersion tensor, and ϕ is the porosity. The ADE is solved by the random walk particle tracking (RWPT) method implemented in the Fortran code RW3D (Fernàndez-Garcia et al., 2005; Henri Table 6. Domain discretization and physical parameters used in all simulations

Parameter	Value
Domain discretization:	
Number of cells, $n_x \times n_y \times n_z$	$120\times 60\times 625$
Cell dimension, $\Delta_x \times \Delta_y \times \Delta_z \ [m \times m \times m]$	$160.0\times100.0\times0.4$
Domain length, $L_x \times L_y \times L_z$ [m×m×m]	$19200.0 \times 6000.0 \times 250.0$
Flow and transport problem:	
Porosity, ϕ [-]	0.3
Average longitudinal hydraulic gradient, i_x [-]	1×10^{-3}
Extraction rate, Q_{out} [m ³ d ⁻¹]	3×10^3

and Fernàndez-Garcia, 2014, 2015). RWPT solves the ADE by moving a large number of particles in successive jumps given by: (e.g., Salamon et al., 2006).

$$\mathbf{x}_{p}(t+\Delta t) = \mathbf{x}_{p}(t) + \Delta t [\mathbf{v}(\mathbf{x}_{p}(t)) + \nabla \cdot \mathbf{D}(\mathbf{x}_{p}(t))] + \sqrt{\mathbf{D}(\mathbf{x}_{p}(t))\Delta t \cdot \xi(t)},$$
(10)

240 where \mathbf{x}_p is the particle position, \mathbf{v} is the velocity vector, and ξ is a normally distributed random variable with zero mean and unit variance.

The detailed discretization of the velocity field described above is capturing the most relevant characteristics affecting the macro-dispersive transport behavior (LaBolle, 1999; LaBolle and Fogg, 2001; Weissmann et al., 2002; Henri and Harter, 2019). Therefore, effects of grid-scale dispersion are assumed to be negligible, i.e., D=0 and Eq. 10 is simplified to 245 $\mathbf{x}_p(t + \Delta t) = \mathbf{x}_p(t) + \Delta t \times \mathbf{v}(\mathbf{x}_p(t))$. This assumption, which potentially impacts NPS management metrics, is further evaluated in Appendix B.

Nitrate transport originating from the water table is simulated using an instantaneous injection of 500,000 particles over the entire top of the domain. Particle transport is tracked using the RWPT algorithm (Eq. 10) for a simulation time of 350 years. In simulations with spatially variable initial contaminant mass loading, the local density of particles reproduces the local

- 250 initial concentration $(c_0(x, y))$ in recharge water at the groundwater table. Histograms of the mean and variance of the initial concentration over the 50 realizations are shown in Figure SM5 if a visualization of the range of values and of the variability is needed. Following the superposition principle, the cumulative mass arrival at wells resulting from an instantaneous injection of mass m_0 can be interpreted as $\dot{m}(t)$, the simulated mass flux at wells resulting from a continuous and temporally constant release of mass m_0 . Well concentrations are computed as $c_w(t) = \dot{m}(t)/Q_{out}$. Flow and transport are simulated for each
- 255 realization of the K-field and water table boundary condition.

2.3 Nonpoint source pollution management metrics

2.3.1 Pollutant travel times

Three relevant nonpoint source (NPS) pollution management metrics are considered to measure the stochastic simulation outcomes: the probability distribution density function of pollutant travel times to wells, the probability distribution of pollutant concentration in wells, and the probability distribution of source locations. The probability distribution

2.3.2 Pollutant travel times

The probability density function of travel times is obtained by recording travel times of particles to the compliance area, that is, the screen of the extraction well, for each particle. We obtain normalized travel times t_i by computing the time required to observe a specified fraction *i* of the total mass that reaches a well over the total 350 year simulation period. For instance, **15** t_5 represents the travel time from the water table to the well for the fifth percentile of the total mass reaching the well in 350 years. Following a stochastic approach, probability density functions (*pdf*s) of travel times t_i are obtained by determining the histogram of t_i in using time series from 150 simulated wells (50 realizations, each with 3 wells). Figure SM6 shows satisfactorily the convergence of the mean and variance of t_{50} . Travel time *pdf*s can represent a useful tool to assess both the expected time of solute arrival at the compliance area and the propagation of uncertainty from the hydraulic conductivity field to pollutant transport (e.g. Dagan and Nguyen, 1989; Cvetkovic et al., 1992; de Barros and Rubin, 2008; Henri et al., 2016).

2.3.3 Pollutant concentrations in wells (breakthrough curves)

The assessment of potential contaminant levels in extraction wells represents a key step in NPS pollution management. Under uncertain flow conditions, managers would benefit from knowing the probability to exceed a threshold concentration such as the maximum contaminant level (MCL) in a given well or in a series of wells. For each of the 150 simulated wells, breakthrough
curves c(t) are obtained. Figure SM7 shows satisfactorily the convergence of the mean and variance of the concentration exceeded by 50% of wells. Their probability distribution P_i(c,t) is obtained as a sample distribution of c(t), where P_i(c,t) is the probability for i% of wells to exceed the contaminant level c at time t.

2.3.4 Capture zones

280

260

NPS pollution management may also require the assessment of the effective source area, i.e., the capture zone or contributing area of the pollution observed in a production well. The spatial variability of hydraulic properties leads to uncertainty about and spatial variability of the source area (e.g. Varljen and Shafer, 1991; Franzetti and Guadagnini, 1996; Riva et al., 1999; Stauffer et al., 2002). In the stochastic framework, the capture zone is assessed by defining the spatial distribution of the probability that a contaminant leaking from the NPS will reach a well (P_w), i.e.

$$P_w(\mathbf{x}_{\text{NPS}}) = \operatorname{Prob}(\mathbf{x}_p(t \in [0, t_{end}]) = \mathbf{x}_w \mid \mathbf{x}_p(t = 0) = \mathbf{x}_{\text{NPS}}), \tag{11}$$

where \mathbf{x}_p is the 3-dimensional location (in the Cartesian coordinate system given by $\mathbf{x} = (x, y, z)$) of a portion of the plume (represented by a particle in this study), \mathbf{x}_w is a location shared with a well screen, and \mathbf{x}_{NPS} is a given location of the NPS. The spatial extension of non-zero probabilities forms then a probabilistic capture zone. The time required for a contaminant leaving a given location of the contributing area to reach the extraction well (or so-called time-related capture zone) is also stochastically analyzed.

290 **3** Upscaling and test cases

3.1 Homogenization of source terms

The NPS metrics from fully heterogeneous simulations are compared to the NPS metrics obtained from a range of upscaled (e.g., Fleckenstein and Fogg, 2008), homogenized simulations that employ effective homogeneous properties rather than the original heterogenous distribution of the *K*, *r*, and c_0 terms. The source terms $(r(x,y), c_0(x,y))$ are homogenized separately for each realization by spatial averaging to obtain $\langle r \rangle$ and $\langle c_0 \rangle$. Histograms of $\langle r \rangle$ and $\langle c_0 \rangle$ show significant variability of the homogenized source terms between realizations (Figures SM3 and SM4). Homogenized recharge rates and source concentrations range from 0.9 to 1.4 m d⁻¹ m⁻² and from 5.0 to 8.5 g m⁻³, respectively. A number of different homogenized models are considered and compared against the reference case:

- a heterogeneous r and heterogeneous c_0 (reference case);
- 300 a heterogeneous r and homogeneous c_0 ;
 - a homogeneous r and heterogeneous c_0 ;
 - a homogeneous r and homogeneous c_0 ;

3.2 Homogenization of the hydraulic conductivity and transport upscaling

To simulate flow and transport in an equivalent homogeneous, upscaled K conditions, we estimate the effective longitudinal
and transverse vertical hydraulic conductivity, K^{*}_x and K^{*}_z, and dispersion, α^{*}_L and α^{*}_{TV}. The transverse horizontal (y-direction)
component of transport is considered negligible given the size of the NPS plume and given that no gradient in y was applied.
Effective parameters in the longitudinal direction (K^{*}_x and α^{*}_L) are determined from the first and second spatial moments of
a plume resulting from an injection of mass in a vertical plane of width 3000.0 m and depth 50.0 m. The same approach
is adopted to estimate the effective parameters in the transverse vertical direction (K^{*}_z and α^{*}_{TV}) by injecting particles in a
horizontal plane covering the entire top of the domain. No extraction is considered in both cases in order to capture the natural behavior of the plume. For each realization of the K-field, the slope of the temporal evolution of the first spatial moment, i.e. the plume center of mass, is used to evaluate the apparent velocities, v^{*}_x and v^{*}_z. After estimation of the gradients from simulated head differentials in the x- and z-directions, Darcy's law is applied to evaluate effective hydraulic conductivities K^{*}_x and K^{*}_z
(Eq. 8). Effective dispersion values (α^{*}_L and α^{*}_{TV}) are similarly obtained by analyzing the slope of the normalized second

spatial moment of particle plume. The importance of representing upscaled dispersion is independently tested. Histograms of resulting of upscaled K_x^* , K_z^* , α_L^* , and α_{TV}^* values as well as the satisfactorily convergence of the mean of the apparent parameters after 50 realizations are shown in Figures SM6 and SM7, SM8 and SM9, respectively.

Furthermore, the cumulative implication of homogenization in aquifer properties and in the source terms c_0 and r is tested. The series of scenarios considered are:

- 320 a heterogeneous K, a heterogeneous r and a heterogeneous c_0 (reference case);
 - an upscaled K, r and c_0 , considering advection only;
 - an upscaled K, a heterogeneous r and c_0 , considering advection only;
 - an upscaled K, a heterogeneous r and c_0 , considering advection and upscaled dispersion.

4 Results and Discussion: Homogenization of source terms

325 The effect of conceptually simplifying recharge, contaminant input concentration, and aquifer heterogeneity on the stochastic description of travel times, well concentrations, and capture zones is here illustrated specifically for the case of quantifying uncertainty about these NPS pollution management metrics at a particular well surrounded by a spatially distributed, but fixed (known) distribution of land use across all realizations (scenario "LU 1"). Alternatively, the effects of homogenization on the analysis of spatial variability across an ensemble of wells in a groundwater basin, where land use is different in each realization (scenario "LU 50"), is further discussed in Section 6.

4.1 Travel time

We analyze the probability distribution pdfs of travel time for the 5%ile (t_5), 50%ile (t_{50}) and 95%ile (t_{95}) mass reaching a well within the 350 year simulation period (Figure 3). These metrics characterize the temporal variability of the early, median, and late mass travel time from of a one-year pollutant (e.g., nitrate) input to the aquifer system. For all simulations, early mass travel times are within a range of 10 to 100 years with an expected value (highest probability) a peak of probability of 50 years (Figure 3a). Late mass travel times are likely to be in the range of 50 to 300 years (Figure 3c), with an expected travel time of a peak probability at about 120 years. These results are 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 (CFC) age data (mean groundwater ages of 10 to 50 years in twice to three times shallower wells than the ones simulated in this study).

340

The homogenization of recharge spatial variability directly affects the flow field in the uppermost part of the aquifer. While the effect is larger than the homogenization of the concentration (see next paragraph), it also has no significant impact on travel time pdfs (Figure 3): The distributions are slightly less spread out over the time axis, with slightly higher and earlier mode (peak of the pdf) and lower probabilities in the tails of the pdf (probability differences at all times < 5%), especially of the late travel time (probability differences at all times around 10%). Previously, Li and Graham (1998) investigated the impact



Figure 3. Probability density function of the time required for 5% (a), 50% (b), and 95% (c) of the total recorded mass to reach a well considering a heterogeneous r and heterogeneous c_0 (reference case, red line); a homogeneous r and heterogeneous c_0 (yellow line); a heterogeneous r and homogeneous c_0 (light blue line); a homogeneous r and homogeneous c_0 (green line). Plain lines refers to the consideration of an identical land-use map for all realizations. For comparison, the red dash lines shows outputs from simulations with a realization-dependent land-use maps.

345 of recharge spatial variability in a more theoretical and simplified 2D heterogeneous aquifer contaminated by a point source under non-pumping conditions. The work highlights that spatial variability in recharge increases spreading, especially in the transverse direction. In our 3D NPS setting, transverse spreading is less relevant (Figure 3) and we do not observe the increase in variability.

350

The homogenization of initial concentrations has no physical impact on travel times since it does not affect the velocity field in the groundwater system. However, the difference in input concentration changes the distribution of the initial mass across the water table. Hence, there are small but discernible differences in the travel time pdfs of variable and homogeneous c both, in the case of spatially variable r (red and blue lines in 3) and in the case of homogeneous r (yellow and green lines in 3).

Importantly, the homogenized representation of r and c has nearly no effect on the time span between early and late arrival times at the well screen (the contrast in the position of the travel time pdfs for t_5 and t_{95}), which represents the age difference

between the youngest and the oldest water captured by the well screen and then mixed during the pumping process (Weissmann et al., 2002; Koh et al., 2018; Henri and Harter, 2019).

4.2 Stochastic capture zone

360

The stochastic capture zone (or source area) is the area characterized by $P_w(x,y) > 0$. Simulation results for the fully stochastic representation of source heterogeneity show that the stochastic capture zone covers an area of about 8000 by 4000 m (about 30% of the simulated domain, containing approximately 300 individual fields), while the zone from where mass is the most likely to reach a well (critical zone, $P_w(x,y) > 0.5$) is more spatially focused (about 3% of the simulated domain, the size of about 30 individual fields, see Figure 4).

As explained above, homogenizing the input concentration does not affect the velocity field and transport processes, and only slightly reduces P_w values of the critical zone (Figure 4b). On the other hand, not accounting for spatial variability in recharge leads to an overestimation of P_w values inside the critical zone (Figure 4c). The same observation is made when both r and c_0 are considered homogeneous (Figure 4d). The location of the critical zone, being controlled by regional flow conditions and well characteristics (extraction rate, depth and length of the screen), is not impacted by the spatial description of source terms. Spatial variability in the recharge is responsible for somewhat more uncertainty (i.e., a decrease of the highest P_w values) in the exact delineation of the capture zone along its margins than what is captured by the homogenization of rc_0 .

- Recharge rates, if considered heterogeneous, are by design correlated to the hydraulic conductivity. Highly conductive material are associated with high recharge rates, which may increase the channeling effect through preferential paths. Accounting for spatial variability in the recharge rate will, therefore, exacerbate the impact of the heterogeneity in the *K*-field, especially near the water table (where recharge is applied), thus increasing the uncertainty about delineating the source area.
- Just as travel time pdfs are little affected (Figure 3), the overall location of the stochastic capture zone is approximated quite well with the homogenized parametrization of concentration and recharge. Consequently, the average travel times required for a particle leaving a given location of the NPS to reach a well also are not dramatically impacted by the spatial representation of the two source terms (Figure 5). The average flow condition is common to all simulations. Since the top of well screens is 100 m deep, the solute transport from the source to the compliance areas occurs mostly at depths far away from the spatially variable top boundary condition, where local change of flow condition at the surface does not impact significantly groundwater fluxes.

4.3 Well NPS pollution concentration

A common characteristic of NPS pollution different from many point source cases is the temporal continuity and consistency in NPS inputs. For example, significant nitrate loading to groundwater began with the introduction of commercial fertilizer just before World War II and has continued since then (Rockstrom et al., 2009; Harter et al., 2017). The long-term consequence of

385

such continuous NPS loading, year-after-year, can be obtained from our simulations by superpositioning breakthrough curves obtained for NPS output in a single year at t = 0. If we neglect long-term trends or year-to-year variations in NPS and assume a constant input of nitrate to the water table, then the stochastic breakthrough curve at the well screen is simply obtained by



Figure 4. Probability of a particle leaving a given grid-cell to reach a well accounting for a heterogeneous r and heterogeneous c_0 (reference case, a); a heterogeneous r and homogeneous c_0 (b); a homogeneous r and heterogeneous c_0 (c); a homogeneous r and homogeneous c_0 (d); a lnK-weighted r and lnK-weighted c_0 (e).



Figure 5. Expected travel time [years] of a particle leaving a given grid-cell and reaching a well for a heterogeneous r and heterogeneous c_0 (reference case, a); a heterogeneous r and homogeneous c_0 (b); a homogeneous r and heterogeneous c_0 (c); a homogeneous r and homogeneous c_0 (d); a lnK-weighted r and lnK-weighted c_0 (e).

computing the cumulative distribution function (CDF) of the concentration pdf (Henri and Harter, 2019) (Figure 6). The CDF plots provide a measure of the expected time at which a given threshold contaminant level (in the *x*-axis) such as the MCL

(10 mg/L for nitrate as nitrogen in the U.S.) will be exceeded with a probability of 90% (P90, left), 50% (P50, center), or 10% (P10, right). In a regional context, these graphs can be interpreted as the time (*x*-axis) after which at least 90%, 50%, and 10% of all wells in the aquifer region exceed a concentration of interest (*y*-axis), respectively.

For the reference scenario (red curve in Figure 6), we observe that the concentration eventually exceeded by 90% of wells is about 4.5 mg N/L, after about 250 years (left graph in Figure 6). Half of the wells will have a concentration exceeding 11
mg N/L (again, after about 250 years). Also in at least half of wells, the onset of rising nitrate levels (to above background levels) will occur no later than 50 years after the start of nitrate loading, reaching levels corresponding to half of the MCL (5 mg N/L) after about 70 years, and reaching the MCL (10 mg N/L) no later than about 150 years (middle graph in Figure 6). The 10% most nitrate contaminated wells will show an onset of nitrate contamination no later than 30 years after the start of NPS pollution, exceed the MCL in less than 70 years, and reach concentrations exceeding 14.5 mg N/L no later than about 150 years (right graph in Figure 6).

400 years (fight graph in Figure 0).

For an individual well, the results indicate that there is a 10% chance for nitrate concentrations to start to rise before 30 years, a 50% chance to rise no later than 50 years, and a 90% chance to rise before 70 years. Similarly, results suggests that the MCL will be exceeded with 10% probability after 70 years and with 50% probability after 140 years.

- These results are consistent with observations of nitrate concentrations in drinking water and irrigation wells in the San Joaquin Valley, the southern half of the Central Valley. In Merced, Stanislaus, Tulare, and Kings County, about 40% of domestic wells (with screen depths not exceeding 100 m) exceed the drinking water standard (Ransom et al., 2013), but only about 10% of the large production wells in the southeastern San Joaquin Valley (the wells represented in this study) exceed the nitrate MCL (10 mg N/L) (Survey et al., 2012, 2013), approximately 70 years after the beginning of extensive fertilizer use in the region. We note that the time scale for these concentration increases is very sensitive to two aquifer parameters: the hydraulic
- 410 conductivity and the average effective porosity. If the regional average K was twice as large as assumed in our model, all times would be half as long. Similarly, if the average regional effective aquifer porosity was 20% larger, travel times would be 20% shorter.

The homogenization of spatial variability in the recharge rate and in the source concentration, while of limited consequence to travel time estimates and to estimates of source area extent, has measurable implications for stochastic well concentration

415 predictions, particularly at the lower margin: Homogenizing the recharge rate only leads to significantly (>40%) underestimating the maximum concentration exceeded by 90% of wells in the intermediate and long term. Homogenization leads to somewhat (\approx 10%) overestimating the concentration exceeded by 10% of wells over the long term, but reproduces well the concentration exceeded by 50% of wells at all time (yellow vs red in Figure 6).

Homogenizing both, recharge and contaminant loading does not affect the predictions quite as much, and in the opposite
direction: The (lower) concentrations exceeded by 90% of wells are overestimated and the (higher) concentrations exceeded
by only 10% of wells are underestimated, while the concentrations exceed by 50% of wells are less than 5% different from the fully stochastic prediction.



Figure 6. Time (y-axis) required for a well to exceed a given concentration (x-axis) with a probability of 90% (left), 50% (middle) and 10% (right) considering a heterogeneous r and heterogeneous c_0 (reference case, red line); a homogeneous r and heterogeneous c_0 (yellow line); a heterogeneous r and homogeneous c_0 (light blue line); a homogeneous r and homogeneous c_0 (green line). For comparison, the red dash lines shows outputs from simulations with a realization-dependent land-use maps.

Li and Graham (1998) stochastically analyze the impact of spatially random recharge rate on transport in a 2D point source setting. Their work concluded that, for those conditions, large variability in – and therefore uncertainty about - recharge increases uncertainty in solute concentration. In our work, we observe the opposite. The difference may be partly due to the 3D non-point source transport, and partly caused by the implicit correlation between the hydraulic conductivity and the recharge rate in our scenarios, which may increase the conditioning of the flow field that leads to the observed decrease of uncertainty relative to the homogenized scenario.

430

Results are also sensitive to a homogenization of only the initial concentrations, which would underestimate all concentrations by about 10% (blue lines in Figure 6). Homogenizing only concentration also leads to an underprediction, by about 20%, of concentrations exceeded by either 90%, 50%, or 10% of well, relative to the fully stochastic land use treatment (green compare blue and red lines in Figure 6).

The results of the homogenization and the differences to treating land use in fully stochastic mode ("50 LUs") are driven directly by three factors: the distribution of land use, including the size of fields relative to the source area, and the distribution

435 of recharge and nitrate leaching among different land uses. As shown in Section 4.2., the extent of the capture zone encompasses hundreds of fields, while the critical capture zone – the core contribution area – encompasses at least 30 fields. For field size much larger than those simulated here, or for a more spatially correlated distribution of crops among fields, homogenization across all land uses in a basin may lead to larger errors due to the smaller number of land use "samples" intersected by the capture zone (Gibbons, 1994). Furthermore, unsaturated zone flow and transport simulations have led to highest contaminant 440 leakage rates in areas of high recharge (almonds, citrus, see Table 5). Homogenizing mass leakage therefore decreases the amount of contaminant in high recharge areas and consequently globally underestimates well concentrations. Outcomes would be different if the highest concentration is associated with the lower recharge rate.

The examples shown here indicate that there may be significant errors in predicting future concentrations exceeded by 90% of wells and by 10% of wells, i.e., the distribution of exceedance probabilities among the ensemble of wells, whereas the concentration exceeded by half of the wells is characterized quite accurately under homogenized land use treatment. Overall, the homogenization of recharge in particular leads to the largest potential errors of NPS pollution management metrics, less so for predicting travel times and capture zone, but significantly so for predicting the distribution of exceedance concentrations across an ensemble of wells.

5 Results and Discussion: Homogenization of K

450 5.1 Travel time

455

In a second step, the implications of upscaling aquifer heterogeneity on the stochastic description of travel times, capture zones and well concentrations are assessed. Probability density functions of early, medium and late travel times are significantly impacted by the full homogenization of the hydraulic conductivity field (Figure 7). A homogenization of both, aquifer and land use random processes (K, r, c_0) drastically reduces the spread of all mass percentile travel times (yellow lines in Figure 7). But the homogenized prediction of modes is quite accurate: The mode of the early mass and median mass travel times (t_5 and t_{50}) are predicted with about 10% accuracy relative to the fully stochastic solution (Figure 7a and b). For the late mass travel times, the mode in the homogenized prediction occurs later than for the case of a fully heterogeneous system (Figure 7c).

Aquifer heterogeneity generates a complex network of well-connected channels but also zones of near-stagnation, all of which controls the spatio-temporal behavior of contaminant plumes across all scales. The effective solute path architecture 460 is, therefore, specific to the K-field realization and highly uncertain. In the stochastic solution, this generates a large range 460 of probable solute travel times (travel time pdfs with large variance) to the well screen that cannot be capture by simulating 460 transport in a homogenized K architecture (Figure 7).

However, the global motion of the plume, characterized by its first spatial moment and the downward movement of the first moment along the the depth interval of the well screen over time would be approximately similar for all realizations, given the geostatistical parameters and regional gradients. Thus, accounting for upscaled advective motion only (obtained from the estimation of the first spatial moment) preserves the large mixing in the well screen (Figure A1), but underestimates the uncertainty on travel times arising from the macro-dispersive effects of heterogeneity. This is captured by the fact that the modes of the early, median, and late mass arrivals are spread over similar time periods (45 years to 140 years), even the prediction based on a completely homogenized representation of both, aquifer and land use processes captures a significant

470 fraction of the age distribution of mass arriving at the well screen. This is due to the significant mixing that occurs in the well screen when the well is being pumped (Weissmann et al., 2002; Henri and Harter, 2019).



Figure 7. Probability density function of the time required for 5% (a), 50% (b), and 95% (c) of the total recorded mass to reach a well considering a heterogeneous K field, heterogeneous r and heterogeneous c_0 (reference case, red solid line); an upscaled K values, averaged r and averaged c_0 accounting for advection only (yellow dashed line); an upscaled K values, heterogeneous r and heterogeneous r_0 accounting for advection and dispersion (green dashed line).

Similar results to those for a fully homogenized representation are found when only the K-field is homogenized, but land use is represented with heterogeneous r and c_0 . The spread of each mass percentile travel time pdf is slightly larger than in the fully homogenized case, but is relatively far from capturing the full extent of the travel time pdfs for the fully heterogeneous simulations (compare blue and red lines in Figure 7).

475

While the homogenization of K removes the controlling process of the macro-dispersive pollutant behavior, the macrodispersive behavior can be approximated by including an upscaled, homogenized dispersion process (Eq. 9) into the simulation (Eq. 10). Using both, homogenized K and a representative, effective macro-dispersion much improves the accuracy of the homogenized prediction and captures significant features of the fully stochastic prediction (green lines in Figure 7). Early mass travel times are slightly underestimated while median and especially late mass travel times are slightly overestimated.

480

Applying second spatial moments from heterogeneous simulations to estimate the macrodispersion of an upscaled homogeneous model, however, assumes that the macro-dispersion process follows a Gaussian process (Dagan, 1990). It has been shown here and in other work (e.g., Dagan, 1984; Cvetkovic et al., 1992) that solute transport in heterogeneous media instead produces significantly skewed plume distributions, with early peak of mass and a long tail. Approximating such a skewed

- 485 distribution with a Gaussian curve that is located at the same center of mass travel time and has the same second spatial moment is know to generate earlier first travel times, a later peak of mass, and later late travel times, consistent with our results. This complexity of upscaling transport from heterogeneous conditions to a simplified homogeneous aquifer using lower spatial moments only has been highlighted before. The results here confirm that for presented here confirm this observation for the case of non-point source contaminations, but also put the macro-dispersive process in relation to the highlight the generation of a quasi-macro-dispersive process through the (vertical) well mixing process. 490

5.2 Time related capture zone

Analogous to the travel time pdfs, the spatial distribution of the stochastic capture zone, i.e., probability for a particle leaving a given location of the NPS to reach a well, is highly impacted by the homogenization of the hydraulic conductivity, much more so than by homogenization of land use processes alone (Figure 8).

- 495 In fully heterogeneous conditions, a wide range of P_w values are distributed over a large portion of the domain surface. However, most of the probabilistic capture zone is characterized by very low P_w values. The critical zone (area of highest probability) is characterized by P_w values of ≈ 0.6 and is centered at a longitudinal distance of about ≈ 2000 meters from the well. In the solution to the equivalent homogeneous parameter and boundary conditions, the uncertainty of the capture zone location is significantly underestimated, with most of the capture zone being characterized by high probability to reach a well
- (Figure 8b). Describing the spatial variability of nitrate mass loading and recharge (with an homogeneous K-field) only adds a 500 moderate degree of uncertainty to the capture zone delineation, (i.e., lower highest P_{iv} values), as expected from travel time pdf results above. Utilizing the alternative homogenized transport modeling approach with a homogenized K and an equivalent macro-dispersion term, unlike for travel time pdfs, does not substantially improve the stochastic prediction of the capture zone (compare Figure 8c and 8d). Furthermore, homogenizing the hydraulic conductivity seems, independently of the description of
- the source terms or of the consideration of dispersion, to mispredict the location of the capture zone: the critical zone is slightly 505 moved downstream, closer to the wells, and the capture zone extends to a small portion of the downstream edge of wells. The most distant part of the critical zone in the homogenized prediction of the capture zone overlaps with the actual location of the critical zone in the fully stochastic solution.

Consistent with these results, the spatial distribution of mean travel time required for the contaminant to reach a well is

510 similarly contracted to a much smaller area that extends downstream from the well, unlike in the fully stochastic representation (Figure 9). The observed gradient of travel times, increasing with the distance from a well, is overestimated when K-fields are upscaled. This leads to higher predicted mean travel times over the entire capture zone for all tested aquifer simplifications.

Expected travel time years of a particle leaving a given grid-cell and reaching a well for a heterogeneous K field, heterogeneous r and heterogeneous c_0 (reference case, a); an upscaled K values, averaged r and averaged c_0 accounting for advection only



Figure 8. Probability of a particle leaving a given grid-cell to reach a well accounting for a heterogeneous K field, heterogeneous r and heterogeneous c_0 (reference case, a); an upscaled K values, averaged r and averaged c_0 accounting for advection only (b); an upscaled K values, heterogeneous r and heterogeneous c_0 accounting for advection only (c); an upscaled K values, heterogeneous r and heterogeneous c_0 accounting for advection only (c); an upscaled K values, heterogeneous r and heterogeneous c_0 accounting for advection only (c); an upscaled K values, heterogeneous r and heterogeneous c_0 accounting for advection only (c); an upscaled K values, heterogeneous r and heterogeneous c_0 accounting for advection only (c); an upscaled K values, heterogeneous r and r

515 (b); an upscaled K values, heterogeneous r and heterogeneous c_0 accounting for advection only (c); an upscaled K values, heterogeneous r and heterogeneous c_0 accounting for advection and dispersion (d).

Simulation outcomes highlight that the set of upscaled K values among the 50 realizations does not cover a range large enough to reproduce the high variability of original contaminant location expected in heterogeneous situation. This indicates that regional hydraulic vertical and longitudinal gradients, common to all simulations, control mostly the behavior of first

- 520 spatial moments of heterogeneous plumes used here to estimate apparent velocities. Thus, contaminant mass reaching the top of the well has little variability — here only to the degree that the homogenization is done individually for each realization =leading to some minor variability in the realization-to-realization variability at the downstream side of the capture zone for the homogenized K between realizations (Figure 8). More uncertainty is observed on the upstream side of the capture zone since it represents mass reaching the bottom of the screen, the vertical position of which is realization dependent.
- Interestingly, the critical zone (high P_w) is predicted to be more downstream than its actual location if K is homogenized using apparent velocities (Figure 8). In case of heterogeneous K, a strong layering effect is observed, due to the superposition of relatively thin layers of highly and poorly conductive materials that stretch the plume longitudinally at large scale and move the capture zone upstream.

Consistent with these results, the spatial distribution of mean travel time required for the contaminant to reach a well is
similarly contracted to a much smaller area that extends downstream from the well, unlike in the fully stochastic representation (Figure 9). The observed spatial variation of the mean travel times, increasing with the distance from a well, is overestimated when *K*-fields are homogenized. This leads to higher predicted mean travel times over the entire capture zone for all tested aquifer simplifications.

5.3 Contaminant levels

- Future concentrations exceeded by only 10% of wells (P_{10}) and those exceeded by half of wells (P_{50}) are captured to within a factor 2 for the transition period between 20 years and 150 years, but agree to within 10% with the fully stochastic simulation results at late time, under near steady-state pollution conditions. The (low) concentration levels exceeded by 90% of wells (P_{90}) differs by a factor 2 or more, at all times, from the fully stochastic solution (Figure 10). Representing the spatial variability of source terms, but using a homogenized K-field improves the prediction of the P_{90} evolution.
- Using the alternative homogenized representation with an equivalent macro-dispersion improves the prediction only at late time (> 150 years) and predicts long-term concentrations for P_{50} and P_{10} very accurately (green lines in Figure 10). But it underestimates the concentrations for all of P_{90} and during the transition time for P_{50} and P_{90} .

The agreement between fully stochastic solutions and the homogenized solutions is in contrast to the seemingly significant differences between homogenized and fully stochastic results observed for travel time distributions of the individual mass percentiles and the capture zone location. That the homogenized prediction is still capable of producing useful results is due to the unique properties of nonpoint source pollution listed earlier: First, the NPS pollution is a continuous process rather than a one-time event, with some interannual variability and slow long-term trends (Hansen et al., 2012; Harter et al., 2017). Second, the mixing of water quality occurring in the well screen greatly controls the observed pollutant levels because of the continuous



Figure 9. Expected travel time [years] of a particle leaving a given grid-cell and reaching a well for a heterogeneous K field, heterogeneous r and heterogeneous c_0 (reference case, a); an upscaled K values, averaged r and averaged c_0 accounting for advection only (b); an upscaled K values, heterogeneous r and heterogeneous c_0 accounting for advection only (c); an upscaled K values, heterogeneous r and heterogeneous r and heterogeneous c_0 accounting for advection only (c); an upscaled K values, heterogeneous r and heterogeneous c_0 accounting for advection only (c); an upscaled K values, heterogeneous r and heterogeneous c_0 accounting for advection only (c); an upscaled K values, heterogeneous r and heterogeneous c_0 accounting for advection only (c); an upscaled K values, heterogeneous r and r



Figure 10. Time (y-axis) required for a well to exceed a given concentration (x-axis) with a probability of 90% (left), 50% (middle) and 10% (right) considering a heterogeneous K field, heterogeneous r and heterogeneous c_0 (reference case, red solid line); an upscaled K values, averaged r and averaged c_0 accounting for advection only (yellow dashed line); an upscaled K values, heterogeneous r and heterogeneous c_0 accounting for advection only (blue dashed line); an upscaled K values, heterogeneous c_0 accounting for advection and bipersion (green dashed line).

loading and because differences in pollutant loading rates for the more permeable soils, across all crops (Table 5) vary within less then one order of magnitude. Third, the composition of the land use and therefore the recharge and mass loading rates vary at a scale that is much smaller than the source area of the well. Hence, any location of the source area will capture a similar overall mass of NPS pollutant over time. Third, the amount of water quality mixing in the well is strongly controlled by the vertical location and length of the well screen and, for typical municipal production wells or agricultural wells, as simulated here, well construction will dominate the range of travel time distributions water and solutes entering the well screen over effects of macro-dispersion. Reproducing the range of average regional gradients potentially observed in a region, and average loading therefore provides critical and important information to reproduce in-well mixing of age and, hence, recorded water quality.

Results show that homogenized K-fields perform more poorly to predict the lowest concentrations (P_{90}) than the highest ones $(P_{50} \text{ and } P_{10})$. From a NPS pollution management perspective, the accuracy of the higher concentrations exceeded by

560 half of wells or even by just 10% of wells is most critical, since they are more likely to exceed the MCL. The homogenized predictions are least accurate during the transition (breakthrough) period when concentrations in the vertically mixed sample obtained from a well are strongly controlled by travel time *pdf*s, which in turn are affected by the heterogeneity in the land use and aquifer dynamics.

Results and Discussion: Regional stochastic analysis 6

- 565 Results and discussion thus far have focused on the uncertainty about predicting concentrations and source area associated with a single well, where land use distribution is heterogeneous, but deterministic (mappable). In the simulations discussed, the land use (but not the soil) was the same across all realizations. In NPS pollution management, an understanding of the variability in concentration evolution over time across the ensemble of wells in a basin, region, or management zone is of equal or more importance than understanding the uncertainty of future pollution dynamics at a particular well. For the regional analysis, the
- 570 conceptual modeling approach is identical to the stochastic analysis of an individual well, except that the land use distribution is also a random variable. To adapt the simulation setup to the regional stochastic analysis, the spatial distribution pattern of crops (i.e., land-use map) across the fixed grid of fields was randomly generated for each realization ("50 LUs"). Thus each realization represents an equi-probable location within a basin that is much larger in extent then the simulation domain. In the regional interpretation of the stochastic results, the range of individual travel times, capture zones, and concentration
- 575 breakthrough curves observed represent the variability across the ensemble of wells in the region, rather than the uncertainty about the outcome at a particular well (ergodicity principle, (Dagan, 1990)).

Adding random land use to the simulations leads to nearly identical travel time pdfs for early and median mass travel times appear and somewhat earlier late mass travel times (compare dashed and plain red lines in Figure 3). Travel times are therefore largely insensitive to the stochastic conceptualization of the land-use spatial variability ("1 LU" compared to "50

- LU"). Similarly, the capture zone area is not sensitive to whether a fixed heterogeneous ("1 LU") or random heterogeneous 580 ("50 LU") stochastic concept is employed (Figure 11). As a result, lowest and highest contaminant levels (P_{90} and P_{10}) are only slightly lower at late times, while P_{50} are similar at any times for both analysis (compare dashed and plain red lines in Figure 6). The similarity in results here is due to the spatial scale of the land use variability, set by the size of the fields, with several dozen of fields occupying the critical area of the capture zone (see above). Given the mixing in the well screen and the
- 585 continuity of NPS pollution, the number of fields in the capture zone is therefore sufficiently large, and the contrast in loading rates sufficiently small, that a single sample of the heterogeneous land use representation ("1 LU") becomes representative of an ensemble of land use patterns. That said, the advantages and disadvantages of the homogenization methods for land use and aquifer properties highlighted above apply equally to the depiction of regional variability in nitrate contamination of large production wells and to the uncertainty of nitrate dynamics in an individual well.
- For instance, results show that homogenized K-fields perform more poorly to predict the lowest concentrations (P_{90}) than 590 the highest ones (P_{50} and P_{10}). From a NPS pollution management perspective, the accuracy of the higher concentrations exceeded by half of wells or even by just 10% of wells is most critical, since they are more likely to exceed the MCL. The homogenized predictions are least accurate during the transition (breakthrough) period when concentrations in the vertically mixed sample obtained from a well are strongly controlled by travel time pdfs, which in turn are affected by the heterogeneity in the land use and aquifer dynamics. 595



Figure 11. Probability of (left) and mean travel time required for (left) a particle leaving a given grid-cell to reach a well accounting for a single randomly generated land-use map for all realizations (top) and for a different randomly generated land-use map for each realization (bottom).

7 Conclusions

A significant body of groundwater flow and transport literature has focused on upscaling flow and transport processes associated with industrial point source pollution. For accidental pollution with pollutants exceeding compliance levels by orders of magnitude, field research has shown that large uncertainties exist in predicting the fate of such contaminant plumes and the inability of upscaled methods to capture site-specific plume behavior. Stochastic methods have been used to characterize such large uncertainties. Here we explore the ability to which homogenized, effective representations of aquifer structure and landscape spatial variability in flow and transport simulations of NPS pollution are capable of accurately predicting pollution management metrics. We use three metrics typically of interest to NPS pollution management: travel time *pdf*'s, stochastic capture zones, and stochastic breakthrough curves. We compare solutions of these metrics for a fully heterogeneous aquifer structure and landscape system with those of a homogenized, upscaled landscape system, those of a homogenized, upscaled aquifer system, and those of a completely homogenized aquifer and landscape system. Within the landscape system, we further distinguish between homogenizing recharge flux and homogenizing pollutant mass flux. The analysis is performed for a typical intensive, irrigated Mediterranean agricultural landscape of orchards, vineyards, and field crops overlying an alluvial aquifer system polluted with nitrate from fertilizer applications. Based on the simulation results presented, we make the following key

- 610 conclusions:
 - Land use, soil, and aquifer heterogeneity lead to large variability in groundwater travel paths, travel times, source location, and therefore well nitrate concentrations across a regional set of wells and, hence, significant uncertainty about pollution dynamics at any one well.

 The impact of continuous landscape pollutant loading to a typical high capacity production well with top of screen at 100 m below the water table is first seen a couple of decades after pollution initiation but is not fully reflected across all wells of a region after one to two centuries until one or two centuries later.

615

620

630

635

640

- With the capture zone of an individual well typically stretching across a diverse sub-set of land use in a region, the homogenization of the recharge and mass loading across the landscape to simulate NPS pollution management metrics can be appropriate, especially for simulating travel time *pdf* s and stochastic capture zones. In this case, nitrate variability between wells is much more affected by aquifer and soil heterogeneity than the heterogeneity in crop patterns across the landscape. This finding may not apply to cases where land use units (fields, orchards) occupy a much larger area or many fields of one crop type are clustered, or for wells with small pumping rates and, hence, small capture zones in those cases the variability in capture zone loading across an ensemble of wells may be ill-represented by a homogenized, regionally averaged recharge and nitrate mass loading.
- Homogenization of the aquifer hydraulic property significantly degrades travel time statistics as well as the stochastic delineation of the capture zone. Accounting for aquifer heterogeneity by utilizing an upscaled macrodispersion only slightly improves predictions of travel time *pdfs* or stochastic capture zones.
 - During the transition period (20 years to 170 years after pollution initialization), simulations using a homogenized representation of the aquifer structure provide aggregated concentration predictions, such as the concentration exceeded by half of wells, that are as much as a factor 2 different from predictions that fully represent aquifer heterogeneity.
 - On the other hand, due to the strong effect of vertical groundwater mixing during the well pumping process and due to the continuity of NPS pollution, an upscaled, homogenized representation of aquifer heterogeneity using an effective hydraulic conductivity produces reliable and useful predictions for the concentration levels exceeded by half of wells and even the higher concentrations exceeded by only 10% of wells, especially in the long-term. These are the wells of most concern in NPS pollution of groundwater.
 - Homogenized approaches may be most useful to predict whether long-term outcomes meet management goals across a
 regional ensemble of wells, but may be less accurate in specifying how quickly such goals may be achievable.

Future work is needed to further understand the role of crop type clustering on landscape homogenization, and the effect of interannual and seasonal loading variability on NPS pollution management metrics. More work is also needed to investigate other forms of partial or full homogenization of aquifer structure on prediction metrics considered here.

Code availability. The online resources located in GitHub (see this link) include the Matlab scripts necessary to reproduce the results.

31



Figure A1. Probability histogram of the simulated screen lengths.

Appendix A: Well screen design

For each realization of the hydraulic conductivity field, 3 extraction wells are implemented. The pumping rate of each well is fixed to 3000 m³/d and the top of the screen is fixed to 100 m. As in real settings, the length of this screen is dependent on the local aquifer properties in order to sustain the total extraction rate. Indeed, pumping effectively occurs through portions of well located in highly conductive aquifer material. To simulate this local K dependence of the well screen length, we are using a rule of thumb stating that 10 cumulative foot (3.05 m) of gravel and sand has to be crossed for each 100 gallon-per-minute (545.1 m³/d) of extraction. The probability histogram (over all realizations) of the simulated screen lengths for each tested extraction rates is shown in Figure A.1.

650 Appendix B: Impact of dispersion

Former studies (LaBolle, 1999; LaBolle and Fogg, 2001; Weissmann et al., 2002) highlighted the insensitivity of transport simulations to local scale dispersivity (α_i, where *i* indicates the transport direction) if aquifer heterogeneity is representing in a finely detailed manner by means of the transition probability method (TPROGS). This insensitivity is explained by the large macrodispersion caused by the well represented facies scale heterogeneity, which renders spreading from local dispersivity insignificant. As a result, fairly small values of α_L are, in this setting, usually adopted. For instance, Weissmann et al. (2002) applied to their transport model (with computational grid similar to the one used in our study) a grid scale longitudinal dispersivity of 0.04 m, which appeared to have a insignificant impact on transport and resulting groundwater age distribution. Values of α_L were chosen to fulfill the magnitude of dispersivity values reported at field sites of scale similar to the computation cells (references in previously cited work). Here, we test the impact of much larger values of dispersivity (1.5 m and 15.0 m) on breakthrough curves recorded at a well. The simulation setup is identical to the one described in the manuscript. Results are shown for a single realization of the hydraulic conductivity field.



Figure B1. Breakthrough curve recorded at a well accounting for advection only (red cruve), for advection and dispersion with a longitudinal dispersivity of 1.5 m (yellow curve), and for advection and dispersion with a longitudinal dispersivity of 15.0 m (blue curve). Transverse horizontal and transverse vertical dispersivities are always, respectively, 1/10 and 1/100 of the longitudinal dispersivity.

Our outputs displays no significant impact on transport of a α_L coefficient of 1.5 m. Increasing grid-scale dispersivity to 15.0 m leads to slightly earlier first travel times, later late travel times and lower contaminant levels observed at intermediate and late times (Figure B1). Therefore, no implications can be expected when accounting exclusively for advection when grid-dispersivity is lower than 1.5 m, as always used in previously cited studies.

Author contributions. CVH and TH designed the Monte Carlo simulations, which was then ran and post-processed by CVH. ED designed and ran the Hydrus simulations. All authors contributed to the results analysis and to the writing of the article.

Competing interests. The authors declare that they have no conflict of interest.

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 purposes only and does not imply endorsement by the SWB.

References

690

- Allen, R. G., Pereira, L. S., Raes, D., and Smith, M.: Crop evapotranspiration- Guidelines for computing crop water requirements, paper 56, FAO Irrigation and drainage, 1998.
- 675 Baram, S., Couvreur, V., Harter, T., Read, M., Brown, P., Kandelous, M., Smart, D., and Hopmans., J.: Estimating Nitrate Leaching to Groundwater from Orchards: Comparing Crop Nitrogen Excess, Deep Vadose Zone Data-Driven Estimates, and HYDRUS Modeling, Vadose Zone J., 15, 2016.
 - Barlow, P. M., A.Leake, S., and Fienen, M. N.: Capture Versus Capture Zones: Clarifying Terminology Related to Sources of Water to Wells, Groundwater, 56, 694–704, 2018.
- 680 Bastani, M. and Harter, T.: Effects of upscaling temporal resolution on groundwater nitrate transport model performance at the regional scale, Hydrogeology Journal, 2020.

Biggar, J. W. and Nielsen, D. R.: Spatial Variability of the Leaching Characteristics of a Field Soil, Water Resour. Res., 12, 78–84, 1976. Carle, S. F.: TProGS: Transition probability geostatistical software, Dep. of Land, Air and Water Resources, Univ. of Calif., Davis, 1999. Carle, S. F. and Fogg, G. E.: Transition probability-based indicator geostatistics, Math. Geol., 28, 453–477, 1996.

- 685 Carle, S. F. and Fogg, G. E.: Modeling spatial variability with one-and multi-dimensional continuous Markov chains, Math. Geol., 29, 891–917, 1998.
 - Conan, C., Bouraoui, F., Turpin, N., de Marsily, G., and Bidoglio, G.: Modeling Flow and Nitrate Fate at Catchment Scale in Brittany (France), J. Environ. Qual., 32, 2026–2032, 2003.
 - Cvetkovic, V., Shapiro, A. M., and Dagan, G.: A solute flux approach to transport in heterogeneous formations: 2. Uncertainty analysis, Water Resour. Res., 28, 1377–1388, 1992.
 - Dagan, G.: Solute transport in heterogeneous porous formations, Journal of Fluid Mechanics, 145, 151–177, 1984.
 - Dagan, G.: Transport in heterogeneous porous formations: Spatial moments, ergodicity, and effective dispersion, Water Resources Research, 26, 1281–1290, https://doi.org/10.1029/WR026i006p01281, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/ WR026i006p01281, 1990.
- 695 Dagan, G. and Nguyen, V.: A comparison of travel time and concentration approaches to modeling transport by groundwater, J. Contam. Hydrol., 4, 79–91, 1989.
 - de Barros, F. P. J. and Nowak, W.: On the link between contaminant source release conditions and plume prediction uncertainty, J. Contam. Hydrol., 116, 2010.

de Barros, F. P. J. and Rubin, Y.: A risk-driven approach for subsurface site characterization, Water Resour. Res., 44, W01 414, 2008.

- Diamantopoulos, E. and Durner, W.: Dynamic Nonequilibrium of Water Flow in Porous Media: A Review, Vadose Zone J., 11, 2012.
 Faunt, C.: Groundwater Availability of the Central Valley Aquifer, California: U.S., Geological Survey Professional Paper, 2009.
 Feddes, R. A., Kowalik, P. J., and Zaradny, H.: Simulation of field water use and crop yield, Simulation monographs, Pudoc, Wageningen, the Netherlands, 1978.
- Fernàndez-Garcia, D., Illangasekare, T. H., and Rajaram, H.: Differences in the scale-dependence of dispersivity estimated from temporal and spatial moments in chemically and physically heterogeneous porous media, Adv. Water Res., 28, 745–759, 2005.
 - Fleckenstein, J. and Fogg, G.: Efficient upscaling of hydraulic conductivity in heterogeneous alluvial aquifers, Hydrogeology Journal, 16, 1239, 2008.
 - Franzetti, S. and Guadagnini, A.: Probabilistic estimation of well catchments in heterogeneous aquifers, J. Hydrol., 174, 149–171, 1996.

Frind, E. O., Molson, J. W., Schirmer, M., and Guiguer, N.: Dissolution and mass transfer of multiple organics under field conditions: The

710 Borden emplaced source, Water Resources Research, 35, 683–694, https://doi.org/10.1029/1998WR900064, https://agupubs.onlinelibrary. wiley.com/doi/abs/10.1029/1998WR900064, 1999.

Gibbons, R. D.: Statistical Methods for Groundwater Monitoring, Wiley; 2 edition, Hoboken, New Jersey, 1994.

- Hansen, B., Dalgaard, T., Thorling, L., Sørensen, B., and Erlandsen, M.: Regional analysis of groundwater nitrate concentrations and trends in Denmark in regard to agricultural influence, Biogeosciences, 9, 3277–3286, 2012.
- 715 Harbaugh, A., Banta, E., Hill, M., and McDonald, M.: MODFLOW 2000 the US Geological Survey Modular ground-water model-user guide to modularization concepts and the ground-water flow process, Open File 00-92, 121pp., Rep. U.S. Geol. Surv., 2000.
 - Harter, T., Lund, J. R., Darby, J., Fogg, G. E., Howitt, R., Jessoe, K. K., Pettygrove, G. S., Quinn, J. F., Viers, J. H., Boyle, D. B., Canada, H. E., DeLaMora, N., Dzurella, K. N., Fryjoff-Hung, A., Hollander, A. D., Honeycutt, K. L., Jenkins, M. W., Jensen, V. B., King, A. M., Kourakos, G., Liptzin, D., Lopez, E. M., Mayzelle, M. M., McNally, A., Medellin-Azuara, J., and Rosenstock, T. S.: Addressing Nitrate
- 720 in California's Drinking Water with a Focus on Tulare Lake Basin and Salinas Valley Groundwater, Report, 78 p., Center for Watershed Sciences, University of California, Davis for the State Water Resources Control Board Report to the Legislature, 2012.
 - Harter, T., K. Dzurella, G. K., Hollander, A., Bell, A., Santos, N., Hart, Q., King, A., J.Quinn, Lampinen, G., Liptzin, D., Rosenstock, T., Zhang, M., Pettygrove, G., and Tomich, T.: Nitrogen Fertilizer Loading to Groundwater in the Central Valley. Final Report to the Fertilizer Research Education Program, Tech. Rep. Projects 11-0301 and 15-0454, California Department of Food and Agriculture and University 2007 June 10 Part 10 Par
- of California Davis, 2017.
 - Henri, C. V. and Fernàndez-Garcia, D.: Toward efficiency in heterogeneous multispecies reactive transport modeling: A particle-tracking solution for first-order network reactions, Water Resour. Res., 50, 7206–7230, 2014.
 - Henri, C. V. and Fernàndez-Garcia, D.: A random walk solution for modeling solute transport with network reactions and multi-rate mass transfer in heterogeneous systems: Impact of biofilms, Adv.in Water Resour., 86, 119, 2015.
- 730 Henri, C. V. and Harter, T.: Stochastic Assessment of nonpoint source contamination: Joint impact of aquifer heterogeneity and well characteristics on management metrics, Water Resour. Res., in press, 2019.
 - Henri, C. V., Fernàndez-Garcia, D., and de Barros, F. P. J.: Assessing the joint impact of DNAPL source-zone behavior and degradation products on the probabilistic characterization of human health risk, Adv.in Water Resour., 88, 124–138, 2016.
 Hillel, D.: Fundamental of Soil Physics, Academic Press, New York, 1980.
- Horn, J. E. and Harter, T.: Domestic Well Capture Zone and Influence of the Gravel Pack Length, Groundwater, 47, 277–286, https://doi.org/10.1111/j.1745-6584.2008.00521.x, https://ngwa.onlinelibrary.wiley.com/doi/abs/10.1111/j.1745-6584.2008.00521.
 x, 2009.
 - Hua, Z. and Harter, T.: Nonpoint source solute transport normal to aquifer bedding in heterogeneous, Markov chain random fields, Water Resour. Res., 42, W06 403, https://doi.org/10.1029/2004WR003808, 2006.
- 740 Jordan, T. E., Correll, D. L., and Weller, D. E.: Relating nutrient discharges from watersheds to land use and streamflow variability, Water Resources Research, 33, 2579–2590, https://doi.org/10.1029/97WR02005, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/ 97WR02005, 1997.
 - Kladivko, E. J., Frankenberger, J. R., Jaynes, D. B., Meeka, D. W., Jenkinson, B. J., and Fausey., N. R.: Nitrate Leaching to Subsurface Drains as Affected by Drain Spacing and Changes in Crop Production System Contribution of the Indiana Agric. Research Programs, Journal of
- 745 Environmental Quality, 33, 1803–1813, 2004.

- Koh, E.-H., Lee, E., Kaown, D., Green, C. T., Koh, D.-C., Lee, K.-K., and Lee, S. H.: Comparison of groundwater age models for assessing nitrate loading, transport pathways, and management options in a complex aquifer system, Hydrological Processes, 32, 923–938, https://doi.org/10.1002/hyp.11465, https://onlinelibrary.wiley.com/doi/abs/10.1002/hyp.11465, 2018.
- Kourakos, G., Klein, F., Cortis, A., and Harter, T.: 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, https://doi.org/10.1029/2011WR010813, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2011WR010813, 2012.

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

- LaBolle, E. M. and Fogg, G. E.: Role of molecular diffusion in contaminant migration and recovery in an alluvial aquifer system, Transp. Porous Media, 42, 155 179, 2001.
- 755 Li, L. and Graham, W.: Stochastic analysis of solute transport in heterogeneous aquifers subject to spatially random recharge, Journal of Hydrology., 206, 16–38, 1998.
 - Loague, K. and Corwin, D. L.: Regional-scale assessment of non-point sourcegroundwater contamination, Hydrol. Process., 12, 957–965, 1998.
 - Logsdon, S. D., Kaspar, T. C., Meek, D. W., and Prueger, J. H.: Nitrate Leaching as Influenced by Cover Crops in Large Soil Monoliths,

Agronomy Journal, 94, 807–814, http://dx.doi.org/10.2134/agronj2002.8070, 2002.
 Nielsen, D. R., Biggar, J. W., and Erh, K. T.: Spatial variability of field-measured soil-water properties, Hilgardia, 42, 215–259, 1973.
 Nolan, B. T., Hitt, K. J., and Ruddy, B. C.: Probability of Nitrate Contamination of Recently Recharged Groundwaters in the Conterminous

United States, Environ. Sci. Technol., 36, 2138–2145, 2002.

770

- Nolan, B. T., Green, C. T., Juckem, P. F., Liao, L., and Reddy, J. E.: Metamodeling and mapping of nitrate flux in the unsaturated zone
- 765 and groundwater, Wisconsin, USA, Journal of Hydrology, 559, 428 441, https://doi.org/https://doi.org/10.1016/j.jhydrol.2018.02.029, http://www.sciencedirect.com/science/article/pii/S0022169418301057, 2018.
 - Perfect, E., Sukop, M. C., and Haszler, G. R.: Prediction of Dispersivity for Undisturbed Soil Columns from Water Retention Parameters, Soil Sci. Soc. Am. J., 66, 696–701, 2002.

Perrone, D. and Jasechko, S.: Deeper well drilling an unsustainable stopgap to groundwater depletion, Nature Sustainability, 2, 773–782, https://doi.org/10.1038/s41893-019-0325-z, https://doi.org/10.1038/s41893-019-0325-z, 2019.

- Ransom, K., King, A., and Harter, T.: Identifying sources of groundwater nitrate contamination in a large alluvial groundwater basin with highly diversified intensive agricultural production, Journal of Contaminant Hydrology, 151, 140 – 154, https://doi.org/https://doi.org/10.1016/j.jconhyd.2013.05.008, http://www.sciencedirect.com/science/article/pii/S0169772213000855, 2013.
- 775 Ritter, W. F. and Shirmohammadi, A.: Agricultural Nonpoint Source Pollution: Watershed Management and Hydrology, CRC Press, Boca Raton, 2000.

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

Rockstrom, J., Steffen, W., and K. Noone, e. a.: A safe operating space for humanity, Nature, 461, 472–475, 2009.

- 780 Rubin, Y.: Applied Stochastic Hydrogeology, Oxford Univ. Press, Oxford, 2003.
 - Rushton, K. and Redshaw, S.: Seepage and groundwater flow Numerical analysis by analogue and digital methods, John Wiley and Sons, New York, 1979.

- Salamon, P., Fernandez-Garcia, D., and Gomez-Hernandez, J. J.: A review and numerical assessment of the random walk particle tracking method, J. Contam. Hydrol., 97, 277–305, 2006.
- 785 Scaap, M. G., Nemes, A., and van Genuchten, M. T.: Comparison of Models for indirect estimation of water retention and available water in surface soils, Vadose Zone J., 3, 1455–1463, 2004.
 - Sisson, J. B. and Wierenga, P. J.: Spatial Variability of Steady-State Infiltration Rates as a Stochastic Process1, Soil Sci. Soc. Am. J., 45, 699–704, 1981.
 - Stauffer, F., Attinger, S., Zimmermann, S., and Kinzelbach, W.: Uncertainty estimation of well catchments in heterogeneous aquifers, Water
- 790 Resour. Res., 38, 1238, 2002.
 - Survey, U. S. G., Burton, C. A., Shelton, J. L., and Belitz, K.: Status and understanding of groundwater quality in the two southern San Joaquin Valley study units, 2005-2006 - California GAMA Priority Basin Project, Tech. rep., USGS, Reston, VA, https://doi.org/10.3133/sir20115218, http://pubs.er.usgs.gov/publication/sir20115218, 2012.
- Survey, U. S. G., Shelton, J. L., Fram, M. S., Belitz, K., and Jurgens, B. C.: Status and understanding of groundwater quality in the Madera, Chowchilla Study Unit, 2008: California GAMA Priority Basin Project, Tech. rep., USGS, Reston, VA,

https://doi.org/10.3133/sir20125094, http://pubs.er.usgs.gov/publication/sir20125094, 2013.

- United States Department of Agriculture: irrigation guide: Natural Resources Conservation Service, information available on World Wide Web, accessed April 14, 2004, at http://www.wcc.nrcs.usda.gov/nrcsirrig/irrig-handbooks-part652.html, National engineering handbook part 652, USDA, 1997.
- 800 van Genuchten, M.: A closed form equation for predicting the hydraulic conductivity of unsaturated soils, Soil Science Society of America Journal, 44, 892–898, 1980.
 - Varljen, M. D. and Shafer, J. M.: Assessment of uncertainty in time-related capture zones using conditional simulation of hydraulic conductivity, Ground Water, 29, 737–748, 1991.
- Vereecken, H., Kamai, T., Harter, T., Kasteel, R., Hopmans, J., and Vanderborght, J.: Explaining soil moisture variability as a function of
 mean soil moisture: A stochastic unsaturated flow perspective, Geophys. Res. Lett., 34, L22 402, 2007.
 - Šimunek, J., van Genuchten, M., and Šejna, M.: Recent developments and applications of the HYDRUS computer software packages, Vadose Zone J., 15, https://doi.org/doi:10.2136/vzj2016.04.0033, 2016.
 - Weissmann, G. S. and Fogg, G. E.: Multi-scale alluvial fan hetero-geneity modeled with transition probability geostatistics in a sequence stratigraphic framework, J. Hydrol., 226, 48–65, 1999a.
- 810 Weissmann, G. S. and Fogg, G. E.: Three-dimensional hydrofacies modeling based on soil surveys and transition probability geostatistics, Water Resour. Res., 35, 1761–1770, 1999b.
 - Weissmann, G. S., Zhang, Y., LaBolle, E. M., and Fogg, G. E.: Dispersion of groundwater age in an alluvial aquifer system, Water Resour. Res., 38, 1198, 2002.

Zektser, I. S. and Everett, L. G.: Groundwater resources of the world and their use, IHP-VI, series on groundwater 6, UNESCO, 2004.

37