

# Predicting the impact of spatial heterogeneity on microbial redox dynamics and nutrient cycling in the subsurface

Swamini Khurana<sup>1</sup>, Falk Heße<sup>2,3</sup>, Anke Hildebrandt<sup>2,4,5</sup>, Martin Thullner<sup>1</sup>

<sup>1</sup>Department of Environmental Microbiology, Helmholtz Centre for Environmental Research – UFZ, Leipzig, 04318, Germany

<sup>2</sup>Department of Computational Hydrosystems, Helmholtz Centre for Environmental Research – UFZ, Leipzig, 04318, Germany

<sup>3</sup>Institute of Earth and Environmental Sciences, University Potsdam, Potsdam, Germany

<sup>4</sup>Institute of Geoscience, Friedrich-Schiller-University Jena, Jena, Germany

<sup>5</sup>German Centre for Integrative Biodiversity Research, Leipzig, Germany

**Correspondence:** Swamini Khurana (swamini.khurana@ufz.de)

## Response to RC2

*We thank the reviewer for the positive feedback with respect to the relevance of the work for the readership of the journal, and for the constructive comments to make the work more accessible. We acknowledge that we have to provide an improved discussion with respect to the applicability of the results for real world systems and also at larger scales. We will update the manuscript addressing this gap, as well as addressing the specific comments made by the reviewer below.*

L1: The term "redox dynamics" is not used in the manuscript (except once when referring to the literature) and I am not entirely sure what the authors want to convey with it.

*We agree with the reviewer and propose updating the title to the following:*

*“Predicting the impact of spatial heterogeneity on microbially mediated nutrient cycling in the subsurface”*

L58: "in this microbial ecosystems...": what does "in this" refer to?

*‘this’ refers to biogeochemical cycles. We agree that the phrasing is ambiguous. We will rephrase the sentence as given below:*

*“...In these biogeochemical cycles, microbial communities play a key role ...”*

L78: "Sufficiently well" for what?

*We apologize for the ambiguous qualifier. We will rephrase this sentence as follows:*

*“Biogeochemical reaction networks have been explored extensively over the past decades with improvement....”*

L81: Please add citations for the statement on biogeochemical reaction networks.  
*Addressed.*

L81-82: It is not clear to me how the sentence starting with "Working with ..." fits into the line of arguments here.

*We agree that the phrasing is ambiguous. Here, we attempted to describe how biogeochemical reaction networks improved over the years, from only physical processes describing element fluxes to moving towards microbial explicit models. We have now rephrased:*

*“Incorporating microbially explicit reaction networks in reactive transport models ....”*

L84: "A straight-forward application of the soil-based biogeochemical model approaches to conditions in deeper subsurface compartments is problematic because the nature of carbon source changes as it travels into the deeper zones." I believe the authors did not specifically look into this- is there a reason for this?

*The reviewer is correct in that we didn't go into a detailed characterization of dissolved organic carbon in the deep subsurface as this was beyond the scope of the study, and studies such as Benk et al (2019) already explore this aspect. This section, on the other hand, primarily motivated the development and parameterization of a reaction network that is appropriate for deep subsurface oligotrophic environments. To develop a reaction network for the deep subsurface, we adapted conceptual approaches from numerous studies (L151-L163), and reparameterized the reaction network (described in Appendix A) which took into account slow reactivity of DOC for example.*

L99: I think using the term "mechanisms" is not great here as I think the manuscript does not address this.

*We acknowledge that the manuscript does not specifically address the mechanistic understanding microbial nutrient dynamics in the subsurface. This section primarily described the shortcoming of field scale studies, in that they do not provide insights into the mechanisms governing microbially mediated nutrient cycling in the subsurface. This statement, even though not addressed in the paper, is a drawback of field scale studies. This is overcome to an extent by pore-scale studies that we describe subsequently. Although we think that using “mechanisms” helps in the flow of argumentation of the text, we will remove it since the reviewer thinks otherwise.*

L196: "established": can you add references for this?

*We understand that the reviewer is concerned about using “established” as a qualifier. While variograms have been used and discussed in numerous works: (Gelhar and Axness, 1983; Johnson and Dreiss, 1989; Webb and Andersen, 1996, Berkowitz, 2002; Dagan et al., 2003; Delhomme, 1979; Heße et al., 2014), we propose to remove the qualifier as it may lead to some confusion. We will thus rephrase the sentence as follows:*

*“... To conceptualize heterogeneity, we used a limited parameter set, i.e., variance in the log normal distribution ....”*

L196: What are variance and anisotropy values used for the base case? When I first looked at Figs. S1 and S2, I was confused because the homogeneous base case was shown in all variance:anisotropy ratios.

*Homogeneous domain is the base case. Homogeneous domains are characterized by the same value of conductivity throughout the domain. Whereas variance and anisotropy are characteristics of heterogeneously distributed properties (hydraulic conductivity in our case). So, the base case (homogeneous domains) is characterized by the same value of hydraulic conductivity at each node (variance is 0, and there is no associated anisotropy).*

*We recognize the source of confusion as we have not explicitly stated this in the manuscript. To rectify this, we will revise the first entry in Table 2 (where the first row refers to the homogeneous domain). We will also update the caption of Fig. S1 and Fig S2:*

“

Figure S1: Flux averaged concentrations of dissolved species in heterogeneous domains (indicated by variance and anisotropy values in the row index) in three types of heterogeneous scenarios (solid lines) compared to that in the homogeneous base case (zero variance and no associated anisotropy, dashed-dot lines) in all flow regimes. The flux averaged concentration profile is the same for a given column (i.e., there is only one homogeneous/base case for comparison in each flow regime).

Figure S2: Spatially averaged concentration profile of the immobile active biomass in heterogeneous domains (indicated by variance and anisotropy values in the row index) in three types of heterogeneous scenarios (solid lines) compared to that in the homogeneous base case (zero variance and no associated anisotropy, dashed-dot lines) in all flow regimes. The spatially averaged concentration profile is the same for a given column (i.e., there is only one homogeneous/base case for comparison in each flow regime).

L250: Based on equation 2, the  $Da$  value should depend on the size of the domain relative to the size of the "heterogeneity". Have the authors looked into this?

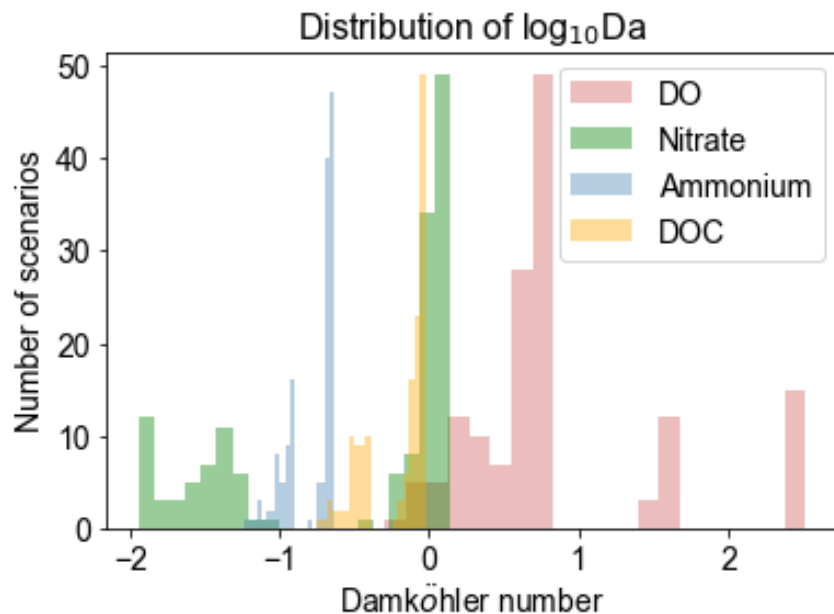
We understand that there is some confusion regarding the calculation of  $Da$  in each systems. We have expanded the section describing its calculation. We hope that it is more comprehensible and easier for readers to adapt for their own studies as well.

L312: "while the removal of TOC was the lowest there...": is this trend related to microbial biomass?

Correct. Please refer to Fig. S9 where we display that inactive mobile biomass contributes substantially to the total biomass, and it is considered in calculation of TOC (L238:241).

L391-392: The bars in Fig. S6 are not linked to redox conditions- is it possible to do so?

We assume that the reviewer is looking for a chemical species-specific distribution of Damköhler number like given below? We will add this as additional figure to the SI since we do not see a way to provide all information in one figure.



L441-451: Given that the prediction of the impact of spatial heterogeneity on redox

regimes is the major posit of this manuscript, I believe this section needs to be improved. A few suggestions for improvement are given in the following few comments.

L444: What is AIC and what do these values mean?

We refer to AIC in the Methods section (L245) as the Akaike Information Criterion (L245). It is an indicator of prediction error of a general linear model. It is a relative criterion commonly used to compare the performance of a collection of models. The model that has a lower value of AIC performs better. We recognize the requirement to explain it better and so we updated the Methods section (L245) with further information on it:

*“We compared the Akaike Information Criterion (AIC) of each model to evaluate the fit of the model. AIC is an indicator of prediction error associated with a general linear model. It is an indicator of relative performance of a group of models; the model with the lowest AIC is concluded to be the one with least prediction error or best performance. With each iteration of the model, we selected the features most influencing the performance of the model and reducing the AIC of the predictions. ...”*

*Additionally, we updated Section 3.5 to explain the results in a better way:*

“

*While conducting the multivariate statistical analysis of change in mass removal of reactive species, we made use of AIC to evaluate governing factors influencing mass removal in a spatially heterogeneous domain. The analysis indicated that AIC was 994 when considering only breakthrough time and chemical species. AIC reduced to -211 when the chemical species, the flow regime, variance in permeability field and the anisotropy of the domain were included as random factors). Please refer to Table S1 for further details. Thus, we concluded that nutrient dynamics is influenced by spatial heterogeneity. Categorizing the systems using  $\log_{10}Da$ , we proposed a linear expression to predict the impact of spatial heterogeneity on nutrient removal. The regression parameters informing this expression are given in Table 5. The results indicated that we may underestimate nutrient removal by 6 times or overestimate it by twice the amount (Fig. 5).*

“

L449: Where in Fig. 5 can I see these under-/overestimations?

*We agree that the caption of Figure 5 can be vastly improved to equip the reader to match the visualization with the text. The Y-axis in Fig. 5 displays these under-overestimations. We will rephrase the figure caption as follows. We hope that this explains the axes and the data referred to in the text better.*

*“Figure 5: Regression analysis: Predicting impact of spatial heterogeneity on chemical species removal in different reaction regimes indicated by  $\log_{10}Da$ . Value on Y-axis indicate the removal of chemical species in heterogeneous domains normalized by that in the corresponding base case. Spatial heterogeneity is plotted on the X-axis, indicated by the breakthrough time in the heterogeneous domain normalized by that in the base case (homogeneous domain). A value of 100% on the Y-axis indicates that the removal of the chemical species is the same as that in the corresponding base case (homogeneous domain). A value of 50% indicates that the removal of the chemical species reduced by half in the corresponding heterogeneous domain. A value of 600 indicates that the removal of the chemical species in the heterogeneous domain was 6 times that in the homogeneous domain.”*

L452: Does this Fig. include data for different reactive species? Also, why are Da numbers given in log10 base (given that per definition Da is already the ln of the concentration ratio between outflow and inflow)? If solutes are consumed in the domain, then  $C_{out}/C_{in}$  is always  $< 1$  and hence the Da per eq. 2 is negative- which will give a complex number when taking log10. Is there something I am missing here?

We understand that it was confusing to follow the calculation of  $Da$ . So we updated the Methods section with the following details. We also acknowledge that there was an error in the formula presented for calculating the  $Da$  earlier ( a factor of -1 was missing); it is now corrected in the updated section:

we used the Damköhler number ( $Da$ ) to indicate the reaction regime for each reactive species.  $Da$  is defined as the ratio of the advective transport time scale and the reaction time scale as described in Eq. 2.

$$Da = \frac{\tau_{transport}}{\tau_{reaction}}, \quad (2)$$

where,  $\tau_{reaction}$  is the characteristic reaction time scale and  $\tau_{transport}$  is the characteristic transport time scale given by the breakthrough time of a conservative tracer in the domain. We adapted this definition and used Eq 3 below to calculate the apparent  $Da$  using values estimable in the field when  $\frac{C_{out}}{C_{in}} > 5\%$ .

$$Da = -\ln \frac{C_{out}}{C_{in}}, \quad (3)$$

with  $C_{in}$  as flux averaged concentration of a reactive species entering the domain, and  $C_{out}$  as flux averaged concentration of the reactive species leaving the domain. In case of  $\frac{C_{out}}{C_{in}} \leq 5\%$ , we used Eq. 4 and Eq. 5 to derive the apparent  $Da$  of the chemical species

$$\tau_{reaction} = \frac{-\ln(0.37)}{-\ln(\frac{C_{y5}}{C_{in}})} \times \tau_{y5}, \quad (4)$$

$$\tau_{reaction} = \frac{\tau_{y5}}{\ln(\frac{C_{y5}}{C_{in}})}, \quad (5)$$

where,  $C_{y5}$  is the concentration of the chemical species at the first cross-section ( $y = y5$ ) when  $\frac{C}{C_{in}} \leq 5\%$ , and  $\tau_{y5}$  is the breakthrough time for a conservative tracer at the same cross-section, i.e.,  $y = y5$ .

$\tau_{transport}$  in this case was the same as the breakthrough time of the conservative tracer in the domain (Eq. 6).

$$Da = \frac{\text{breakthrough time}}{\frac{\tau_{y5}}{\ln(\frac{C_{y5}}{C_{in}})}}, \quad (6)$$

Thus, we were able to characterize reaction dominant system where  $Da > 1$ . We took the logarithm of  $Da$  to the base 10 ( $\log_{10}Da$ ) to characterize the regime for each reactive species in each domain.

For a scalable relationship addressing impact of spatial heterogeneity on reactive species removal, we conduct a simple linear regression analysis of species removal vs. residence time (both in relative units to the homogeneous reference cases) for different  $\log_{10}Da$  ranges.

“

L461: I think it would be beneficial to explicitly state the range of the scenarios.

We agree with the reviewer that it will be beneficial to refer to the range of scenarios that we have described earlier in the text. We will revise the sentence as below:

“ ...This approach allowed us to generate a wide range of spatially heterogeneous domains (with variance of the log normal distribution of conductivity varying from 0.1 to 10, and anisotropy varying from 2 to 10), which is not possible experimentally. ... “

L463: "correlation length": this is not discussed in detail in the results section and I am not sure how it fits in here.

We agree with the reviewer here. Since we didn't vary the correlation length (kept it constant in all the scenarios), we don't need to mention it in this section. We will remove it from the sentence.

L507-508: "We establish that the persistence of microbial species in the domain is governed by the presence of the appropriate carbon source and electron acceptor, ..." I

am wondering how microbial species are linked to carbon sources and electron acceptors in the model. Is species distribution independently modelled or could this finding in part be a result of the way the reaction network and model are set up?

*In Appendix A, we describe the rate expressions that we used in the reaction network. We adapted Michaelis Menten kinetics for the rate of microbial respiration (section A.3.1.) with carbon substrate concentration (DOC) and electron acceptor concentration (DO for aerobic degraders, nitrate for nitrate reducers and so on). Additionally, we link microbial growth (section A.3.2) with the rate of respiration, ammonium availability and yield coefficient. So microbial biomass is a result of the reaction network.*

*We did not explicitly specify the distribution of either chemical species or microbial species. The species distribution and results thereof that we present are at steady state conditions, which the model itself reaches given uniform species distribution as initial conditions. Thus, the distribution of species evolved due to the spatial heterogeneity of the domain.*

L627: Can the authors elaborate on the significance of this results for environmental systems? For example, when and where do they expect these heterogeneities to be most significant?

*We make a reference to the geological settings which can be represented using the simulated random fields in L464-L467. We further add scenarios in this passage where we expect the results of the studies to be applicable.*

“

*We expect advection dominated systems to be impacted by spatial heterogeneity because spatial heterogeneity had a higher impact on the transport profiles in these systems. These are typically systems that are shallow, less compacted (in case of alluvial sediments), or fractured rock systems. Furthermore, the shallow subsurface also receives bioavailable and reactive organic matter with the incoming water which enables a relatively high microbial activity. In contrast, in the deep subsurface microbial activity is lower and rather relies on inputs from the matrix material, which is ubiquitous and doesn't rely on transport for access. We expect additional studies exploring the impact of varying concentrations of chemical species, parameters relevant to these ecosystems or subject sites to add to the evidence generated by our study that the impact of spatial heterogeneity on subsurficial reactive systems may be predicted using field estimated indicators such as breakthrough time,  $Pe$  and  $Da$ .*

“

L654: Or underestimate nutrient removal six-fold as stated in L443?

*Yes, we agree that it may result in underestimation of nutrient removal. We discussed this in detail (L621:624), where we clarify that the 6-fold increase in removal is likely due to the small domain size. The low  $Da$  range refers to nitrate removal in the fast flow regime, and nitrate removal in the base case was  $\sim 1-2 \mu M$ . Thus, a 6-fold increase actually means that the removal increased to  $6 \mu M$ , which is still low in absolute terms. Thus, we do not make reference to it again in the Summary section (L654).*

Technical corrections:

L18: "used" instead of "undertake"

*Noted.*

L83: "group" instead of "groups"

*Noted.*

L115: "attempted" instead of "attempt"

*Noted.*

L123: "Disentangle" from what? I think describe/define would be better.  
*Noted.*

L271: I think the "removal" before "impact" should be deleted.  
*Noted.*

L424: Why are some words bold?  
*Noted: It was a formatting error.*

L446: Seems like a repetition of L 443f.  
*Noted. We propose to update the section as follows:*

*“While conducting the multivariate statistical analysis of change in mass removal of reactive species, we made use of AIC to evaluate governing factors influencing mass removal in a spatially heterogeneous domain. The analysis indicated that AIC was 994 when considering only breakthrough time and chemical species. AIC reduced to -211 when the chemical species, the flow regime, variance in permeability field and the anisotropy of the domain were included as random factors). Please refer to Table S1 for further details. Thus, we concluded that nutrient dynamics is influenced by spatial heterogeneity. Categorizing the systems using  $\log_{10}Da$ , we proposed a linear expression to predict the impact of spatial heterogeneity on nutrient removal. The regression parameters informing this expression are given in Table 5. The results indicated that we may underestimate nutrient removal by 6 times or overestimate it by twice the amount (Fig. 5).”*

L553-555: A verb is missing in this sentence.  
*This is possibly a confusion between “nitrate reducers” and “nitrate reduces”. We will rephrase.  
“It must be noted though that the concentration of nitrate decreases when and where the concentration of DO is below 15  $\mu M$  (Fig. S1).”*

L609: Delete "towards".  
*It looks like there is some confusion with the current phrasing of the sentence. We have rephrased this as follows and we hope this alleviates the confusion.*

*“For regimes where  $-1 < \log_{10}Da$ , 0, first order kinetics may be substituted with zero order kinetics”*

Tables: Captions should be above tables, not below.  
*Noted.*

Figures: It would be helpful to have different symbols for the three flow regimes so that the figures are readable in black and white.

*Noted.*