

*We use the following format for comments and responses:*

*R3C1 = Reviewer 3, comment 1. R3AR1 = Reviewer 3, Author Response 1.*

R3C1 - The manuscript "Organic matters, but inorganic matters too: column examination of elevated mercury sorption on low organic matter aquifer material using concentrations and stable isotope ratios" uses careful mercury breakthrough column experiments to determine the transport of mercury in low organic matter aquifer material. The experiments and analytical work are well done, and I commend the authors on their careful work. The information contained within the manuscript, particularly the role of inorganic materials in mercury transport, will be of interest to the wider mercury scientific community.

**R3AR1 – We thank the reviewer for their time and overall positive sentiments towards the manuscript.**

R3C2 - However, the lack details of the solute transport modelling methods raises more questions and concerns that need to be addressed, as detailed below.

**R3AR2 – We will attempt to address and allay these concerns below and within the resubmitted manuscript.**

### **Specific Comments**

R3C3 - The methods detail the experimental methods well but do not detail the fitting of the solute transport model to the data. The lack of such section (and associated results) causes confusion as the breakthrough curves presented in figure 2 are dependent on time, while the equation presented in Table S6.1 is not dependent on time. How did the authors fit the equation to the data? Was the advection-dispersion equation fit first using the chloride tracers to estimate conservative transport parameters (i.e., hydrodynamic dispersion of the media) then the Freundlich equation fit using the Hg data? Or were the presented equations in S6 fit to the data without considering flow (i.e., a statistical fit). If the latter, then the resulting parameters have no physical meaning and are only the summation of multiple processes occurring (advective transport, surface complexation, mineral (clay) matrix diffusion, reduction to Hg0, desorption, etc.), thus not representing the Freundlich isotherm as suggested in the text (this point also applies to the flushing analysis and the exponential decay function). I also note that the Freundlich equations presented in S6 deviates from the common form ( $q_e = K_f \cdot C_e^{1/n}$ ), where  $q_e$  is the solid phase concentration,  $K_f$  is the Freundlich equilibrium constant,  $C_e$  is the aqueous concentration and  $n$  is a fitting constant determined by linearizing the equation. There was no citation for the specific Freundlich equation used in the text, nor for the reasoning behind dividing the breakthrough curve into two separate analyses. The resulting units of the proposed curves are mg/L<sup>2</sup> but the Freundlich isotherm units ( $K_f$ ) are mg/g (mass/mass).

Fitting the advection-dispersion equation with a Freundlich or Langmuir (as discussed below) isotherm is straightforward in freely available programs such as RETC or Hydrus-1D. There needs to be a clear section on the modelling procedure, including assumptions and equations, since the results and discussion rely on these methods. This needs to be added to the main manuscript prior to publication.

R3AR3 – Firstly, we want to clarify that the curves presented are breakthrough curves and not sorption isotherms. So reference to sorption isotherms functions does not apply here. A sorption isotherm would plot the solid phase concentration against the concentration within the eluate. We cannot plot sorption isotherms using our data as we only have ~50% breakthrough and equilibrium solid-phase concentrations. We did not have sufficient mass of solid phase materials to conduct enough experiments to sacrifice sufficient columns during breakthrough to create sorption isotherms (we note again here that we used all available solid phase materials in these experiments; none of this material remains). We could create sorption isotherms theoretically based on what we would expect in the solid phase from the eluate concentrations (i.e., using data from Table 2). However, this would essentially be plotting the liquid-phase concentration data against itself (the “theoretical” solid-phase concentrations are derived directly from the liquid-phase concentration). This approach would add nothing to the study; and hence, we do not consider it valid to derive sorption isotherms for these experiments.

Nevertheless, the Freundlich and exponential decay functions (/models) were fitted empirically. The purpose of this was to create a fit that displayed the behaviour of the data (not attempting to define mechanism from the function). Considering the comments from reviewer 1 and 3, we are prepared to take two course of action: (1) simply remove both sets of functions and let the data itself display the “S-shaped” breakthrough profile and “exponential decay-like” desorption profile; or (2) leave the fitted functions in, and more clearly define that these are empirically (and not mechanistically) fitted functions that serve to assist the readers in understanding the data. To be very clear, we derive the mechanisms of sorption and desorption from the concentration data and the stable isotope data, not the fitted functions. We are happy for the associate editor to specify which method they deem most appropriate.

Our preferred option is option (2) (leaving the fitted functions in). With this option, we will update the first sentence of Section 3.1.1 and the caption of Figure 2 to note the functions are empirically fitted more clearly:

“As expected, the uptake of the HgCl<sub>2</sub> solution to the solid phase aquifer materials followed an S-shaped breakthrough curve best described by a Freundlich function (Figure 2; note these are empirically fitted functions).”

“...Sorption curves were fitted with Freundlich functions (blue dashed lines), and desorption curves were fitted with exponential decay functions (red dashed lines).

These functions were empirically (not mechanistically) fitted to the data as these plots are not sorption isotherms (see Section S6 for details of fitting functions)."

If the editor prefers option (1) (to remove the functions) we will remove the functions and references to the specific named functions. The reason that this is our less preferred option is that this is purely for naming purposes; this has no impact on our description of the sorption and desorption mechanisms (based on concentration and stable isotope data themselves). Using these empirically (well)-fitted functions ( $R^2 > 0.99$  in all cases) makes descriptions more efficient and the story more concise.

### Other Specific Comments

R3C4 - L27-28 The Freundlich model describes sorption not breakthrough curves, which is mediated by water flow. This needs to be more clearly stated here, perhaps adding "of the" in-between sorption and breakthrough.

R3AR4 – Yes, we agree. As we described in R3AR3, the Freundlich function was empirically fitted empirically and not for mechanism (not a sorption isotherm).

R3C5 - L61 comma splice after solubility. Regardless of this small grammatical issue, perhaps changing "have" to "having" will help the readability of the sentence.

R3AR5 – This will be changed to "...but are relatively stable ores that have very low solubility and bioavailability..."

R3C6 - L201-202 What is the total volume taken for analysis (was it the same as the waste 10mL or was the 15mL tube mentioned in 2.2.3 filled) and how does this compare to the total pore volume? If the ratio between the sample volume and the pore volume is large, then there needs to be some consideration that a sample taken at a given time represents a range of times rather than a specific time as assumed in the analysis. The larger the ratio the less precise your results. This comment also applies to the Hg isotope samples.

R3AR6 – 5 mL of sample was collected, which represents <25% of one pore volume. The following sentence was updated in Section 2.1 to reflect this:

"10 mL of eluate was allowed to flow off into a waste vessel before 5 mL of sample was collected periods for analysis (this applied to all analyses)."

We also note that we consider the different volumes sampled between columns as the x-axis uncertainty in the boxplot data (Figure 2).

R3C7 - L296-297 The definition of effective porosity is the proportion of total void space that is capable of transmission fluid under advective fluxes. In most cases,

this value is close to the total porosity but some media it can be quite a bit lower. Given your packed columns, I suspect that your effective porosity is closer to your total porosity as you defined. However, this assumption needs to be explicitly stated or measurements of effective porosity (e.g., soil air content at -100 mb) presented to confirm your assumptions.

R3AR7 – We added the sentence “effective porosity (...; assumed to be equal to total porosity)”

R3C8 - L301-302 Given the well-known soil texture and artificially packed nature of the columns, the effective porosity can be relatively accurately estimated using freely-available pedotransfer functions (Rosetta — ISMC (soil-modeling.org)). Such estimations would allow for KD to be estimated on all columns that achieved 50% breakthrough. These values can then be compared to your estimated values from columns C1.1-1.3 and C2.1-2.3.

R3AR8 – One of the major goals of this study was to simulate the original contamination to understand mechanisms of Hg dynamics in low OM aquifer materials and then derive Kd and Rd values from these column experiments for future modelling work (these experimentally derived Kd and Rd values are completely lacking). We deem the additional modelling scenarios suggested here by reviewer 3 outside the scope of this work and a part of the aforementioned future modelling work that applies the data obtained from these experiments.

R3C9 - L311 missing closing parenthesis.

R3AR9 – This will be fixed in the updated manuscript.

R3C10 - L303-316 How much Hg could be sorbed to the walls of the syringe? Would accounting for this improve your Hg mass balance (increasing eluate concentration from ≈91% for instance)? I note you discuss this briefly, but this may be worth a small batch experiment to see if the amount of sorbed Hg onto the plastic might be significant.

R3AR10 – As we have stated, there is no more of this aquifer material remaining. We had a limited supply of this material (it was donated to us from our colleagues who made the drill core), and it has been consumed. Hence, additional batch experiments with these materials is not possible. Simply doing this “batch” scenario without any materials (zero flow) is an entirely different scenario to a column of packed solid-phase materials with materials filling “roughened” surfaces of the syringes and impacting/limiting flow. We do not consider that such experiments would reveal any additional information that would benefit the study.

R3C11 - L364-365 The determination of theoretical max sorption condition suggest that the sorption characteristics can be fit with a Langmuir isotherm rather than a

Freundlich isotherm model. Based on the discussion and proposed multi-mechanism for Hg desorption (outer-sphere complexation vs mineral matrix diffusion), a multi-site Langmuir adsorption isotherm may best describe the actual processes rather than the simpler Freundlich isotherm. The authors seem to use a 2-site exponential decay function in the Desorption phase of EXP2 but no discussion of this is presented nor the rational. By using the Freundlich isotherm it is assumed that both processes that remove Hg from the liquid phase (complexation and mineral matrix diffusion) are occurring at the same rate and have the same potentials. The flushing phase of your column experiment and the isotopic results, with the current analysis, suggest that this is not the case. See Swenson and Stadie (2019, 10.1021/acs.langmuir.9b00154) for a good overview of Langmuir isotherms.

R3AR11 – We refer the reviewer to the explanation given in R3AR3.

R3C12 - L574 I appreciate the back-of-the-envelope calculations that really contextualize the magnitude of the potential Hg<sub>0</sub> production at the contaminated site, however there needs to be a more explicit description of the mathematics and the values used in the calculation either here or in the SI.

R3AR12 (Please note this is the same response provided in R1AR10) – Figure S9.6 and Table S9.1 will be added to the SI showing the details (including the percentage of the integrated area of the peaks; 0.11% Hg(0)) for Hg(0) and Hg(II) peaks from the mean PTD curves for all columns from EXP2. The text referring to the *back-of-the-envelope* calculations has been updated as follows:

"If we conservatively assume conservative values for a mean depth of contamination of 2 m (aquifer ≈3-4 m depth; Bollen et al., 2008; McLagan et al., 2022), mean THg concentration of 2 mg kg<sup>-1</sup> (solid phase THg concentration of 2-162 mg kg along the contaminated aquifer; Bollen et al., 2008), the fraction of Hg(0) produced per day is 0.01 – 0.001% of the THg (based off 0.1% Hg(0) peak integration of total peak area of mean PTD curve from EXP2; see Section S9) based off data from Bollen et al., 2008; McLagan et al., 2022), and the same bulk density and flow rates as in our experiments, we can produce a back-of-the-envelope estimate of the mass of Hg(0) produced and potentially lost from the aquifer to overlying soils."

R3C13 - L592 Given the reactive transport focus of the paper, I suggest replacing "three-dimensional spread" to the more appropriate terms "longitudinal and transverse dispersion".

R3AR13 – This will be updated in the next version of the manuscript.

R3C14 - L598 I really like this point!

R3AR14 – We thank the reviewer for their kind words. We do believe this also links directly to the next comment (R3C15).

R3C15 - L600-615 The lower, or even different, KD for the higher concentration EXP2 is somewhat surprising, as the KD is the linear partitioning coefficient, which assumes equilibrium between the liquid and solid phases. Assuming there is no saturation of adsorption sites, KD should be close if not equal in both experiments since the materials, packing, and flow rates are the same. However, there was no explanation of these values or their differences beyond stating other literature values. Are the differences in KD due to slight variations in clay content (experimental error) or is there another explanation? I suggest that there needs to be a bit more discussion here to explain these values more mechanistically.

R3AR15 – We did provide an explanation for the difference in Kd values between EXP1 and EXP2:

"The difference in RD and KD values between EXP1 and EXP2 (Table 3) indicate stock solution concentration is a factor in the transport of mercury within these columns. The elevated stock solution concentrations may be undermining the assumption of equal accessibility to sorption sites; REDUCED KINETICS AT HIGHER CONCENTRATIONS (USEPA, 2004)."

The part in ALL CAPS will be added for clarification purposes within these sentences.

However, as we note it was our intention to simulate the original contamination, and we would also refer to the reviewers support of our statement (R3C13) in L598:

"Considering the high concentrations of Hg that have been observed within this and other Hg contaminated aquifers (Katsenovich et al., 2010; Lamborg et al., 2013; Demers et al., 2018), it is critical that we do not isolate our study of Hg transport dynamics to low concentration experiments that meet assumptions for theoretical sorption (batch and column) experiments."