Response to Reviewer Will Wolheim –

We thank the reviewer for their comments and constructive feedback. We respond directly below each comment (Italic) in normal font.

WW2. 1: *I appreciate the addition of an example of how the curvature parameter (now called Curvmax) was calculated (Figure B1). Based on this example, it looks like a linear relationship (in log-log space) would be adequate. Why add the complexity of more degrees of freedom? But that would also mean no curvature. It seems like a metric like AIC would help identify what the best fits are for each site. Given the fits in this figure, how does Local Curvature vary so much in opposite directions and sign? It was unclear which of the fits was used to identify Curvmax in this particular example. If the estimate is so variable, it seems like the noise in the data will contribute greatly to a Curvmax. An objective approach (like AIC) for selecting curvature should be used. This should be very clear in the ms given this is a new approach that is being argued will be helpful.*

We thank the reviewer for these remarks. Several studies (eg. Moatar et al., 2017; Diamond and Cohen, 2018 and Marinos et al., 2020) show that characterizing NO3- log(C)-log(Q) relationships as linear can imply an information loss (l. 88). For example, Moatar et al., 2017 found that 44% of the investigated catchments have higher low flow slopes than high flow slopes for the NO3- log(C)-log(Q) relationship. In Marinos et al., 2020 NO3- log(C)-log(Q) relationships are more accurately represented with a piece wise regression model than a linear regression for 32 of 33 study sites. Based on these and other studies we argue that more degrees of freedom are necessary to characterize NO3- C-Q behaviour in the log-log space.

The Curvmax metric, introduced in this study, complements rather than replaces the established linear regression model, as it allows to quantify the shape of log(C)-log(Q) relationship without the assumption of a fixed form (l.117-121). The Curvmax metric can theoretically range from 0 to + or – infinity, with CurvMax= 0 indicating no bending in the log(C)-log(Q) relationship. Because we study the degree of (non-)linearity rather than choosing between a linear or non-linear model, AIC would not be helpful here.

We explain in l. 124-134 of the manuscript how CurvMax is calculated: "A smoothed spline, f, is iteratively fitted with increasing degrees of freedom (df) to capture the general log(C)-log(Q) shape accurately but avoid overfitting (Fig. B1). Initially, df = 3 and the log(Q) region of the largest instantaneous change is identified as Q_m±0.05 with $Q_m = arg\max_{\log Q}|f"|$. Then, df is increased until, at df=i, the log(Q) corresponding to the largest instantaneous change is not within the initial Q_m region anymore. Consequently, *Curv_{max}* is calculated for a smoothed spline fit, f, with df = i-1 as $\begin{cases} \max_{\log Q} f" \ if \left|\max_{\log Q} f"\right| \geq \left|\min_{\log Q} f"\right| \\ \min_{\log Q} f" \ if \left|\max_{\log Q} f"\right| < \left|\min_{\log Q} f"\right| \end{cases}$." Because of the nature of the smoothed spline fits, local curvature can vary, especially with higher degrees of freedom. We added a sentence in the caption of Fig. B1 explaining that df=5 was chosen as the final fit.

We examined the robustness of the CurvMax metric by selecting subsamples of observed noisy C-Q data from 444 French stations (Dupas et al., 2019) without replacement but with overlap (l.143-149). The results of this assessment in Fig. B4 shows that Curvmax tends to be robust (Sect. 3.1).

WW2. 2: *I think the results also indicate, as discussed by the authors, that linearity may occur (no bentness), across a wide range of uptake velcoity values (Figure B6). Figure B6 suggests most logC vs. LogQ relationships will likely be linear or close to linear when using observational data sets, which likely have enough noise to make it hard to distinguish the slight curvature that may occur. Thus, bending may not be observed, even if network scale uptake is high. The key for determining whether network scale uptake is high is knowing what the loading (vf=0) scenario is. Use of conservative tracer (chloride or specific conductance) may help with this interpretation (if you have those in your observational data sets).*

We thank the reviewer for this comment. Figure B6 mainly shows that the Curvmax metric can remain nearly constant under varying vf values if vf is not close to 0. We showed in Sect 3.1 and Fig. B4, where we calculate CurvMax for NO3- log(C)-log(Q) relationships of 444 French stations that Curvmax ranges between -5.25 and 3.88 (l.353). Moatar et al., 2017 found that for NO3- 44 % of the studied catchments can be classified as bended and 33 % as linear. If in our case a slight curvature occurs we would capture a CurvMax close to 0 thereby not classifying the C-Q relationship as bended or linear. We found that 77 % of the French stations have CurvMax ≤ 0 which is similar to the findings of Moatar et al., 2017. We acknowledge that conservative tracers can help to interpret CurvMax which can be an interesting approach in future studies. We added this aspect to the manuscript in the outlook in l. 718.

WW2. 3: *I think more succinct discussion is warranted on what circumstances (across the vf, aw, watershed size and other parameter space) curvature could be evident in observational data sets. I realize the paper is about bentness, but in the end that is just an abstraction. Ultimately bentness will be helpful to understand when river network scale uptake is important, and clearly bentness alone is insufficient for those purposes. I think the regression tree is meant to address this, but I find it very confusing to understand. Perhaps another approach or conceptual figure would help.*

We like to thank the reviewer. To solve this concern we would need more direct measurements of uptake together with observations of C and Q to make this linkage not only in the model but also in the data-reality world. However this would be outside of the scope of this paper where we offer a first insight in what can cause C-Q bending, using conceptual modelling. We indeed use the regression trees as well as Fig. 6 and Fig. B7 to link curvmax values to other parameters. To help with the interpretation of the decision tree we added a sentence in the caption stating that the trees are read from top to bottom, following the binary splits.

WW2. 4: *I am still puzzled by the results presented in Figure 2, even after reading the response to reviewers about my earlier question. The percent removed is very low (<5%) throughout the time series. The "no uptake" scenario results in C ~3mg/L, and observations fluctuate between 0.5 and 3mg/L. Based on this, I would expect to see removal fluctuating between 0 and 85%. And again (as with Figure B6), given the noise in the data in Figure 2b, would a nonlinear fit be selected over a linear fit using something like AIC to select the most parsimonious model in the lower panel?*

We thank the reviewer for this comment. The concentration time series in Fig. 2a is observed at the outlet of a catchment. We agree that if we would consider the percentage load removed solely at the outlet grid cell, removal would fluctuate between 0 and 85%. However, the percent load removed is simulated at the network scale. Which means it considers the incoming and outgoing loads at each network gridcell and is displayed in Fig. 2a as a median value. As we do not choose between linear or non-linear models but rather investigate the "degree of non-linearity" in this study, computing AIC would not help here (also see WW2. 1).

WW2. 5: *Regarding the use of uptake velocity in the analysis (comment Ed. 6) I strongly agree that it is the correct term to use. It is equivalent to piston velocity for gases, or settling velocity for sediments.*

We thank the reviewer for reaffirming this.

WW2. 6: *I think some of the other reviewers comments are valid regarding the density of results being presented. There is also a lot of information in some of the figures, and it is hard to identify the main the result that should be gleaned from it (Figure 3). More statements in the text guiding the reader as to the gist of the results would be helpful, especially in the topic sentence of paragraphs, to better tell the story, and make the paper more influential. The discussion in 3.3 in particular gets really difficult to follow.*

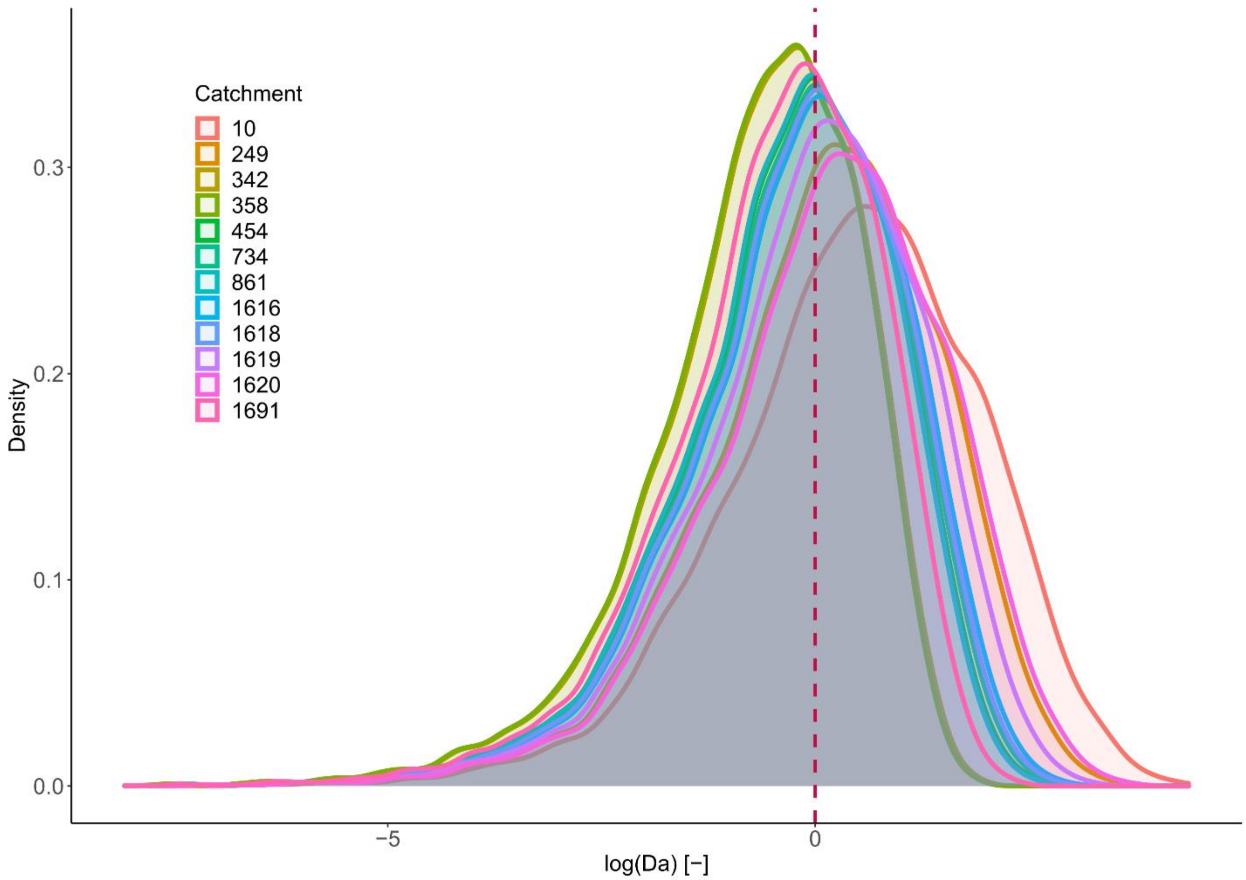
We thank the reviewer for helping to improve the readability of the manuscript. We revised the discussion in Sect 3.3.

WW2. 7: *What is the difference between the Figure B* series and the appendix figures? It would be simpler to have all in appendix for navigation purposes*

We thank the reviewer. We follow HESS guidelines for manuscript composition here: "Additional figures, tables, as well as technical and theoretical developments which are not critical to support the conclusion of the paper, but which provide extra detail and/or support useful for experts in the field and whose inclusion in the main text would disrupt the flow of descriptions or demonstrations may be presented as appendices. These should be labelled with capital letters: Appendix A, Appendix B etc. Equations, figures and tables should be numbered as (A1), Fig. B5 or Table C6, respectively. Please keep in mind that appendices are part of the manuscript whereas supplements (see below) are published along with the manuscript." We would therefore keep the Appendix figures (B-series) in the Appendix and the supporting figures in the supporting information (S-series).

WW2. 8: *L497. Most Da do not seem to be around 1.*

The Monte Carlo output of log(Da) is displayed for each of the catchments below. Here we see that the simulations are distributed around Da equal to 1 (log(Da)=0). We added this plot to the supporting information in the revised manuscript as Figure S3.



WW2.9: *L546. I don't agree with the statement "lower Curvmax" is related to high Lr.percent. It implies this condition is needed. If uptake velocity is very high, there is no bending, yet network scale removal can be very high. Same with L585, vf is clearly also important. Please modify wording.*

We agree with the reviewer that this wording can be misleading. In the first case we changed the sentence to "…lower curvmax can be related to higher Lr.perc" in the second case we have changed the verb to 'correlate'.

WW2.10: *Conclusions. I recommend making clearer what the take home messages are, rather than bringing in results again. E.g. L689, I think the main point is that bending is only evident under certain geomorphological conditions. I also disagree with statement under this bullet that vf does not influence bending. Curvmax=0 both when there is no uptake, and when uptake is really high (vf high), meaning you can only see bending at intermediate vf. In bullet 4, I think the takehome should be that Curvmax helps with estimating Lr.perc, but only under certain conditions.*

We thank the reviewer for these remarks. We state that bending is shaped primarily by geomorphological parameters while vf has a secondary influence. In Fig. 6 we demonstrate that if we select simulations corresponding to ranges for low, medium and high vf the associated distributions of CurvMax do not differ much and vice versa. Which means you can see bending for low, medium or high vf. We agree with the suggestion of the reviewer on the take home message of the 4th bullet point and included this sentence accordingly in the revised manuscript.

WW2.11: *L707. I did not find interpreting the Cart tree simple.*

We changed this sentence in the revised manuscript to remove the word “simple”.