Response to reviewer #2:

General remarks:

1. The proposed methodology have a significant interest among the present challenges in karst hydrology for flow and transport modeling. The manuscript is of interest for the scientific community. Nonetheless, I found some important methodological aspect that, in my point of view, should be addressed before potential publication.

Reply: We thank the reviewer for the helpful comments. See below our replies to the specific comments.

Specific comments to the paper:

1. Line 62: the role of the pore connection should be also mentioned, porosity is an important factor in flow and transport as well as the continuity between pores. This notion should appears in the introduction, as well as tortuosity, which then used in the proposed methodology.

Reply: We agree. We revise the text accordingly: *"Furthermore, the connectivity of the different porosities often results in a fracture-cave network, which dominates the flow structures in karst systems (Zhang, 2022)"*. The following was added to the reference list in the revised text: Zhang, X., Huang, Z., Lei, Q., Yao, J., Gong, L., Sun, S., and Li, Y.: Connectivity, permeability and flow channelization in fractured karst reservoirs: A numerical investigation based on a two-dimensional discrete fracture-cave network model, Adv. Water Resour., 161, 104142, https://doi.org/10.1016/j.advwatres.2022.104142, 202

2. Line 104: "well-defined spatial catchment" sounds unclear, are the boundaries well defined with the topography and/or geological setting?

Reply: We revise the text to clarify this point: *"This high-alpine karst system has been thoroughly studied and offers a catchment with a well-defined spatial boundary."*.

3. Line 117 "ansatz" is a German word, this appears several time across the manuscript.

Reply: Ansatz is a word in German, but it has been extensively used in the context of physical and mathematical problems and is well-accepted as an international term in the scientific literature.

4. Line 117: “water parcels” the notation sounds unclear is it like “water bodies” between flow zone and dead zone or is it a quantity of water that should be considered as an analogy with an amount of particle ? This point can be clarified in the paragraph.

Reply: Our definition of parcel, in this context, is the classical definition of a fluid parcel in continuum mechanics (Lagrangian approach). We edit the text to reflect this: *“(i.e., infinitesimal volumes of water)”*.

5. Line 123: “water parcels” and “tracer parcels” sounds unclear, a more detailed explanation might help the reader to understand how the analogy is done in the manuscript.

Reply: We revise the text accordingly to clarify this point (see also comment #4 above): *“The ansatz asserts that the accumulation and release of water parcels in the various volumes in the karst system resemble the accumulation and release of "parcels" of a chemical tracer (i.e., infinitesimal volumes of tracer) over time in a porous medium.”*.

6. Fig 1. (c) and (d) does the volumetric discharge considers a constant concentration and does the concentration variation assumes a constant discharge ? In my point of view, discharge and concentration cannot be separated so the two might appears on the graph, then the final time series that the author are using will appears more clear to the reader. At some point, the present version bring some confusion with the widely used tracer BTC.

Reply: We present the analogy of measured discharge and measured tracer concentration separately as each of them can be measured and presented as a time series. The

clarifications made in the text for comments #4 and #5 (see above) and #7 (see below) will help clarify this point.

7. Line 172: "C(t) is reinterpreted as translated to a water volume" please clarify the underlying hypothesis to move from concentration curve to fluxes curve.

Reply: We revise the text to clarify this point: *"The concentration at a specific point is analogous to the moisture content, and the classical C(t) breakthrough curve is analogous to the (volumetric) amount of water per time reaching the domain outlet (or measurement plane)."*.

8. Fig 3. I don't think we can consider the fit as satisfactorily. The data do not exhibit a normal distribution shape, also the fitted distribution provides negative distances. The fitted distribution should then at least be a truncated distribution to avoid negative values.

Reply: During development of the simulation, different distributions were examined and the end results were similar. In our findings, we concluded that a normal distribution describes the statistical distribution of distances in this case quite well. We revise the text to address this: *"A normal distribution was chosen as a simplified representation of the distance distribution; preliminary simulation results were similar for different skewed distributions."*. Note that we do set negative values to zero as stated in the text: "(physically unrealistic, negative sampled values are set to 0)".

9. Fig 5. How is the curvature of the likelihood distribution fixed? Is it a working hypothesis or is it derived from somewhere else?

Reply: This is an example for SF_I=0.01% as stated in the text. We edit the revised caption as it was indeed missing and might be confusing for the reader.

10. Line 283: ‘optimization is achieved’ How is the optimization performed? Which methodology/algorithm? I would recommend also to write “model parameter estimation” rather than “optimization”.

Reply: We revise the text accordingly: *“Given the presence of multiple model parameters (refer to Table 1), optimization is achieved by applying a bound constraint version of the Matlab fminsearch function (D’Errico, 2024) to minimize the Root Mean Squared Error (RMSE) between observed and simulated discharge using different combinations of parameter values. The 2016 dataset was first utilized for model parameter estimation”*.

11. Line 293: Why using NSE values that tends to favorize large values while the focus seems to be on tailing? Other performance criteria for model evaluation would be more suitable, or another option would be to compute NSE on a variable transformation such as 1/Q or squared_root(Q)

Reply: While the focus of this study was indeed fitting the tails, this is the first implementation of the CTRW-PT for modelling karst aquifer discharge. As such, we included a performance criterion that is widely used in hydrological studies (and was specifically used for the karst system in question) to support the overall validity of our model. We have added another performance criterion, often used in KarstMod and revise the text accordingly (The BE scores were added to the figure captions.): *“The Nash-Sutcliffe efficiency (NSE) and modified balance error (BE) were calculated for the optimized simulations, as a measure of the goodness of fit. The NSE and BE are the performance criteria utilized, for example, by the widely used KarstMod software (Frank et al., 2021). They are defined as the normalized variant of the mean squared error and the relative bias of the simulated and observed flow durations, respectively:*

$$NSE = 1 - \frac{\sum(x_s(t) - x_o(t))^2}{\sum(x_o(t) - \mu_o)^2} \qquad (12)$$

$$BE = 1 - \left|\frac{\sum(x_o(t) - x_s(t))}{\sum x_o(t)}\right| \qquad (13)$$

”. We also revise the text in the methods to clarify to the reader that other criteria exist as well: *“However, it should be noted that the NSE has limitations when there is large variability in the data, and in some cases other performance criteria may be more relevant*

for different datasets (see Cinkus et al., 2023a for a comparison of different performance criteria)". Please note that Cinkus et al., 2023a was missing from reference list and was added to the revised text as well: Cinkus, G., Mazzilli, N., Jourde, H., Wunsch, A., Liesch, T., Ravbar, N., Chen, Z., and Goldscheider, N.: When best is the enemy of good - critical evaluation of performance criteria in hydrological models, Hydrol. Earth Syst. Sci., 27, 2397–2411, https://doi.org/10.5194/hess-27-2397-2023, 2023a.

12. Table 1: Please add the min and max of the investigated parameter space for the model parameter estimation. Also, please give more information about how the optimized value is estimated (see one of the previous comment) + the estimated tortuosity appears pretty high compared with the literature (e.g. reference below) could you discuss a little on that ? Is it a realistic value for the catchment? Jouves, J., Viseur, S., Arfib, B., Baudement, C., Camus, H., Collon, P., Guglielmi, Y., 2017. Speleogenesis, geometry, and topology of caves: A quantitative study of 3D karst conduits. Geomorphology 298, 86–106. https://doi.org/10.1016/j.geomorph.2017.09.019 Collon, P., Bernasconi, D., Vuilleumier, C., Renard, P., 2017. Statistical metrics for the characterization of karst network geometry and topology. Geomorphology 283, 122–142. https://doi.org/10.1016/j.geomorph.2017.01.034

Reply: We investigated a wide range of the fitting parameters and prefer to focus Table 1 on the identified results and not on the optimization process; we therefore decided not to include the min-max of the parameters. An explanation of how the optimization was achieved was added in the revised text (see comment #10).

We agree that the estimated tortuosity in the system presented may be considered slightly high, but it is still within a reasonable range for a karst system; see, for example, Assari and Mohammadi (Assari, A., Mohammadi, Z. 2017 Assessing flow paths in a karst aquifer based on multiple dye tracing tests using stochastic simulation and the MODFLOW-CFP code. Hydrogeo. J. 25 (6). 1679-1702), which describes values between 1.1 and 3.9. Indeed, in some of the cases the tortuosity is higher than 1.6 in the references provided by the reviewer. Therefore, we can consider that the karst system described in this study is

rather tortuous. Furthermore, in many karst papers such as those noted by the reviewer, tortuosity is calculated for the branch scale, while in our model the tortuosity is a catchment scale factor. In the case of applying the CTRW-PT for other karst systems, the fitting can benefit from a more detailed characterization of the geometry of the entire system. We revise the text to explain this point: *"The optimized tortuosity factor of 1.6 found for the Disnergschroef system is somewhat higher than that found in some cases (~1.2-1.4, e.g., Jouves et al., 2017; Collon et al., 2017), but well within the range (1.1-3.9) reported for karst systems (e.g., Assari and Mohammadi, 2017). The higher value can be attributed to the morphology of the specific system, and also to the fact that while tortuosity is often calculated at the cave branch scale (e.g., Jouves et al., 2017; Collon et al., 2017), the CTRW-PT model uses a catchment scale tortuosity factor. The variability of tortuosity in different karst morphologies should therefore be recognized when considering different modeling scenarios."*. Please note that the two papers suggested by the reviewer and the Assari and Mohammadi (2017) reference were added to the reference list in the revised text.

13. Line 324: Is it correlation coefficient or correlation pic derived from cross correlation function analysis? What is the time lag considered for the correlation coefficient or what is the lag response obtained based on precipitation-discharge cross correlation function?

Reply: The correlation coefficients considered here are without any time lag, as our aim is to show the variation between the three weather stations. They are different for the three stations although they are located a few kms away from one another.

14. Line 328: This sentence is not useful. I recommend to delete.

Reply: We agree and remove this sentence in the revised text.

15. Line 374: As previously mentioned in my comments, other performance criteria would be more suitable to evaluate the predictive performance of the model regarding high flow and low flow periods respectively. Computing NSE by removing the high value is not a suitable justification for the model improvement on low flow period. Indeed, by skipping the high discharge value to compute the NSE you are changing the benchmark of NSE (the mean of observed time series) so comparison of NSE is not straightforward in that case.

Reply: In light of the comment made by the reviewer, we removed the NSE calculation for the low flow data. Furthermore, we omitted this comparison from the manuscript as it does not help to convey the point of the CTRW-PT fitting of the long tails of the data. Regarding the choice to use NSE, see #11 above.

16. Line 377: Are there some evidence of piston effect with temperature and/or conductivity?

Reply: Yes, and in the following paragraph we state that Frank et al (2021) have reported this.

17. Fig 8. Is not very informative as "sensitivity analysis". Including a sensitivity analysis on the model parameter regarding the model performance would be much more informative. Among the model parameters which one are the more sensitive in the model?

Reply: The figure presents a sensitivity analysis on the new parameters introduced to the model in this study. There are many papers that have already studied sensitivity analysis of different CTRW parameters, but none have done so for the new parameters (as they are new). We edit the caption in the revised text to convey the exact parameters shown in the sensitivity analysis.