Author response to Referee #2 comments:

We thank reviewer 2 for the detailed and careful review of our work. We hereby provide our point by point responses how the comments by referee #2 will be addressed in the revised manuscript.

Best,
Hanwu Zheng (on behalf of all co-authors)

Anonymous Referee #2
This manuscript applies a large-scale tracer-aided modeling (TAM) approach to disentangle ecohydrological processes in the heavily managed Middle Spree catchment (MSC), Germany, an evapotranspiration-dominated region facing strong anthropogenic pressures. By integrating stable water isotopes ($\delta^{18}O$ and $\delta^{2}H$) with streamflow into the distributed STARR model and calibrating with a multi-objective NSGA-II algorithm, the study evaluates runoff generation, groundwater contributions, and evapotranspiration (ET) partitioning across four sub-catchments (Berste, Wudritz, Vetschauer, Dobra). The key contribution lies in showing how streamflow–isotope trade-offs emerge as diagnostic signals of epistemic errors from unrecorded human impacts, such as irrigation or mining legacies. While isotope inclusion sometimes reduced discharge simulation performance, it significantly improved process representation such as subsurface mixing. Overall, the study demonstrates that even sparse seasonal isotope datasets can provide critical constraints in TAM for complex, human-altered hydrological systems, offering new insights into ecohydrological partitioning and informing future water management under anthropogenic and climatic pressures. From a reader's perspective not deeply familiar with isotope tracer methods, I have several comments and suggestions for clarification.

Reply: *We thank reviewer 2 for the detailed and careful review of our work and acknowledgment of the novel contribution showing the value of even coarse isotope data for insights into ecohydrological functioning.*

Points for the Authors to Consider
1.Clarifying the added value of isotopes
The added value of incorporating isotopes over other hydrological variables remains somewhat unclear. For instance, while the introduction emphasizes human influences, isotope integration did not appear to improve the model's ability to capture these anthropogenic effects, which raises questions about the practical contribution of isotopes in this context.

Reply: *Thanks for the comments. The human influences clearly increase complexity of processes representation in the catchments investigated, and we emphasise that the conflicts between streamflow and isotopes in the calibration could be the results of these human factors. We still provide a valuable qualitative way to capture these anthropogenic effects, and other factors, such as model system errors and uncertainty of input dataset, could also be part of the reasons for the conflicts, which were discussed in the discussion. The main value of using isotopes in this context is that isotopes contribute insights into pathways, storage and ages of ecohydrological fluxes EVEN in such heavily impacted systems. In our revision, we will clarify the value of incorporating isotopes in the introduction and discussion section. We would also argue that isotope integration did not improve the model's ability to capture these anthropogenic effects by much IS actually an important and novel contribution – as to our knowledge not many people if anyone has shown this so clearly before. We are aware that human influences may be masked in discharge-only calibrations, since these usually yield seemingly acceptable performance due to the large freedom of these models. The slightly decreased model performances after integrating isotopes suggest potential anthropogenic effects: water withdrawn during summer led the model to re-present faster runoff processes which contradict the longer flow path indicated by the observed flattened isotope variations. Additionally, the inconsistent ET partitioning between the model and RS products indicates the stronger fractionation processes, possibly due to irrigation processes. These conflicts itself provide deeper insights to the catchment functioning* *and we will make it clearer in the revision. We will include such argumentation in the discussion of the revision.*

How would the results compare if ET data were used in a multi-objective calibration of the STARR model?
Reply: *Isotope variations in the catchment are primarily governed by mixing and fractionation processes, which is why isotopes have this unique role in the calibrations. First, isotopes are particularly useful for ET partitioning, whereas there is no clear evidence that ET estimates would provide a similar contribution in calibrations. Second, isotopes can effectively constrain runoff partitioning. Remote sensing ET potentially could provide information on temporal and spatial pattern over the whole catchment, and potentially better constrain the equifinality in some processes, as such a large isotope dataset for calibration are rarely available. We will add this in the discussion. Since there is a large amount of literature regarding the calibrations based on (mainly remote sensed) ET, we will extend our discussion in this aspect.*

Could the process descriptions be refined to more clearly illustrate the unique role isotopes play relative to other potential data sources?
Reply: *Yes, sure. We pointed out that the ability of isotopes lies in constraining water partitioning, in addition to providing information on flow paths and storage dynamics. This can not usually be replicated by other datasets, such as ET, soil moisture or discharge. We will highlight this unique role by comparing with other potential data sources in the discussion.*

2.Improving figure clarity and linkage to discussion
Figures 5–8 combine multiple dimensions (temporal, spatial, and calibration metrics), making them information-rich but sometimes challenging to interpret. The figure captions and related explanations in the text could more directly highlight the core message of each figure. Including a short statement of motivation or the specific hypothesis addressed by each figure would help guide readers and improve accessibility. Moreover, because the figures are complex and the key messages are not always clearly highlighted, the subsequent discussion section becomes less convincing. Readers may find it difficult to fully trust the discussion, as the results and the interpretations are not always tightly aligned. Strengthening the clarity of figures and explicitly linking their core findings to the corresponding discussion points would improve the manuscript's overall persuasiveness.
Reply: *We apologise for the confusion the figures caused. Including these short statements of motivation for each figure is a great suggestion. We will improve clarity of figures and clearly explain the information contained in the figures. We will also refer more to the specific figures in the discussion section.*

Specific Comments
Lines 127 and 140: Please clarify the meaning of SE and m.a.s.l.
Reply: *Sorry for the confusion, "SE" is southeast, while "m.a.s.l" means meters above sea level. We will clarify these abbreviations accordingly*

Lines 240–243: Rainfall inputs are provided at daily resolution, whereas precipitation isotope inputs are monthly. How does this temporal inconsistency affect the results, and is this assumption reasonable?
Reply: *We actually pointed out the resolution inconsistency and the temporally coarse resolution could bring uncertainties. However, daily rainfall inputs but with monthly precipitation isotope inputs are quite common in the isotope-aided modelling works, this inconsistency is mostly inevitable due to data limitations. We are confident the monthly isotope data are still valuable as one of the key characteristics of isotopes is their integrative character, i.e. integrating signals on functioning over time (and space).*

Lines 249–251: Although a citation is provided, the manuscript would benefit from more detail on the isotope observations. Were these instantaneous grab samples, or integrated/accumulated values?
Reply: *They were instantaneously collected "grab" samples. We will add information in the methods on how we conducted sampling and processed the data.*

Table 3 (Scheme 1): Please clarify whether the calibration was performed jointly across all basins, or if each basin was calibrated independently.
Reply: *The scheme1 was calibrated jointly across all basins, and this was explained in lines 294-295.*

Figure 3: Why are only δ^2H time series presented, while $\delta^{18}O$ observations and simulations are not shown? It would also help readers unfamiliar with isotope applications if key concepts such as LMWL and VSMOW were briefly explained.
Reply: *Since the ^{18}O and ^2H have the similar fractionation and mixing processes, normally we use only one variable to constrain the model and just show the variation of the used parameter. We will explain LMWL and VSMOW in the caption of Figure 3.*

Figure 4: KGE is used for isotopes and NSE for streamflow. Why not use the same performance metric for both, to improve comparability?
Reply: *Sorry for the confusion, since we had only seasonal isotope data and we wish to reduce the potential bias brought by difference at single data point, we used KGE for isotopes. This is widely reported in the literature (i.e. that KGE is a better performance measure for isotopes and their dynamics). Interestingly, in our case, both KGE or NSE showed not too much difference and we used both (but we had not reported this). We will replace NSE with KGE in the calibration in the revision to exclude any potential misleading impacts from NSE.*

Table 4: The description of Table 4 appears in the first paragraph of the Results, though the table is first referenced in Section 3.2.2. Consider relocating the description for consistency.
Reply: *We agree and will correct it accordingly.*