Author's response to the Reviews on the Manuscript "Hysteresis between groundwater and surface water levels indicates the states of hydrological turnover affecting solute transport and redox processes" (egusphere-2025-1674)

Dear Dr. Hongkai Gao,

Thank you very much for the constructive handling of our manuscript and for providing the reviewers' comments. We appreciate the generally positive assessment as well as the thoughtful and detailed feedback from both reviewers in this round. In particular, we acknowledge the additional conceptual insights provided by Reviewer 3.

We have carefully considered all comments and concerns raised, and we have revised the manuscript, accordingly, providing detailed explanations and implementing all necessary changes.

<u>Response to Reviewer #2:</u>

We thank Reviewer 2 for the careful reading of the manuscript and the helpful comments, which improved the clarity and presentation of the study.

Specific comments

> ***- Table 1: You need to cite it in the text. I observed that for some events, total precipitation was very low (<5 mm) and durations were long (>24 hours), yet they led to relatively high runoff coefficients (e.g., the first event). Do you have any explanations for these unexpected values? Perhaps such runoff volumes are associated with wet antecedent moisture conditions (this information could be added to the table). Furthermore, I think you computed the runoff coefficient as the ratio between the runoff volume/stormflow (numerator) and the precipitation (denominator); please revise the caption, as it is currently unclear.***

We thank the Reviewer for pointing out this issue and for drawing attention to the events with low total precipitation but comparatively high runoff coefficients. We agree that these values appear unexpected. In accordance with the Reviewers concerns we double checked our baseflow separation. We now have changed to a new baseflow separation filter method (Lyne and Hol- lick,1979) instead of utilizing a graphical separation. The results reflect a consistent variation of comparably high runoff coefficients under saturated conditions, and low ones during dry periods. Over all the coefficients are reasonable low now for a forest catchment and differences between events are lower. Exceptions are e. g. convective events, which can lead to comparably high runoff coefficients as well (Event 03.08.2020). We took the duration out of the table, since it was the time of observed response in the hydrograph and not the rain event duration. We added total runoff in mm as well as event runoff (mm) after baseflow separation. Based on that we added the base flow index as a recognizable hydrological metric for the events. As suggested, we added usable field capacity, derived from the near DWD Weather station.

We added the following sentences:

Line 155-157:

> *"Based on the recorded hydrograph data, 68 events were selected for hysteresis analysis using the h-index developed by Zuecco et al. (2016). Corresponding HT measurements (n= 28) were available for 20 of these events (Table 1)."*

Line 169:

"Sampling was conducted under various baseflow conditions (Table1; BFI)."

We have revised the caption of Table 1 accordingly and added an explicit explanation of how the runoff coefficient was calculated.

Line 163:

*"**Table 1.** Characteristics of runoff events occurring prior to one or multiple tracer experiments: event precipitation, Total Runoff [mm], Event Runoff [mm], Baseflow Index [BFI] as 1- (Total Runoff/ Event Runoff), and the runoff coefficient calculated as the ratio of direct event runoff volume to precipitation-derived input volume (runoff coefficient = event runoff volume / (precipitation × catchment area)), Soil as field capacity in %."*

In addition, we now include information on antecedent moisture conditions to provide better context for events with seemingly high coefficients.

Line 282-283:

"Antecedent moisture conditions provide the boundary for discharge generation (Table 1), the Olewigerbach is during both seasons a highly reactive stream with increased discharge amplitudes (Figure 4 b)."

- The revised discussion is very long. Its structure can be improved by having two subsections, one about hysteresis dynamics and hydrologic turnover, and a second one about solute transport and redox conditions. In addition, you should write whether your results lead you to accept or reject your research hypotheses.

We appreciate this helpful suggestion. In the revised manuscript, we now have restructured the Discussion into two focused subsections (Hysteresis dynamics and hydrologic turnover/ solute transport and redox conditions). We agree that this improves readability and aligns the interpretation more closely with our research objectives.

Position of Figure 10 was changed accordingly

We added an explanatory sentence regarding Hypothesis (3)

Line 464-467:

"However, point-scale measurements of DO or pH in riparian groundwater would still represent spatially mixed signals, because small-scale heterogeneity in flow paths and reaction zones produces biogeochemical hot spots that are difficult to resolve with bulk sampling alone (Frei et al., 2012; Knorr et al., 2009; Peiffer et al., 2021; Alewell et al., 2008)."

In addition, we now explicitly state how our results relate to the proposed hypotheses within the Discussion chapter (Reviewer 3 Bullet points).

Line 439-444:

Hypothesis I & II

"Our results support Hypothesis (1): the direction and magnitude of the h-index exhibit a clear and consistent seasonal pattern, with clockwise hysteresis dominating under winter gaining conditions and counterclockwise hysteresis prevailing during summer and drought losing conditions. We also accept Hypothesis (2), since hydrological turnover (HT), as an independent measure of exchange, strongly co-varied with h-index values. Periods of intensified HT aligned with negative h-index

values under steep losing gradients, whereas low HT occurred under positive, or near-zero gradients associated with gaining conditions."

Line 484-487:

Hypothesis III:

"[...], supporting Hypothesis (3) Counterclockwise hysteresis under summer losing conditions coincided with elevated DOC, Fe^{2+} and Mn^{2+} at the near-stream well, indicating that hysteresis-driven exchange promotes redox fluctuations and the formation of riparian redox hot spots. Overall, the observed SW–GW level hysteresis reflects the system's exchange-driven redox response and should therefore be interpreted as a diagnostic indicator, not as a mechanistic explanation."

Technical corrections

We sincerely thank the reviewer for the remarks on technical corrections. We changed the manuscript accordingly:

- Line 51: 'stream' instead of 'steam'.

- Figure 1: Please add an indication of the water flow direction or of the catchment outlet.

We changed Figure 1, adding the outlet and flow direction.

- Line 143: There is a missing verb between 'to' and 'HT'.

We added the verb "estimate "

- Line 140: 'From the hydrograph' is not very clear; maybe you meant 'Based on hydrograph analysis'.

(Line 155) We revised the sentence to "based on the recorded hydrograph data"

- Line 207: Table number is missing.

This was a relic of a previous version of the manuscript, we apologize and excluded the reference.

- Line 208: The year is missing in Macherey & Nagel.

(Line 197) Macherey & Nagel is the Laboratory filter brand. We made it clear is the Text.

- Lines 208-209: 'ion chromatography' repeated twice. You can write 'Nitrate, potassium, and silica were determined by ion chromatography'.

(Lines 198-199) We changed that accordingly

- Equation 2: Please define Qi-1. Does it refer to the upstream location?

Yes, Q_i is the downstream location and Q_{i-1} the upstream location, the difference is ΔQ used to determine Q_{Gain}.

We changed the sentences to clarify the notation.

(Lines 207 & 208)

- Figure 5: Please increase the font size of the titles of the y axes.

We increased the font size of Figure 5

- Line 298: Here, it is unclear whether the positive and negative index values refer to HT or to the hysteresis index.

Line (277) We refer to the index values in that sentence. We made it clear in the text.

- Figure 6: By writing the significance codes, you do not need to add '95% confidence interval'. I guess 'Holm-Bonferroni' is the correction method; please integrate the caption.

We changed the captions accordingly

- Line 317: What do you mean by 'elevated'? Do you refer to the size of the hysteresis?

Yes, we mean elevated in the sense of hysteresis loop size. We changed the sentence as follows

Line (295)

"h-index values are substantially larger during drought conditions compared to winter, indicating an increase in hysteresis loop size."

- Line 322: Please add the verb or complete the sentence.

We rephrased the sentence.

- Lines 348-354: The acronym SW is rarely used in the previous text. Please be consistent throughout the manuscript.

We changed the acronym in the Cluster description

- Line 359: 'identifiyed exchange paterns'. Please revise the English.

We revised the English:

Lines (340-342)

"This separation enables redox tracers (DOC, NO_3^-, Fe, Mn) to capture the biogeochemical responses associated with the occurring exchange patterns"

- Lines 359-361: This sentence belongs to the methods; please move it.

We moved the sentence to the method section.

- Line 379: 'of' is missing before 'iron and manganese ions'.

We included the missing word.

- Lines 383-384: In this sentence, the English needs to be revised.

We revised the English of that paragraph as follows:

Lines (363-367)

"Manganese and iron are simultaneously present at the groundwater wells, often in similarly high concentrations, and show clear temporal patterns. During summer, concentrations of both elements were higher at GW1 than at GW2, suggesting intensified redox activity closer to the stream. These conditions coincided with hydraulic gradients indicative of losing conditions and counterclockwise hysteresis, indicating that stream-derived organic carbon inputs may stimulate microbial activity in the near-stream sediments."

<u>Response to Reviewer #3:</u>

We sincerely thank Reviewer 3 for the detailed and insightful comments. The suggestions were highly valuable in sharpening the conceptual framework and improving the clarity of our interpretations.

1. **Title and conceptual framing**

   The title does not fully reflect the scope and evidence presented in the manuscript. It implies that a hysteresis index derived from groundwater and surface-water time series can reveal hydrological turnover and redox processes. However, this connection remains insufficiently supported. Hysteresis is not a cause of these processes; rather, it is typically a result of incomplete understanding of complex system interactions or of decisions not to explicitly model these interactions. In other words, a multidimensional functional relationship—often linked to state variables—is being projected into a two-dimensional representation (Gharari and Razavi 2018 - Journal of Hydrology). Observing hysteresis does not establish causality, even if the patterns appear related. Causal mechanisms must be explicitly modeled or robustly justified. Because hysteresis indices represent effects rather than causes, interpreting them in terms of solute transport, redox dynamics, or hydrological turnover (or residence time, which I assume the authors mean) requires significant caution.

We appreciate the reviewer's concern regarding the scope of the title.

Nevertheless, we respectfully disagree with the reviewer's interpretation. The observed hysteresis between groundwater and surface water levels is neither an artifact of dimensionality reduction nor the result of unmodelled complexity. It arises from physically interpretable and measurable differences in the timing and magnitude of hydraulic responses between stream water level and near-stream groundwater. These patterns reflect event-scale dynamics of storage, transmissivity, hydraulic gradients, and bank storage, all well-established physical controls on GW–SW exchange (Cardenas, 2015; Brunner et al., 2009; Wondzell, 2011; Gu et al., 2012). Their seasonal variability produces consistent and reproducible hysteresis behaviour at our site.

We do not claim that hysteresis causes hydrological turnover or redox dynamics. Instead, hysteresis (as the result of the complex process interplay described above) serves as an integrative descriptor of system response under different antecedent conditions. This interpretation is consistent with previous works that uses hysteresis to understand event-scale GW–SW interactions and their dependence on storage (Zuecco et al., 2016; Lloyd et al., 2016; Gelmini et al., 2022). Hydrological turnover quantifies the fraction of stream water engaged in hyporheic exchange, while the hysteresis index captures the temporal offset between groundwater and stream responses; assessing whether these independent metrics co-vary across seasons is therefore hydrologically coherent.

Accordingly, we interpret hysteresis not as a mechanistic driver of HT or redox processes but as an event-scale diagnostic of system state. This diagnostic link is supported by our data: counterclockwise hysteresis consistently coincides with steep losing gradients and high HT, whereas clockwise hysteresis aligns with winter gaining conditions and low HT. These patterns are physically plausible and reproducible and thus form our interpretation of system behaviour.

2. **Need for a causal framework**

   Building on the previous point, I strongly encourage the authors to articulate a clear conceptual model describing the processes that give rise to hysteresis in groundwater–surface water interactions, including the relevant water-balance components and solute-transport mechanisms. This is especially important given the strong conclusions drawn regarding solute transport. The authors have addressed some of these mechanisms in their earlier 2024 WRR paper; integrating that conceptual clarity here would strengthen the argument considerably.

We appreciate the reviewer's request for a causal framework. In response to the reviewer's comment, we have added a concise conceptual framework at the end of the Introduction section. This framework explicitly links seasonal hydraulic boundary conditions, storage states, and exchange processes to the observed hysteresis patterns and hydrological turnover and provides the conceptual basis for the hypotheses tested in this study.

Lines 103-120:

> *"In this study, we aim to evaluate whether the hysteresis between groundwater and stream water levels at the event scale reflects seasonal changes in system state and hydrological turnover, building on the conceptual framework introduced in Bäthke & Schuetz (2024). We conceptualize stream–groundwater hysteresis as an integrated hydraulic response to seasonally and event-scale varying storage states, transmissivity, hydraulic gradients, and bank storage within the riparian corridor (Cardenas, 2015; Brunner et al., 2009; Wondzell, 2011; Gu et al., 2012). Hysteresis is interpreted as a diagnostic indicator linking hydraulic boundary conditions to hydrological turnover (HT) and associated biogeochemical responses.*
>
> *During summer and drought conditions, low groundwater levels and steep losing hydraulic gradients promote stream water infiltration into the streambed and banks. Reduced transmissivity and limited antecedent moisture lead to delayed groundwater responses relative to stream stage fluctuations, associated with enhanced bank storage activation and high relative HT.*
> *Conversely, during winter and wet conditions, elevated groundwater tables and reduced gradients favor gaining or weakly losing conditions. Increased hydraulic connectivity leads to more synchronous stream - groundwater responses, with the exchange dominated by groundwater discharge rather than bank storage-driven turnover.*
> *Within this study, we analyze whether seasonal shifts in hysteresis direction and magnitude reflect transitions between storage-controlled and exchange-controlled GW–SW interaction. Thus, we hypothesize that (1) the direction and magnitude of hysteresis between groundwater and stream water levels are indicative of seasonal hydrological states, and (2) consequently, the direction and magnitude of the hysteresis is related to hydrological turnover. thus facilitating (3) redox fluctuations creating oxic-anoxic transition in the riparian zone, shaping redox hot spots."*

Additionally, we clarify the diagnostic role of hysteresis more explicitly in the revised Discussion section:

Line 383:

> "The diagnostic use of hysteresis utilizes lagged hydraulic responses as indicators of system state rather than causal mechanisms."

Line 486:

> "Overall, the observed SW–GW level hysteresis reflects the system's exchange-driven redox response and should therefore be interpreted as a diagnostic indicator, not as a mechanistic explanation."

3. **Hysteresis indices and prediction under change**

   Similar concerns arise regarding the manuscript's discussion of prediction under climate change. It is not clear how climate change is incorporated into the analysis. Moreover, using hysteretic behavior as a predictive tool for climate-change impacts is questionable. Robust climate-change assessments generally require mechanistic models capable of representing future and potentially unprecedented system behavior. Relying solely on hysteresis indices for such predictions appears insufficient.

We thank the reviewer for raising this concern. We fully agree that hysteresis indices are not a mechanistic prediction tool for climate-change impacts. However, our manuscript does not seek to predict climate-change responses, nor do we propose hysteresis as a forecasting tool.

Instead, our reference to climate change is limited to discussing how long-term shifts in hydrological regimes (e.g., increased frequency of low-flow or drought conditions) may alter the occurrence of the seasonal system states identified in this study. In this context, hysteresis is used solely as a diagnostic indicator of present-day system behaviour, not as a mechanistic predictor. We now clarify this explicitly in the revised Discussion.

Accordingly, we have removed language that could be interpreted as predictive and now focus exclusively on how the identified seasonal relationships between HT and GW-SW hysteresis may help in interpretation of potential trends in future observational data, rather than forecasting them.

Lines 428-430:

> *"This description of present-day summer and drought system states may support the interpretation of possible HT process shifts if hydrological regime changes occur in the future."*

4. **Hypotheses framing**

   The hypotheses presented in the introduction are rather general and, in some cases, trivial. It is unclear whether hysteretic behavior is required to address them, or whether the available data would yield similar insights without invoking hysteresis or linking groundwater and surface water to multiple storage components. The manuscript refers to "addressing a gap," but the nature of this gap remains vague and difficult to comprehend.

We thank Reviewer 3 for the critical view on our hypotheses. With the help of Reviewer 1 and Reviewer 2, we have rephrased and clarified our hypotheses and addressed them in the Introduction (Lines 116-120), the Discussion (Lines 464–467; 484–487), and the Conclusion (rewritten in response to comment 6) to provide a clearer structure to the manuscript.

We agree with Reviewer 3 that the manuscript previously referred to "addressing a gap" without clearly explaining the nature of this gap. We have revised the sentence accordingly and added an explanatory statement.

Lines 83-96:

> *"Despite the recognition of hydrological turnover (HT) as the cumulative expression of bidirectional GW-SW exchange (Payn et al., 2009; Covino et al., 2011; Mallard et al., 2014), its temporal variability and the range of HT magnitudes that can occur within a single catchment and even within one stream reach is insufficiently quantified (Jähkel et al., 2022; Jimenez-Fernandez et al., 2022; Bäthke & Schuetz, 2024). A key limitation in HT monitoring is methodological: direct HT quantification requires tracer-based dilution gauging and therefore provides only snapshots of exchange, while GW–SW interactions are known to shift substantially with changing storage states, hydraulic gradients, and event forcings across seasons (e.g., Ward et al., 2013, 2019)."*
>
> *Here, we address this gap for a headwater stream reach by combining process-based HT measurements with high-frequency stream–groundwater hysteresis metrics. We focus on hysteresis between near-stream groundwater levels and stream stage because this signature captures the relative timing of groundwater and streamflow responses and has been shown to provide a sensitive, event-scale diagnostic indicator of changing hydrological connectivity and storage activation under contrasting antecedent conditions (Zuecco et al., 2019; Gelmini et al., 2022). This site-specific approach allows us to assess whether event-scale hysteresis reflects changing exchange states at the reach scale and thereby improves our understanding of the spatio-temporal dynamics of groundwater–surface water exchange at the study site.*
>
> *By quantifying hysteresis loops and HT over two seasons, we determine the impact of groundwater-surface water exchange variability on solute transport and stream chemistry."*

5. **Novelty relative to earlier work**

   The manuscript cites the authors' 2024 WRR publication, which addressed similar conceptual questions. It would be helpful for the authors to explicitly state the added value or unique contribution of the present manuscript relative to that earlier work. What new insights does this study offer?

We appreciate the reviewer's request for clarification regarding the unique contribution of this manuscript compared to our earlier WRR publication (Bäthke & Schuetz, 2024). The two studies address related themes, but with distinct objectives and conceptual focus. The WRR paper developed a seasonal, catchment-scale framework describing how groundwater storage states control groundwater–surface water exchange, using silica as a conservative tracer.

The present manuscript advances this framework in three key ways: (1) it utilizes an event-based, quantitative hysteresis metric (h-index) as a diagnostic indicator of GW–SW interaction dynamics at high temporal resolution; (2) it empirically links hysteresis behavior to independently measured hydrological turnover (HT), providing a cross-validation of exchange processes and (3) it extends the interpretation to redox-sensitive solutes, demonstrating how event-scale exchange dynamics may shape biogeochemical responses in the near-stream groundwater.

6. **Diffused Discussions and Conclusions**

   I recommend adding a concise, bullet-point conclusion section that clearly states both general and specific takeaways. At present, the conclusions and discussion are too general and could almost be written without reference to the specific analyses performed. What should readers take away from this manuscript? And are the hypotheses stated in the introduction explicitly addressed and resolved by the end?

We took the reviewers' valued advice and have rewritten the conclusion, incorporating bullet points that formulate clear takeaways for the reader (Lines 490-536).

The revised manuscript after the first review iteration presents a restructured Discussion and Conclusion. As requested by Reviewer 2, we reorganized the Discussion into two subsections:

4.1 Hysteresis dynamics and hydrologic turnover and *4.2 Solute transport and redox conditions*.

As mentioned above (see response to comment 4), we added sentences explicitly referring to the hypotheses of this work at the end of both discussion subsections.