Authors' response to Reviewer 1 [hess-2025-1530-RC1]

We thank the reviewer for his evaluation of our manuscript and his many helpful comments (hess-2025-1530). Below we address the reviewer's comments (full text) indented by arrows and coloured in blue. We appreciate the efforts by the reviewer, which will help to improve our manuscript.

General comments

The authors define Fnew as water younger than approximately 16 days, based on their fortnightly sampling interval. While the authors note that fortnightly sampling limits Fnew resolution, a brief discussion on how the chosen sampling frequency may bias or underrepresent fast-response components would strengthen the interpretation of Fnew.

→ You are right to mention that fortnightly sampling limits the Fnew interpretation in terms of fast-response components, which can represent an issue in particularly reactive streams, e.g., in impermeable layer catchments with high contents of marls or claystone. During high streamflow events, much of the hydraulic response will occur within that two-week window timeframe, which would be interesting to investigate at higher temporal resolution. Once either the infiltration capacity or the storage capacity of the soils in these catchments has been reached, overland flow processes can occur very quickly, within hours or even minutes. This could be essential for studies focusing on trigger mechanisms of intense short-term events, such as flash floods caused by highly intensive convective rains. This study could serve as a guide to identify streams that would benefit from high-frequency sampling or monitoring campaigns. We will make sure to include this aspect in the discussion.

While the manuscript rightly identifies bedrock geology as a key control on flow partitioning and Fnew, the analysis is limited to areal lithological classification. Crucial structural factors—such as bedrock depth, regolith thickness, and the presence of fractures or faults—are not considered, despite their known influence on subsurface storage and connectivity. If such data are available, they should be integrated into the analysis; otherwise, their absence should be acknowledged as a limitation.

→ Thank you for the pertinent remark. Factors such as the bedrock depth, regolith thickness, and the presence of fractures or faults, or even soil properties, are important controls on subsurface storage and connectivity, which were not explicitly considered here. The main reason for this was that such data was only available for specific catchments, where physiographic controls on catchment functions had been previously investigated in detail, e.g., in the Weierbach, Wollefsbach, and Huewelerbach catchments (Wrede et al., 2015; Martínez-Carreras et al., 2016; Douinot et al., 2022; Kaplan et al., 2022). The subsurface structure was mentioned as a control on catchment functions in these anterior studies, e.g., cracks and fissures in the schistose Weierbach catchment, creating substantial storage volumes and explaining the dominance of subsurface flow (Angermann et al., 2017). The results and conclusions from these studies are taken as reference and briefly summarized under section 2.1. entitled "Physiographic characterisation of the nested catchments".
Another argument is that we estimated the maximum storage capacity of the catchments, which implicitly contains information on the bedrock depth and retention volume, or regoltih thickness. One could also argue that these factors will matter less in catchments where the infiltration rates are low and the bedrock is impermeable, i.e., marly and claystone catchments. They are however important controls in the schistose catchment with a highly weathered front, which might have a thickness that

varies due to the presence of fractures and fissures. We will acknowledge this as a limitation, particularly in the Colpach catchment, where the depth of the regolith is largely unknown.

The manuscript emphasizes the dominant role of bedrock geology in shaping hydrological response, however, the potential confounding influence of co-varying catchment attributes (e.g., slope, land cover, elevation) is not adequately resolved.

→ Thank you for bringing this to our attention. The confounding influence of co-varying catchment attributes was addressed under section 4.1. in the following paragraph:

"Other important controls that have been reported to affect catchment functions include topography, precipitation, soil properties or vegetation (Von Freyberg et al., 2018; Lutz et al., 2018, Floriancic et al., 2023), yet in our study area, many of these parameters were correlated (Fig. S5) or similar across the different catchments. Across the 12 examined catchments the underlying bedrock geology played an important role in their shaping landscapes thereby affecting topography, soil properties or vegetation. ~~Considering only characteristics that were independent from each other~~, we found that isotopically inferred F_{new} during times of high streamflow was positively correlated to the percentage of marls ($r = 0.88$, $p = 0.001$), and to the percentage of grassland area ($r = 0.61$, $p = 0.036$, Fig. 8). The negative correlation with forest area ($r = -0.59$, $p = 0.044$) may also relate to the geological properties, since the two catchments with the highest percentage of forest area also consist almost exclusively of schist or sandstone, which are generally associated with low Fnew. One could expect higher infiltration rates with root structures in forest, e.g. as reported in weathered layer catchments (Angermann et al., 2017), and soil consolidation or artificial drainage in agricultural areas triggering fast overland and shallow subsurface flow (Loritz et al., 2017), leading to correlations between F_{new} and forest and agriculture land use."

We could also refer to Pfister et al. (2002), where they had already investigated to what extent factors such as drainage density, catchment shape, catchment area, specific slope, percentage of less permeable substrate, and land use may control stormflow across 18 nested catchments in the Alzette basin. It turned out that the percentage of impermeable substrate was the main control factor.

Although we acknowledge that forests could have an influence on the streamflow here, we could stress the importance of land use more. Upon second examination, a clear linear relation exists between grassland, or forest fractions, and the fractions of new water at high streamflow rates – if we disregard the two points with the highest F_{new} (Fig. 8). We will mention that. The two points correspond to impermeable layer catchments, where one could argue that F_{new} at high stream flow rates can be expected to be high, no matter the land use. Also, we will remove the statement "Considering only characteristics that were independent from each other" because it was motivated based on the results of the correlation matrix in Figure S5, which require a cautious interpretation as there were only 12 points (i.e., the number of nested catchments). Our argument was that the underlying bedrock geology contributes to shape catchment attributes, but it remains interesting to show and discuss the influence of these attributes following your advice. Thus, we propose to expand Figure 8 and add the elevation range, the mean slope and the maximum slope of the catchments, to include the topography in our discussion. Results suggest while a relation with the mean catchment elevation seems to exist (Fig. 8e), there is none with the elevation range. This seams to be an indication that the previous relation with mean elevation is just the result of the spatial distribution of the catchments with similar properties, as the hydro-lithological clusters appear clustered (Fig. 8e). For the slope, we find a significant negative relation of the fraction of new water at high streamflow rates with the mean slope of the catchments only. This result is rater surprising, as one would expect higher F_{new} with steeper slopes. One explanation is that the impermeable layer catchments in our study have slopes that are less steep than aggregated catchments, or the small sandstone-dominated Huewelerbach catchment. However, the discrepancy between mean and maximum slope values already shows that these results are strongly dependent on how these parameters are computed.

The manuscript would benefit from a clearer explanation of the methodological constraints of the Ensemble Hydrograph Separation (EHS) approach. Specifically, does EHS impose minimum requirements on sampling interval, time series length, or end-member stability?

→ By definition, EHS assumes that the sampling interval is constant, as it compares isotopic concentrations at a certain timestep with previous observations without considering the exact moment of sampling. Still, this constraint is not very strict in the sense that EHS can still be applied to timeseries with slightly unequal timesteps without becoming unstable. In this case, the best approximation is to take the mean sampling interval as the definition for the fraction of new water, as was done in this study. Gaps in isotopic measurements need to be filled with NA values, while zero values can be filled in for missing streamflow or precipitation values, without affecting the computations. This step is important, because it affects the calculations if EHS assumes a direct link between two consecutive samples in the record, which are few sampling intervals apart. There is no minimum requirement for the timeseries length per se, however, larger datasets yield more significant results, with the error margins becoming smaller in comparison to the magnitude of the detected signal in either the F_{new} profiles or the TTDs. The size of the dataset also becomes increasingly important when sub-setting the data for the computation of profiles or TTDs at different states of the catchment. For the profiles, we tried to work with at least ~30 observations, but in the end, that decision is in the hands of the user. The greatest issues we faced were when the order of the observations was not respected, e.g., by making mistakes in alignments of the streamflow and precipitation isotopic samples. Otherwise, the EHS method appears to be quite robust and broadly applicable. Please refer to Kirchner et al. (2019) to read the underlying principles of EHS, or Kirchner and Knapp (2020) for the practical guide on how to use EHS, with scripts provided in R and Matlab, and instructions on data requirements.

Specific comments

Line 163: The phrase “i.e.i.e.” is repeated. Please delete the redundant “i.e.” to correct the typo.

→ Thank you, we will remove it.

Line 221: Please clarify how α limits evaporation.

→ “Evapotranspiration” is probably the correct term here, as α will predominantly represent the reduction of transpiration with decreasing water availability. We hope this answers your question.

Line 224: The manuscript defines Smax as the highest 0.5% of daily catchment storage values, please justify why this specific quantile was chosen

→ We thought taking a quantile would be more robust than taking the absolute maximum of storage values, but we must admit that there was no specific reason other than that. We might consider taking the maximum value instead.