

Thank you for your recognition of our manuscript and your valuable comments. Below are the detailed responses to each comment.

1. The baseflow separation methods seem to generate distinct BFIs and event contrasting trends in certain time periods. How do you tackle the different separation results?

Reply: Regarding the discrepancies between baseflow results derived from different separation methods, our core criterion for evaluation is to compare the Nash-Sutcliffe Efficiency (NSE) and Kling-Gupta Efficiency (KGE) coefficients between each method's output and the source data. Notably, the NSE and KGE coefficients we calculated were specific to runoff periods excluding those with precipitation recharge. During these periods, river discharge is solely replenished by baseflow—thus, there should be a clear consistency between total river discharge and baseflow volume.

However, after precipitation recharge ceases, some surface runoff may still be en route to the river. This means river discharge during this phase is sourced not only from baseflow but also from residual surface runoff. To mitigate this issue, we specifically selected small watersheds with an area of less than 2,500 square kilometers. For such small watersheds, the travel time of surface runoff to the river is relatively short, which further minimizes the impact of surface runoff on the accuracy of baseflow separation results.

Thus, we believe using these two coefficients is a viable approach for assessing the reliability of baseflow separation methods. The specific assessment methods and the process for excluding precipitation recharge periods can be found in the Methods section of this manuscript.

2. In Figure 7, the authors compared BFI in karst regions and the global average. However, the global average is impacted by the coverage ratio of karst regions. Consider replace the global mean with non-karst area mean.

Reply: Your comment is critical, and we will revise Figure 7 in the revised version of the manuscript to better achieve the effect of comparison with global-scale data.

Specifically, we chose the mean baseflow curve of global non-karst regions as the comparison dataset for the karst baseflow curve in Figure 7. This choice is based on the following consideration: the number of catchment datasets for global non-karst regions exceeds 6,000, while the dataset used in this study to calculate karst baseflow values comprises 1,375 catchments. Due to this difference in dataset size, the difference between the global average baseflow curve and the global non-karst average baseflow curve is minimal (less than 1%). We therefore opted to use the global non-karst baseflow curve as the comparison dataset.

3. Figure 11 shows the factor impact on BFI, but this does not reveal if the impact is positive or negative. Consider including analysis such as SHAP.

Reply: Your consideration is valid. Initially, we only focused on the magnitude of the impact of different factors on the Base Flow Index (BFI), without fully incorporating their impact characteristics. In the revised version of the study, we will incorporate SHAP values (SHapley Additive exPlanations) to refine and supplement the analysis of how different factors influence the BFI, including their specific impact characteristics.