We are sincerely grateful to the reviewers for their positive assessment and the constructive suggestions provided. In accordance with their guidance, we have carefully revised the manuscript point-by-point and addressed other related issues identified during our review. The specific modifications made in response to each comment are detailed as follows.

L27: What is the baseflow index? Please clarify!

Reply: Let us first clarify the concept of baseflow. Baseflow refers to the sustained and stable component of total river discharge, distinct from the rapid and highly variable surface runoff driven by episodic events such as heavy rainfall or snowmelt. It is primarily sourced from slow groundwater recharge (e.g., from shallow soil moisture or deep confined aquifers). During dry seasons when precipitation recharge diminishes or ceases, baseflow becomes the critical water source that sustains river flow and supports aquatic ecosystems.

The Baseflow Index (BFI) is defined as the ratio of baseflow volume to total streamflow volume over the same period. As a key metric for quantifying the extent of groundwater contribution to river discharge, the BFI often reflects the stability of the watershed's hydrological cycle.

L36: Please provide the full name of the BFIs.

Reply: BFI stands for "Baseflow Index." This omission occurred as we streamlined the abstract for brevity, and we acknowledge that the full term should be provided upon its first use. We will add the definition of BFI in the subsequent revision.

L45-46: "as a slow recharge component of ..."and "as a hydrological stabiliser"wore both presented. Please rephrase this sentence.

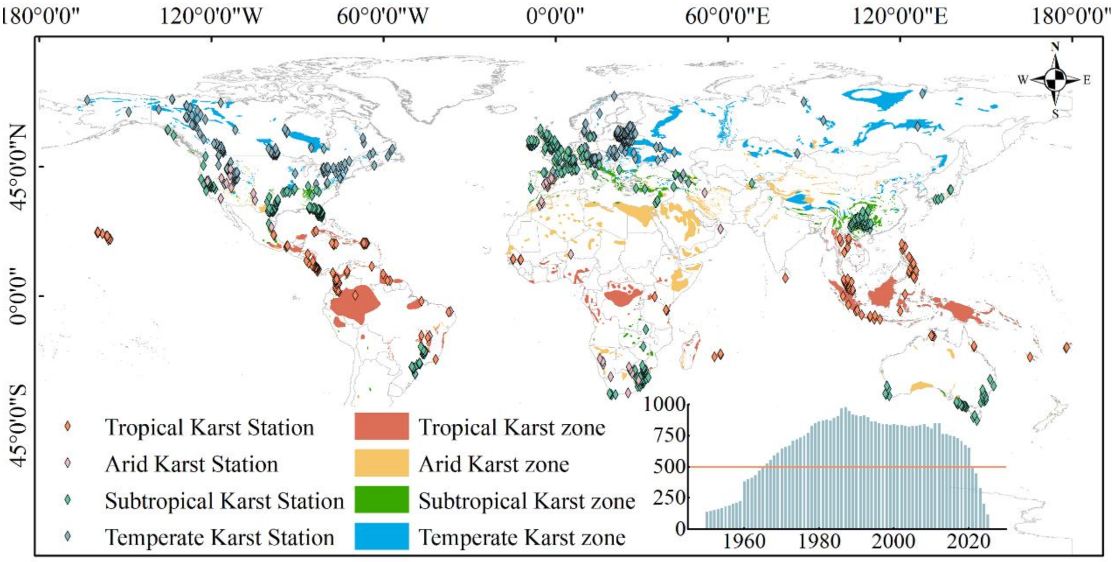
Reply: The sentence has been revised to: "Baseflow, as a slow recharge component from groundwater to runoff, plays a central role as a hydrological stabiliser." This new formulation accurately presents both the origin and the significance of baseflow in a logical progression, thereby addressing the lack of clarity in our initial phrasing. We have corrected this issue and will carefully scrutinize the entire manuscript to prevent similar oversights.

L60-66: This sentence was too tedious and long. Please improve this sentence.

Reply: We have revised this section to describe the regional differences in baseflow more concisely:" At a regional scale, BFI spatial patterns vary significantly. Studies show a higher BFI in the eastern parts of both the United States and India compared to their western regions. In China, the Yellow River Basin exhibits a pattern of high-low-high from upstream to downstream, whereas the Wei River Basin shows a gradual decline"

L108: For the daily-scale runoff data from 1375 watersheds within the karst region, the datasets for how many years?

Reply: The data used in this study were compiled from publicly available datasets across multiple countries, covering the period from 1880 to 2024. However, significant data gaps exist prior to 1960 and in recent years, which limits the reliability of a robust global assessment of karst baseflow characteristics. To address this, we established a screening criterion: only years with at least 500 effectively monitored basins globally were included in the analysis. The annual distribution of valid data volume is shown in the statistical subplot in the lower-right corner of Figure 1 (below). The red horizontal line in this subplot represents the threshold of 500 valid basins.



L122: 1412 watersheds? You can directly present the 1375 stations since 37 stations were not used in present study.

Reply: We will correct this in the manuscript to state that 1,375 data points were used. The discrepancy arose because our initial screening yielded 1,412 basins. However, during subsequent calculations, we identified that some basins suffered from severe data gaps (e.g., river flow interruption), including instances of zero recorded flow for two consecutive years. To prevent potential bias, these basins with extensive missing data were excluded, resulting in the final set of 1,375 valid data points.

186-190: Eight methods to calculate the baseflow should be described in detail. Alternatively, you can add a Table to exhibit these eight methods.

Reply: We thank the reviewer for this valuable comment. We acknowledge that the

Methodology section lacks sufficient explanation of the background and principles of each baseflow separation method used. In response, we have consolidated all the methods employed in this study into the table below, which summarizes both the background and fundamental principles of each method. To improve the manuscript's completeness, we will either incorporate detailed explanations of each method into the main text or include this table as supplementary material.

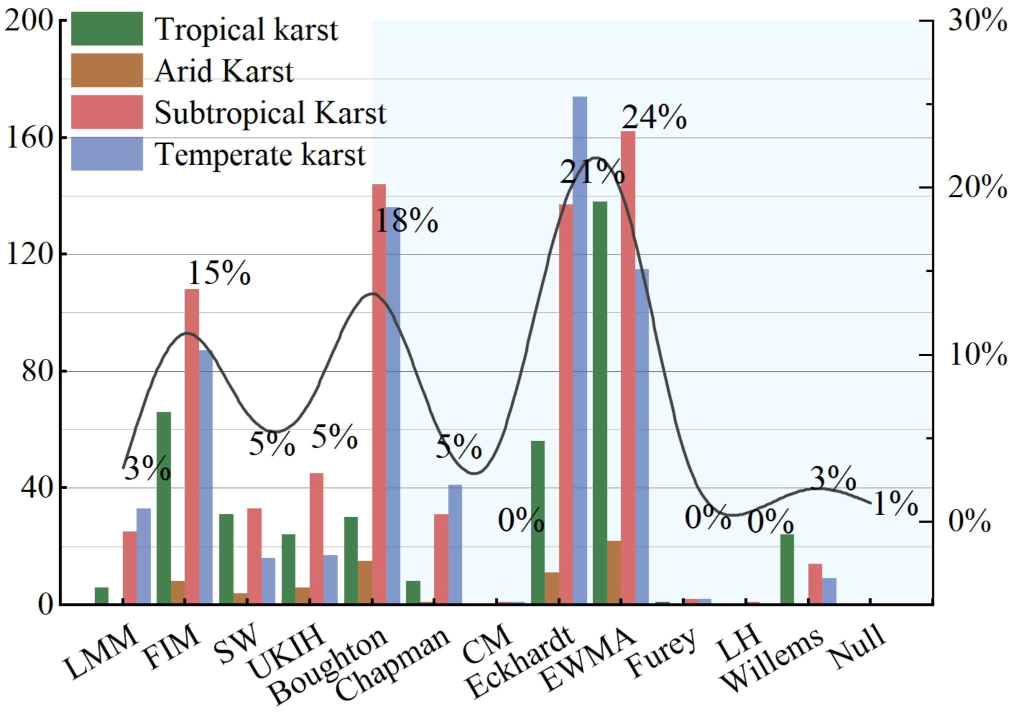
Graphical Methods	
Name (Abbreviation)	Description
Fixed Interval Method (FIM)	Proposed by Pettyjohn & Henning in 1979 and first introduced as a core method within the HYSEP program. Its principle involves segmenting the hydrograph into consecutive fixed-time intervals and taking the minimum flow within each interval as the baseflow.
Local Minimum Method (LMM)	Integrated into the HYSEP program by Sloto & Crouse in 1996 as a standard graphical separation technique. The method works by identifying local minimum points in the flow time series to demarcate the separation between baseflow and surface runoff.
Sliding Window Method (SW)	Also proposed by Sloto & Crouse within the HYSEP program, this method improves upon the rigidity of the Fixed Interval Method. Its principle is to traverse the hydrograph using a sliding window of fixed width, dynamically calculating the minimum flow within each window and assigning it as the baseflow value at the window's center point.
UK Institute of Hydrology (UKIH)	Originally developed by the UK Institute of Hydrology in 1980 and later refined by Wels et al., who also developed its computational program. It is a baseflow separation method that incorporates precipitation thresholds and flow response, dynamically adjusting the baseflow threshold to identify the separation point between rainfall events and baseflow.

Digital Filtering Methods	
Name (Abbreviation)	Description
Boughton Method (Boughton)	Proposed by Boughton in 1993 as a representative single-parameter recursive filtering method. It calculates the baseflow at the current timestep based on the baseflow value from the previous timestep and the total flow value at the current timestep.
Chapman-Maxwell Filter Method (CM)	An enhancement of the Chapman filter by Chapman & Maxwell in 1996, which improves accuracy by dynamically adjusting the

	recession constant. It computes baseflow by treating it as a weighted average of the concurrent total streamflow and the baseflow from the preceding timestep.
Chapman Filter Method (Chapman)	Proposed by Chapman in 1991 to address the issue of unrealistically constant baseflow at the end of recession periods found in the Lyne-Hollick method. Its principle involves calculating baseflow as a weighted average of the current total streamflow and the previous timestep's baseflow.
Exponential Weighted Moving Average (EWMA)	Introduced to hydrology by Vogel & Kroll in 1992 for Baseflow Index (BFI) calculation. The method estimates baseflow by applying exponential weighting to smooth the streamflow time series data.
Eckhardt Filter Method (Eckhardt)	Proposed by Eckhardt in 2005, this is a two-parameter recursive filtering method. It estimates baseflow by evaluating the maximum values of the recession constant and the maximum baseflow index.
Furey Digital Filter Method (Furey)	Proposed by Furey in 2001, based on a physical-statistical model of hillslope hydrological processes. Its principle involves estimating baseflow by considering the recession constant and a calibrated parameter.
Lyne-Hollick Digital Filter Method (LH)	First introduced by Lyne & Hollick in 1979, it is one of the earliest recursive digital filter methods. The principle involves a two-pass filtering process to estimate baseflow.
Willems Digital Filter Method (Willems)	Proposed by Willems in 2009, based on a linear reservoir model and least squares optimization. It estimates baseflow by calculating it as a weighted average of the baseflow from the previous timestep and the total flow at the current timestep.

2: The colors for these four karst regions were to similar. Please improve the color.

Reply: We have adjusted the colors for the different categories to improve their distinguishability in the figure, as shown below.



3-4: Significant difference test should be added.

Reply: We have supplemented the significance tests for Figures 3 and 4. Using the Kruskal-Wallis test, we confirmed the statistical significance of the differences in both the KGE and NSE coefficients among the 12 baseflow separation methods. Accordingly, we will enhance the main text by adding a discussion on the performance differences between different types of methods, along with further interpretation of the effectiveness of each separation method. In the figure below, the letters denote the results of multiple comparisons based on the Mann-Whitney U test, while the horizontal lines at the bottom of the figure indicate significant differences between the graphical methods and digital filtering methods. Methods sharing the same letter show no significant difference at the $p < 0.05$ level, whereas methods with different letters are statistically significantly different. Methods assigned multiple letters do not differ significantly from multiple groups.

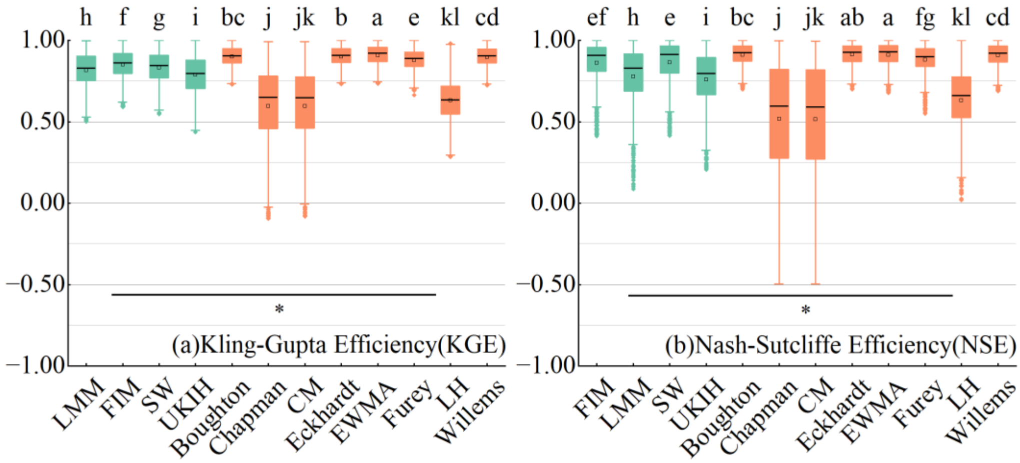
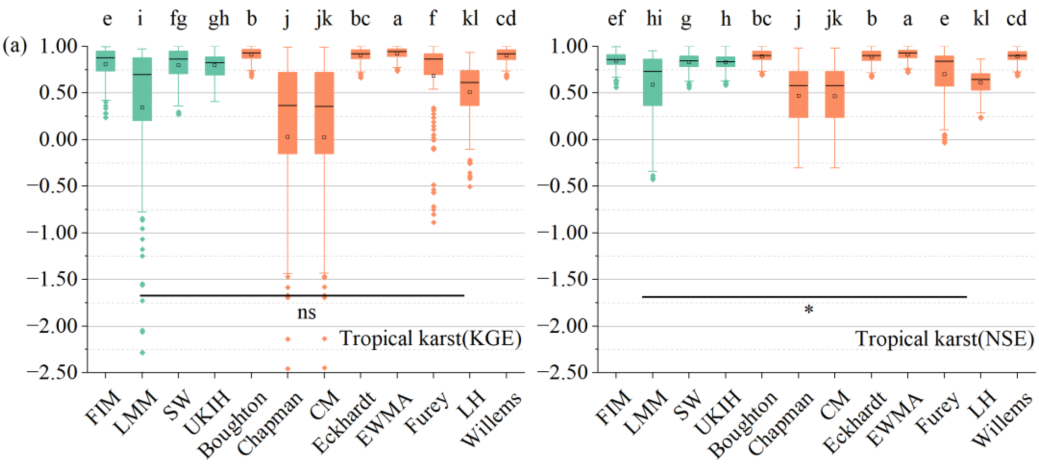


Figure 3. Comparison of KGE coefficients (a) and NSE coefficients (b) for the 12 baseflow separation methods. The X-axis represents each separation method, and the Y-axis indicates the value of the coefficients. Green color in the plot denotes the graphical method, while orange denotes the digital filtering method. The letters above the boxes indicate significant differences among the different baseflow separation methods, while the horizontal lines in the lower part of the figure represent significant differences between the graphical method and the digital filtering method. The black line inside the boxplot denotes the mean value, with upper and lower limits set at 1.5 times the interquartile range (IQR). Values exceeding this range are considered outliers and are marked as dots at the top and bottom of the boxplot.



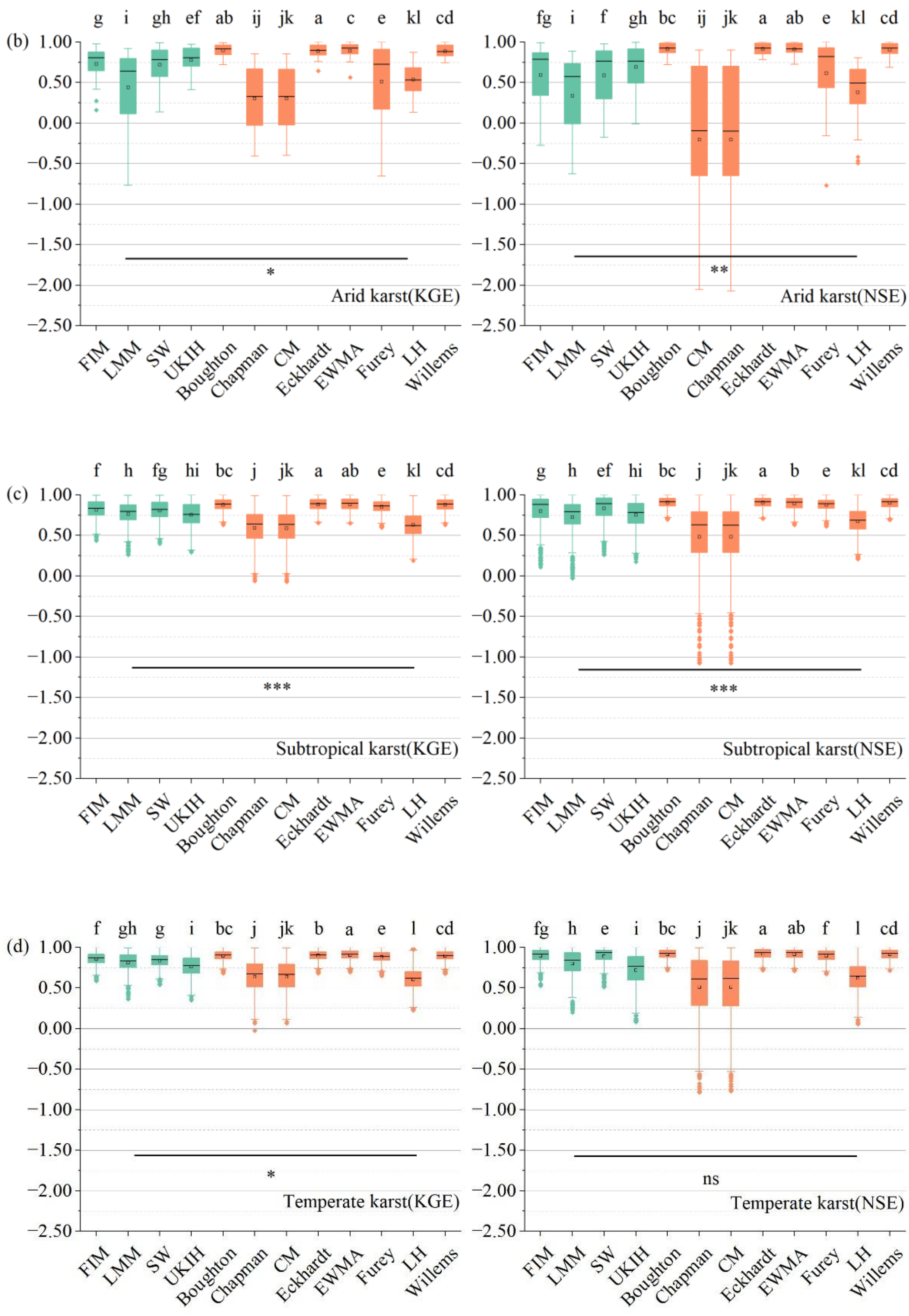


Figure 4. Comparison of KGE coefficients (left) and NSE coefficients (right) for karst regions in different climatic zones (as labeled in the bottom-right corner of each subplot). The X-axis represents each separation method, and the Y-axis indicates the coefficient values. The letters above the boxes indicate significant differences among the baseflow separation methods, while the horizontal lines in the lower part of the figure denote significant differences between the graphical and digital filtering method groups. Green color in the plot denotes the

graphical method, and orange represents the digital filtering method. The black line inside each boxplot indicates the mean value, with the upper and lower limits set at 1.5 times the interquartile range (IQR). Data points beyond this range are considered outliers and are marked as dots at the top and bottom of the boxplot.

8: Please provide the P value.

Reply: We performed the Mann-Kendall test on the data in Figure 8 using the pymannkendall library. The results reveal a statistically significant increasing trend in the baseflow characteristics, with a p-value of 0.00002 (< 0.05). Furthermore, we supplemented this analysis with a linear regression trend test. The results show a Durbin-Watson statistic of approximately 1.8 (falling within the acceptable range of 1.5-2.5, indicating no significant autocorrelation and thus reliable results), and a p-value of approximately 0.00003 (< 0.05). Both methods confirm that the increasing trend observed in the experimental results is statistically reliable and not a chance occurrence. The revised figure is presented below.