Supplement of Hydrol. Earth Syst. Sci., 29, 4395–4416, 2025
https://doi.org/10.5194/hess-29-4395-2025-supplement

Hydrology and Earth System Sciences
Open Access
EGU

Supplement of

Unveiling the impact of potential evapotranspiration method selection on trends in hydrological cycle components across Europe

Vishal Thakur et al.

Correspondence to: Vishal Thakur (thakur@fzp.czu.cz)

S1. Citation Statements

This section presents the original excerpts from the research articles utilized in our study. We excluded citations for the datasets, models, and PET method formulations. We organize the excerpts alphabetically and order any repeated excerpts based on their appearance.

Ajami et al. (2017): AWAP potential evapotranspiration is calculated based on the Priestley–Taylor equation (Raupach et al., 2009).

Anabalón and Sharma (2017): For most of the datasets, the correlations in water–limited environments were higher between AET and PP. In energy-limited regions, the results were mixed: MOD16, GLEAM v3a, E2O-ORCH, and E2O-PG presented higher correlations between AET and PET trends, while in the remaining datasets, the correlations were higher between AET and PP trends.

Anabalón and Sharma (2017): To compare the directions and statistical significance of the trends found in the different datasets, the Data Concurrence Index(DCI), introduced by Anabalón [2016] was used. This index compares the total number of positive significant trends versus the total number of negative significant trends, as expressed below. where ND is the number of databases with information, ST_i is the magnitude of the trend, and hi is a binary variable that indicate if there is a significant trend (1) or not (0) for database i.The DCI is calculated on every cell with information, therefore it allows one to assess the spatial distribution of the differences in the trends.

Anabalón and Sharma (2017): Figure 1d shows a completely different picture of the multidecadal PET trends when compared with those derived for the short-term period. Most of the world exhibits mid to high concurrence of increasing PET, particularly Europe, South America, and Northern Africa. The regions under null DCI are reduced considerably, and the negative DCI zones are concentrated in a few regions in the tropical range. Table 4 shows the land surface coverage of the long-term PET DCI.

Anabalón and Sharma (2017): For most of the datasets, the correlations in water–limited environments were higher between AET and PP. In energy-limited regions, the results were mixed: MOD16, GLEAM v3a, E2O-ORCH, and E2O-PG presented higher correlations between AET and PET trends, while in the remaining datasets, the correlations were higher between AET and PP trends.

Aouissi et al. (2016): The methods integrated in SWAT to estimate PET (PM, HA, PT) had little influence on monthly streamflow and actual ET predictions, because runoff is well induced by precipitation and the period of high ET does not correspond to the wet season.

Bai et al. (2016): In runoff modeling, different PET inputs can produce similar runoff simulations in both nonhumid and humid

regions. However, when estimating AET and WSC in humid regions, different PET inputs will result in significantly different AET and WSC simulations, which should be noted by model users.

Bai et al. (2016): The reasons for the insensitivities of runoff simulations to PET inputs are attributed to model parameter calibration, which is able to eliminate the discrepancies in model inputs and generate similar runoff simulation results.

Bai et al. (2016): Different PET inputs mainly affect AET and WSC in humid regions, and the sums of WSC and AET values are similar across different PET drivers in humid regions, as are the runoff values.

Bai et al. (2016): The water balance components in nonhumid regions are mainly controlled by precipitation rather than PET, resulting in a good agreement between water balance components, even though the PET inputs are considerably different.

Beck et al. (2017): The calibrated parameter sets for the donors were subsequently transferred to 0.5° grid cells with similar climatic and physiographic characteristics, resulting in parameter maps for HBV with global coverage. For each grid cell, we used the 10 most similar donor catchments, rather than the single most similar donor, and averaged the resulting simulated Q, which enhanced model performance.

Berghuijs et al. (2017): The absolute elasticity of runoff to potential evaporation changes is always lower than the elasticity to precipitation changes. The global pattern indicates that for 83% of the land grid cells runoff is most sensitive to precipitation changes, while other factors dominate for the remaining 17%. This dominant role of precipitation contradicts previous global assessments, which considered the impacts of aridity changes as a ratio.

Birhanu et al. (2018): The hydrological models' performance was satisfactory for each PET input in the calibration and validation periods for all of the tested catchments. The five hydrological models' performance were found to be insensitive to the 12 PET inputs because of the SCE-UA algorithm's efficiency in optimizing model parameters. However, the five hydrological models' parameters in charge of transforming the PET to actual evapotranspiration were sensitive and significantly affected by the PET complexity.

Boeing et al. (2024): Figure 3. (A) yearly TWS anomalies 1766–2022, colored bars denote the length of the recovery periods. (B) and (C) depict yearly P and E flux anomalies. The thick line in all subplots is a 30 year running mean to depict long–term trends. The points denote significant positive or negative 30-year trends including the year where the point is marked and 29 preceding years. The mean climatology for the anomalies in (A)–(C) is calculated based on the 1766–2022 period. Note that the y-axes have different limits to enhance the visibility.

Bruno and Duethmann (2024): Variations in EC were related with variations in precipitation and radiation, with a potentially increasing influence of precipitation after 2000s.

Caloiero et al. (2018): At the annual scale (Fig. 2a and Table 1), a prevalent negative rainfall trend has been detected. In fact,

about 34%, 30% and 22% of the grid points of the study area showed a negative trend at significance levels SL = 90%, 95% and 99%, respectively. By contrast, a positive rainfall trend was evidenced in about 30% (SL = 90%), 25% (SL = 95%) and 17% (SL = 99%) of the grid points. From a spatial point of view (Fig. 2b and Table 1), considering SL = 95%, the negative trend mainly involved the Mediterranean basin, with the highest negative values ($\leq$ 20 mm / 10 years) detected in Portugal, Spain, Italy, Bosnia, Montenegro, Greece and in the extreme eastern side of the study area. High negative values (down to -16 mm / 10 years) were also detected in Croatia, Hungary, Turkey, Morocco, Algeria and Libya. A marked negative trend was further detected in Northern Ireland ($\leq$ 20 mm / 10 years) and in the northern side of Norway (down to -16 mm / 10 years). Conversely, high rainfall increases (> 20 mm / 10 years) were observed in central–northern Europe—in particular in Norway, Sweden, France, the United Kingdom and Estonia—while smaller increases (up to 20 mm / 10 years) were detected in Germany, Denmark, the Benelux countries, Finland and Éire. A positive trend (> 20 mm / 10 years) was also evidenced in the Mediterranean basin, specifically in Tunisia, Italy and along the Atlantic coast of Spain.

Clerc-Schwarzenbach et al. (2024): The overall deterioration in model performance for calibration with the Caravan dataset indicates that the quality of the forcing data from ERA5-Land is lower than the quality of the data that are available for the US, Brazil, and Great Britain. As the ERA5-Land data are coarser than most data in the CAMELS datasets (Table 1), this was, to a certain extent, expected. These results mean that one should be cautious when using the Caravan precipitation data instead of more reliable (e.g. station-based) precipitation data because the conclusions may be affected by the lower data quality of the forcing data in the Caravan dataset. In our opinion, ERA5-Land precipitation data should only be used for catchments for which there are no alternative data (so that these catchments can still be included in large-sample studies).

Devia et al. (2015): In general, rainfall-runoff models are the standard tools used for investigating hydrological processes. A large number of models with different applications ranges from small catchments to global models has been developed. Each model has got its own unique characteristics and respective applications. Some of them are comprehensive and uses the physics of underlying hydrological processes and are distributed in space and time.

Guo et al. (2017): The two PET models consistently indicated temperature to be the most important variable for PET, but showed large differences in the relative importance of the remaining climate variables. In particular for the Penman–Monteith model, wind and relative humidity were the second-most important variables for dry and humid catchments, respectively, whereas for the Priestley-Taylor model solar radiation was the second-most important variable, with the greatest influence in warmer catchments. This information can be useful to inform the selection of suitable PET models to estimate future PET for different climate conditions, providing evidence on both the structural plausibility and input uncertainty for the alternative models.

Hanselmann et al. (2024): Jensen-Haise estimates a relatively stable high PET from mid-June to the end of July, with the highest values of 1.3 mm/day in mid-July.

Hua et al. (2020): For the entire arid zone, the PM method overestimated the potential evapotranspiration by 33.7 mm per year, with an overestimation of 29.0 mm from July to August. The most significant overestimation was evident in the mountainous and plain nonvegetation areas, in which the annual mean overestimation reached 5% and 10%, respectively; during July, there was an estimation of 10% and 20%, respectively. Although the annual evapotranspiration of the plains with better vegetation coverage was slightly underestimated, overestimation still occurred in July and August, with a mean overestimation of approximately 5%.

Jung et al. (2013): Runoff sensitivity can readily explain the conflicting directions of AET changes under similar precipitation change. Under increasing precipitation, AET decreases when runoff is increasing more rapidly than precipitation based on the water balance. Conversely, AET increases when runoff is decreasing more rapidly than precipitation.

Kingston et al. (2009): The latitudinal structure of baseline (1961–90) PET (Figure 1a) for all methods consists of a primary (secondary) peak in the northern (southern) subtropics as a result of seasonal maxima in temperature and solar radiation. However, the magnitude of PET differs substantially between methods (by up to 600 mm), with differences between methods not consistent across latitudes. For example, Jensen-Haise provides the highest estimate of PET at 20°N, but the lowest PET between 50–60°S and N.

Kuentz et al. (2017): Table 1. Catchment descriptors and the original source of information. Type of descriptor is indicated in brackets after variable name (T: topography; LC: land cover; S: soil type; G: geology; C: climate). Variables marked with grey color were removed from the analysis because no significant correlation was found between these and the flow signatures (see Sect. 2.3).

Kumar et al. (2013): We transfer hydrologic model parameters from a single central European basin, to 80 candidate basins within the United States. The regionalization is further tested across a range of climatic and land-cover conditions to identify potential biases in transferability.

Liu et al. (2022): ETp and ETa in most parts of the world indicated an increasing trend ($Z > 0$), but the spatial distribution of the increasing trend of the two was quite different. ETp had a remarkable increasing trend in the Ural Mountains on the border between Asia and Europe, the Appalachian Mountains and Mount Whitney in North America (α=0.05, $Z > 1.96$) (Fig. 4d), while ETa in these areas did not show an increasing trend but even declined (Fig. 4e). ETa had a marked increasing trend in the Amazon Plain, Congo Basin, South Asia and East Asia. Since water storage changes (TWC and NWC) are water losses caused by evapotranspiration, there is a strong negative correlation with ETp, ETa and Er, but the human water changes (HWC)

are positively correlated, notably in Peru, Congo, Brasilia, and Indonesia (Fig. S6e in supplementary information) where the large-scale irrigation system and irrigation water accelerate ETa

Lu et al. (2005): The commonly used PET methods for this comparison study gave a wide range of values, showing differences in PET across the southeastern United States among six methods as high as 500 mm/yr.

Maček et al. (2018): Table 3. Slope (k) of the linear regression for the radiative and aerodynamics components of the ET0 and average relative contribution of radiative and aerodynamics components to the total ET0.

Markonis (2025): Moreover, the study relies on a process-based model (GLEAM) for evaporation data, which carries its own assumptions and limitations. These factors highlight the need for further research to validate this framework in broader contexts and integrate observational data, for example, FLUXNET, to overcome current model-related limitations. Still, this study provides preliminary evidence that similar statistical properties, such as clustering, may exist for evaporation across all climate types (Figures S3-S4 in Supporting Information S1). Furthermore, the observed associations between energy sources (shortwave and longwave radiation) and water availability (precipitation) appear to be consistent across different climates (Figure S5 in Supporting Information S1), suggesting the potential generality of this framework for understanding extreme evaporation dynamics.

Markonis et al. (2024): We estimate annual terrestrial precipitation at approximately $114{,}000 \pm 9400\,\mathrm{km3}$, with about 70% falling over tropical, subtropical and temperate regions. Our results highlight substantial inconsistencies, mainly, over the arid and the mountainous areas. To quantify the overall discrepancies, we utilize the concept of dataset agreement and then explore the pairwise relationships among the datasets in terms of "genealogy", concurrency, and distance. The resulting uncertainty-based partitioning demonstrates how precipitation is distributed over a wide range of environments and improves our understanding on how their conditions influence observational fidelity.

Markonis et al. (2021): Here, using a multidimensional machine learning–based clustering algorithm and the hydrologic reconstruction of European drought, we determine the dominant drought types and investigate the changes in drought typology. We report a substantial increase in shorter warm-season droughts that are concurrent with an increase in potential evapotranspiration. If shifts reported here persist, then we will need new adaptive water management policies and, in the long run, we may observe considerable alterations in vegetation regimes and ecosystem functioning.

Markonis et al. (2019): we scrutinize a new global, high-resolution precipitation data set, namely, the Multi-Source Weighted-Ensemble Precipitation v2.0, to determine changes in the precipitation distribution over land during 1979–2016. To this end, the fluctuations of wet days precipitation quantiles on an annual basis and their interplay with annual totals and number of wet days were investigated. The results show increase in total precipitation, number of wet days, and heavy events over land, as

suggested by the intensification hypothesis.

Massari et al. (2022): The climate of the specific basin was defined based on the aridity index, which we calculated as the ratio between long-term average annual potential evaporation and precipitation (both from ERA5) (Unesco, 1979).

Massari et al. (2022): We relied on multidecadal records of streamflow and precipitation for more than 200 catchment areas across various European climates, which distinctively show the emergence of similar periods of exacerbated runoff deficit identified in previous studies, i.e. runoff deficit on the order of -20 % to -40 % less than what expected from precipitation deficits. The magnitude of this exacerbation is two to three times larger for basins located in dry regions than for basins in wet regions, and is qualitatively correlated with an increase in annual evaporation during droughts, in the order of +11 % and +33 % over basins characterized by energy-limited and water-limited evaporation regimes, respectively. Thus, enhanced atmospheric and vegetation demand for moisture during dry periods induces a nonlinear precipitation-runoff relationship for low-flow regimes, which results in an unexpectedly large decrease in runoff during periods of already low water availability.

Massari et al. (2022): Runoff exacerbation is related to an increase in evaporation occurring under two defined and concurrent preconditions: (1) water storage can support ET during the drought period, and (2) there is a sufficient vapour-pressure deficit (mainly driven by the temperature increase) to generate evaporation.

Mazzoleni et al. (2019): Precipitation datasets belonging to Class 2 outperformed the other datasets in basins with Tropical and Temperate-Arid climate (e.g. Congo, Mississippi and Godavari), while Class 3 datasets showed the highest NSE values in Temperate and Temperate-Cold basins (e.g. Danube, Rhine and Volga). In addition, datasets from Class 3 gave the best performances at basin outlets in case of dense precipitation monitoring networks, as in the Danube and Rhine basins.

Ndiaye et al. (2024): Observed and reanalysis data are used to calculate daily PET values according to 21 methods. Combinatory methods and temperature-based methods provide the most accurate estimation. However, the aerodynamic methods underestimate PET and some temperature and radiation-based methods (Heydari and Heydari and Priestley-Taylor) overestimate it. With regard to the sensitivity of GR models to the different PET estimation methods, all three GR models showed an ability to readjust the estimation errors of the PET methods.

Rakovec et al. (2022): We identified the 2018–2020 drought event as a new benchmark having an unprecedented intensity that persisted for more than 2 years, exhibiting a mean areal coverage of 35.6% and an average duration of 12.2 months. What makes this event truly exceptional compared with past events is its near-surface air temperature anomaly reaching +2.8 K, which constitutes further evidence that the ongoing global warming is exacerbating present drought events.

Oudin et al. (2005): The second step confirmed that the four rainfall runoff models considered here do not significantly benefit from PE. This is quite disconcerting from a modelling point of view, because a better streamflow simulation is expected when

using a more relevant evaporative demand. Rainfall-runoff models are poorly responsive to observed classical Penman PE

Oudin et al. (2005): The first step of our study consisted in confirming this suspicion over a large and climatically varied catchment sample, as well as for several representative rainfall-runoff models. The conclusion is that looking for daily observed PE data as inputs into a rainfall-runoff model is not necessary: a long-term average regime curve will serve as well. One of the purposes of this article was to generalize previous results found in the literature and our conclusions substantiate these findings.

Pfeifroth et al. (2018): Box plot of the regional trends (W/m2/decade) of surface solar radiation at the station locations (white) based on stations, (red) SARAH-2, and (blue) CLARA-A2 data, including mean trends (diamonds). Outliers are shown as dots. Regions: CE = central east, CW = central west, NW = northwest, S = south, and N = north. For each region the plot area is divided into trends for the full time period 1983–2015 (for stations and SARAH-2) and trends for the time period 1992–2015 (for stations, SARAH-2, and CLARA-A2).

Pimentel et al. (2023): Figure 2. Global distribution of best PET formula according to: (a) MOD16 PET (115,151 catchments), (b) MOD16 AET (115,151 catchments), and (c) streamflow observations (4,367 catchments).

Pimentel et al. (2023): Hargreaves was the best PET formula in 50% of the catchments, most of them located in the Amazonas, central Europe, and Oceania; Jensen-Haise was better for catchments in northern latitudes (36%). Finally, Priestley-Taylor was the best formula for India and latitudes above 65° North.

Pimentel et al. (2023): The assessment against independent streamflow gauges shows a clear predominance of Priestley-Taylor as the best formula to use as a compromise world-wide (Figure 3c), as it minimizes the RE in about 56% of the analyzed catchments (Table 3, column 7). Jensen-Haise is a suitable alternative PET formula in central North America, South Africa and some spots in central Europe, central Asia, and Australia.

Reaver et al. (2022): The Budyko framework posits that a catchment's long-term mean evapotranspiration (ET) is primarily governed by the availabilities of water and energy, represented by long-term mean precipitation (P) and potential evapotranspiration (PET), respectively.

Seiller and Anctil (2016): The choice of the PET formula moderately affected the calibration efficiency, except for some formulas leading to much larger PET quantities than others, but hydrological models, during their calibration process, adapted their parameters so that efficiency was maximized.

Seiller and Anctil (2016): Combinational formulas lead to quite similar PET quantities and shapes (annual distributions) in both catchments, probably because of structures all emanating from Penman's concept. In contrast, temperature-based formulas provide highly contrasting PET quantities and shapes, from the highest (E12, MOHYSE) to the lowest (E10, Romanenko)

in both catchments. Finally, radiation-based formulas (black) are not as widely spread as the temperature based ones. E17 (Jensen-Haise) is the highest radiation-based PET on both catchments. As expected, FUT PET quantities are higher.

Shah et al. (2023): Figure 1 shows the antecedent meteorological conditions for soil moisture flash drought across Europe by quantifying the proportion of flash drought having precipitation deficit (i.e., $\Delta P < 0$) and above normal evaporative demand (i.e., $\Delta PET > 0$) for three pentads before onset (i.e., $t-3$ to $t-1$) and at onset pentad (i.e., t) for all flash droughts that occurred during 1950–2019.

Shi et al. (2023)b: In specific, JH produced excellent or good estimation of annual ETp across all stations and Tu produced good or fair estimation of annual ETp across all stations.

Tang et al. (2023): The uncertainty shows a notable scale effect. Four scales are investigated according to their spatial (S) and temporal (T) dimensions: ST-both based on all GRUs and time steps, ST-none after spatial and temporal averaging, S-only after temporal averaging, and T-only after spatial averaging (defined in Section 2.4). For the basins and scales of this study, the uncertainty magnitude rank is ST-none < S-only < T-only < ST-both. For T-only, the Spearman correlation coefficients between basin areas and uncertainty reduction compared to ST-both are above 0.7 for most variables, indicating larger spatial scales will have a more evident spatial averaging effect (i.e., smaller T-only uncertainties). For S-only, uncertainty reduction compared to ST-both is substantial but distinct for different variables. The response speed to meteorological inputs could decide the uncertainty reduction caused by temporal averaging: quick-response variables often show larger uncertainty reduction compared to slow-response variables (e.g., baseflow vs. quickflow, and state variables vs. flux variables). As the result, the temporal averaging can change the relative magnitude between variables (i.e., ST-both vs. S-only). This finding shows that averaging simulations in certain ways may lead to an underestimation of the true uncertainty, and modeling on coarser spatiotemporal scales could have quite different uncertainty performance compared to finer scales.

Tarek et al. (2020): The lack of precipitation should therefore become apparent in the simulated hydrograph; however, the streamflow is higher for ERA5 than for the observations, when the opposite would normally be expected. It is important to note that the hydrological model can adapt its mass balance by adjusting the potential evapotranspiration scaling, which it has clearly done in this case. The difference in hydrological modelling then comes from the temporal distribution of precipitation, and it can be seen that the ERA5 winter precipitations are relatively lower in winter than for the rest of the year.

Teuling et al. (2019): In much of central-western Europe, our results show an increase in evapotranspiration of the order of 5 %–15 % between 1955–1965 and 2005–2015, whereas much of the Scandinavian peninsula shows increases exceeding 15 %. The Iberian Peninsula and other parts of the Mediterranean show a decrease of the order of 5 %–15 %. A similar north–south gradient was found for changes in streamflow, although changes in central-western Europe were generally small. Strong de-

creases and increases exceeding 45 % were found in parts of the Iberian and Scandinavian peninsulas, respectively. In Sweden, for example, increased precipitation is a larger driver than large-scale reforestation and afforestation, leading to increases in both streamflow and evapotranspiration. In most of the Mediterranean, decreased precipitation combines with increased forest cover and potential evapotranspiration to reduce streamflow. In spite of considerable local- and regional-scale complexity, the response of net actual evapotranspiration to changes in land use, precipitation, and potential evaporation is remarkably uniform across Europe, increasing by $\sim$ 35–60 km3 yr-1, equivalent to the discharge of a large river. For streamflow, effects of changes in precipitation ($\sim$ 95 km3 yr-1) dominate land use and potential evapotranspiration contributions ($\sim$ 45–60 km3 yr-1).

Thornthwaite (1948): The vegetation of the desert is sparse and uses little water because water is deficient. If more water were available, the vegetation would be less sparse and would use more water. There is a distinction, then, between the amount of water that actually transpires and evaporates and that which would transpire and evaporate if it were available. When water supply increases, as in a desert irrigation project, evapotranspiration rises to a maximum that depends only on the climate. This we may call "potential evapotranspiration," as distinct from actual evapotranspiration

Voisin et al. (2008): Globally and in every continent, simulated runoff is much more sensitive to precipitation differences than evapotranspiration.

Wang et al. (2023): First, the hydrological model can often adapt to some inaccurate rainfall inputs showing good streamflow simulation to some extent. The overall rainfall BIAS helps in predicting hydrological model adaptability to rainfall data sets. Furthermore, the hydrological model adaptability has an adaptable threshold of overall rainfall BIAS, which is -140 mm for our study (total rainfall is about 518 mm). Second, hydrological model adaptability is not only affected by the overall rainfall BIAS but also significantly influenced by the event-based rainfall BIAS. How the event-based rainfall bias of rainstorms shapes the overall rainfall bias affects hydrological model adaptability. In particular, the rainfall accuracy during the heaviest rainstorm which results in peak flow has a large impact. The hydrological model would be unable to adapt well to the large overall rainfall bias that is mainly caused by a significant underestimation of the severest rainstorm.

Xiang et al. (2020): Early applications of ETp were in meteorology and hydrology, but later its applications were extended to other fields (Parajuli et al., 2019; Zhang et al.,2019a,b). For example, in agronomy it was related to crop water requirements and irrigation scheduling (Yao et al., 2020; Feng et al., 2019).

Zhang et al. (2016): The dominant contributor to ET variability is ET0 under energy-limited condition (aridity index $\leqslant 0.76$), whereas the dominant contributor is precipitation under equitant ($0.76 <$ aridity index $\leqslant 1.35$) and water-limited conditions (aridity index $\geqslant 1.35$). The contribution of ET0 to ET variability decreases with the aridity index, whereas the contribution of precipitation to ET variability increases with the aridity index.

Zhang et al. (2016): Fig 2. The long-term hydroclimatological characteristics of the 282 catchments based on Budyko hypothesis. All the catchments are classified into three climate conditions: energy-limited condition (where aridity index (ET0/P) $\leqslant$ 0.76), equitant condition (0.76 < aridity index $\leqslant$ 1.35) and water-limited condition (aridity index $\geqslant$ 1.35).

Zhang et al. (2016): ET0 is estimated by the Penman formula (Penman, 1948, Shuttleworth, 1993), which is showed to be the most appropriate form when considering a changing climate (Donohue et al., 2010a, McVicar et al., 2012a).

Zhang et al. (2019): Fig 5. The global distribution of relative contributions of PCRU, ETMODIS, and R in explaining TWSC.

Zhao et al. (2013): There are two groups of evapotranspiration estimation methods in hydrological models: one first estimates separately water surface evaporation, soil evaporation and vegetable transpiration, and then integrates them to get the basin evapotranspiration depending on the land use pattern. The other one first estimates potential evapotranspiration (ETp) and then converts it into actual evapotranspiration (ETa) applying the Soil Moisture Extraction Function. In this paper the first type methods are called the classification gathering methods, while the second – the integrated converting methods.

S2. Seasonal changes in hydrological cycle components

The summer season (DJF) exhibits the smallest changes across each hydrological component (Figure S10). For PET, temperature-based methods indicate a strong positive trend compared to combination-type methods in energy-limited and mixed catchments. Blaney-Criddle demonstrates the largest trend among all methods for energy-limited and mixed catchments, while Hargreaves-Samani presents the largest trend in water-limited catchments. Only the Thornthwaite method indicates a negative trend for water-limited catchments. For AET, the trend patterns for each PET method are similar to those observed for PET. In energy-limited and mixed catchments, temperature-based and energy-based methods reflect positive changes at the median level, while combination-type methods indicate no change in AET. In water-limited catchments, combination-type methods present the most pronounced changes, around 0.1 mm seas^{-1} year^{-1}, whereas Thornthwaite indicates a negative trend (Figure S10). Almost all PET methods indicate similar changes in Q across catchment categories. For TWS, energy-limited catchments appear insensitive to the PET method selection, while in mixed catchments, all PET methods show negative changes, with Hargreaves-Samani demonstrating the most negative trend and Blaney-Criddle reflecting the greatest variability. In water-limited catchments, all PET methods indicate negative changes except for Blaney-Criddle and the Milly-Dunne method.

In the spring season (MAM), all PET methods demonstrate positive changes in PET for each catchment category (Figure S6). The Hargreaves-Samani method exhibits the highest trend, followed by Blaney-Criddle, across all catchment types for PET. The overall variability in PET methods decreases from energy-limited to mixed, and then to water-limited catchments. For AET (Figure S6), all PET methods reflect a positive trend in each catchment category. In energy-limited and mixed catchments, Blaney-Criddle and Hargreaves-Samani reveal the highest trends, with Hargreaves-Samani leading in energy-limited catchments. Radiation and combination-type methods demonstrate greater variability compared to temperature-based methods in energy-limited catchments, while in mixed catchments, temperature-based methods reveal greater variability. In water-limited catchments, the Blaney-Criddle method exhibits larger variability. Q indicates almost no change among PET methods in energy-limited catchments (Figure S6). However, in water-limited catchments, Blaney-Criddle indicates the least change, while other PET methods exhibit approximately 0.5 mm seas^{-1} year^{-1}. In water-limited catchments, Blaney-Criddle reflects the least change, while the highest trend (0.4 mm seas^{-1} year^{-1}) is indicated by the Thornthwaite method. Overall, energy-limited methods reflect higher trends compared to temperature and combination-type methods. TWS remains largely insensitive to PET method selection for energy-limited catchments. For mixed catchments, Hargreaves-Samani demonstrates the highest negative trend compared to the median of other PET methods (Figure S6). For water-limited catchments, all PET meth-

ods demonstrate positive trends for the TWS hydrological component. The variability of energy and combination-type methods is less than that of temperature-based methods in mixed catchment categories. Among the temperature-based methods, Blaney-Criddle exhibits higher variability in mixed and water-limited catchments. In water-limited catchments, Blaney-Criddle shows the highest positive trend, followed by the energy-based Milly-Dunne method.

In the autumn season (SON), all PET methods reflect trends similar to those in the summer season, but with reduced magnitudes (Figure S8). The Jensen-Haise method records the highest trends, with median values of approximately 0.2, 0.3, and 0.5 mm $seas^{-1}$ $year^{-1}$ for energy-limited, mixed, and water-limited catchments, respectively. For AET, all PET methods indicate a positive trend compared to the median. water-limited catchments exhibit greater variability in response to PET methods than energy-limited and mixed catchments. In mixed catchments, nearly zero trends are observed across all PET methods, with Blaney-Criddle and Jensen-Haise demonstrating higher variability. Q shows little sensitivity to trends in PET method across all catchment categories, with slight variability observed among the methods within each catchment category. For TWS, energy-limited and mixed catchments are largely unaffected to the PET method selection, except for Jensen-Haise, which shows a negative trend. Although there is slight variation among PET methods within each category, mixed behavior is observed in water-limited catchments. Some temperature-based methods, such as Blaney-Criddle, McGuinness-Bordne, and Hargreaves-Samani, indicate positive trends, while others reflect negative trends. All combination-type methods, like Penman-Monteith and Penman-Monteith[CO_2], demonstrate negative trends.

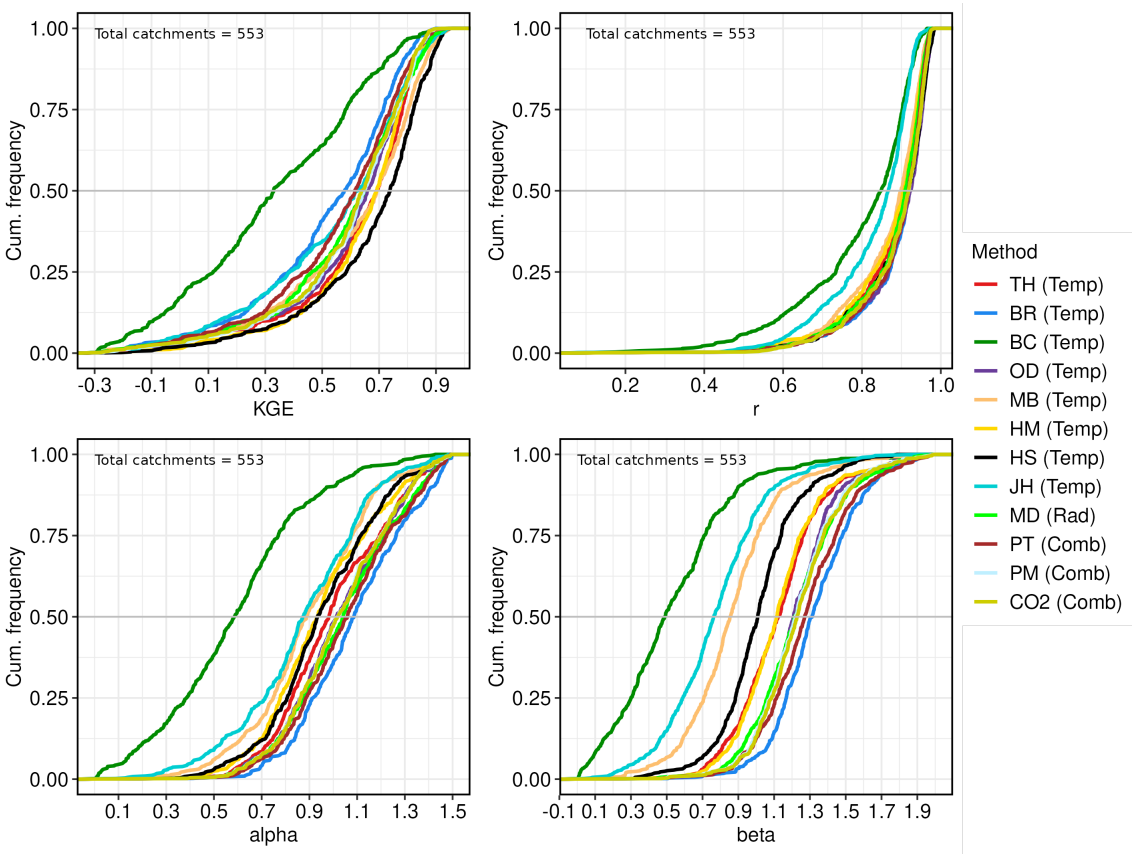

Figure S1. Evaluation of the hydrological model (mHM) performance using simulated monthly streamflow across 553 catchments, forced with EM-Earth meteorological data. The figure presents the cumulative frequency distributions of the Kling–Gupta efficiency (KGE) and its three components: correlation (r), variability ratio (alpha), and bias ratio (beta), providing insights into the model's performance across different PET methods.

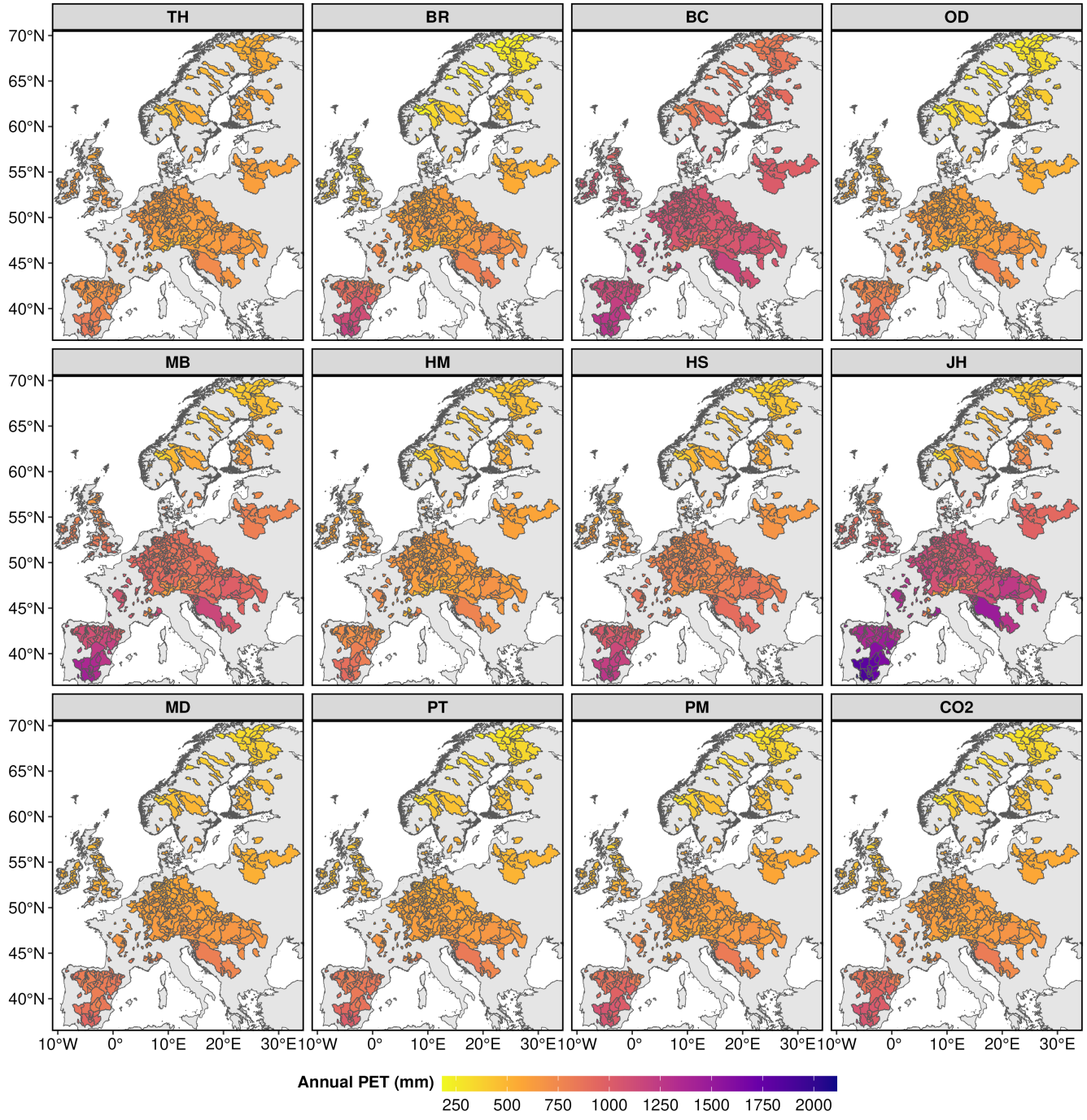

Figure S2. Average annual PET across European catchments for each PET method. Abbreviations used for different PET methods are TH: Thornthwaite, BR: Baier-Robertson, BC: Blaney-Criddle, OD: Oudin, MB: McGuinness-Bordne, HM: Hamon, HS: Hargreaves-Samani, JH: Jensen-Haise, MD: Milly-Dunne, PT: Priestley-Taylor, PM: Penman-Monteith, CO_2: Penman-Monteith[CO_2].

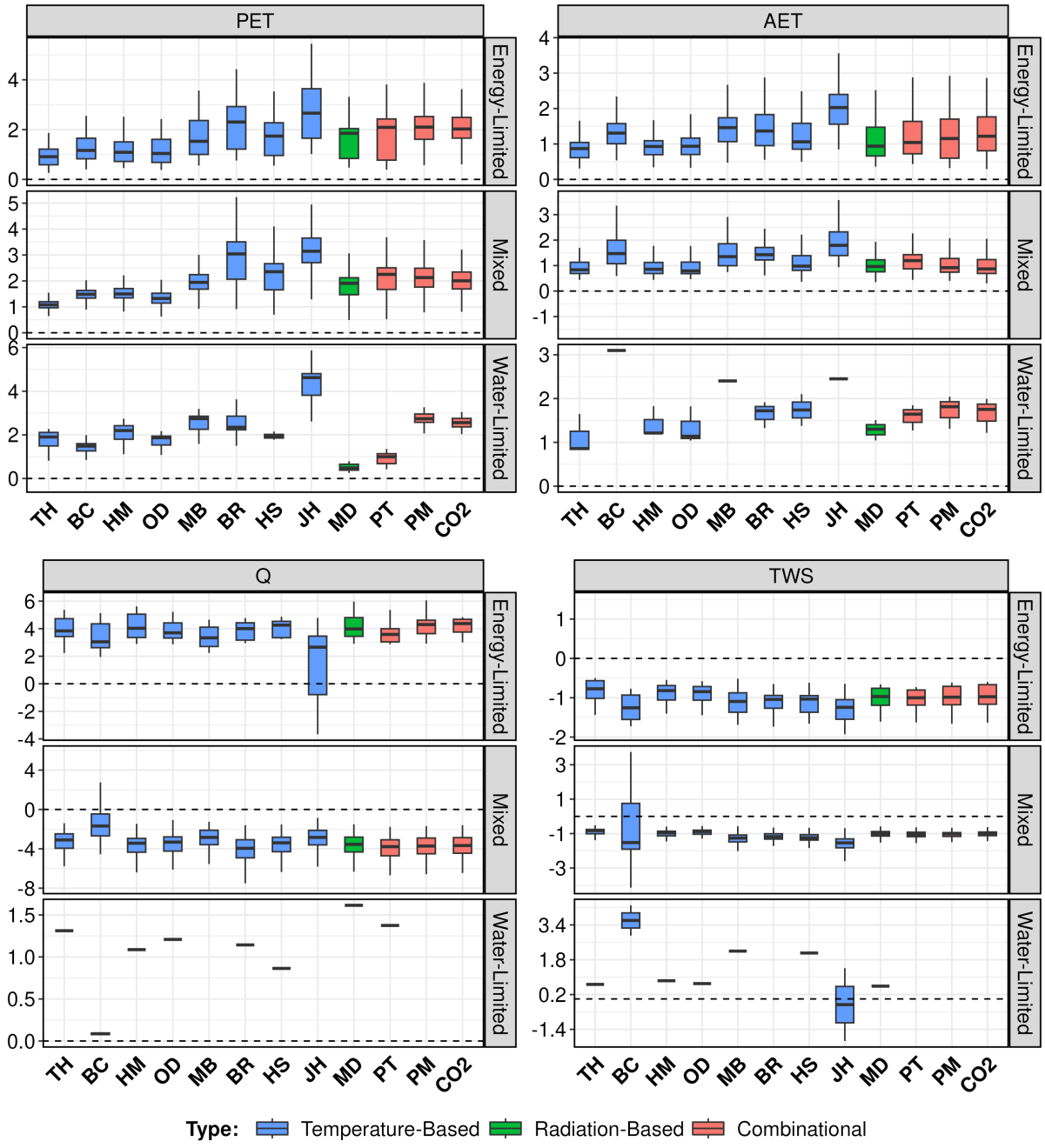

Figure S3. Boxplots represent significant annual trends (mm year^{-1}) of different PET methods for PET, AET, Q, and TWS across energy-limited and mixed catchments. The whiskers represent the 10th and 90th percentiles, and the box encompasses the 25th and 75th percentiles, with the median represented by middle line of the box. Abbreviations used for different PET methods are TH: Thornthwaite, BR: Baier-Robertson, BC: Blaney-Criddle, OD: Oudin, MB: McGuinness-Bordne, HM: Hamon, HS: Hargreaves-Samani, JH: Jensen-Haise, MD: Milly-Dunne, PT: Priestley-Taylor, PM: Penman-Monteith, CO_2: Penman-Monteith[CO_2].

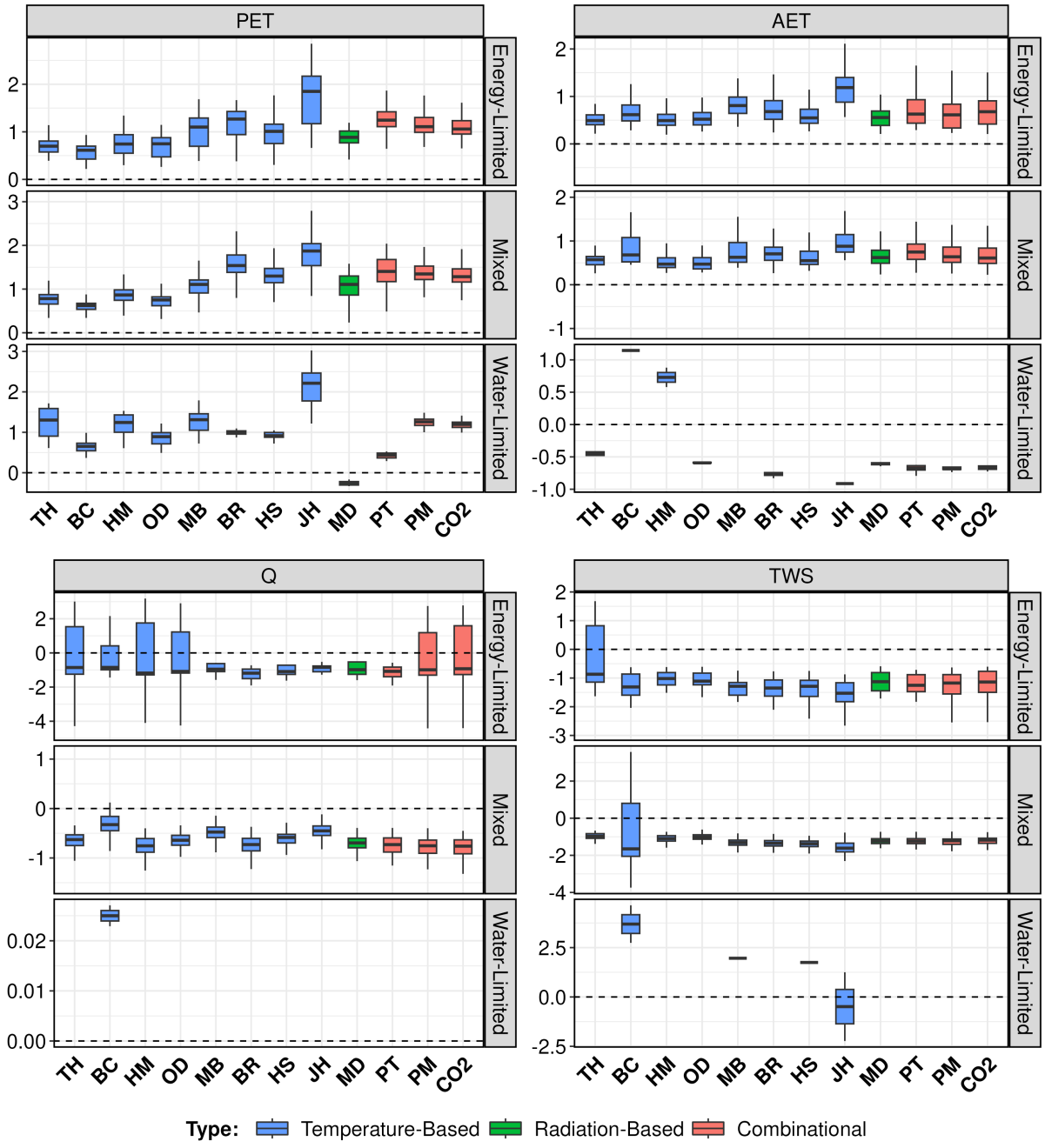

Figure S4. Boxplot represents significant seasonal (summer season (JJA)) trends of different PET methods for PET, AET, Q, and TWS across categories of catchment. The whiskers represent the 10th and 90th percentiles, and the box encompasses the 25th and 75th percentiles, with the median represented by the black line within the box. Abbreviations used for different PET methods are TH: Thornthwaite, BR: Baier-Robertson, BC: Blaney-Criddle, OD: Oudin, MB: Mcguinnes-Bordne, HM: Hamon, HS: Hargreaves-Samani, JH: Jensen-Haise, MD: Milly-Dunne, PT: Priestley-Taylor, PM: Penman-Monteith, CO_2: Penman-Monteith[CO_2]. Trend units are in mm seas^{-1} year^{-1}.

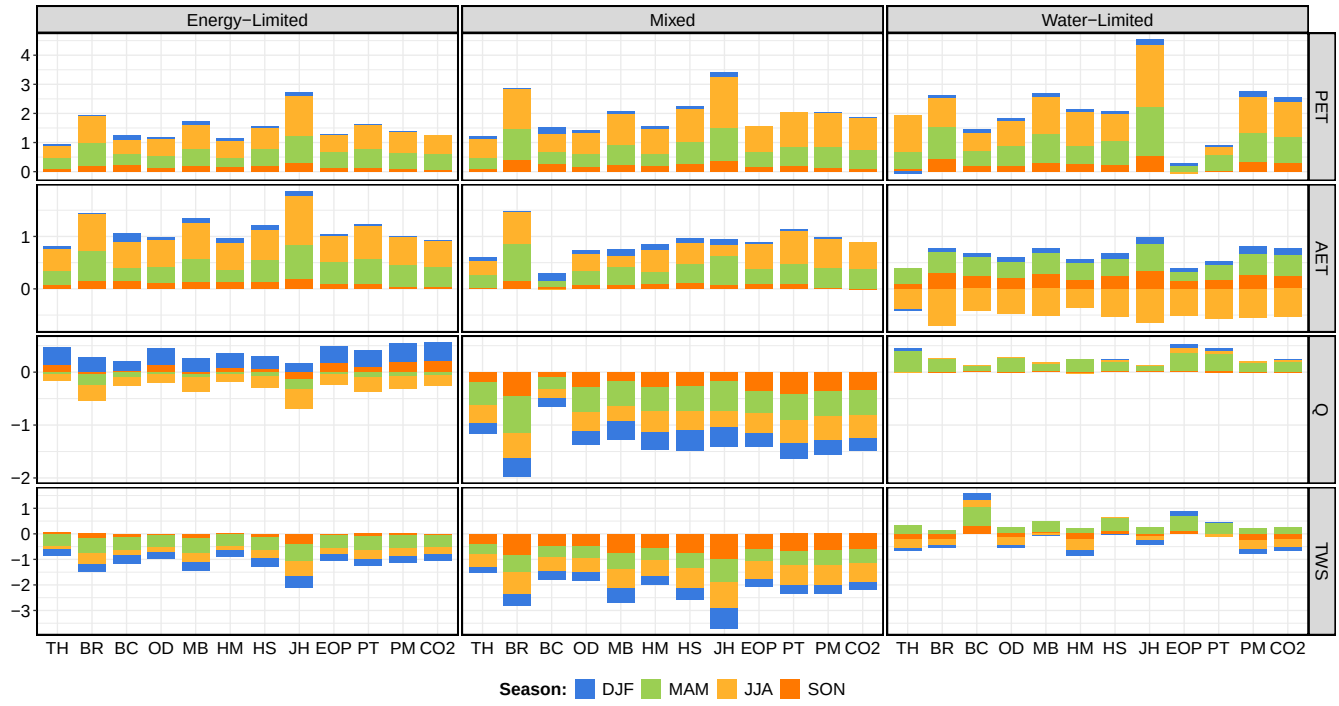

Figure S5. Seasonal trend contribution in annual trend for each catchment category for PET and all the hydrological components (AET, Q, and TWS). Abbreviations used for hydrological cycle components are as follows: PET, AET, Q, and TWS represent potential evapotranspiration, actual evapotranspiration, runoff and, total water storage, respectively. Abbreviations used for different PET methods are TH: Thornthwaite, BR: Baier-Robertson, BC: Blaney-Criddle, OD: Oudin, MB: McGuinness-Bordne, HM: Hamon, HS: Hargreaves-Samani, JH: Jensen-Haise, MD: Milly-Dunne, PT: Priestley-Taylor, PM: Penman-Monteith, CO_2: Penman-Monteith[CO_2]. For seasons DJF, MAM, JJA, and SON indicate summer, spring, summer, and autumn, respectively. Y-axis units are in mm seas^{-1} year^{-1}.

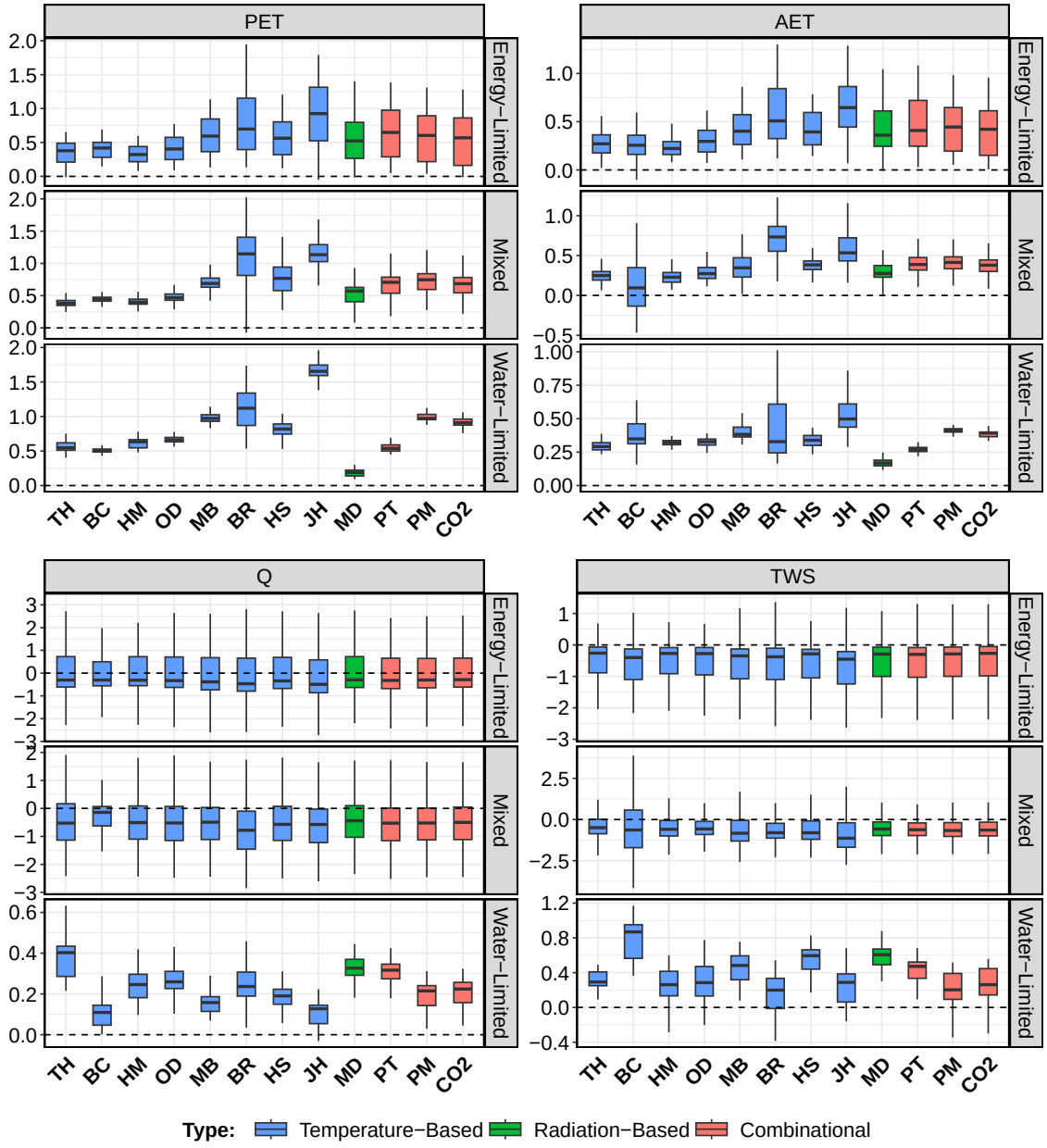

Figure S6. Boxplot represents the seasonal (spring season (MAM)) trends of different PET methods for AET, and Q across three categories of catchment: energy-limited, mixed, and water-limited. The whiskers represent the 10th and 90th percentiles, and the box encompasses the 25th and 75th percentiles, with the median represented by the black line within the box. Trend units are in mm seas^{-1} year^{-1}.

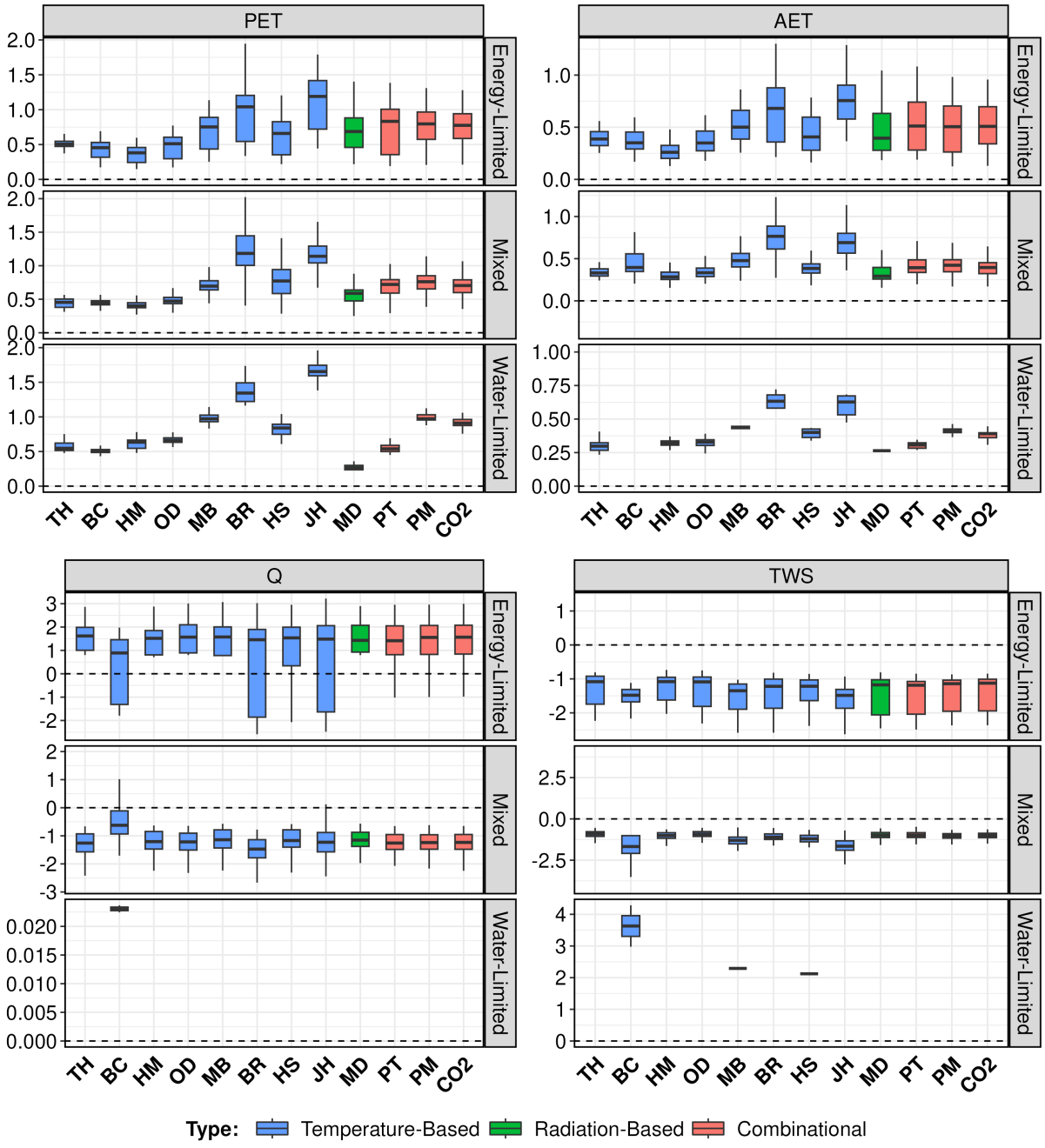

Figure S7. Boxplot represents significant seasonal (spring season (MAM)) trends of different PET methods for PET, AET, Q, and TWS across categories of catchment. The whiskers represent the 10th and 90th percentiles, and the box encompasses the 25th and 75th percentiles, with the median represented by the black line within the box. Abbreviations used for different PET methods are TH: Thornthwaite, BR: Baier-Robertson, BC: Blaney-Criddle, OD: Oudin, MB: Mcguinnes-Bordne, HM: Hamon, HS: Hargreaves-Samani, JH: Jensen-Haise, MD: Milly-Dunne, PT: Priestley-Taylor, PM: Penman-Monteith, CO_2: Penman-Monteith[CO_2]. Trend units are in mm seas^{-1} year^{-1}.

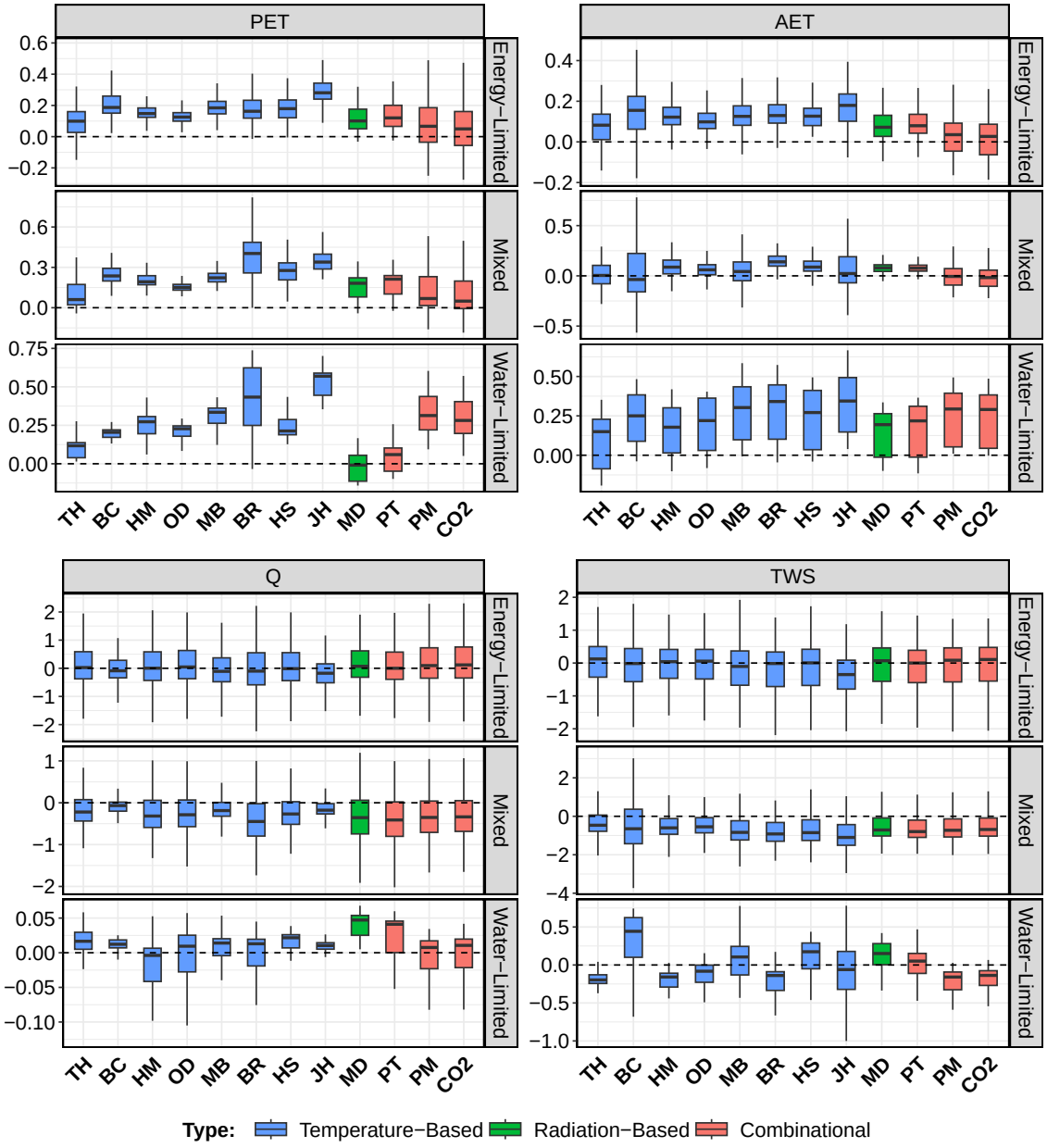

Figure S8. Boxplot represents the seasonal (autumn season (SON)) trends of different PET methods for AET, and Q across three categories of catchment: energy-limited, mixed, and water-limited. The whiskers represent the 10th and 90th percentiles, and the box encompasses the 25th and 75th percentiles, with the median represented by the black line within the box. Trend units are in mm seas^{-1} year^{-1}.

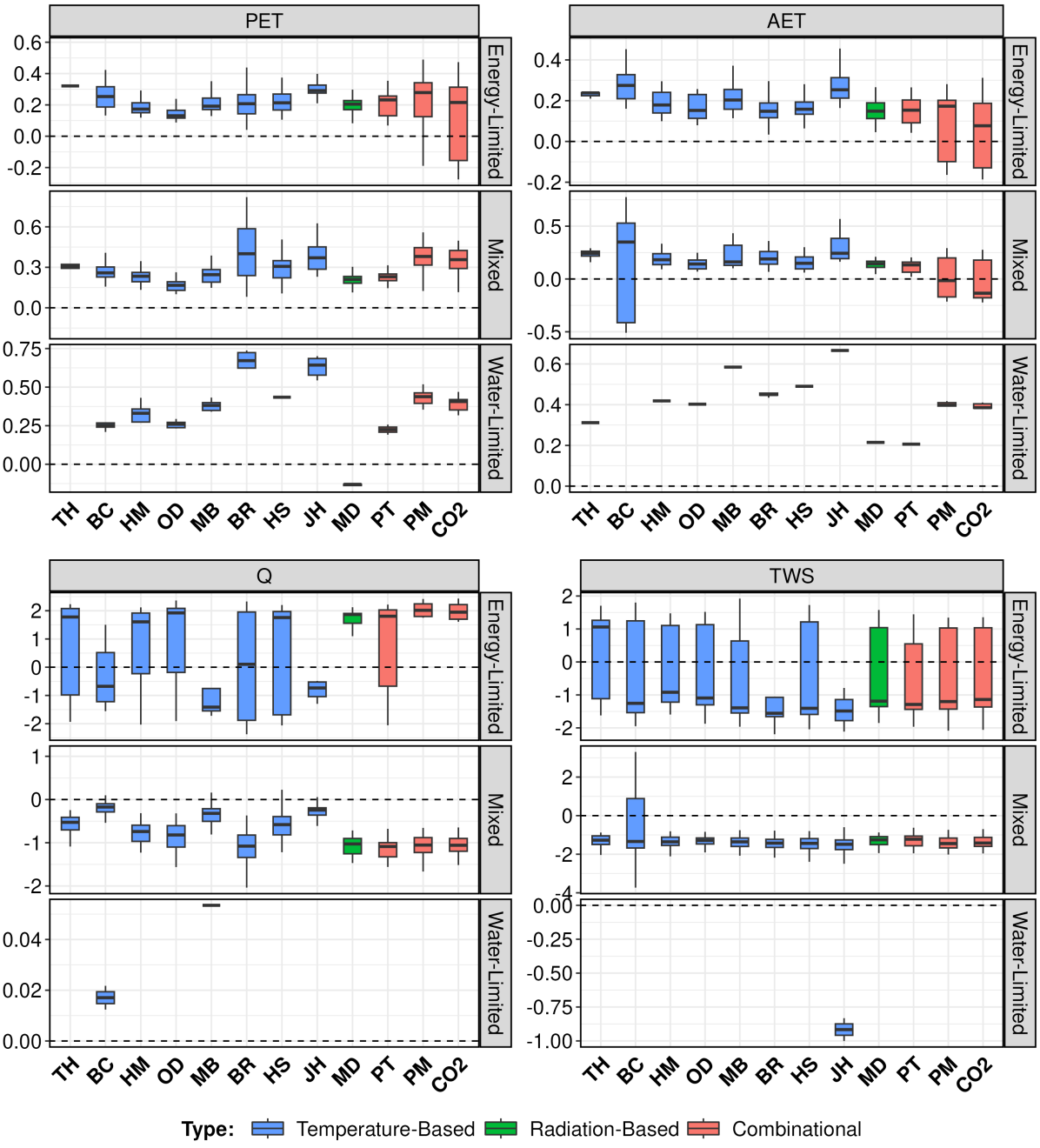

Figure S9. Boxplot represents significant seasonal (autumn season (SON)) trends of different PET methods for PET, AET, Q, and TWS across categories of catchment. The whiskers represent the 10th and 90th percentiles, and the box encompasses the 25th and 75th percentiles, with the median represented by the black line within the box. Abbreviations used for different PET methods are TH: Thornthwaite, BR: Baier-Robertson, BC: Blaney-Criddle, OD: Oudin, MB: Mcguinness-Bordne, HM: Hamon, HS: Hargreaves-Samani, JH: Jensen-Haise, MD: Milly-Dunne, PT: Priestley-Taylor, PM: Penman-Monteith, CO_2: Penman-Monteith[CO_2]. Trend units are in mm seas^{-1} year^{-1}.

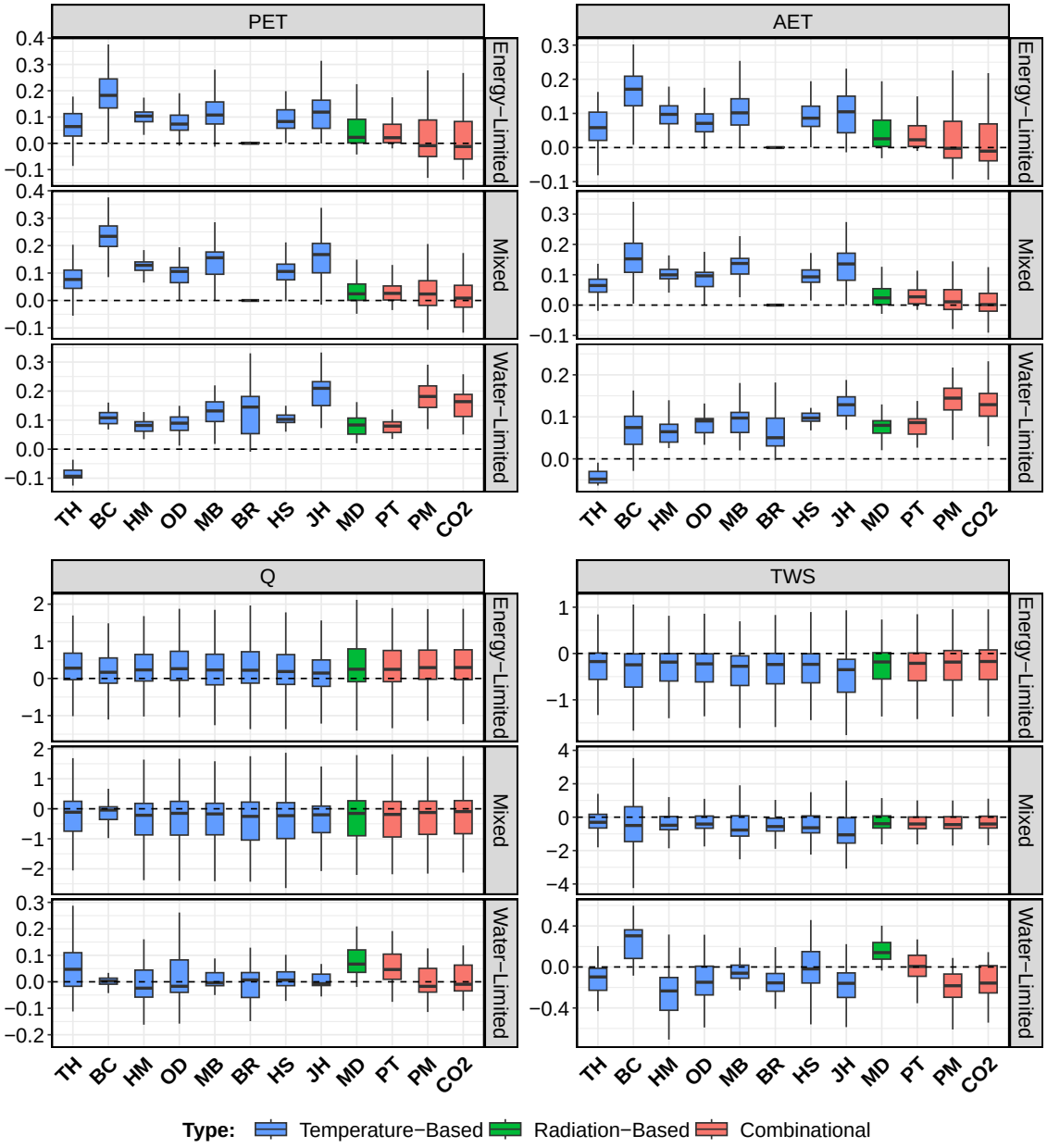

Figure S10. Boxplot represents the seasonal (winter season (DJF)) trends of different PET methods for AET, and Q across three categories of catchment: energy-limited, mixed, and water-limited. The whiskers represent the 10th and 90th percentiles, and the box encompasses the 25th and 75th percentiles, with the median represented by the black line within the box. Trend units are in mm seas^{-1} year^{-1}.

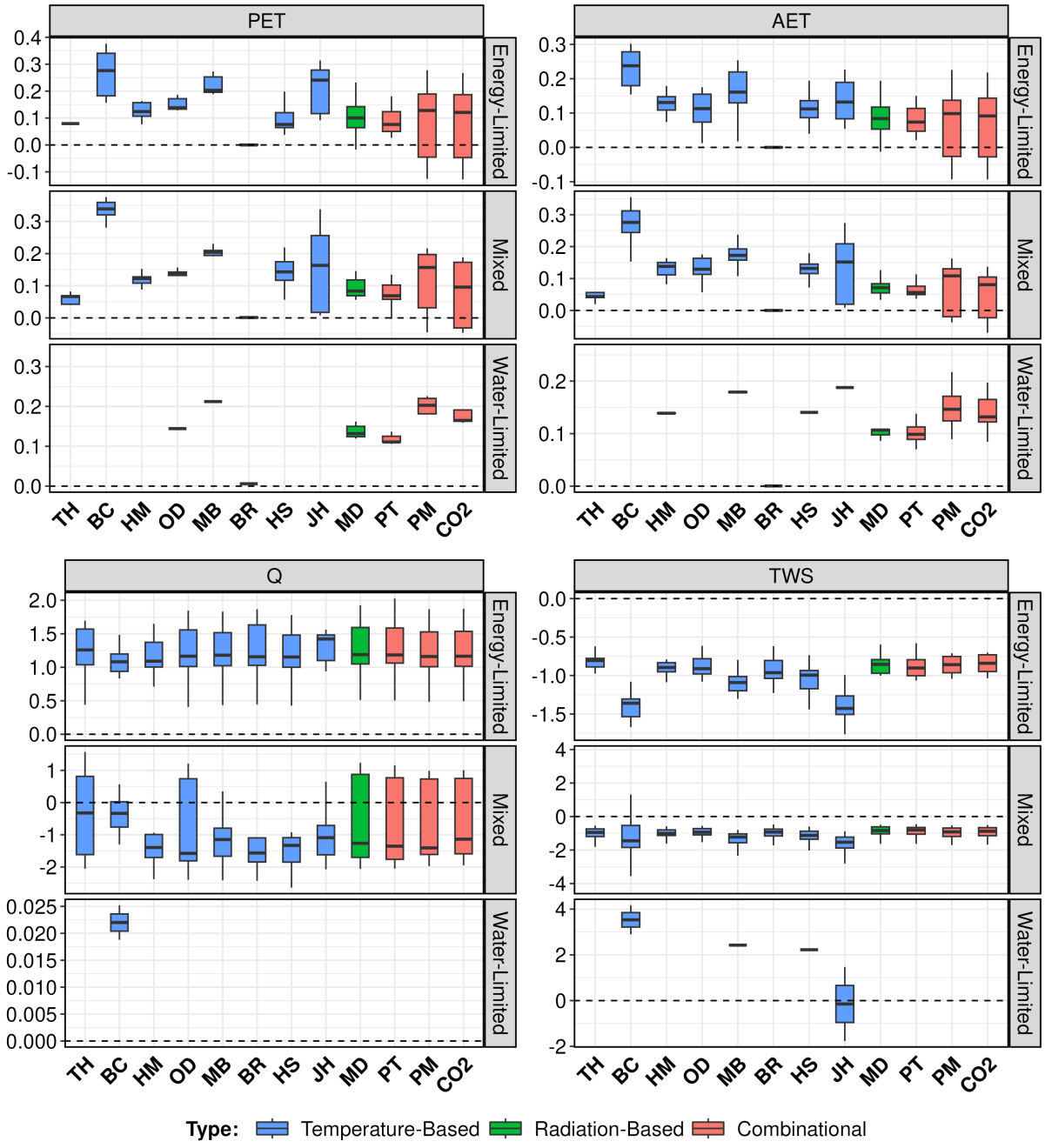

Figure S11. Boxplot represents significant seasonal (winter season (DJF)) trends of different PET methods for PET, AET, Q, and TWS across categories of catchment. The whiskers represent the 10th and 90th percentiles, and the box encompasses the 25th and 75th percentiles, with the median represented by the black line within the box. Abbreviations used for different PET methods are TH: Thornthwaite, BR: Baier-Robertson, BC: Blaney-Criddle, OD: Oudin, MB: Mcguinnes-Bordne, HM: Hamon, HS: Hargreaves-Samani, JH: Jensen-Haise, MD: Milly-Dunne, PT: Priestley-Taylor, PM: Penman-Monteith, CO_2: Penman-Monteith[CO_2]. Trend units are in mm seas^{-1} year^{-1}.

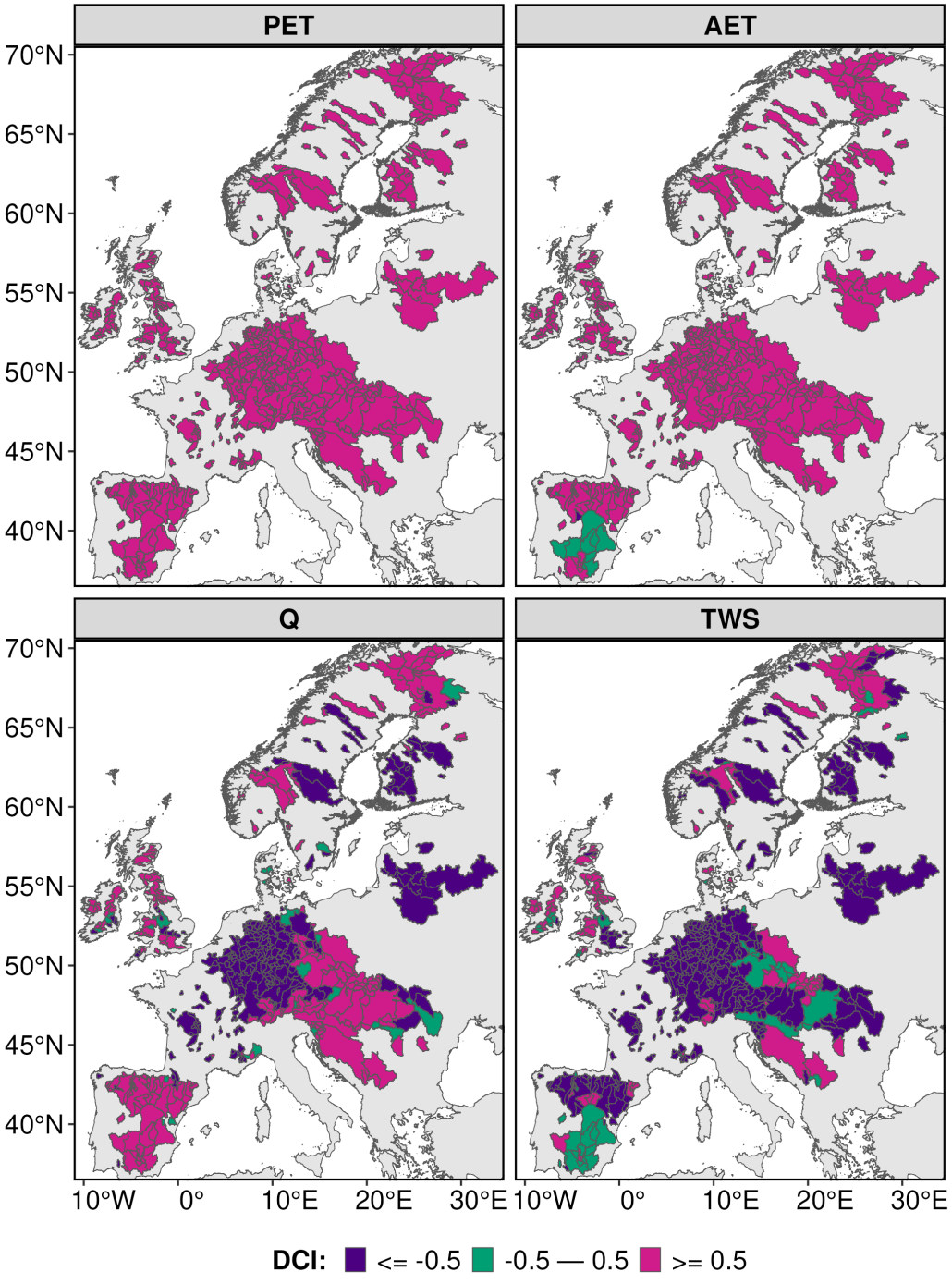

Figure S12. Spatial distribution of annual scale data concurrence index (DCI) for PET, AET, Q, and TWS. Considering eight PET methods namely Blaney-Criddle, Baier-Robertson, Hamon, Hargreaves-Samani, Milly-Dunne, Priestley-Taylor, Penman-Monteith and Modified Penman-Monteith accounts CO2.

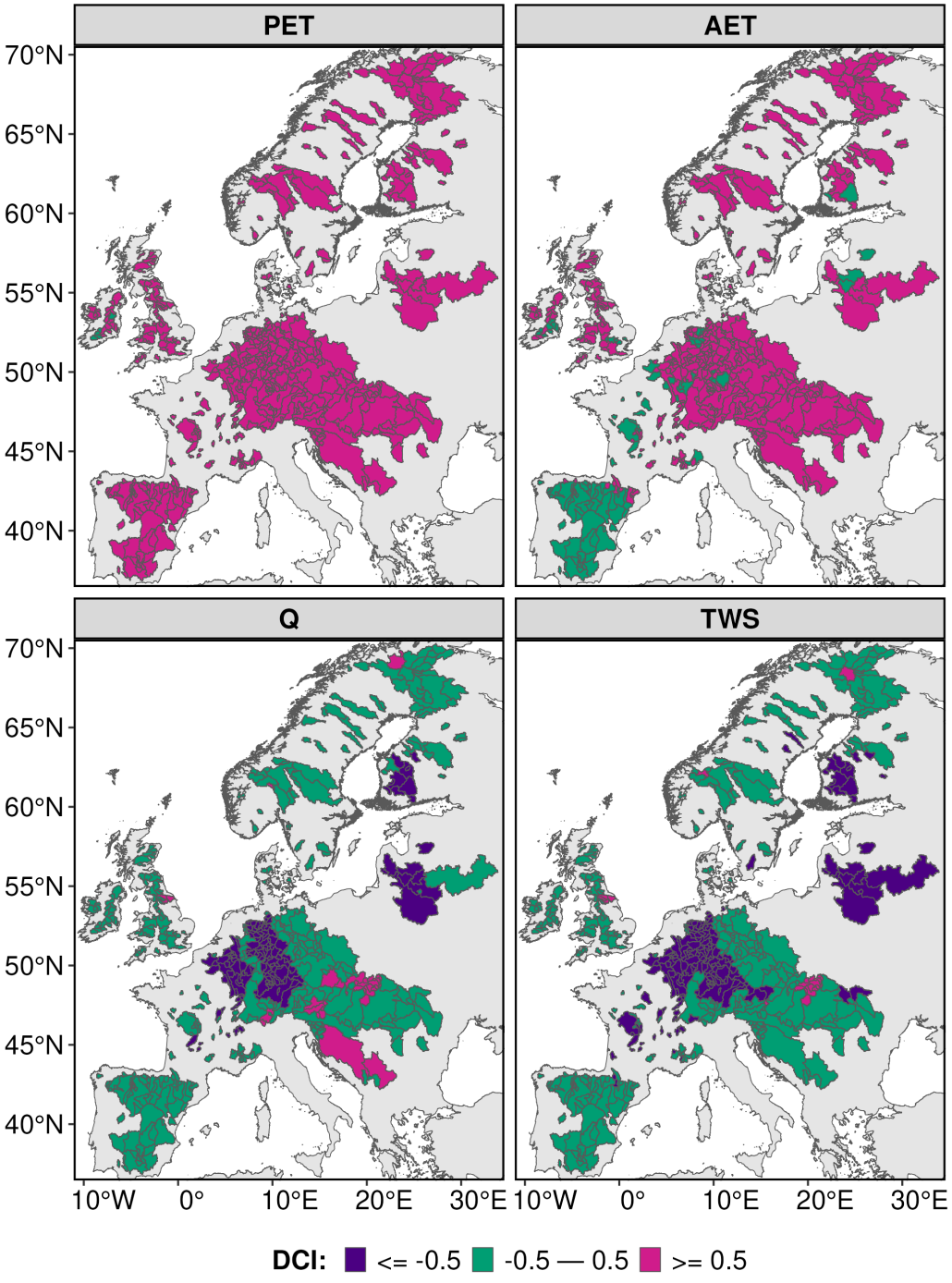

Figure S13. Spatial distribution of annual scale data concurrence index (DCI) for PET, AET, Q, and TWS. The DCI value at each shown catchment corresponds to eight significant trends (one radiation-based, three combinational, and four from temperature-based methods). Where DCI represents data concurrence index, PET represents potential evapotranspiration, AET represents actual evapotranspiration, Q represents runoff at the outlet of the catchment and TWS represents total water storage.

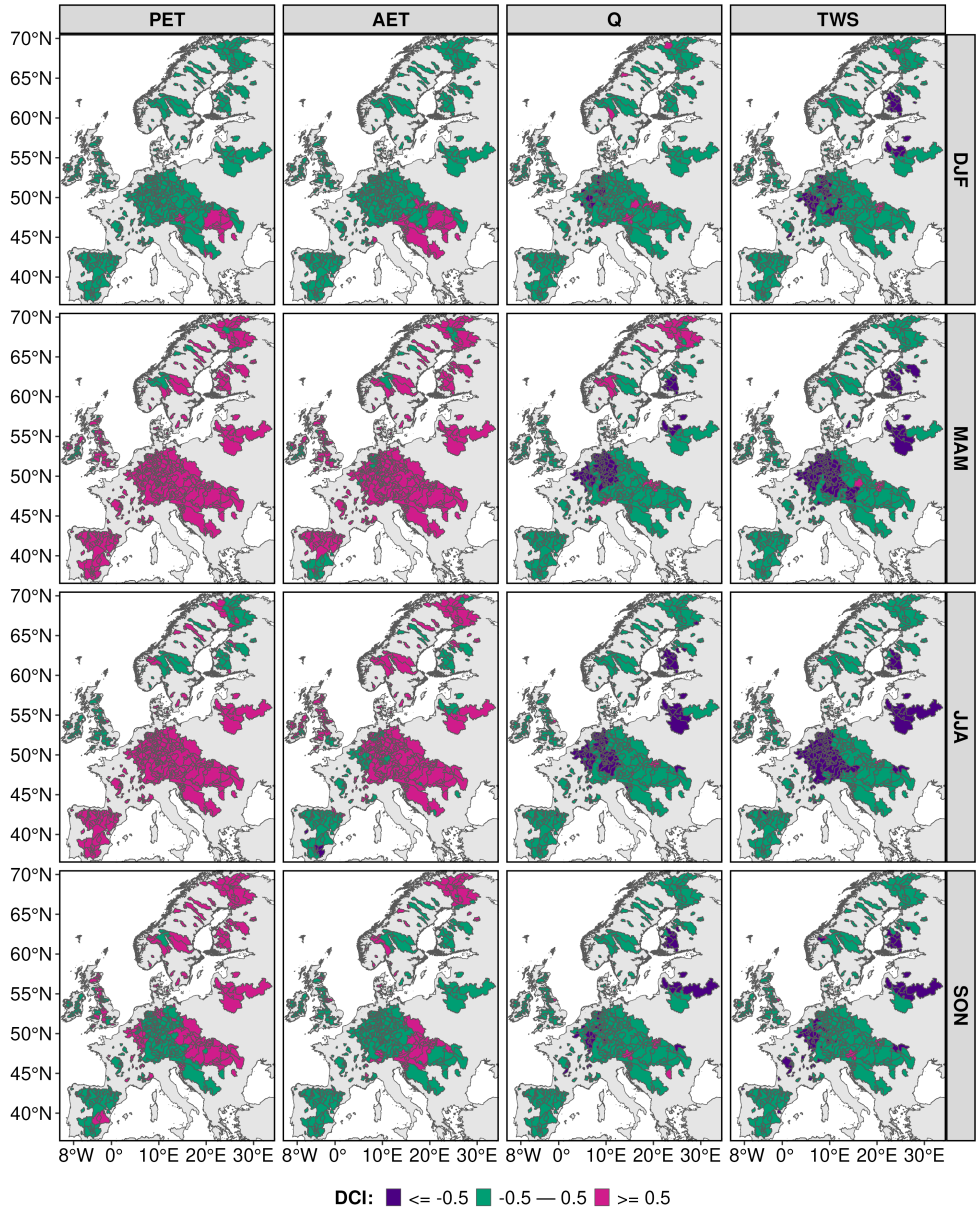

Figure S14. Spatial distribution of seasonal scale (winter (DJF), spring (MAM), summer (JJA), and autumn (SON)) DCI for PET, AET, Q, and TWS. The DCI value at each shown catchment corresponds to eight significant trends (one radiation-based, three combinational, and four from temperature-based methods). Where DCI represents data concurrence index, PET represents potential evapotranspiration, AET represents actual evapotranspiration, Q represents runoff at the outlet of the catchment and TWS represents total water storage.

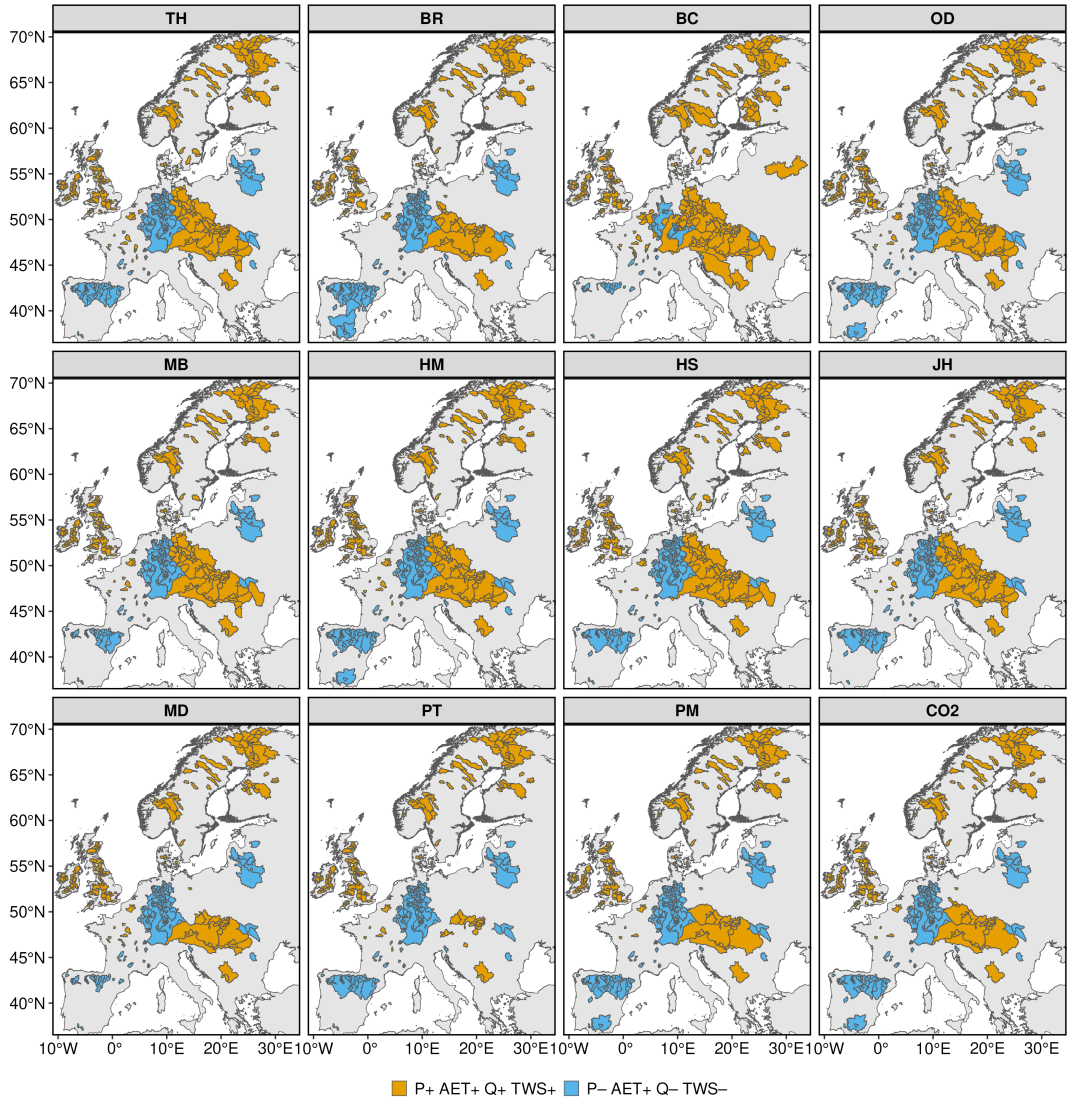

Figure S15. Spatial distribution of catchments corresponding to two hydrological cycle patterns for each PET method. Yellow catchment denotes an increase in all hydrological cycle components. Blue catchments represent a decrease in P, Q, and TWS and an increase in AET. Abbreviations used for different PET methods are TH: Thornthwaite, BR: Baier-Robertson, BC: Blaney-Criddle, OD: Oudin, MB: McGuinness-Bordne, HM: Hamon, HS: Hargreaves-Samani, JH: Jensen-Haise, MD: Milly-Dunne, PT: Priestley-Taylor, PM: Penman-Monteith, CO_2: Penman-Monteith[CO_2].

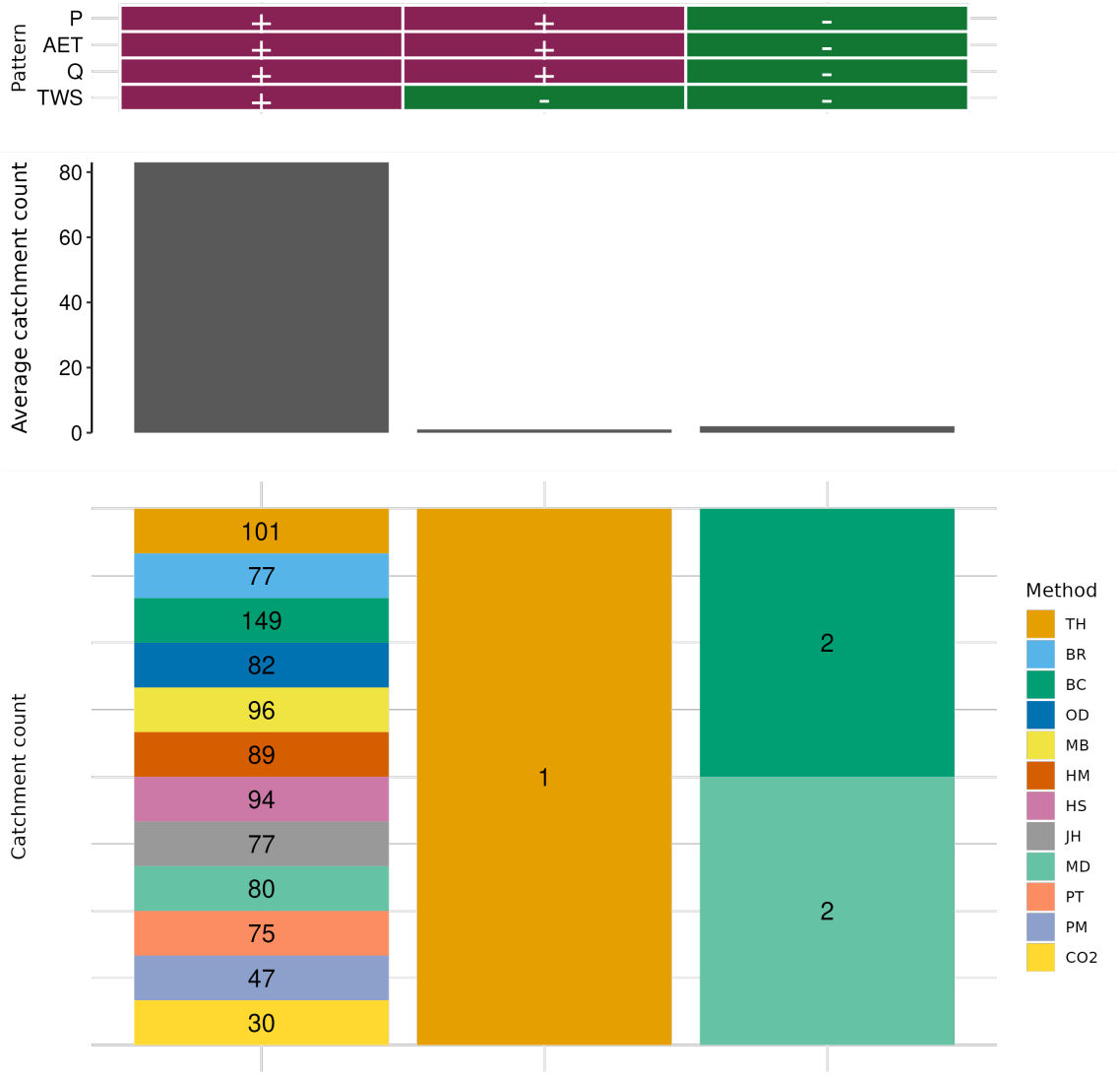

Figure S16. Pattern of hydrological cycle components and the corresponding influence of PET methods at the annual scale, considering only statistically significant trends. The first panel represents different patterns of hydrological cycle components. Each vertical column in this table corresponds to one pattern of hydrological cycle. For example, the first column is filled with '+' signs, indicating that all hydrological components (P, AET, Q, and TWS) exhibit positive changes. The '+' and '-' signs denote positive and negative changes in the respective components. In the second panel, each bar represents the average number of catchments for each hydrological cycle pattern. The third panel shows the number of catchments associated with each PET method for the corresponding hydrological cycle patterns. The color of each cell represents a specific PET method, and each column aligns with the hydrological cycle pattern represented in the corresponding column of Panel One. Where P is precipitation, AET is actual evapotranspiration, Q is runoff and TWS is total water storage. Abbreviations used for different PET methods are TH: Thornthwaite, BR: Baier-Robertson, BC: Blaney-Criddle, OD: Oudin, MB: McGuinness-Bordne, HM: Hamon, HS: Hargreaves-Samani, JH: Jensen-Haise, MD: Milly-Dunne, PT: Priestley-Taylor, PM: Penman-Monteith, CO_2: Penman-Monteith[CO_2].

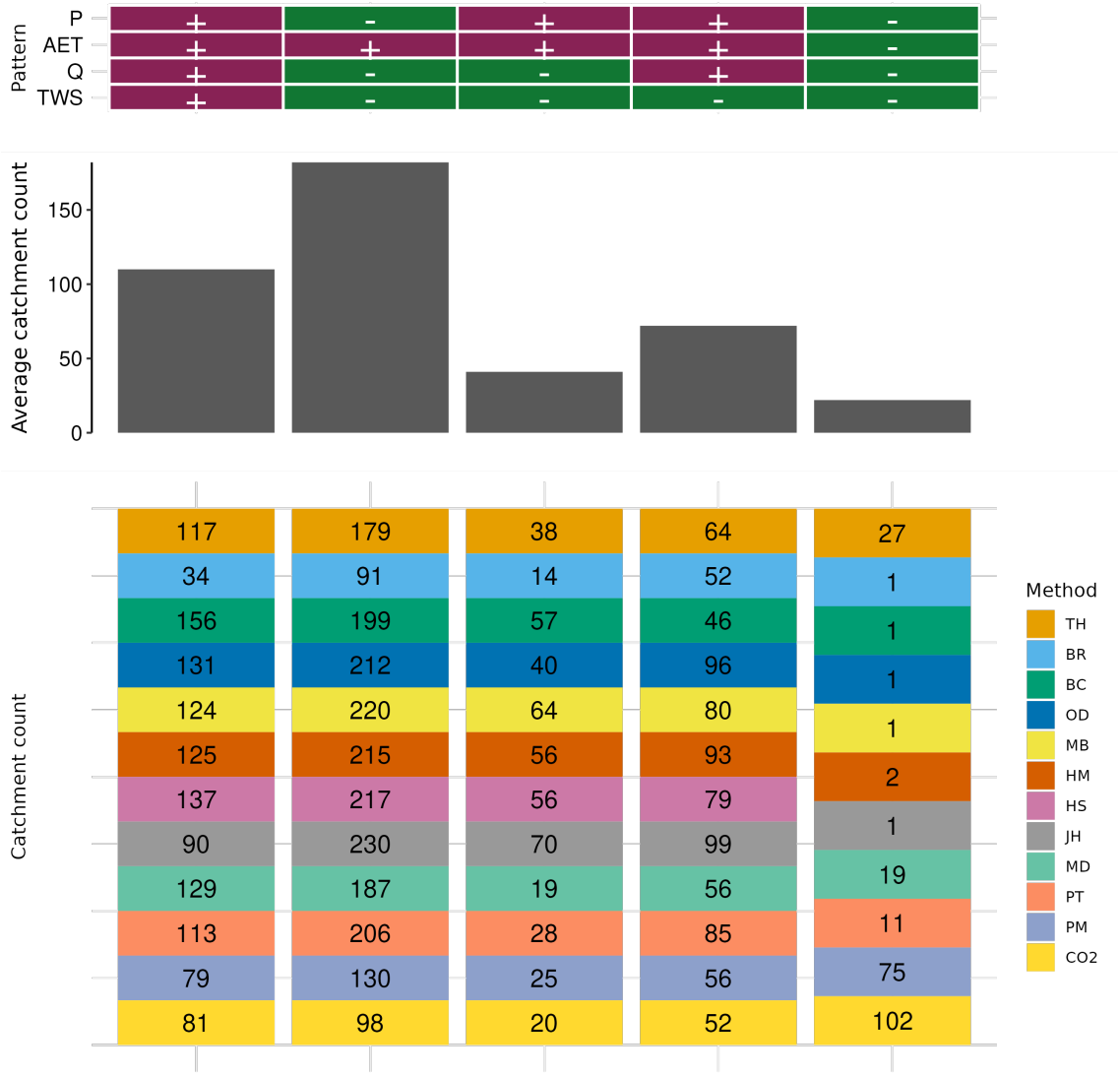

Figure S17. Pattern of different hydrological cycle components and the corresponding influence of PET methods for winter season (DJF). The first panel represents different patterns of hydrological cycle components. Each vertical column in this table corresponds to one pattern of hydrological cycle. For example, the first column is filled with '+' signs, indicating that all hydrological components (P, AET, Q, and TWS) exhibit positive changes. The '+' and '-' signs denote positive and negative changes in the respective components. In the second panel, each bar represents the average number of catchments for each hydrological cycle pattern. The third panel shows the number of catchments associated with each PET method for the corresponding hydrological cycle patterns. The color of each cell represents a specific PET method, and each column aligns with the hydrological cycle pattern represented in the corresponding column of Panel One. Where P is precipitation, AET is actual evapotranspiration, Q is runoff and TWS is total water storage. Abbreviations used for different PET methods are TH: Thornthwaite, BR: Baier-Robertson, BC: Blaney-Criddle, OD: Oudin, MB: McGuinness-Bordne, HM: Hamon, HS: Hargreaves-Samani, JH: Jensen-Haise, MD: Milly-Dunne, PT: Priestley-Taylor, PM: Penman-Monteith, CO_2: Penman-Monteith[CO_2].

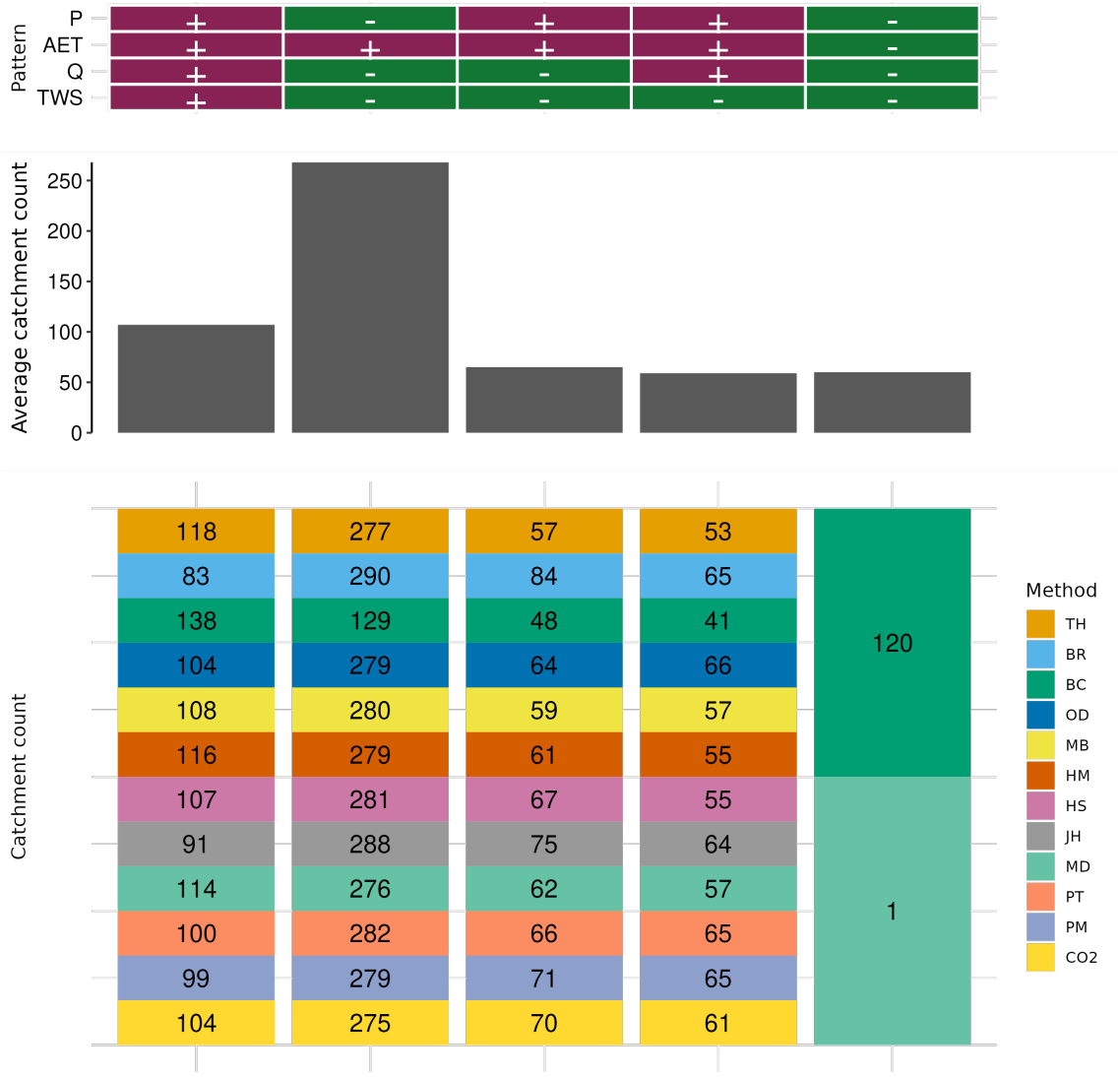

Figure S18. Pattern of different hydrological cycle components and the corresponding influence of PET methods for spring season (MAM). The first panel represents different patterns of hydrological cycle components. Each vertical column in this table corresponds to one pattern of hydrological cycle. For example, the first column is filled with '+' signs, indicating that all hydrological components (P, AET, Q, and TWS) exhibit positive changes. The '+' and '-' signs denote positive and negative changes in the respective components. In the second panel, each bar represents the average number of catchments for each hydrological cycle pattern. The third panel shows the number of catchments associated with each PET method for the corresponding hydrological cycle patterns. The color of each cell represents a specific PET method, and each column aligns with the hydrological cycle pattern represented in the corresponding column of Panel One. Where P is precipitation, AET is actual evapotranspiration, Q is runoff and TWS is total water storage. Abbreviations used for different PET methods are TH: Thornthwaite, BR: Baier-Robertson, BC: Blaney-Criddle, OD: Oudin, MB: McGuinness-Bordne, HM: Hamon, HS: Hargreaves-Samani, JH: Jensen-Haise, MD: Milly-Dunne, PT: Priestley-Taylor, PM: Penman-Monteith, CO_2: Penman-Monteith[CO_2].

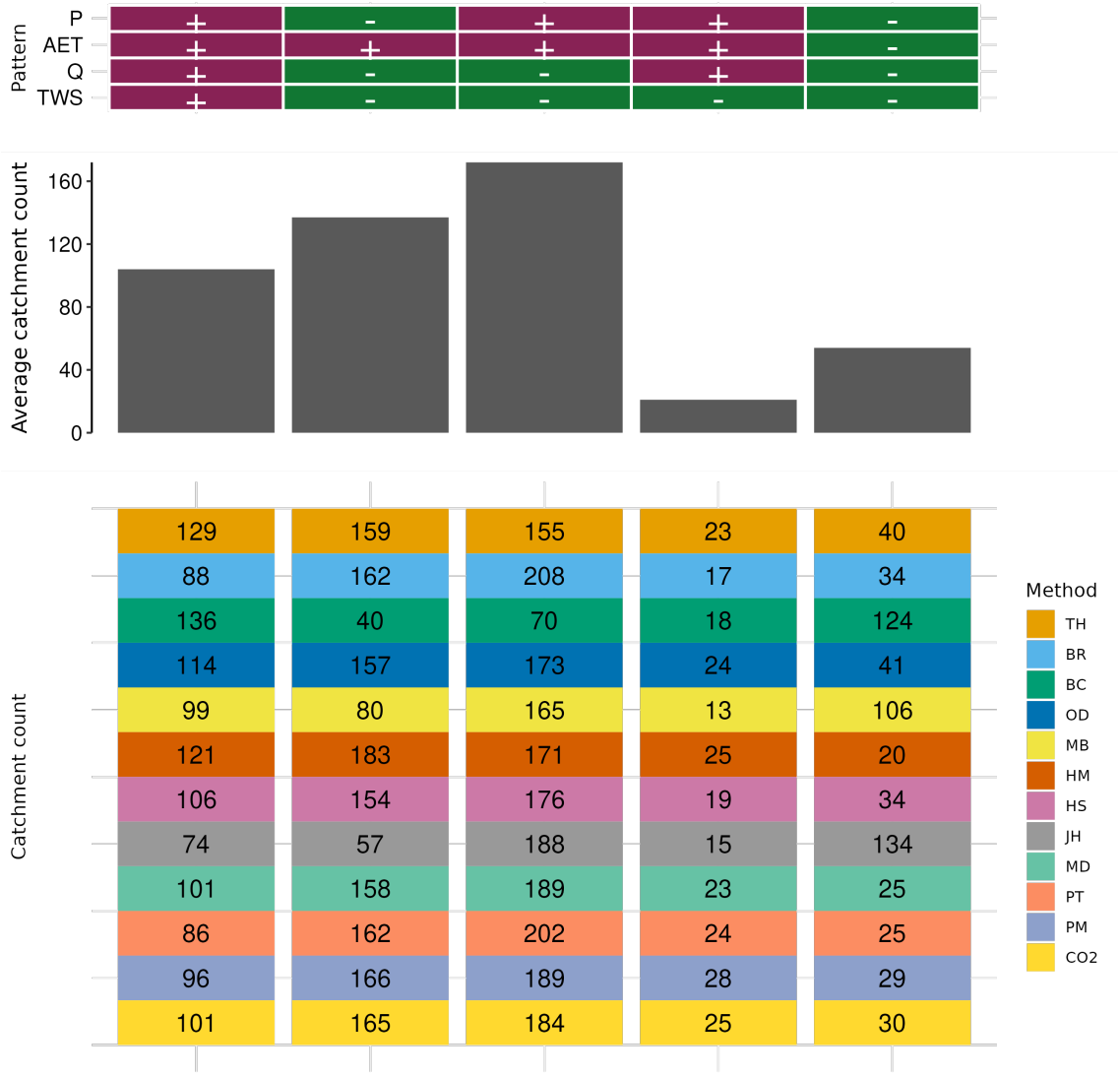

Figure S19. Pattern of different hydrological cycle components and the corresponding influence of PET methods for summer season (JJA). The first panel represents different patterns of hydrological cycle components. Each vertical column in this table corresponds to one pattern of hydrological cycle. For example, the first column is filled with '+' signs, indicating that all hydrological components (P, AET, Q, and TWS) exhibit positive changes. The '+' and '-' signs denote positive and negative changes in the respective components. In the second panel, each bar represents the average number of catchments for each hydrological cycle pattern. The third panel shows the number of catchments associated with each PET method for the corresponding hydrological cycle patterns. The color of each cell represents a specific PET method, and each column aligns with the hydrological cycle pattern represented in the corresponding column of Panel One. Where P is precipitation, AET is actual evapotranspiration, Q is runoff and TWS is total water storage. Abbreviations used for different PET methods are TH: Thornthwaite, BR: Baier-Robertson, BC: Blaney-Criddle, OD: Oudin, MB: McGuinness-Bordne, HM: Hamon, HS: Hargreaves-Samani, JH: Jensen-Haise, MD: Milly-Dunne, PT: Priestley-Taylor, PM: Penman-Monteith, CO_2: Penman-Monteith[CO_2].

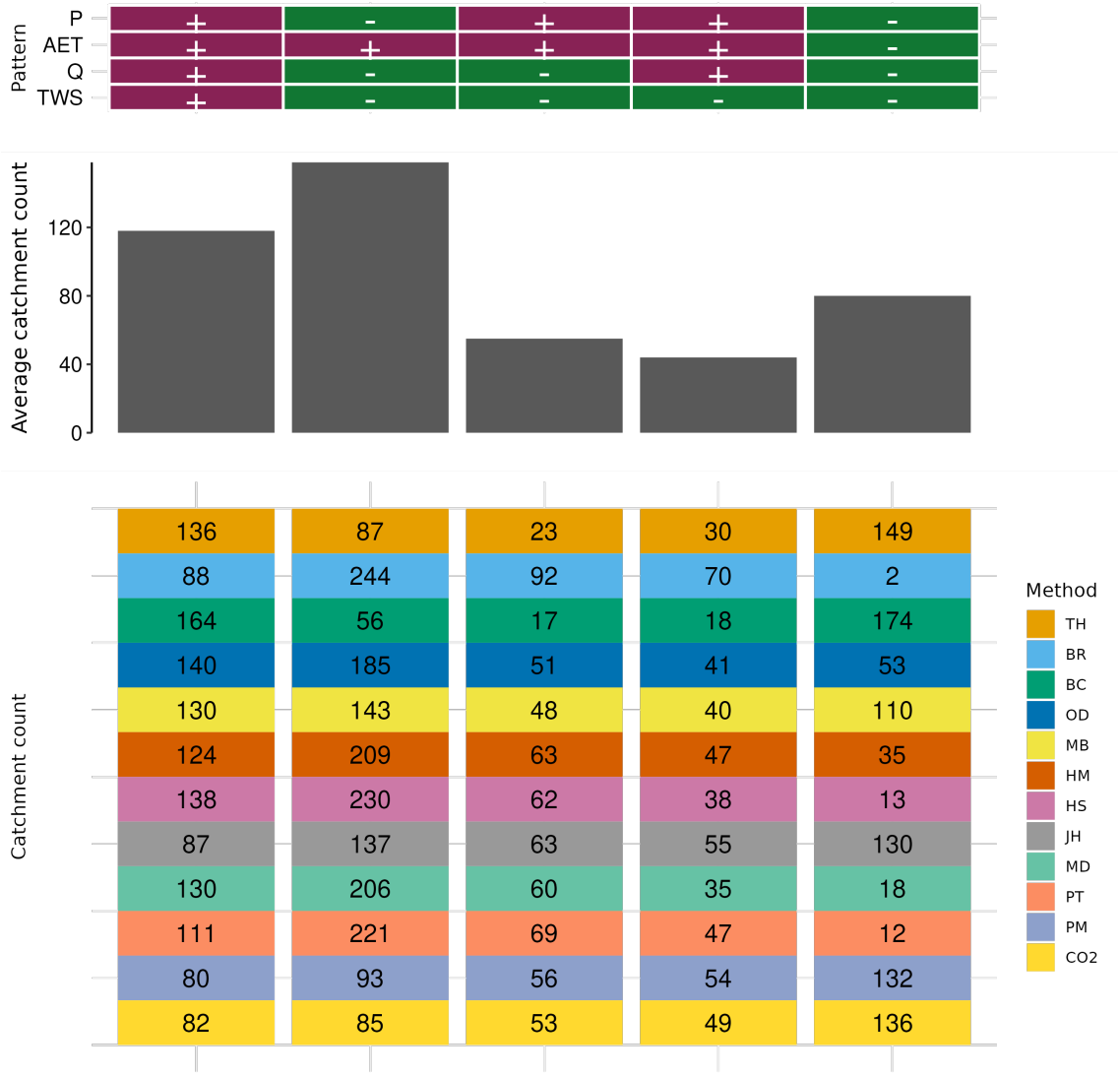

Figure S20. Pattern of different hydrological cycle components and the corresponding influence of PET methods for the autumn season (SON). The first panel represents different patterns of hydrological cycle components. Each vertical column in this table corresponds to one pattern of hydrological cycle. For example, the first column is filled with '+' signs, indicating that all hydrological components (P, AET, Q, and TWS) exhibit positive changes. The '+' and '-' signs denote positive and negative changes in the respective components. In the second panel, each bar represents the average number of catchments for each hydrological cycle pattern. The third panel shows the number of catchments associated with each PET method for the corresponding hydrological cycle patterns. The color of each cell represents a specific PET method, and each column aligns with the hydrological cycle pattern represented in the corresponding column of Panel One. Where P is precipitation, AET is actual evapotranspiration, Q is runoff and TWS is total water storage. Abbreviations used for different PET methods are TH: Thornthwaite, BR: Baier-Robertson, BC: Blaney-Criddle, OD: Oudin, MB: McGuinness-Bordne, HM: Hamon, HS: Hargreaves-Samani, JH: Jensen-Haise, MD: Milly-Dunne, PT: Priestley-Taylor, PM: Penman-Monteith, CO_2: Penman-Monteith[CO_2].

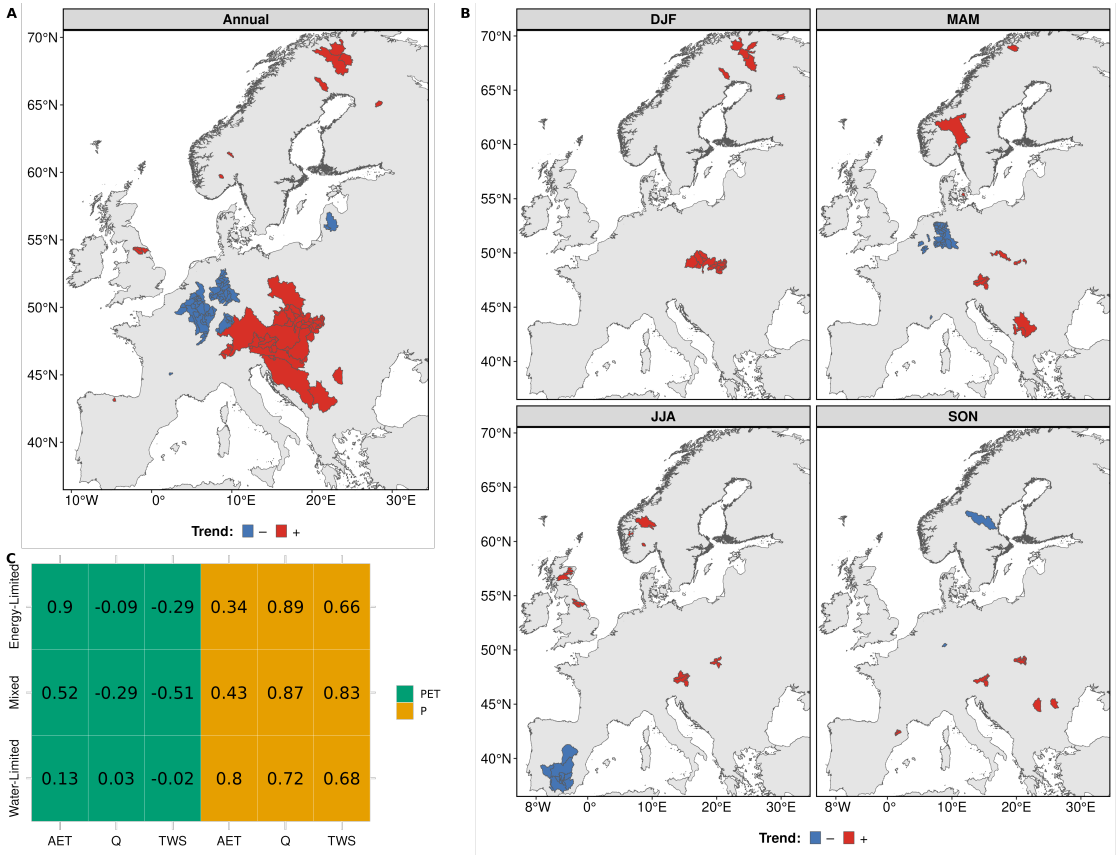

Figure S21. Spatial variation of statistically significant precipitation (P) trends and their relationship with other hydrological components across catchment categories. Panel A represents the spatial distribution of increasing and decreasing annual P trends. Panel B illustrates the seasonal variation in increasing and decreasing P trends. Panel C represents the median correlation between P and AET, Q, and TWS, as well as between PET and AET, Q, and TWS, for each catchment category across all PET methods.

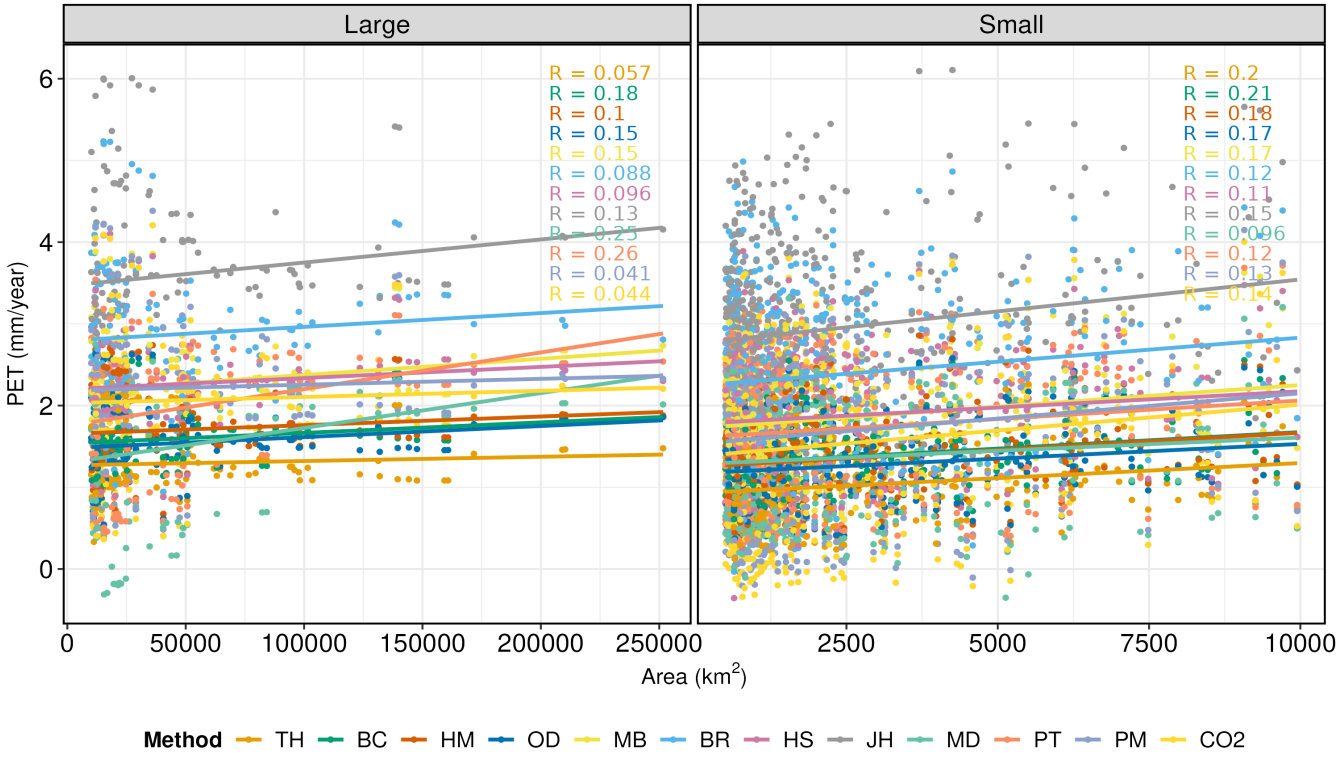

Figure S22. Correlation between PET and small (area < 10000 km^2) and large basins (area $\geq$ 10000 km^2). Abbreviations used for different PET methods are TH: Thornthwaite, BR: Baier-Robertson, BC: Blaney-Criddle, OD: Oudin, MB: McGuinness-Bordne, HM: Hamon, HS: Hargreaves-Samani, JH: Jensen-Haise, MD: Milly-Dunne, PT: Priestley-Taylor, PM: Penman-Monteith, CO_2: Penman-Monteith[CO_2].

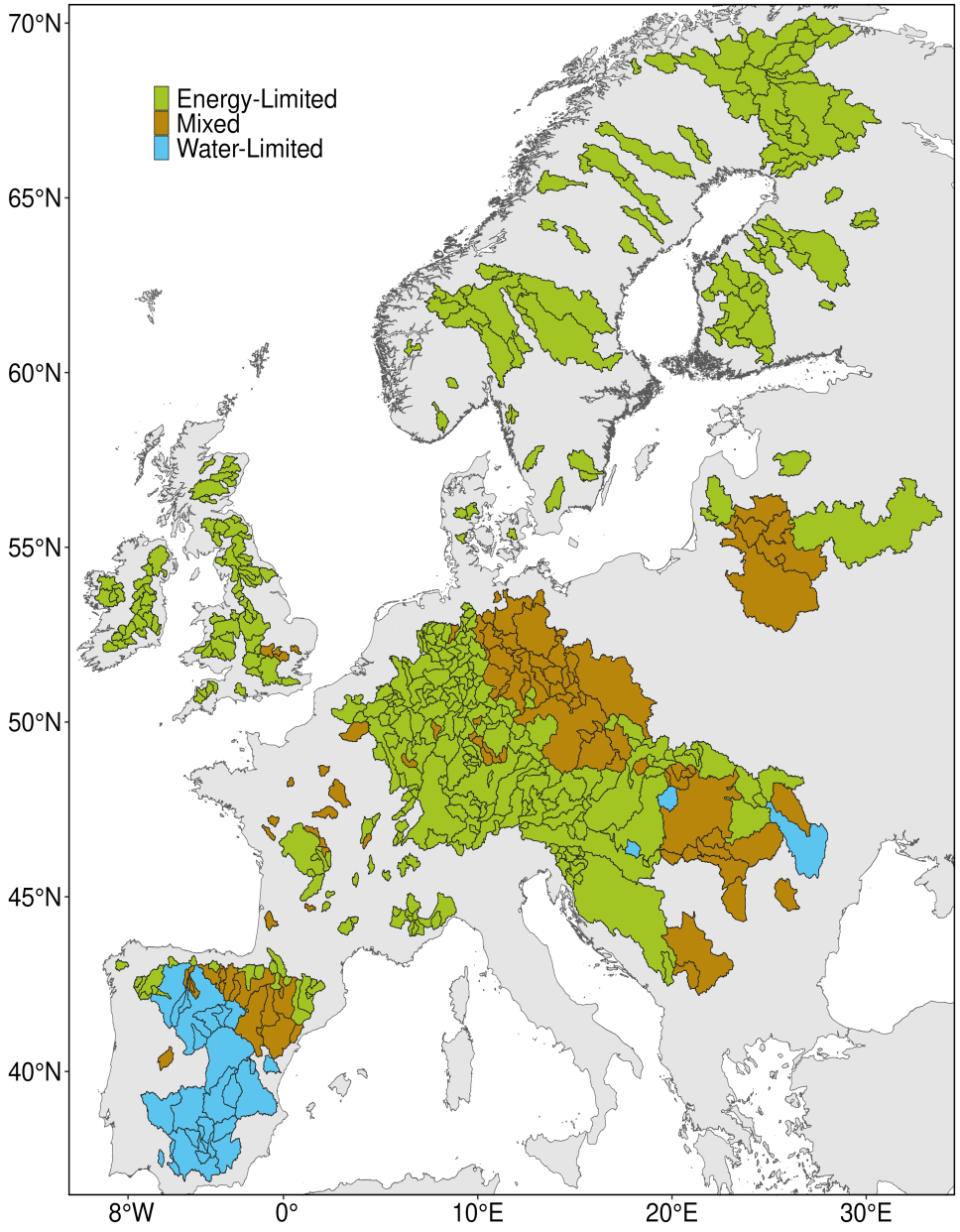

Figure S23. Catchment classification by excluding three PET methods namely Jensen-Haise, McGuinness-Bordne and Blaney-Criddle

Table S1. Table represents the number of catchments in each category under different scenarios. For instance, the "All" column represents the inclusion of all PET methods for catchment classification, the "MB" column represents all methods except MB, and the "MB-BC" column includes all PET methods except for MB and BC. Abbreviations used for different PET methods are TH: Thornthwaite, BC: Blaney-Criddle, MB: McGuinness-Bordne, HM: Hamon, JH: Jensen-Haise.

Category	All	JH	BC	TH	MB–BC	BC–JH	MB–JH	MB–JH–BC
Energy-Limited	189	197	235	189	235	306	197	417
Mixed	330	322	284	325	284	213	322	102
Water-Limited	34	34	34	39	34	34	34	34

Table S2. Average number of basins with statistically significant trends, for each hydrological component, basin type (Energy-Limited, Mixed, Water-Limited), and PET method type (temperature-based, radiation-based, combinational methods).

Components	Basin Categories	Temperature	Radiation	Combinational
PET	Energy-Limited	176	140	126
	Mixed	329	283	288
	Water-Limited	34	19	34
AET	Energy-Limited	165	179	149
	Mixed	191	258	273
	Water-Limited	2	3	3
Q	Energy-Limited	28	28	28
	Mixed	158	149	148
	Water-Limited	1	1	1
TWS	Energy-Limited	36	35	37
	Mixed	179	162	160
	Water-Limited	1	1	0

References

Ajami, H., Sharma, A., Band, L. E., Evans, J. P., Tuteja, N. K., Amirthanathan, G. E., and Bari, M. A.: On the non-stationarity of hydrological response in anthropogenically unaffected catchments: an Australian perspective, Hydrology and Earth System Sciences, 21, 281–294, https://doi.org/10.5194/hess-21-281-2017, 2017.

Anabalón, A. and Sharma, A.: On the divergence of potential and actual evapotranspiration trends: An assessment across alternate global datasets, Earth's Future, 5, 905–917, https://doi.org/10.1002/2016EF000499, 2017.

Aouissi, J., Benabdallah, S., Lili Chabaâne, Z., and Cudennec, C.: Evaluation of potential evapotranspiration assessment methods for hydrological modelling with SWAT—Application in data-scarce rural Tunisia, Agricultural Water Management, 174, 39–51, https://doi.org/10.1016/j.agwat.2016.03.004, 2016.

Bai, P., Liu, X., Yang, T., Li, F., Liang, K., Hu, S., and Liu, C.: Assessment of the Influences of Different Potential Evapotranspiration Inputs on the Performance of Monthly Hydrological Models under Different Climatic Conditions, Journal of Hydrometeorology, 17, 2259–2274, https://doi.org/10.1175/JHM-D-15-0202.1, 2016.

Beck, H. E., Van Dijk, A. I. J. M., De Roo, A., Dutra, E., Fink, G., Orth, R., and Schellekens, J.: Global evaluation of runoff from 10 state-of-the-art hydrological models, Hydrology and Earth System Sciences, 21, 2881–2903, https://doi.org/10.5194/hess-21-2881-2017, 2017.

Berghuijs, W. R., Larsen, J. R., van Emmerik, T. H. M., and Woods, R. A.: A Global Assessment of Runoff Sensitivity to Changes in Precipitation, Potential Evaporation, and Other Factors, Water Resources Research, 53, 8475–8486, https://doi.org/10.1002/2017WR021593, 2017.

Birhanu, D., Kim, H., Jang, C., and Park, S.: Does the Complexity of Evapotranspiration and Hydrological Models Enhance Robustness?, Sustainability, 10, 2837, https://doi.org/10.3390/su10082837, 2018.

Boeing, F., Wagener, T., Marx, A., Rakovec, O., Kumar, R., Samaniego, L., and Attinger, S.: Increasing influence of evapotranspiration on prolonged water storage recovery in Germany, Environmental Research Letters, 19, 024 047, https://doi.org/10.1088/1748-9326/ad24ce, 2024.

Bruno, G. and Duethmann, D.: Increases in Water Balance-Derived Catchment Evapotranspiration in Germany During 1970s–2000s Turning Into Decreases Over the Last Two Decades, Despite Uncertainties, Geophysical Research Letters, 51, e2023GL107 753, https://doi.org/10.1029/2023GL107753, 2024.

Caloiero, T., Caloiero, P., and Frustaci, F.: Long-term precipitation trend analysis in Europe and in the Mediterranean basin, Water and Environment Journal, 32, 433–445, https://doi.org/10.1111/wej.12346, 2018.

Clerc-Schwarzenbach, F., Selleri, G., Neri, M., Toth, E., van Meerveld, I., and Seibert, J.: Large-sample hydrology – a few camels or a whole caravan?, Hydrology and Earth System Sciences, 28, 4219–4237, https://doi.org/10.5194/hess-28-4219-2024, 2024.

Devia, G. K., Ganasri, B., and Dwarakish, G.: A Review on Hydrological Models, Aquatic Procedia, 4, 1001–1007, https://doi.org/10.1016/j.aqpro.2015.02.126, 2015.

Guo, D., Westra, S., and Maier, H. R.: Sensitivity of potential evapotranspiration to changes in climate variables for different Australian climatic zones, Hydrology and Earth System Sciences, 21, 2107–2126, https://doi.org/10.5194/hess-21-2107-2017, 2017.

Hanselmann, N., Osuch, M., Wawrzyniak, T., and Alphonse, A. B.: Evaluating potential evapotranspiration methods in a rapidly warming Arctic region, SW Spitsbergen (1983–2023), Journal of Hydrology: Regional Studies, 56, 101 979, https://doi.org/10.1016/j.ejrh.2024.101979, 2024.

Hua, D., Hao, X., Zhang, Y., and Qin, J.: Uncertainty assessment of potential evapotranspiration in arid areas, as estimated by the Penman-Monteith method, Journal of Arid Land, 12, 166–180, https://doi.org/10.1007/s40333-020-0093-7, 2020.

Jung, I. W., Chang, H., and Risley, J.: Effects of runoff sensitivity and catchment characteristics on regional actual evapotranspiration trends in the conterminous US, Environmental Research Letters, 8, 044 002, https://doi.org/10.1088/1748-9326/8/4/044002, 2013.

Kingston, D. G., Todd, M. C., Taylor, R. G., Thompson, J. R., and Arnell, N. W.: Uncertainty in the estimation of potential evapotranspiration under climate change, Geophysical Research Letters, 36, 2009GL040 267, https://doi.org/10.1029/2009GL040267, 2009.

Kuentz, A., Arheimer, B., Hundecha, Y., and Wagener, T.: Understanding hydrologic variability across Europe through catchment classification, Hydrology and Earth System Sciences, 21, 2863–2879, https://doi.org/10.5194/hess-21-2863-2017, 2017.

Kumar, R., Livneh, B., and Samaniego, L.: Toward computationally efficient large-scale hydrologic predictions with a multiscale regionalization scheme: LARGE-SCALE HYDROLOGIC PREDICTIONS, Water Resources Research, 49, 5700–5714, https://doi.org/10.1002/wrcr.20431, 2013.

Liu, Y., Jiang, Q., Wang, Q., Jin, Y., Yue, Q., Yu, J., Zheng, Y., Jiang, W., and Yao, X.: The divergence between potential and actual evapotranspiration: An insight from climate, water, and vegetation change, Science of The Total Environment, 807, 150 648, https://doi.org/10.1016/j.scitotenv.2021.150648, 2022.

Lu, J., Sun, G., McNulty, S. G., and Amatya, D. M.: A COMPARISON OF SIX POTENTIAL EVAPOTRANSPIRATION METHODS FOR REGIONAL USE IN THE SOUTHEASTERN UNITED STATES, Journal of the American Water Resources Association, 41, 621–633, https://doi.org/10.1111/j.1752-1688.2005.tb03759.x, 2005.

Markonis, Y.: On the Definition of Extreme Evaporation Events, Geophysical Research Letters, 52, e2024GL113 038, https://doi.org/10.1029/2024GL113038, 2025.

Markonis, Y., Papalexiou, S. M., Martinkova, M., and Hanel, M.: Assessment of Water Cycle Intensification Over Land using a Multisource Global Gridded Precipitation DataSet, Journal of Geophysical Research: Atmospheres, 124, 11 175–11 187, https://doi.org/10.1029/2019JD030855, 2019.

Markonis, Y., Kumar, R., Hanel, M., Rakovec, O., Máca, P., and AghaKouchak, A.: The rise of compound warm-season droughts in Europe, Science Advances, 7, eabb9668, https://doi.org/10.1126/sciadv.abb9668, 2021.

Markonis, Y., Vargas Godoy, M. R., Pradhan, R. K., Pratap, S., Thomson, J. R., Hanel, M., Paschalis, A., Nikolopoulos, E., and Papalexiou, S. M.: Spatial partitioning of terrestrial precipitation reveals varying dataset agreement across different environments, Communications Earth & Environment, 5, 217, 2024.

Massari, C., Avanzi, F., Bruno, G., Gabellani, S., Penna, D., and Camici, S.: Evaporation enhancement drives the European water-budget deficit during multi-year droughts, Hydrology and Earth System Sciences, 26, 1527–1543, https://doi.org/10.5194/hess-26-1527-2022, 2022.

Mazzoleni, M., Brandimarte, L., and Amaranto, A.: Evaluating precipitation datasets for large-scale distributed hydrological modelling, Journal of Hydrology, 578, 124 076, https://doi.org/10.1016/j.jhydrol.2019.124076, 2019.

Maček, U., Bezak, N., and Šraj, M.: Reference evapotranspiration changes in Slovenia, Europe, Agricultural and Forest Meteorology, 260-261, 183–192, https://doi.org/10.1016/j.agrformet.2018.06.014, 2018.

Ndiaye, P. M., Bodian, A., Dezetter, A., Ogilvie, A., and Goudiaby, O.: Sensitivity of global hydrological models to potential evapotranspiration estimation methods in the Senegal River Basin (West Africa), Journal of Hydrology: Regional Studies, 53, 101 823, https://doi.org/10.1016/j.ejrh.2024.101823, 2024.

Oudin, L., Michel, C., and Anctil, F.: Which potential evapotranspiration input for a lumped rainfall-runoff model?, Journal of Hydrology, 303, 275–289, https://doi.org/10.1016/j.jhydrol.2004.08.025, 2005.

Pfeifroth, U., Sanchez-Lorenzo, A., Manara, V., Trentmann, J., and Hollmann, R.: Trends and Variability of Surface Solar Radiation in Europe Based On Surface- and Satellite-Based Data Records, Journal of Geophysical Research: Atmospheres, 123, 1735–1754, https://doi.org/10.1002/2017JD027418, 2018.

Pimentel, R., Arheimer, B., Crochemore, L., Andersson, J. C. M., Pechlivanidis, I. G., and Gustafsson, D.: Which Potential Evapotranspiration Formula to Use in Hydrological Modeling World-Wide?, Water Resources Research, 59, e2022WR033 447, https://doi.org/10.1029/2022WR033447, 2023.

Rakovec, O., Samaniego, L., Hari, V., Markonis, Y., Moravec, V., Thober, S., Hanel, M., and Kumar, R.: The 2018–2020 Multi-Year Drought Sets a New Benchmark in Europe, Earth's Future, 10, e2021EF002 394, https://doi.org/10.1029/2021EF002394, 2022.

Reaver, N. G. F., Kaplan, D. A., Klammler, H., and Jawitz, J. W.: Theoretical and empirical evidence against the Budyko catchment trajectory conjecture, Hydrology and Earth System Sciences, 26, 1507–1525, https://doi.org/10.5194/hess-26-1507-2022, 2022.

Seiller, G. and Anctil, F.: How do potential evapotranspiration formulas influence hydrological projections?, Hydrological Sciences Journal, 61, 2249–2266, https://doi.org/10.1080/02626667.2015.1100302, 2016.

Shah, J., Kumar, R., Samaniego, L., Markonis, Y., Hanel, M., Attinger, S., Hari, V., and Rakovec, O.: On the role of antecedent meteorological conditions on flash drought initialization in Europe, Environmental Research Letters, 18, 064 039, https://doi.org/10.1088/1748-9326/acd8d3, 2023.

Shi, L., Wang, B., Liu, D. L., Feng, P., Cleverly, J., Li, L., Zhang, G., and Yu, Q.: Performance of potential evapotranspiration models across different climatic stations in New South Wales, Australia, Journal of Hydrology: Regional Studies, 50, 101 573, https://doi.org/10.1016/j.ejrh.2023.101573, 2023.

Tang, G., Clark, M. P., Knoben, W. J. M., Liu, H., Gharari, S., Arnal, L., Beck, H. E., Wood, A. W., Newman, A. J., and Papalexiou, S. M.: The Impact of Meteorological Forcing Uncertainty on Hydrological Modeling: A Global Analysis of Cryosphere Basins, Water Resources Research, 59, e2022WR033 767, https://doi.org/10.1029/2022WR033767, 2023.

Tarek, M., Brissette, F. P., and Arsenault, R.: Evaluation of the ERA5 reanalysis as a potential reference dataset for hydrological modelling over North America, Hydrology and Earth System Sciences, 24, 2527–2544, https://doi.org/10.5194/hess-24-2527-2020, 2020.

Teuling, A. J., de Badts, E. A. G., Jansen, F. A., Fuchs, R., Buitink, J., Hoek van Dijke, A. J., and Sterling, S. M.: Climate change, reforestation/afforestation, and urbanization impacts on evapotranspiration and streamflow in Europe, Hydrology and Earth System Sciences, 23, 3631–3652, https://doi.org/10.5194/hess-23-3631-2019, 2019.

Thornthwaite, C. W.: An Approach toward a Rational Classification of Climate, Geographical Review, 38, 55, https://doi.org/10.2307/210739, 1948.

Voisin, N., Wood, A. W., and Lettenmaier, D. P.: Evaluation of Precipitation Products for Global Hydrological Prediction, Journal of Hydrometeorology, 9, 388–407, https://doi.org/10.1175/2007JHM938.1, 2008.

Wang, J., Zhuo, L., Han, D., Liu, Y., and Rico-Ramirez, M. A.: Hydrological Model Adaptability to Rainfall Inputs of Varied Quality, Water Resources Research, 59, e2022WR032 484, https://doi.org/10.1029/2022WR032484, 2023.

Xiang, K., Li, Y., Horton, R., and Feng, H.: Similarity and difference of potential evapotranspiration and reference crop evapotranspiration – a review, Agricultural Water Management, 232, 106 043, https://doi.org/10.1016/j.agwat.2020.106043, 2020.

Zhang, D., Liu, X., Zhang, Q., Liang, K., and Liu, C.: Investigation of factors affecting intra-annual variability of evapotranspiration and streamflow under different climate conditions, Journal of Hydrology, 543, 759–769, https://doi.org/10.1016/j.jhydrol.2016.10.047, 2016.

Zhang, Y., He, B., Guo, L., Liu, J., and Xie, X.: The relative contributions of precipitation, evapotranspiration, and runoff to terrestrial water storage changes across 168 river basins, Journal of Hydrology, 579, 124 194, https://doi.org/10.1016/j.jhydrol.2019.124194, 2019.

Zhao, L., Xia, J., Xu, C.-y., Wang, Z., Sobkowiak, L., and Long, C.: Evapotranspiration estimation methods in hydrological models, Journal of Geographical Sciences, 23, 359–369, https://doi.org/10.1007/s11442-013-1015-9, 2013.