Supplement of Hydrol. Earth Syst. Sci., 26, 6339–6359, 2022
https://doi.org/10.5194/hess-26-6339-2022-supplement

Hydrology and Earth System Sciences
Open Access
EGU

Supplement of

River flooding mechanisms and their changes in Europe revealed by explainable machine learning

Shijie Jiang et al.

Correspondence to: Shijie Jiang (shijie.jiang@ufz.de)

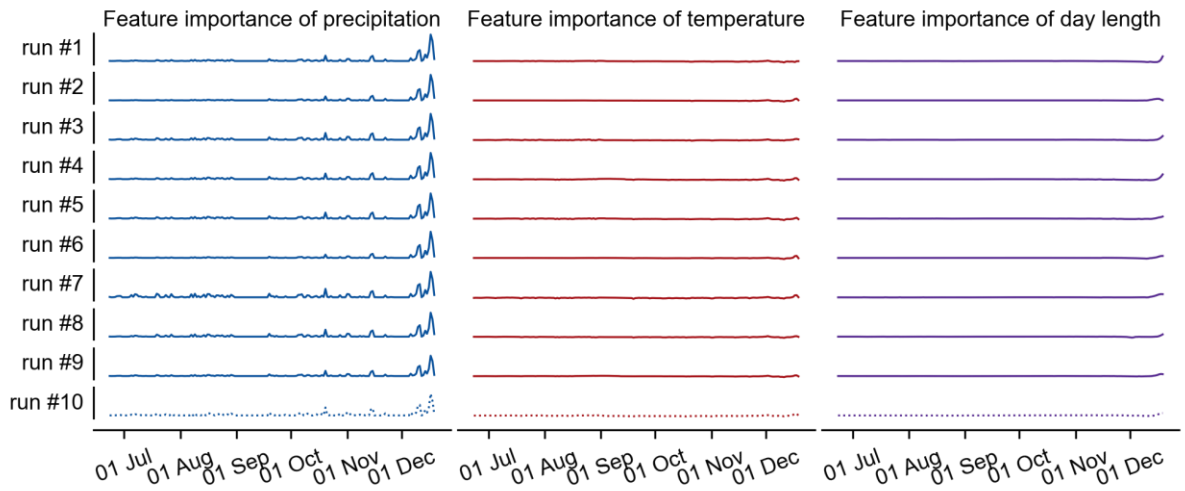

Figure S1: The integrated gradient (IG) scores of precipitation, temperature, and day length extracted from the 10 independent models for predicting the peak discharge that is illustrated in Fig. 2a in the main text. The models were trained and tested in 10-fold cross-validation, where the target discharges in run #10 was in the testing dataset (indicated by dashed lines), and the target discharges in the remaining runs were in the training dataset (solid lines). All the y-axes use the same scale for better comparisons. The average of the 10 sequences across different models generates the heatmaps shown in Fig. 2b in the main text.

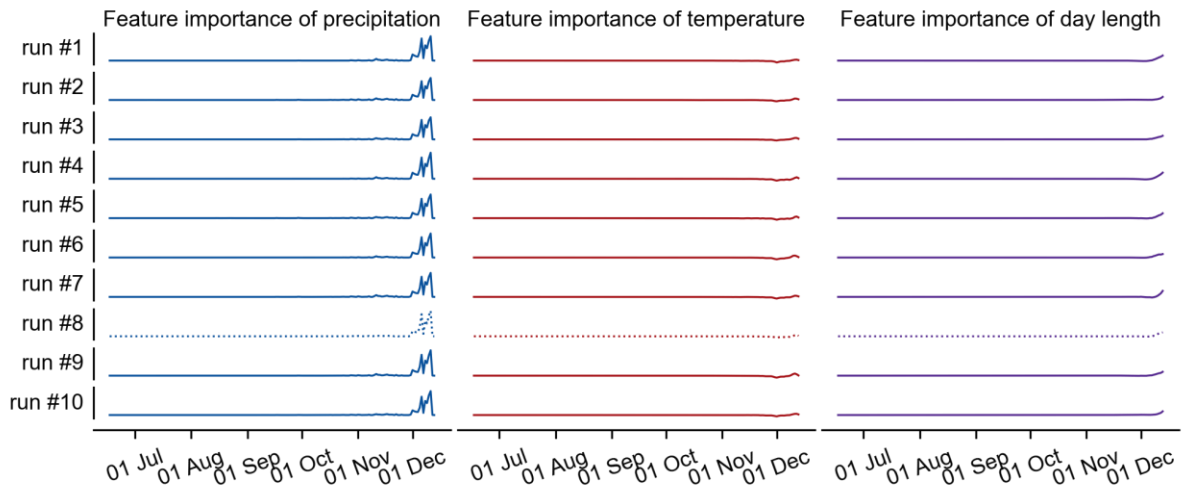

Figure S2: The same case as in Fig. S1 to generate heatmaps illustrated in Fig. 4a in the main text.

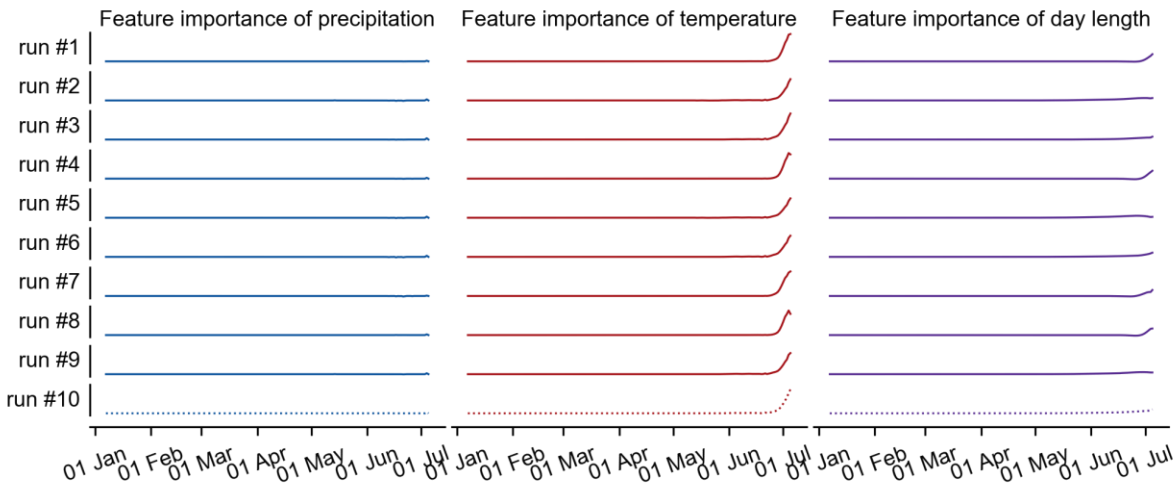

Figure S3: The same case as in Fig. S1 to generate heatmaps illustrated in Fig. 4b in the main text.

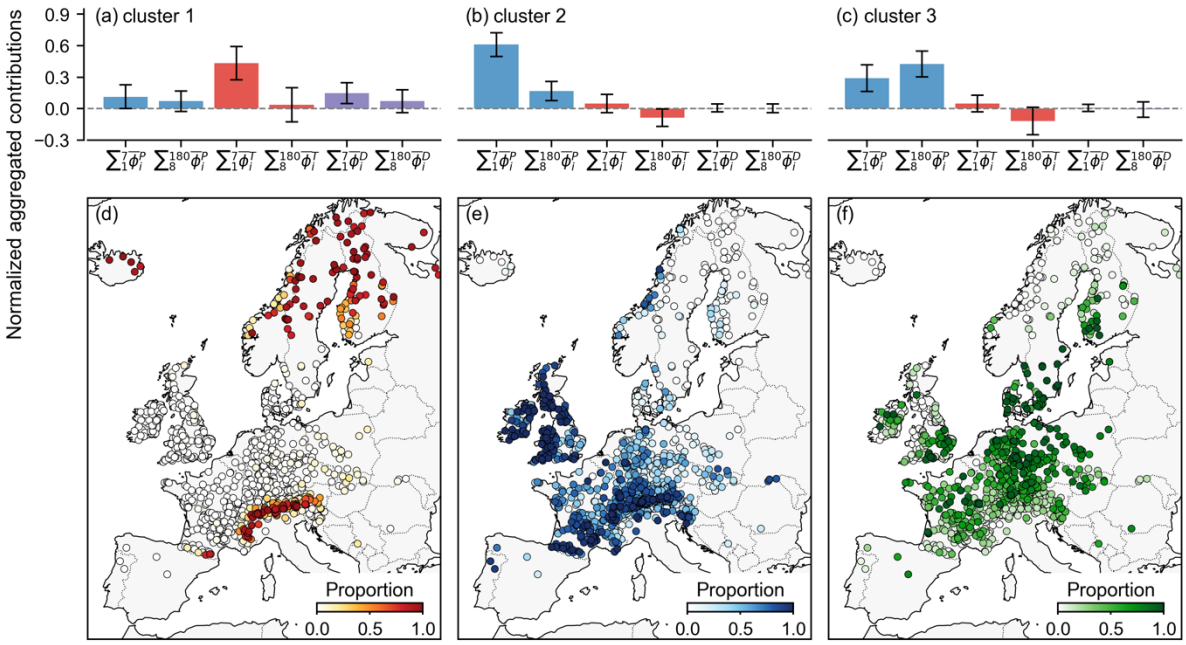

Figure S4: The same case as in Fig. 5 in the main text, but we used the IG scores based on the peaks in testing datasets alone to perform cluster analysis. The events identified with snowmelt, recent precipitation, or antecedent precipitation as the primary causes account for 15.7%, 49.0%, and 35.3% of all the 53,968 annual maximum peak discharges, which is only slightly different from using the averaged IG scores from the 10-fold models for individual peaks.

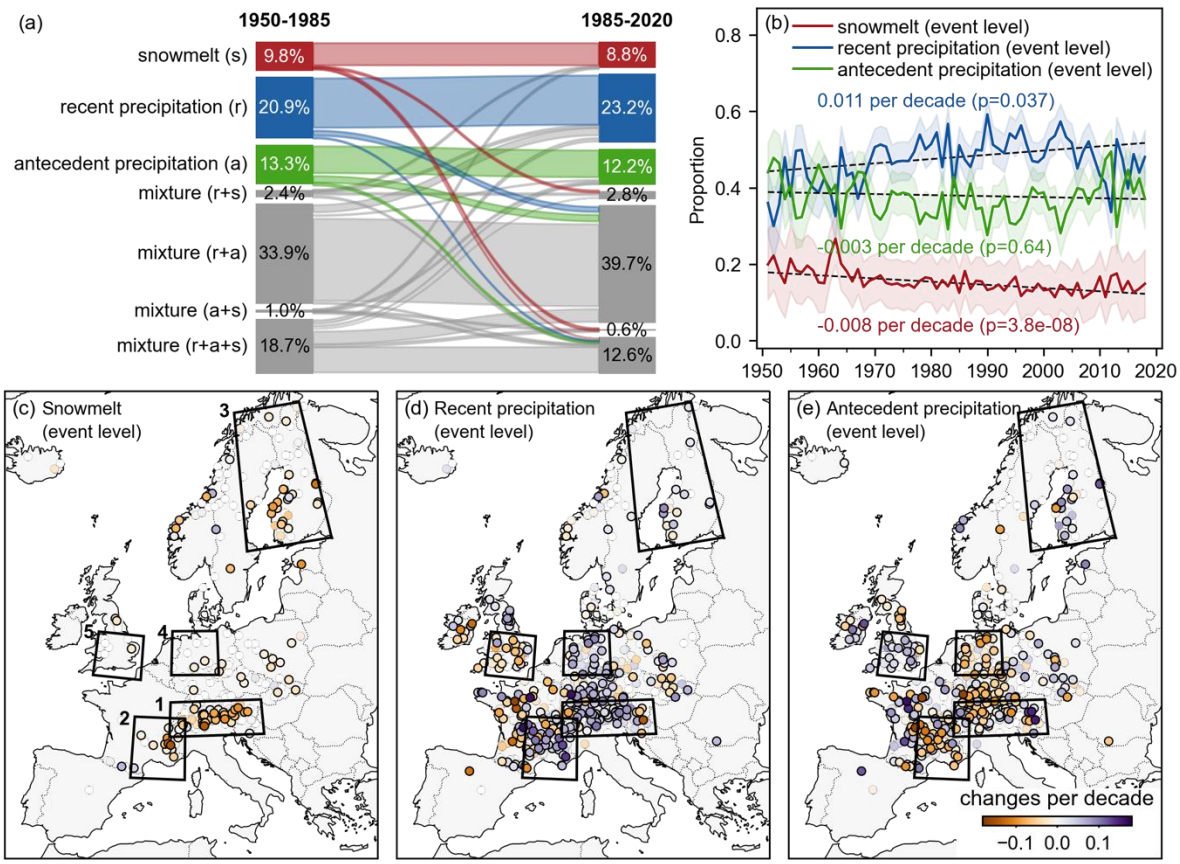

Figure S5: The same case as in Fig. 7 in the main text, but we used the IG scores based on the peaks in testing datasets alone to perform cluster analysis and trend analysis.

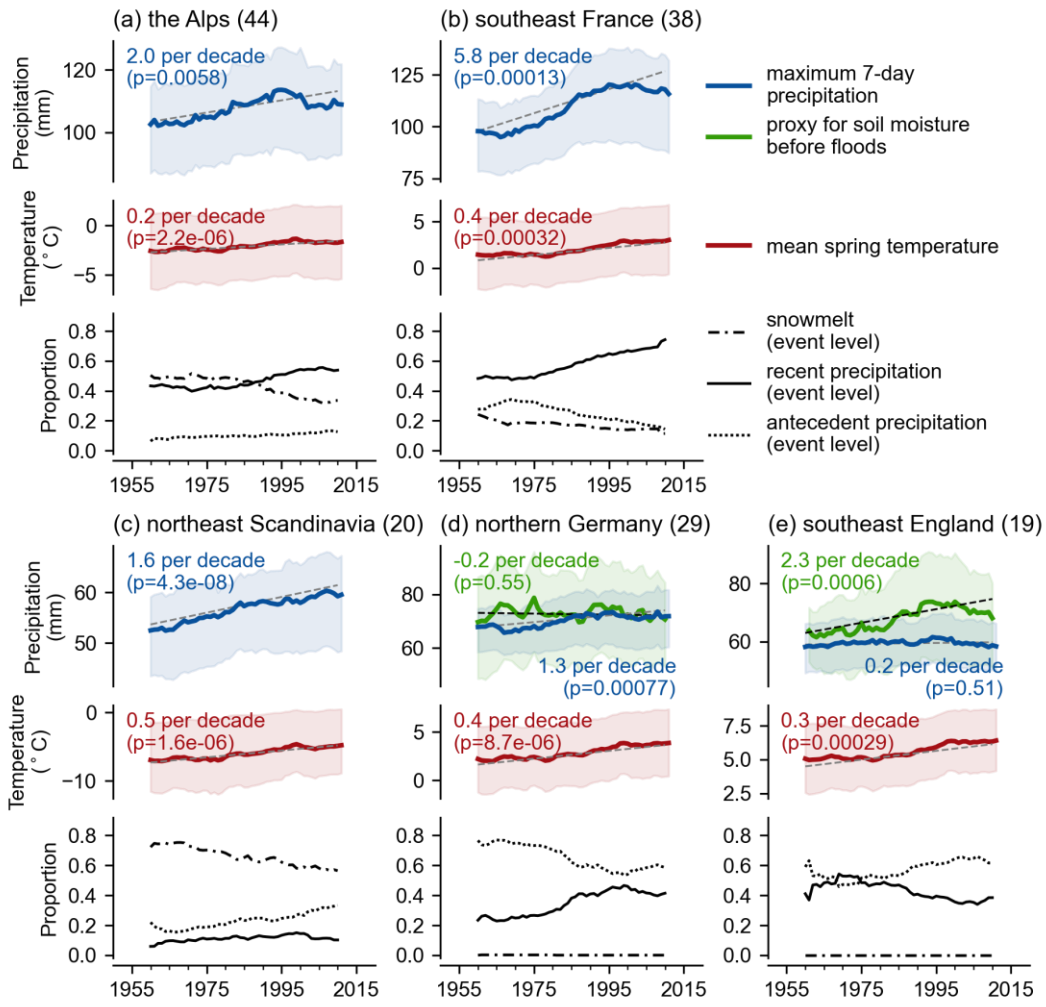

Figure S6. The temporal changes of the event-level mechanisms in relevant catchments within the five selected regions (see Fig. 7c in the main text), as well as the changes in average extreme precipitation, mean spring temperatures, and antecedent soil moisture conditions prior to flooding. The notions are the same in Fig. 8 in the main text, except for the shades that further illustrate the 25th and 75th percentiles of the yearly changes in respective meteorological drivers (smoothed by a 20-year moving average window as well).

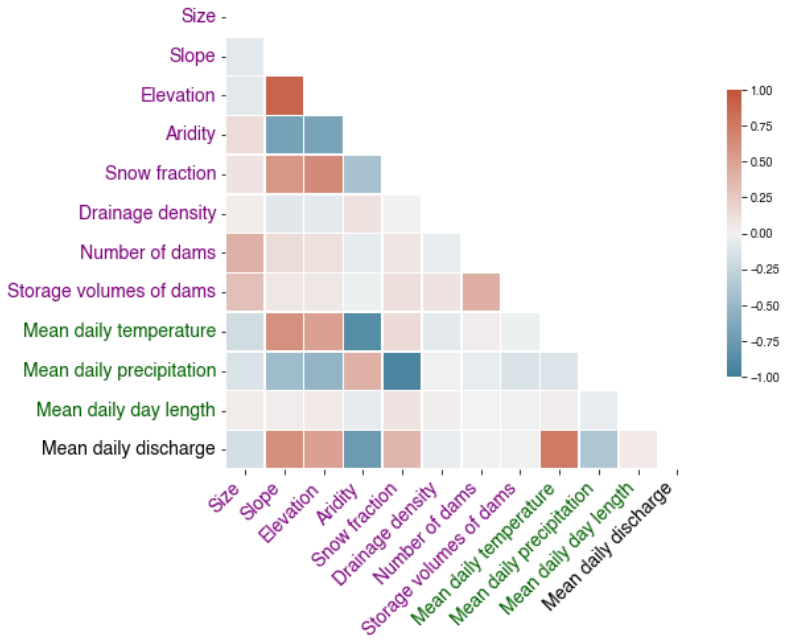

Figure S7. The Pearson correlation heatmap between some common static catchment attributes (the first eight attributes, written in purple), daily mean meteorological drivers (in green), and daily mean river discharges (in black) for the 1,077 catchments in the main text (see Fig. 1). The catchment size, slope, elevation, drainage density, number of dams, and storage volumes of dams were derived from the Global Streamflow Indices and Metadata Archive (GSIM, https://doi.org/10.1594/PANGAEA.887477), the aridity index and snowfall fraction were calculated from the catchment-averaged precipitation and temperature. The daily mean meteorological drivers include the daily mean value of catchment-averaged precipitation, temperature, and day length during 1950–2020. Daily mean river discharges were calculated by using the available discharge records during 1950–2020 and they have been represented in mm/d to exclude the effect by catchment size. Each grid represents the correlation between two variables across the 1,077 catchments, with the dark red or dark blue color denoting strong positive or negative correlations.

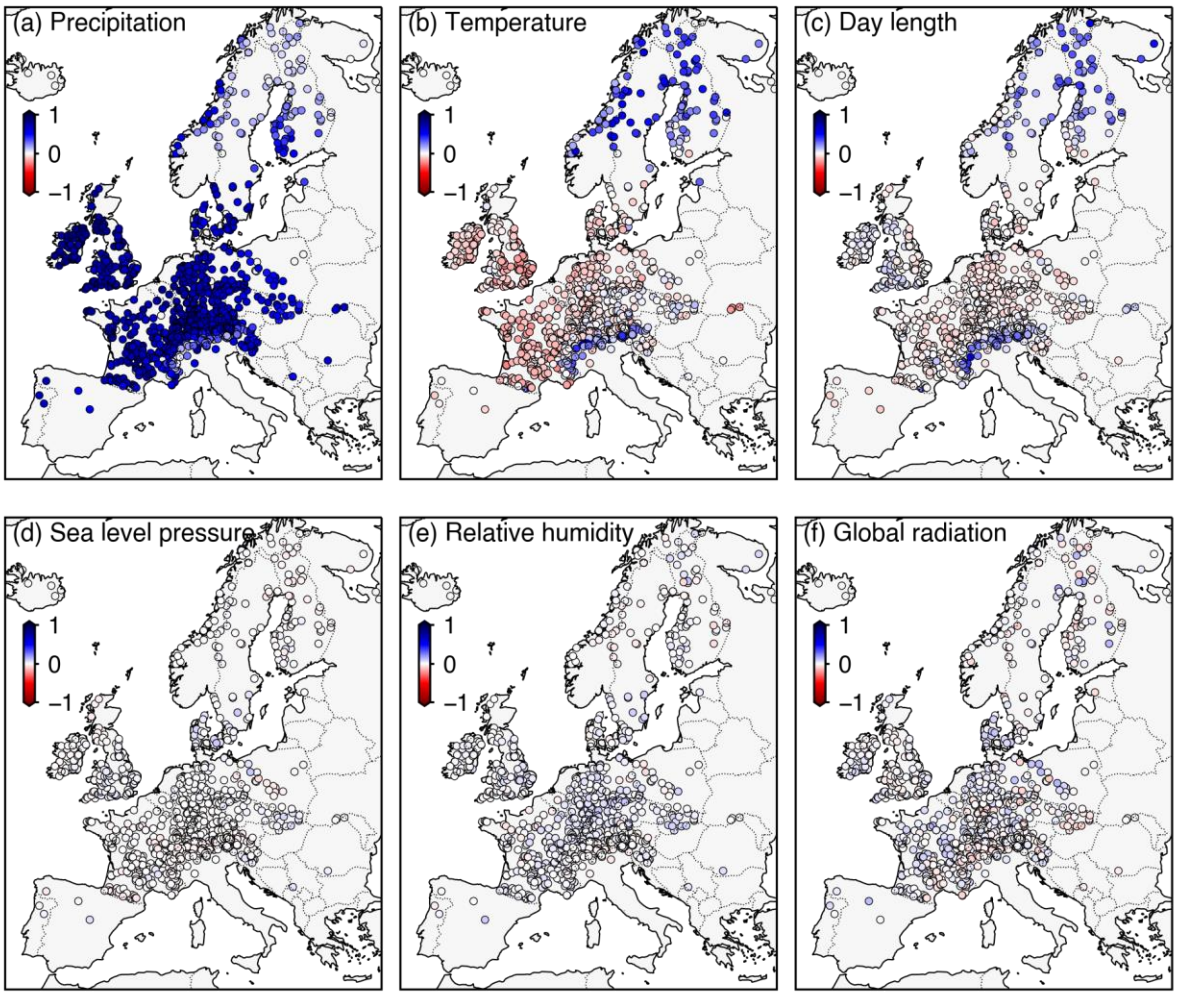

Figure S8. The average normalized aggregated contribution for respective input variables in each catchment if sea level pressure, relative humidity, and global radiation were added into the model in the main text in addition to precipitation, temperature, and day length. The sea level pressure, relative humidity, and global radiation were retrieved from E-OBS dataset (cited in the main text), and they were processed as catchment-averaged time series by the method described in the main text. Model settings and interpretations are also the same as in the main text, where only precipitation, temperature, and day length are used. Here the aggregated contribution indicates the contributions of a variable in all the 180 days to the target peak discharges (i.e., $\Sigma_1^{180}\bar{\phi}_i$, see the notions in Section 2.3 in the main text). The aggregated contribution for each variable has been normalized per peak discharge for comparability. The color reflects the average value of the normalized aggregated contribution for respective input variables in each catchment. Darker colors indicate that the specific variables have a greater impact (either positive or negative) on peak discharges. The figure implies that the sea level pressure, relative humidity, and global radiation are less important features.

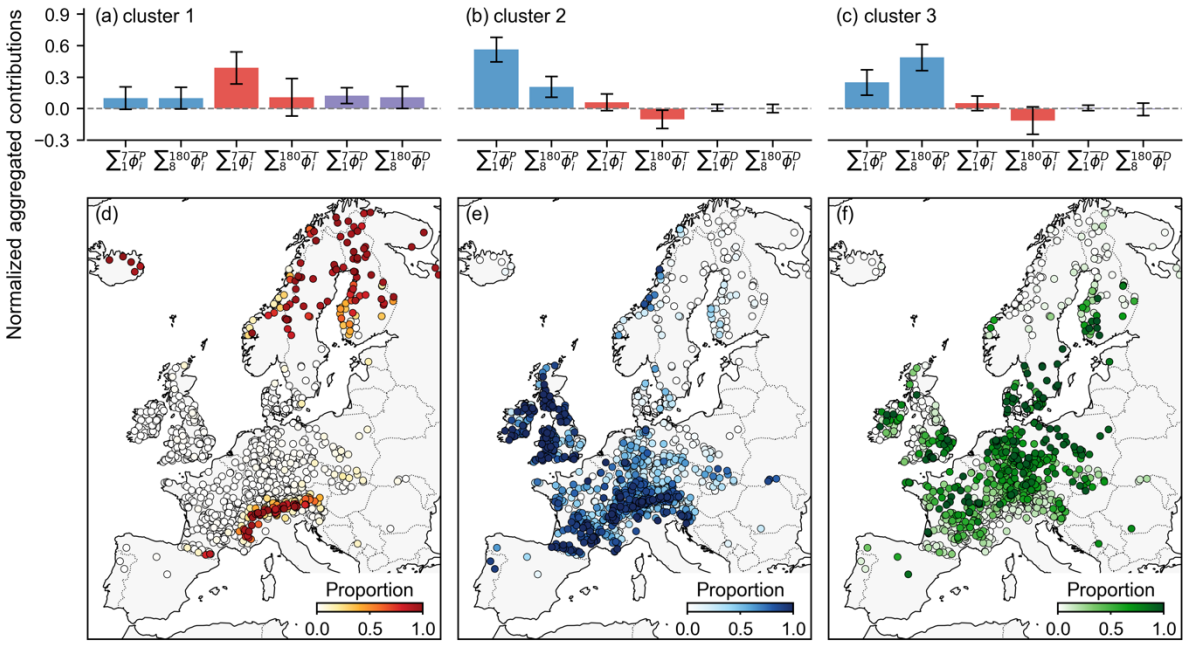

Figure S9: The same case as in Fig. 5 in the main text, but a 5-day window was used to separate contributions of a variable in recent days and an earlier antecedent period.

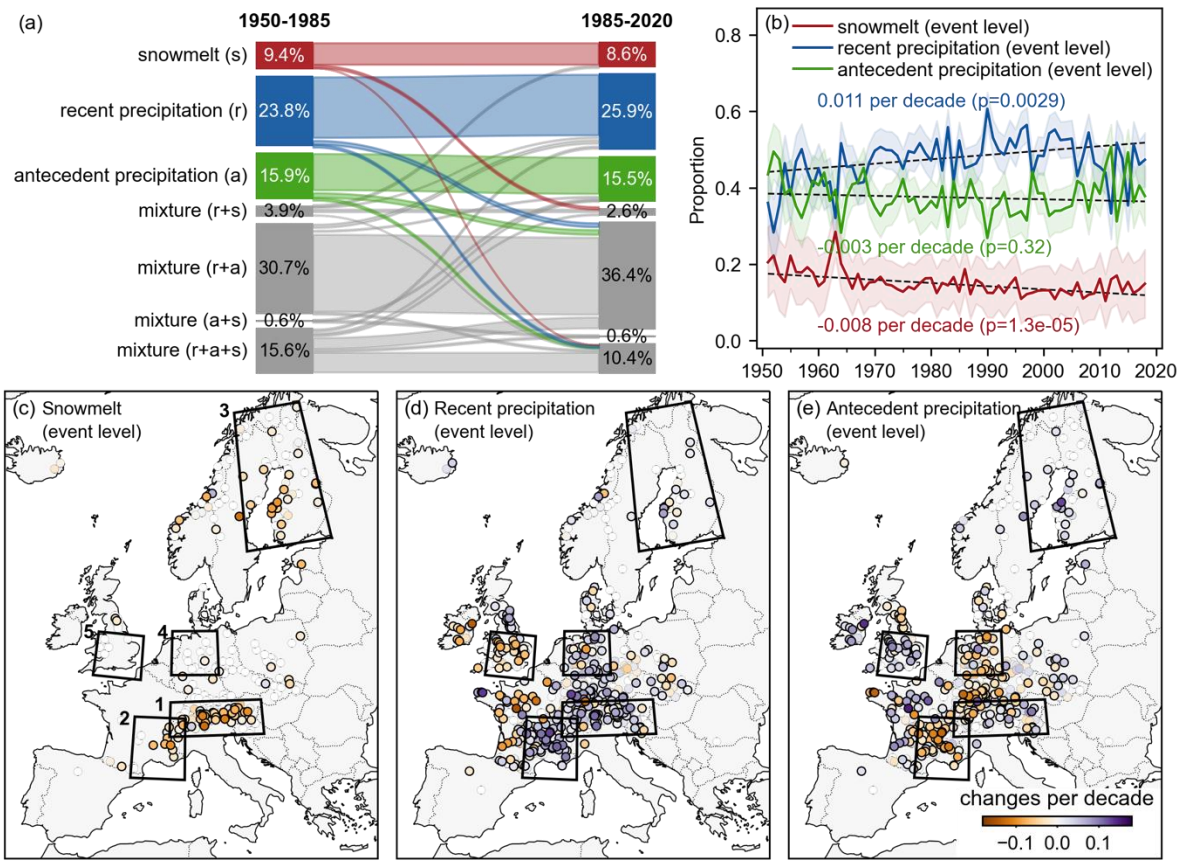

Figure S10: The same case as in Fig. 7 in the main text, but a 5-day window was used to separate contributions of a variable in recent days and an earlier antecedent period.