Supplement of Hydrol. Earth Syst. Sci., 29, 4153–4178, 2025
https://doi.org/10.5194/hess-29-4153-2025-supplement


Hydrology and Earth System Sciences
Open Access

*Supplement of*

# Comparison of high-resolution climate reanalysis datasets for hydro-climatic impact studies

**Raul R. Wood et al.**

*Correspondence to:* Raul R. Wood (raul.wood@slf.ch)

# S1. Supplementary figures for precipitation and temperature metrics

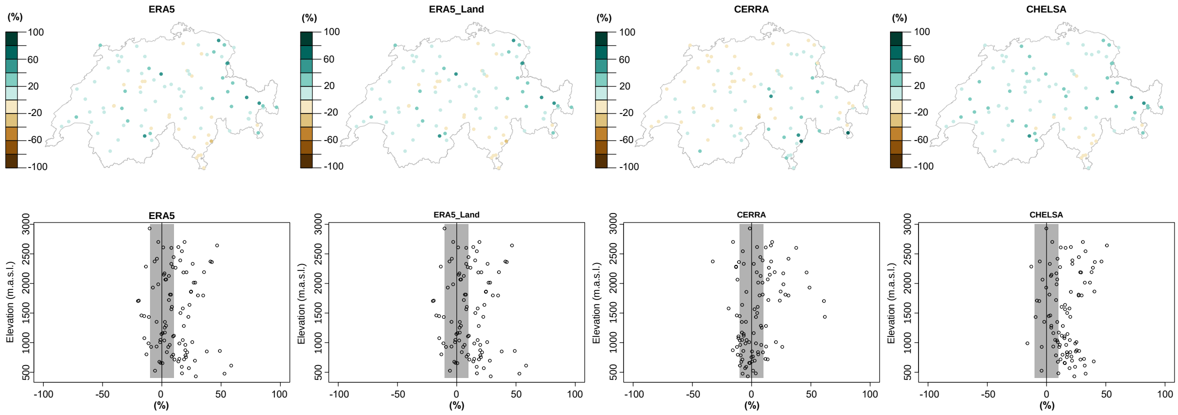


Figure S1: Comparison of the mean daily precipitation at the catchment level of the four reanalysis datasets (ERA5, ERA5-Land, CERRA, and CHELSA) with the gridded observations and elevation dependence of the biases. The first row shows relative biases (in %) at the catchment level for ERA5, ERA5-Land, CERRA, and CHELSA (left to right) compared to observations. The second row shows relative biases (x-axis) sorted by catchment elevation (y-axis); the grey box indicates relative biases within ± 10%.

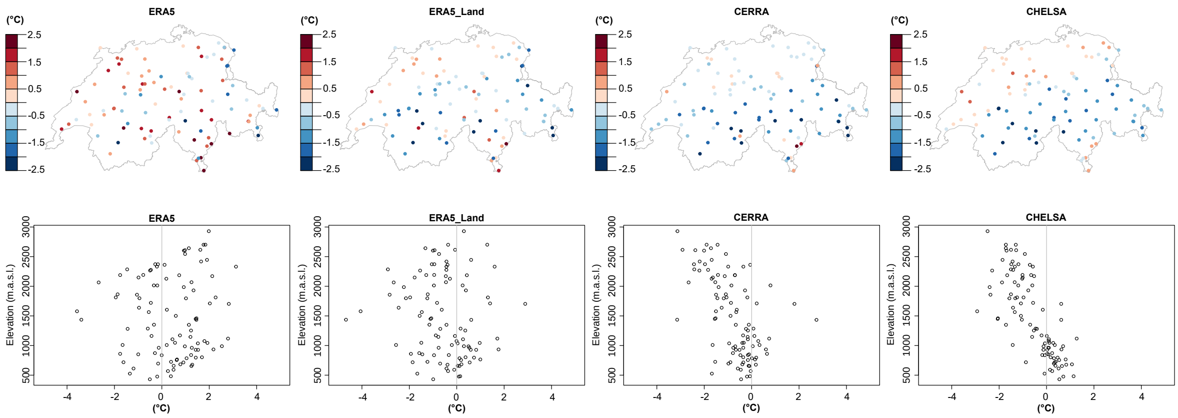


Figure S2: Comparison of the mean daily temperature at the catchment level of the four reanalysis datasets (ERA5, ERA5-Land, CERRA, and CHELSA) with the gridded observations and elevation dependence of the biases. The first row shows absolute biases (in °C) at the catchment level for ERA5, ERA5-Land, CERRA, and CHELSA (left to right) compared to observations. The second row shows absolute biases (x-axis) sorted by catchment elevation (y-axis).

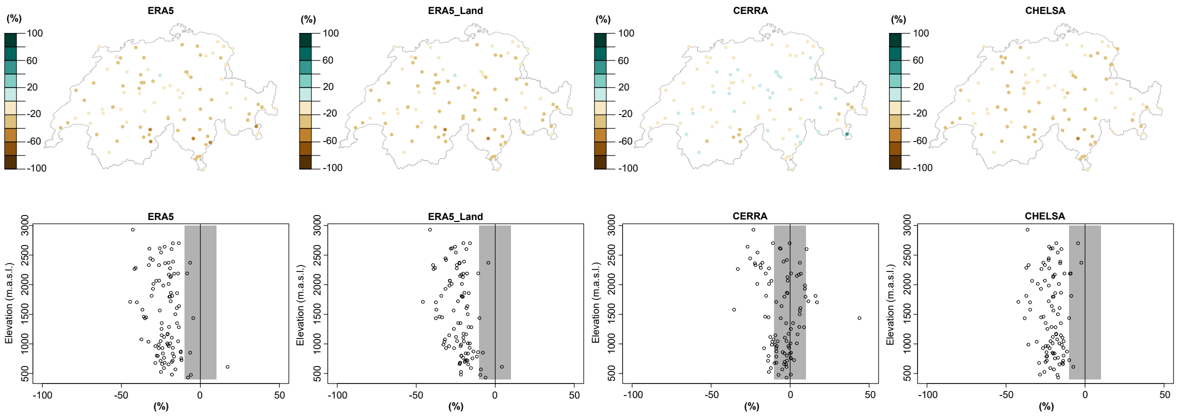


Figure S3: Comparison of the mean annual maximum 1-day precipitation (rx1d) at the catchment level of the four reanalysis datasets (ERA5, ERA5-Land, CERRA, and CHELSA) with the gridded observations and elevation dependence of the biases. The first row shows relative biases (in %) at the catchment location for ERA5, ERA5-Land, CERRA, and CHELSA (left to right) compared to observations. The second row shows relative biases (x-axis) sorted by catchment elevation (y-axis); the grey boxes indicate relative biases within ± 10%.

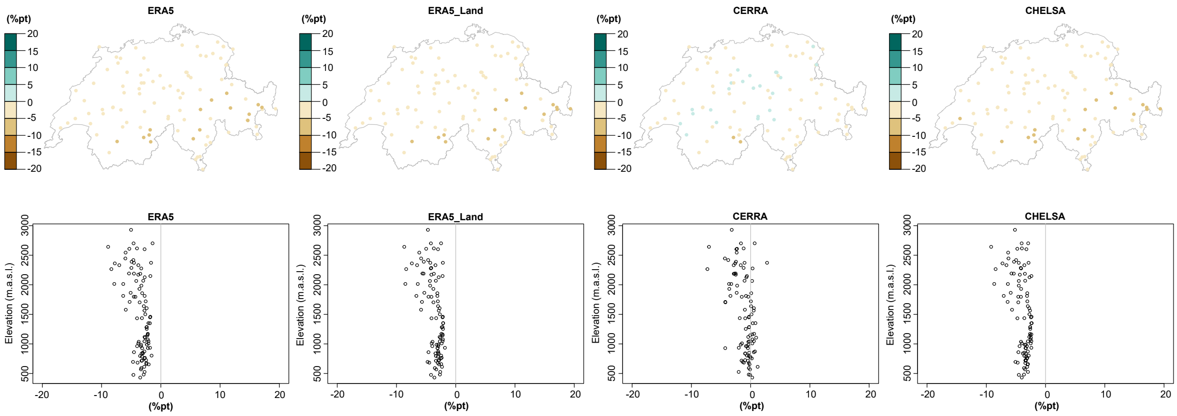


Figure S4: Comparison of the mean fraction of precipitation due to extremely wet days (r99ptot) at the catchment level of the four reanalysis datasets (ERA5, ERA5-Land, CERRA, and CHELSA) with the gridded observations and elevation dependence of the biases. The first row shows absolute biases (in percentage points) at the catchment level for ERA5, ERA5-Land, CERRA, and CHELSA (left to right) compared to observations. The second row shows absolute biases (x-axis) sorted by catchment elevation (y-axis).

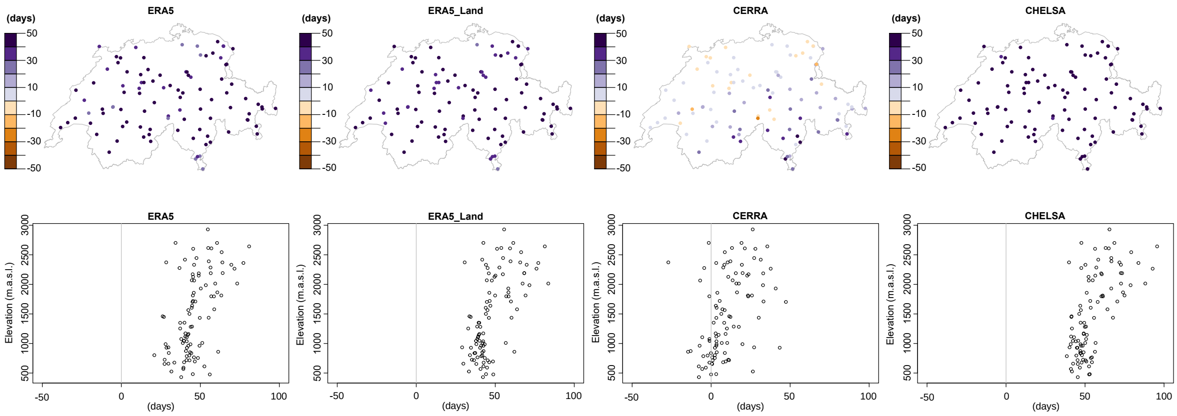


Figure S5: Comparison of the mean number of wet days (wetdays) at the catchment level of the four reanalysis datasets (ERA5, ERA5-Land, CERRA, and CHELSA) with the gridded observations and elevation dependence of the biases. The first row shows absolute biases (in days) at the catchment location for ERA5, ERA5-Land, CERRA, and CHELSA (left to right) compared to observations. The second row shows absolute biases (x-axis) sorted by catchment elevation (y-axis).

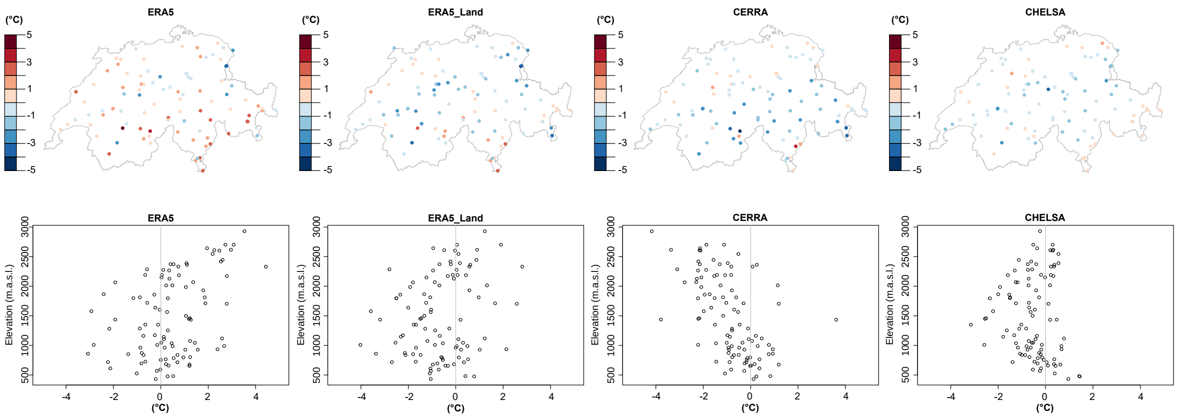


Figure S6: Comparison of the mean annual maximum of daily mean temperature (tg_max) at the catchment level of the four reanalysis datasets (ERA5, ERA5-Land, CERRA, and CHELSA) with the gridded observations and elevation dependence of the biases. The first row shows absolute biases (in °C) at the catchment level for ERA5, ERA5-Land, CERRA, and CHELSA (left to right) compared to observations. The second row shows absolute biases (x-axis) sorted by catchment elevation (y-axis).

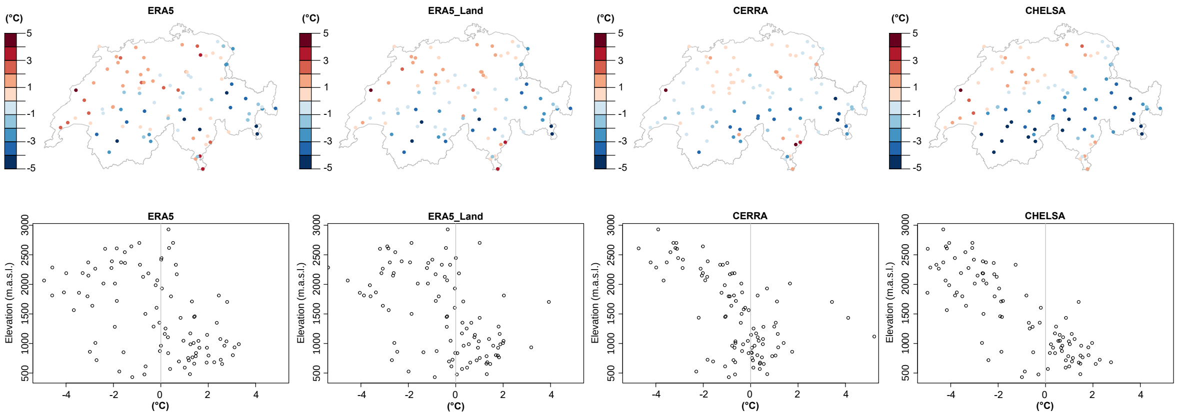


Figure S7: Comparison of the mean annual minimum of daily mean temperature (tg_min) at the catchment level of the four reanalysis datasets (ERA5, ERA5-Land, CERRA, and CHELSA) with the gridded observations and elevation dependence of the biases. The first row shows absolute biases (in °C) at the catchment level for ERA5, ERA5-Land, CERRA, and CHELSA (left to right) compared to observations. The second row shows absolute biases (x-axis) sorted by catchment elevation (y-axis).

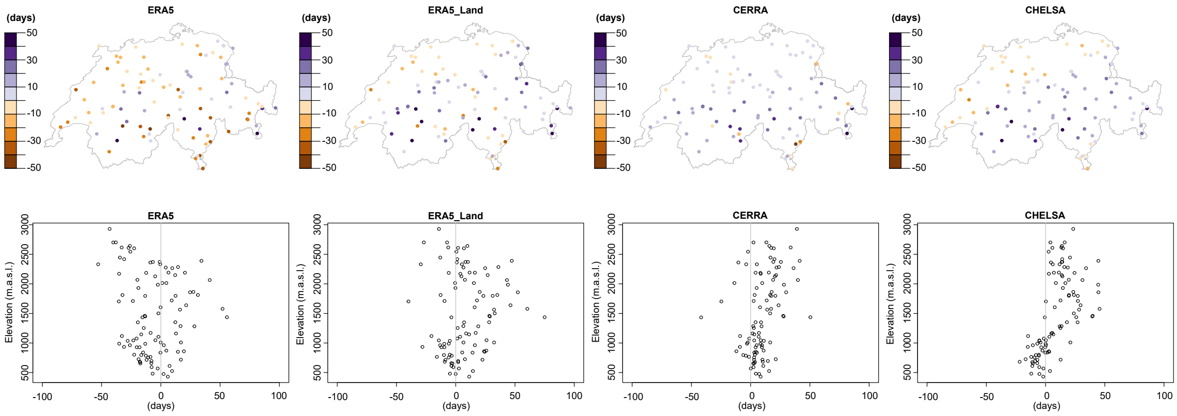


Figure S8: Comparison of the mean number of cold days (colddays) at the catchment level of the four reanalysis datasets (ERA5, ERA5-Land, CERRA, and CHELSA) with the gridded observations and elevation dependence of the biases. The first row shows absolute biases (in days) at the catchment level for ERA5, ERA5-Land, CERRA, and CHELSA (left to right) compared to observations. The second row shows absolute biases (x-axis) sorted by catchment elevation (y-axis).

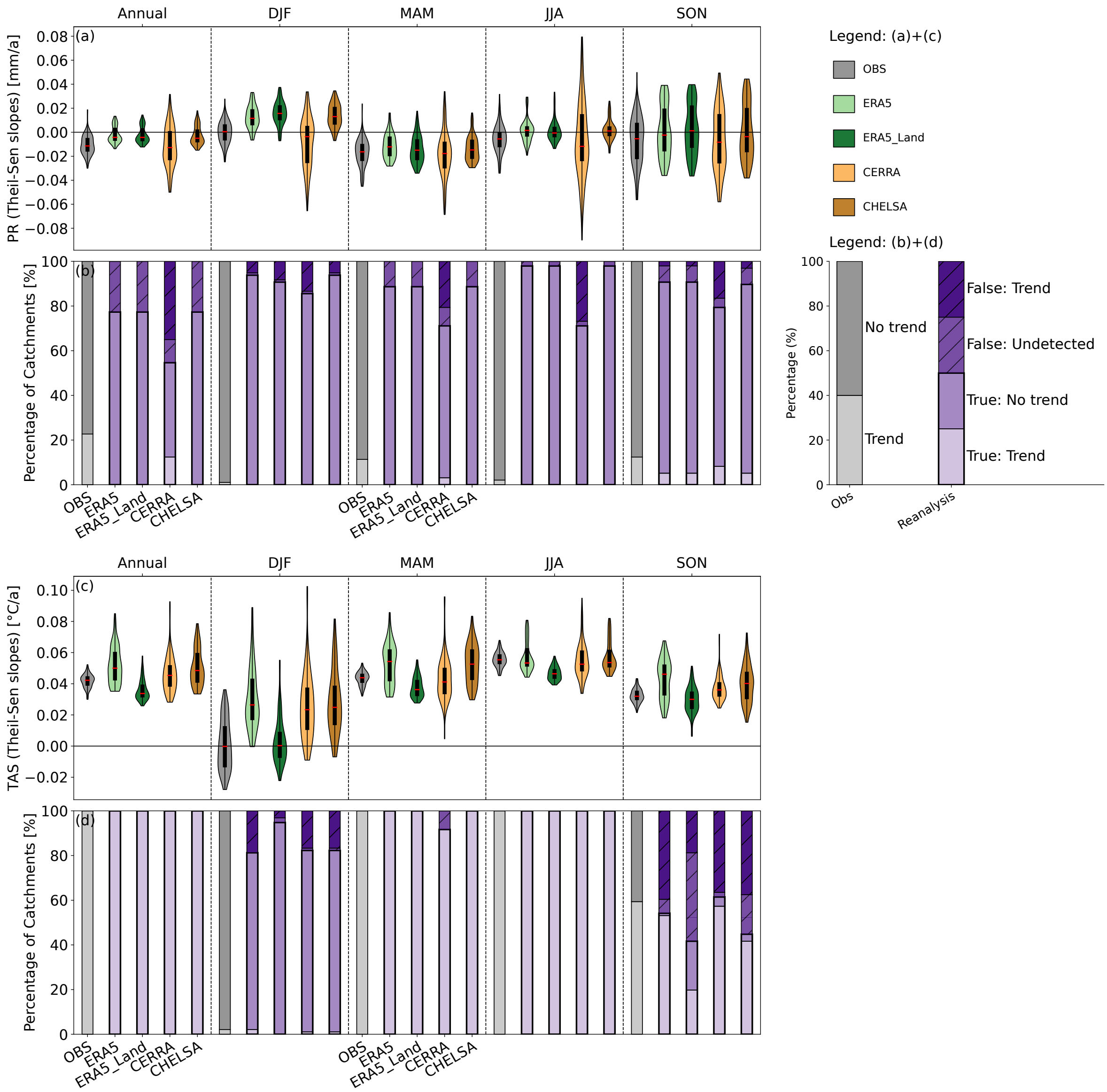


Figure S9: Comparison of trends in precipitation and temperature metrics. Violin plots show trend magnitudes (theil-sen slopes; 1986-2020) in observations (grey) and reanalysis datasets (colors) across all catchments for mean precipitation (a) and temperature (c) at annual and seasonal scales. Stacked barplots show the correct spatial matching of significant trends or no trends in precipitation (b) and temperature (d) by the reanalysis datasets (purple bars) compared to observations. Grey bars show the percentage of catchments with a significant trend (light grey) and no trend (dark grey) in the observations. Purple bars without hatching show the percentage of catchments in the reanalysis datasets with a correct matching of a significant trend (lightest coloring) or no trend (medium light). Hatched bars show the percentage of undetected (medium dark) or false trends (darkest coloring). Each bar represents the respective trends from the aligned violin plots.

# S2. Supplementary figures for the drought analysis

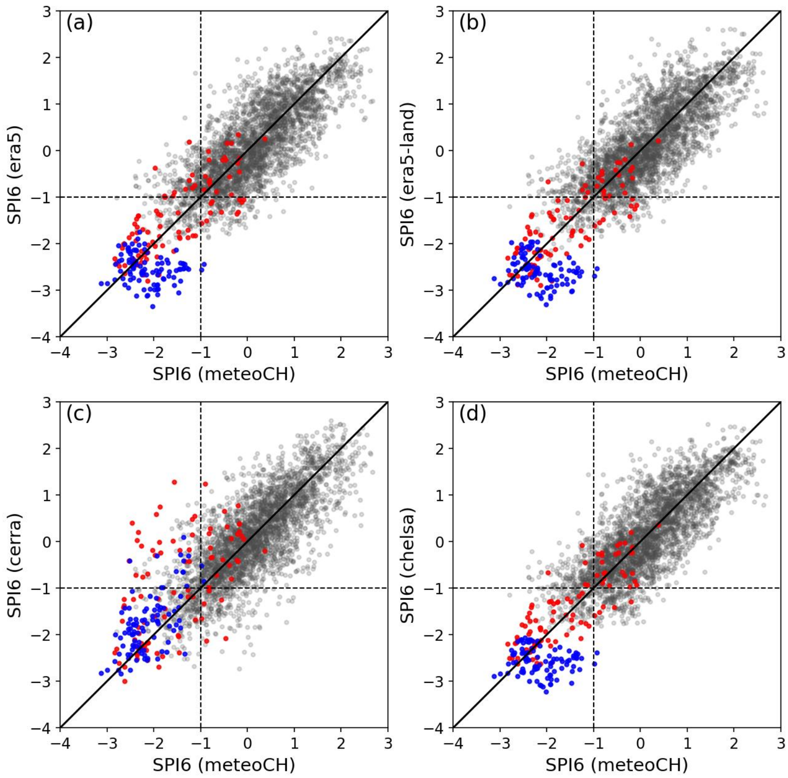


Figure S10: Comparison of SPI6 catchment values (March-August) of all years by the four reanalysis datasets ((a) ERA5, (b) ERA5-Land, (c) CERRA, and (d) CHELSA) compared to gridded observations. Observations on the x-axis and the respective reanalysis dataset on the y-axis. Colored dots show the SPI6 values of the 2003 (blue) and 2018 drought events (red); grey dots all other years. The solid black line indicates the 1:1 line and the dashed lines indicate SPI6 values of -1.

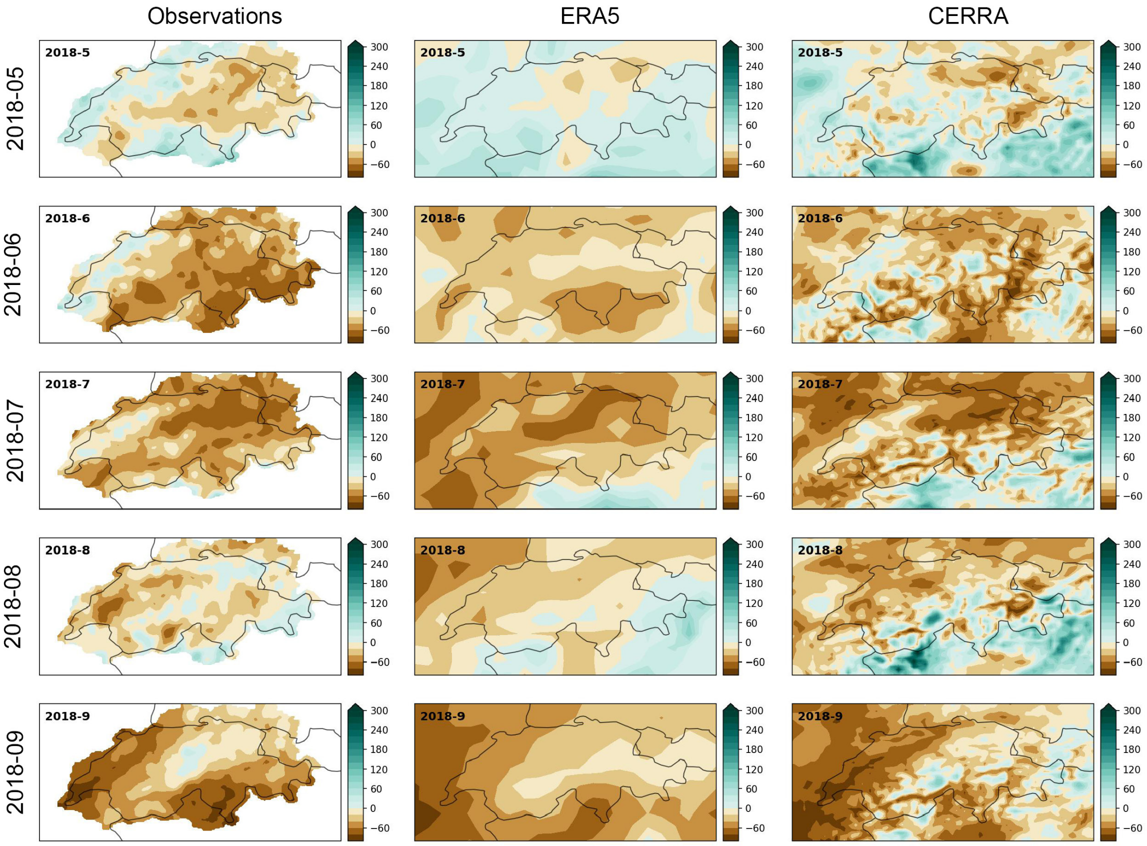


Figure S11: Maps of monthly relative precipitation anomalies (May-September) in observations, ERA5 and CERRA over Switzerland compared to 1986-2020. Observed monthly precipitation anomalies (in %) (left column), ERA5 anomalies (middle), and CERRA anomalies (right) compared to the long-term average monthly precipitation sums for the months May (top) to September (bottom). Green shading indicates above normal and brown colors below normal monthly precipitation.

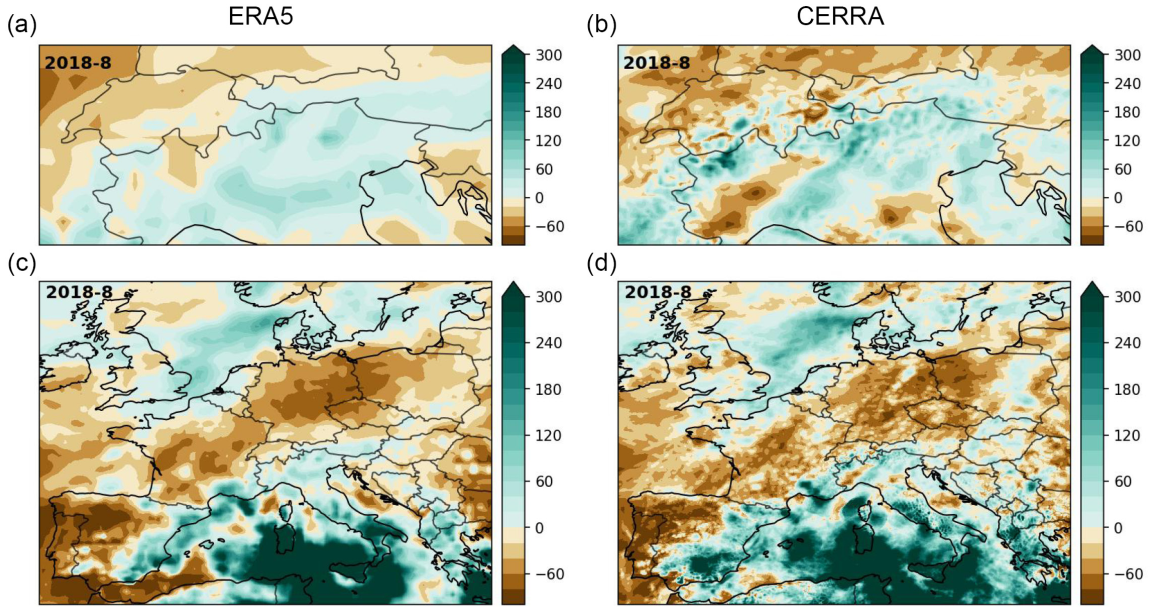


Figure S12: Maps of August precipitation anomalies in ERA5 and CERRA over the Alps and Europe. (a) ERA5 precipitation anomalies (%) over the Alps; (b) CERRA precipitation anomalies (%) over the Alps; (c) ERA5 precipitation anomalies (%) over Europe; (b) CERRA precipitation anomalies (%) over Europe. Relative anomalies compared to the long-term average monthly precipitation sums (1986-2020). Green shading indicates above normal and brown colors below normal monthly precipitation.

## S3. Supplementary figure for the extreme precipitation event analysis

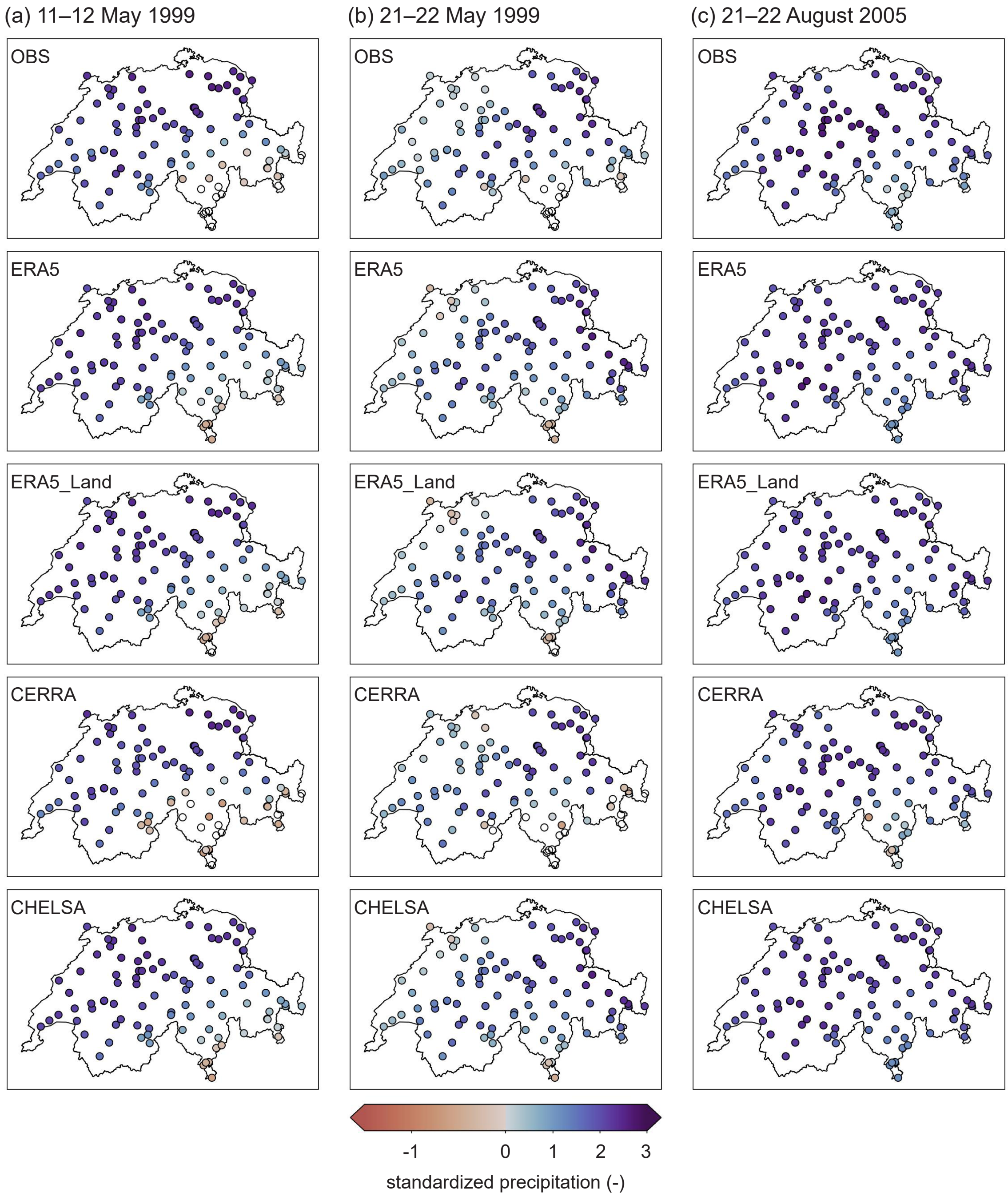


Figure S13: Standardized 2-day catchment precipitation sums of the three extreme precipitation events in the observations and reanalysis datasets. (a, left column) shows the standardized values for the first event (11–12 May 1999), (b, center column) for the second (21–22 May 1999), and (c, right column) for the third event (21–22 Aug 2005). Each colored dot represents a single catchment.

## S4. Supplementary material for the summary of results

In the following figures we have summarized all biases from the four reanalysis datasets (ERA5, ERA5-Land, CERRA, CHELSA) compared to the gridded observations. The biases are reflective of the results shown in the main manuscript. For each mean, extreme and temporal variability metric of precipitation/temperature the median bias (absolute or relative) in high (>2000 m), mid (1000-2000 m) and low (≤1000 m) elevation catchments, and all catchments is given. The biases in metrics are divided into three categories of positive and negative biases; the respective bounds of bias categories are given in the figure legends. For the two extreme event types -meteorological drought (droughts) and 2-day heavy precipitation events (floods)- we quantify the three performance metrics: detection, intensity and extent. The *detection* is representative of the percentage of catchments in the observations that are also correctly detected in the reanalysis dataset. That means the optimal detection is 100% of catchments in the observations also trigger a detection (drought: SPI6<-1, floods: SPI-2day > +1.5) in the reanalysis. We subtract the percentage value by 100% (optimal value) to receive the underestimation in detection in percent. The *intensity* gives the relative bias in drought deficit (drought) and 2-day precipitation sum (floods) as the median bias for the subset of catchments that show an extreme event in the observations. The bias in *extent* is calculated as the ratio of *the number of catchments in the reanalysis with event detection* divided by *the number of catchments in the observations with event detection* minus 1; multiplied by 100 to get percentages. This gives in percentage the over-/underestimation of the drought/flood extent. For the biases in trends, we determine the correct detection of *significant* (i.e. agreement in sign and significance) and *non-significant* observed trends separately. That means the optimal detection is 100% of catchments in the observations also trigger a detection of (non-)significant trends in the reanalysis. We subtract the percentage value by 100% (optimal value) to receive the underestimation in detection in percent. We evaluate the agreement in trends in high (>2000 m), mid (1000-2000 m), and low (≤1000 m) elevation catchments, and all catchments separately. Further, we evaluate the bias in the spatial *extent* of significant trends defined as the ratio of *the number of catchments in the reanalysis with significant trends* divided by *the number of catchments in the observations with significant trends* minus 1; multiplied by 100 to get percentages. For both the trend agreement and the extent, we exclude all cases where less than 10 catchments in the observations show a significant trend.

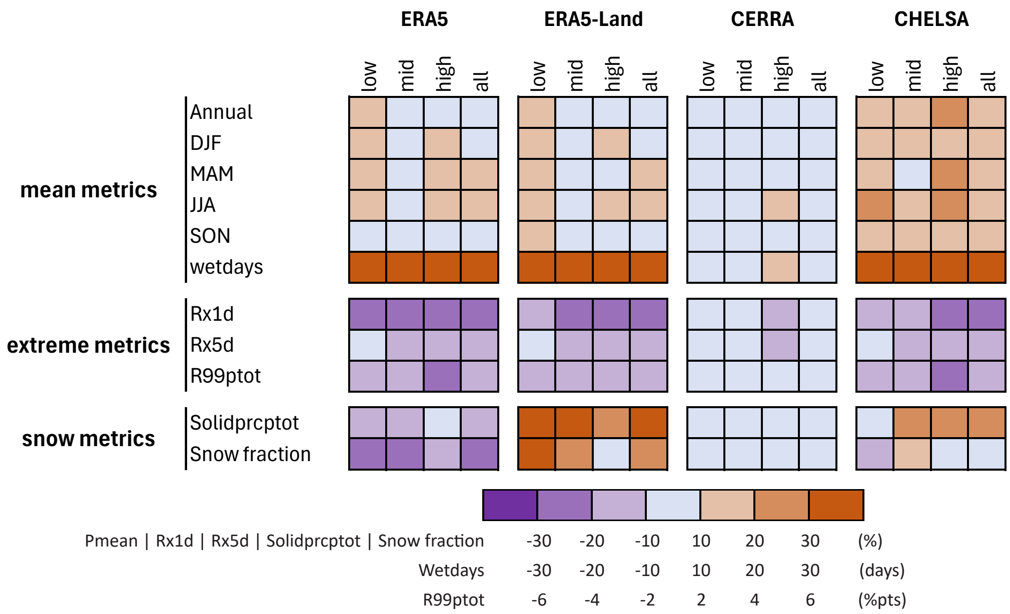


Figure S14: Summary of biases of precipitation metrics (mean and extreme) and snow metrics. For each metric the median bias (absolute or relative) in low (≤1000 m), mid (1000-2000 m) and high (>2000 m) elevation catchments, and all catchments is given. The biases in metrics are divided into three categories of positive (orange) and negative biases (purple); the respective bounds of bias categories are given in the figure legends.

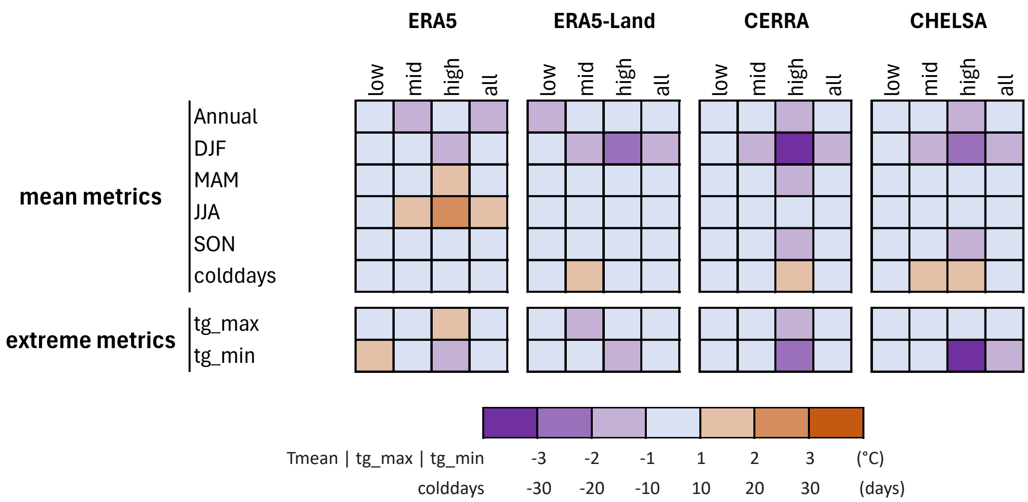


Figure S15: Summary of biases of temperature metrics (mean and extreme). For each metric the median bias (absolute) in low (≤1000 m), mid (1000-2000 m) and high (>2000 m) elevation catchments, and all catchments is given. The biases in metrics are divided into three categories of positive (orange) and negative biases (purple); the respective bounds of bias categories are given in the figure legends.

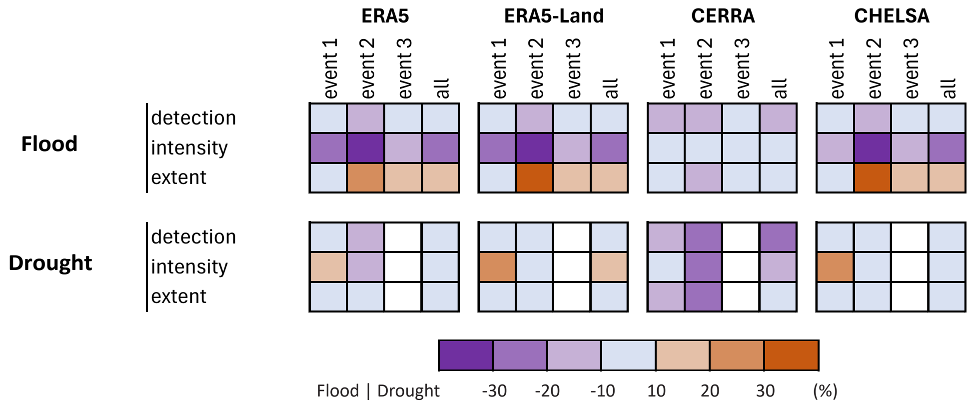


Figure S16: Summary of biases of two meteorological drought events and three heavy precipitation events. For each event the performance metrics detection, intensity, and extent are given.

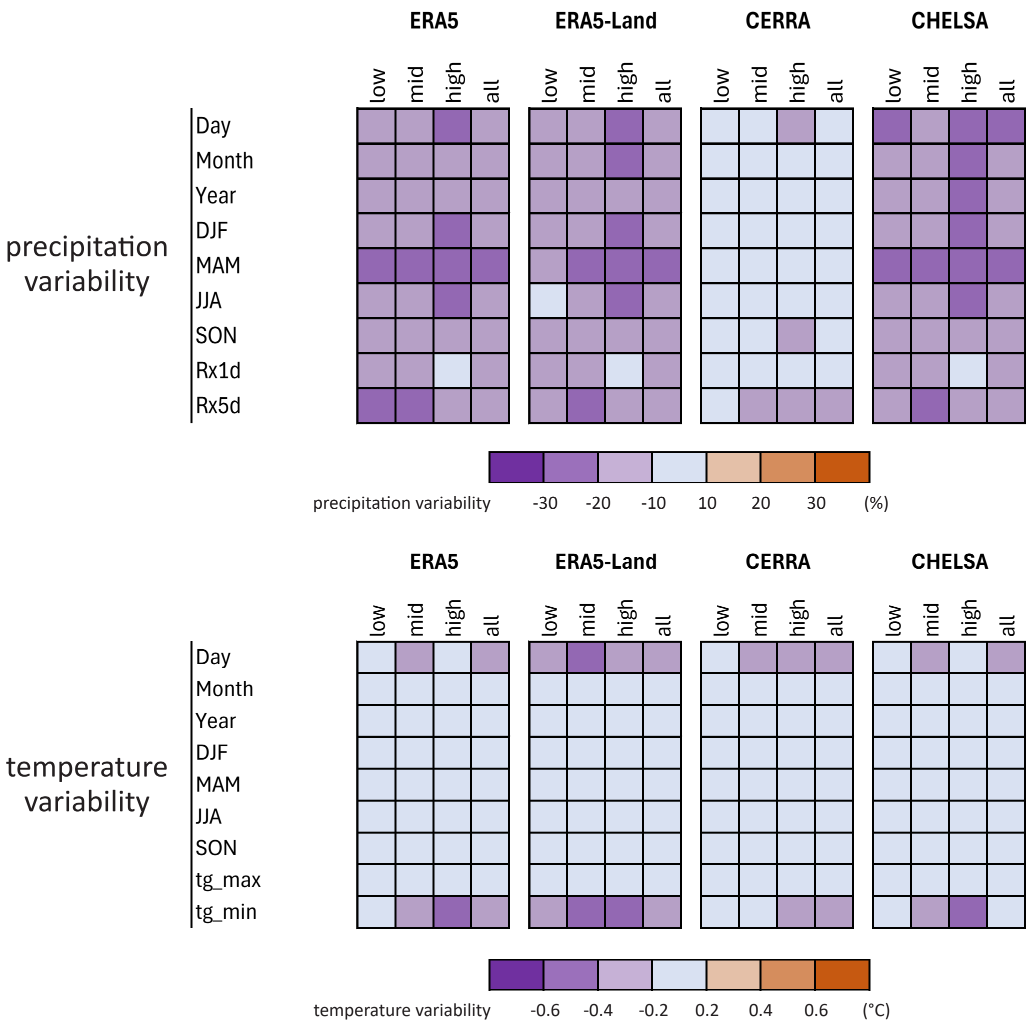


Figure S17: Summary of biases of precipitation and temperature variability. For each metric the median bias (absolute or relative) in low (≤1000 m), mid (1000-2000 m) and high (>2000 m) elevation catchments, and all catchments is given. The biases in metrics are divided into three categories of positive (orange) and negative biases (purple); the respective bounds of bias categories are given in the figure legends.

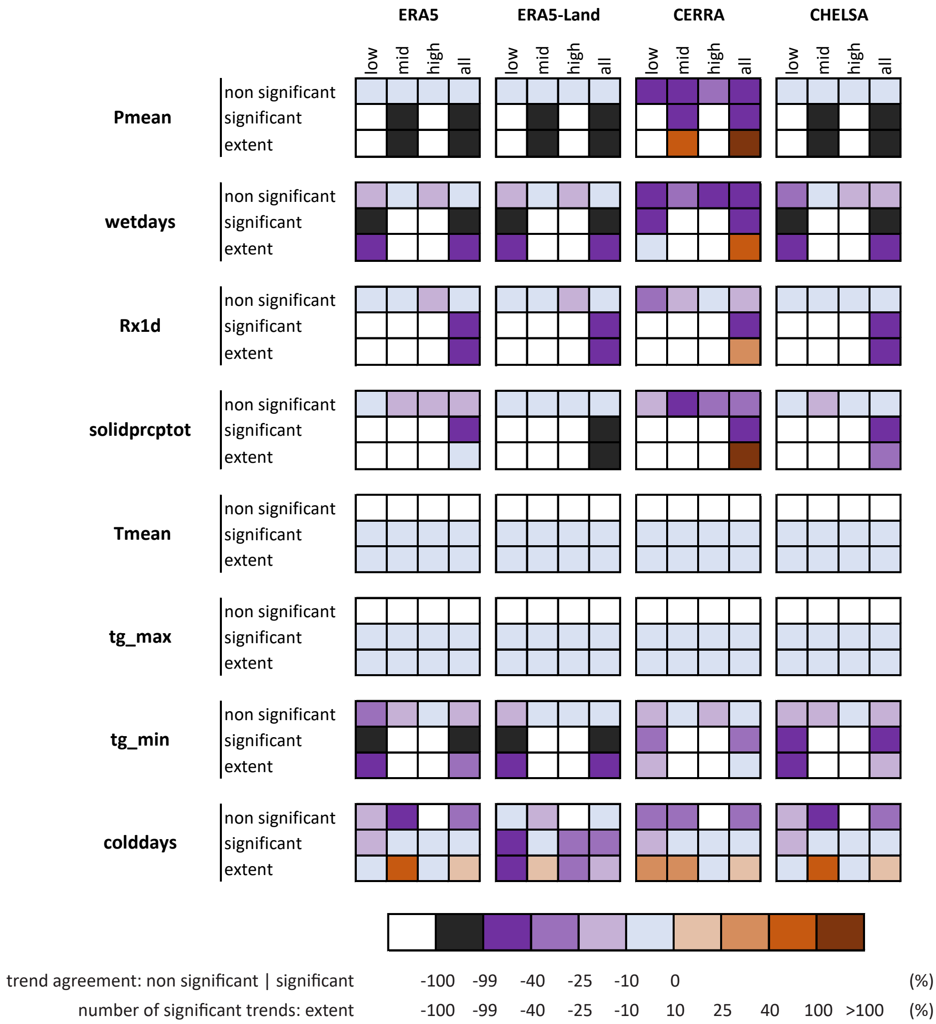


Figure S18: Summary of trend agreement and extent of significant trends. For each metric the agreement of the reanalysis and the observations on *non-significant* and *significant* (i.e. sign and significance) trends is given. Further, the bias in the overall number of significant trends (*extent*) is given. The trends in low (≤1000 m), mid (1000-2000 m) and high (>2000 m) elevation catchments, and all catchments are evaluated. Positive biases in orange-brown and negative biases purple-black; the respective bounds of bias categories are given in the figure legends. Black shading indicates that there is no agreement between reanalysis and observations on significant trends, or in the case of the extent that the reanalysis has no catchments with significant trends. White shading indicates that these cases are excluded from the comparison, because less than 10 catchments in the observations show a significant trend.