Supplementary material

HESS Opinions: More efforts and scientific rigour are needed to attribute trends in flood time series

Bruno Merz, Sergiy Vorogushyn, Steffi Uhlemann, Jose Delgado, Yeshewa Hundecha

Abbreviations:

AMS	Annual maximum flow
LU	Land-use
MK	Mann-Kendall
NAO	North Atlantic Oscillation
OLS	Ordinary least square linear regression
POT	Peak-over-threshold series

Table 2: Selected studies on changes in flood behavior containing attribution statements. The table provides a short characterization of the studies and briefly summarizes how they have approached the ingredients of trend attribution, namely proof of consistency, proof of inconsistency and provision of confidence statement.

Paper	**Characterization of study** (**O**: object of study; **D**: data used; **M**: methods used; **R**: results; **A**: attribution statement)	**Proof of consistency**	**Proof of inconsistency**	**Reliability statement**
Bormann et al. (2011)	**O:** Trends in annual maximum flow series (AMS), flood frequencies and water stage series in German rivers. Quantification of effects of river engineering. **D:** 78 gauges Germany throughout Germany; different record lengths across gauges. **M:** Petitt test and standard normal homogeneity test; 2 trend tests: Mann-Kendall (MK) without pre-whitening; Ordinary least-square linear regression (OLS). Specific gauge analysis. **R:** Heterogeneous spatial patterns with increasing and decreasing trends. Difference in detected trends between 2 trend methods. **A:** Trends at 19 (OLS) / 18 (MK) gauges out of 78 are forced by upstream mechanisms like climate or land-use (LU) change. Simultaneous decreasing trends at 6 gauges are attributed to river training/dam construction.	**Based on qualitative reasoning** • Flood trends are in 19 (OLS) / 18 (MK) cases attributed to upstream mechanisms such as change in climate and land use without distinguishing between both. Increasing trends at gauges across Germany are connected to changes in regional climate including increasing frequency in flood-generating circulation patterns. This statement is inferred by referencing other published works without analyzing the consistency between the studied time periods of trends in discharge and regional climate. Additionally, the flood trends are associated with increase in runoff yields due to urbanization and intensified agriculture based on the references to the published works without a proof of consistency between investigated catchments and time periods. • Authors associate the decreasing trends at 6 gauges with construction of upstream dams without a quantitative proof of the dam effect.	**Partly based on qualitative reasoning, partly missing:** • Specific gauge analysis proves inconsistency with in-stream mechanisms, i.e. river training impact at the gauge location. It cannot prove the absence of river training impact on discharges and water levels originating upstream of the gauge. • Inconsistency proof with climate forcing or land-use changes is missing.	**No**

Cunderlik & Burn (2004)	**O:** Trends in monthly maximum flows, mean monthly air temperature, monthly maximum 1-day precipitation, monthly maximum precipitation events in pristine catchments (less than 5% LU change) in climatologically and physiographically homogeneous region British Columbia, Canada. **D:** 26 discharge gauging stations with average length of 38 years (1960-1999); 25 air temperature stations (average length 88-90 years); 30 precipitation stations (average length 83-84 years), some dating back to 1895; 1st set between 1960-1999 for comparison with gauging stations; 2nd subset of 19 temperature and 18 precipitation stations 1920-1999. **M:** MK test for trend detection; regional homogeneity and regional trend significance; combined local and regional trend significance index. **R:** Different increasing and decreasing trends in flow and climatic data at different months of the year. **A:** Positive trends in spring flow attributed to positive trends in temperature (earlier snow melt) and slight increase in precipitation. Decreasing maximum monthly flow in June attributed to positive temperature trends in April/May. Negative trends in autumn flow attributed to the increase in mean monthly temperatures in summer.	**Based on quantitative reasoning:** The authors carry out a cross-correlation analysis on residuals of maximum monthly flows and climate variables after removing all serially dependent components. The cross-correlation analysis is performed with lag-0 and lag-1 variables.	**Based on qualitative reasoning:** The authors state that the investigated catchments are pristine and less than 5% area experienced LU change. If the impact of human interventions in the river system and catchments on flood behavior is negligible, the selection of catchments qualifies as proof of inconsistency.	**No**
Hannaford and Marsch (2008)	**O:** Flood trends in largely undisturbed catchments distributed across entire UK. Relationships between flood trends and NAO (North Atlantic Oscillation). **D**: UK benchmark network of near natural catchments, 2 standard periods: 87 stations covering 1969-2003, 30 catchments covering 1959-2003; 6 stations >55 years for comparison **M:** OLS, MK, correlation, FS test; multiple flood indicators. Comparison of different time periods. **R:** Positive trends in flood indicators observed in maritime-influenced, upland catchments in the north and west of UK. **A:** Suggest that changes in flood behavior are climate driven. The high-flow trends in western UK since 1960s are likely caused by the shift towards more prevalent positive NAO index.	**Based on semi-quantitative reasoning:** • Reference to changes in extreme precipitation (no quantification). • Correlation analysis between NAO index and winter high flow series shows some (regional) influence of NAO index on flood generation (NAO shift and its influence on precipitation not shown, reference to literature).	**Partly based on qualitative reasoning, partly missing:** • For small catchments (benchmark network), the inconsistency of the detected changes with LUC or river training is proven by choosing undisturbed catchments. • Not given for larger and disturbed catchments.	**No**

Mudelsee et al. (2003)	**O:** Gradual trends in flood frequency for gauge Dresden (Elbe River) and gauges Krosno, Eisenhüttenstadt (Oder River) **D:** Systematic gauge measurements for 80-150 years; historical sources to derive time series of flood occurrence rates for period 1000-2000. **M:** Systematic discharge observations combined with historical information; reconstruction of time series of impact-related magnitude scale (minor, strong, exceptionally strong flood); trend detection based on Gaussian kernel estimation and bootstrap simulation. **R:** Decrease in winter flood frequency; no change in summer floods, no upward trends in occurrence of extreme floods. **A:** Decrease in winter flood frequency partly attributed to fewer events with strong freezing and river ice. Decreasing trends in winter floods coincide with decrease of 25-year maxima of winter precipitation. Reductions in river length, construction of reservoirs, deforestation had minor effects on flood frequency change.	**Based on qualitative reasoning:** • The authors argue that river ice may enhance floods severely and that the rate of river freezing reduced in recent time. Additionally, it is argued that increased winter temperatures reduce the occurrence of frozen soils which may enhance winter floods. • The authors argue that a decreasing trend in winter floods is consistent with a decreasing trend in the occurrence of 25-year maxima of winter precipitation as inferred from the gridded 2.5°x3.75° monthly precipitation dataset. However, the strength of the causal link based on such crude spatial and temporal resolution is arguable.	**Based on quantitative reasoning (concerning reservoirs) and qualitative reasoning (concerning LU):** • In simple quantification approach, the authors account for the effect of increases in reservoir size over time. They conclude that the reservoirs are too small to influence the occurrence of heavy floods. • The authors qualitatively argue that LU changes would induce upward trends and therefore cannot be responsible for decreasing trends. • Effects of reduction in river length are not studied.	**No**
Novotny and Stefan (2007)	**O:** Trends in multiple flood indicators for gauging stations throughout Minnesota, USA; correlation with changes in climate **D**: 36 gauging stations in 5 river basins in Minnesota (end year 2002, min. length 50 years, no heavy anthropogenic discharge controls in the stream); precipitation data were pooled and averaged (10 to 20 stations) for 9 climate divisions (time period not given) **M:** Exclusion of highly regulated stations. MK including pre-whitening; FS test; multiple flow indicators (e.g. high flow days, extreme flow days), multi-period trend testing, trend testing on 5-year, 10-year and 25-year running averages to detect changes in trend over time, spectral density functions for detecting periodicities, comparison of mean annual precipitation trend patterns with flow trend patterns, correlation of flow series per station with precipitation series (average over corresponding climate division). **R:** Similar temporal patterns of the different stream flow statistics. Trends are not monotonic but periodic. Snowmelt flood statistic has not changed for northern parts. Peak flows due to rainfall (summer) are increasing and coincide with more frequent heavy rains (reference to literature); generally stronger trends in all flow statistics in 80s/90s compared to entire study period. **A:** In some cases trends were found in the annual maximum precipitation that are consistent with those of the summer peak. Based on the high correlations obtained, the authors concluded that trends in the flow indicators are caused by corresponding changes in precipitation.	**Based on quantitative and qualitative reasoning:** • Investigated changes in annual precipitation (average of 10-20 stations per climate division) and their correlation with the flow statistics, however annual precipitation might be a weak measure of extreme flow. • Qualitative comparison of spatial patterns of trends in rainfall and discharge.	**No.** Comparison to literature to infer influences of river regulation, farming, logging and urbanization. Highly unsystematic, no conclusions can be drawn.	**No**

Petrow & Merz (2009)	**O:** Gradual trends in flood time series (different indicators) for many catchments throughout Germany. **D:** AMS and other flood indicators for 145 catchments (> 500 km^2) for joint time period 1951-2002. **M:** MK test, including pre-whitening; FS tests **R:** Trends detected for considerable fraction of basins. Trends mostly increasing. Marked differences in spatial and seasonal trend patterns. Basins with trends were spatially clustered. **A:** Suggest that observed flood trends are climate-driven (no statement on contribution of climate variability vs. climate change).	**Based on qualitative reasoning:** • Authors argue that coherent (in terms of spatial and seasonal patterns) flood trends at many sites across large geographical regions points to climatic driver. • Qualitative comparison of flood trends with changes in atmospheric circulation patterns supports hypothesis of climatic driver. Comparison based on referencing other published works.	**Based on qualitative reasoning:** • Local anthropogenic effects may influence at-site behavior, but they are not expected to cause coherent seasonal changes over large areas. • Detected changes are not caused by LU change, since LU change is expected to show dependence on scale.	**No**
Pinter et al. (2006)	**O:** Gradual trends in peak winter (November-April) discharge (POT) and frequencies of high flows at gauge Cologne/Rhine in Germany. Basin cumulative precipitation sums over 10-days prior to the peak derived from data at 31 climate stations and frequencies of precipitation events per decade exceeding certain thresholds in the period 1900-2000. **D:** Stage data and peak discharge for gauge Cologne for 1900-2000. Precipitation data from 31 stations in catchment. **M:** Linear regression specific gauge analysis. **R:** Significant increase in winter flood peaks and flood frequencies for floods between 5000 and 7500 m^3/s consistent with significant increase of cumulative basin precipitation sums and frequencies of precipitation events (40-120 mm). The authors found a significant increase in runoff yield for the precipitation sums between 60 and 80 mm in the basin gauged at Cologne. **A:** Increase in flood magnitude and frequency is partly driven by the increase in magnitude of flood-causing precipitation. River training at the reach from Maxau to Cologne contributed little or nothing to these changes. Increase in runoff yield coincides with massive LU changes and change in agricultural practices.	**Based on quantitative and qualitative reasoning:** • The authors establish a link between flood-causing precipitation and flood indicators (magnitude and frequency) via a correlation analysis Flood-causing precipitation is determined as the basin precipitation sum over 10-days prior to a flood peak. However, it is open to which extent the basin precipitation integrated over 10 days and over the basin area of 144232 km^2 is an appropriate measure for flood peaks. • Changes in runoff yield are associated with increase in urban areas, decrease in agricultural acreage, shift in the crop mix, changes in soil drainage and soil degradation due to agricultural practices by referencing other literature sources. No quantitative description and prove of cause-effect relationships are provided.	**Partly based on quantitative analysis, mostly qualitative reasoning:** River training measures between Maxau and Cologne contributed little or nothing to changes in flood magnitudes and frequencies: However, specific stage (stages associated with constant discharge) analysis cannot prove the absence of impact of river training in the entire catchment on changes in flood discharges and frequencies. It only provides a statement for the reach of investigation.	**No**
Robson et al. (1998)	**O:** Gradual trends in 2 UK national flood time series, representing magnitude and frequency of flood occurrence at national scale. **D:** Different time periods (mainly since 1940); POT and AMS data from 890 gauges in UK pooled across all sites. **M:** OLS, normal scores regression, Spearman's correlation; smoothing techniques to quantify fluctuations. **R:** No trends in national flood behavior detected, however, year-to-year variation shows fluctuations. **A:** Year-to-year fluctuations are driven climatically.	**Based on semi-quantitative reasoning:** Similar fluctuations in and highly significant correlation between UK annual average rainfall and average number of floods per year. However, authors point to problem that annual average rainfall is "rather crude measure of flood-producing potential".	**Based on qualitative reasoning:** Climatic variation is most obvious source of fluctuations over time scales of 5-10 years. Other factors tend to change in slow and progressive way.	**No**

Van der Ploeg and Schweigert (2001)	**O:** Attribute increase in flood peak runoff in East Germany to changes in agriculture practices after 1950. **Data:** Series of flood stages and discharges at gauges Dresden/Elbe and Wittenberge/Elbe 1900-2000. **M:** Frequency distribution of river stages at Dresden and Wittenberge 1900-2000. Comparison of peak discharges for 10 most extreme floods in Wittenberge to corresponding peak discharges in Dresden. Comparison of agricultural practices before and after World War II with indicators such as average farm size and power of tractors. No correlation between the statistics. **R:** Floods that exhibit a difference of >1000m³ between Dresden and Wittenberge have only started to occur past 1950 (control period 1900-2000). Farm sizes and machinery weight increased after WWII. **A:** The observed increase in peak discharges in Wittenberge coincides with a change in agricultural management, i.e. large scale farms, heavy machinery, extensive subsurface drainage that leads to increased runoff formation during heavy rainfall.	**Based on speculation:** No proof of the effect of agriculture on runoff (only cites study on river Rhine). No correlation between runoff and any agricultural variable. Hypothesize on the coinciding phenomena only.	**No**	**No**
Villarini et al. (2011)	**O:** Investigate the validity of the stationarity assumption in the flood peak records over part of Central Europe. **D:** Series of daily averaged discharge of 55 gauging stations, Central Europe (Germany - 32, Switzerland - 13, Czech Republic - 6, Slovakia - 4), record length at least 75 years. **M:** Derivation of annual and seasonal maxima series. Step changes (mean and variance) are tested with Pettitt test. Gradual trends are tested using MK and Spearman tests in blocks separated by the time point of step changes. **R:** 16 stations with change points in mean; 4 in variance. If change points are accounted for, most gradual trends are no longer significant. Abrupt changes, rather than monotonic trends are responsible for changes in flood behavior. **A:** Detected step changes can often be associated with anthropogenic effects, such as construction of dams and reservoirs and river training. It is difficult to detect a possible climate change signal in the flood peak record, due to profound changes in the catchments.	**Based on qualitative reasoning:** Authors compared the detected change points and human interventions in rivers/catchments reported in literature. This comparison is not consistent for the case of the Elbe River, since the intervention dates and regions in the cited studies (Helms et al. 2002, Socher et al. 2008) do not coincide with detected change points.	**No**	**No**