Water Resources Research®

RESEARCH ARTICLE
10.1029/2021WR030049

Complex High- and Low-Flow Networks Differ in Their Spatial Correlation Characteristics, Drivers, and Changes

Manuela I. Brunner[1,2] **and Eric Gilleland**[2]

[1]Institute of Earth and Environmental Sciences, University of Freiburg, Freiburg, Germany, [2]Research Applications Laboratory, National Center for Atmospheric Research, Boulder, CO, USA

Key Points:

- We propose and use a tail dependence measure to map and compare complex networks of high and low flows in Central Europe at a seasonal scale
- Low flows are related more strongly and over longer distances than high flows and relationships are strong in spring and weak in summer
- Seasonal flow correlation is shaped by spatial dependence in drivers with varying importance of precipitation, evaporation, and snowmelt

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

M. I. Brunner,
manuela.i.brunner@gmail.com

Citation:

Brunner, M. I., & Gilleland, E. (2021). Complex high- and low-flow networks differ in their spatial correlation characteristics, drivers, and changes. *Water Resources Research*, *57*, e2021WR030049. https://doi.org/10.1029/2021WR030049

Received 24 MAR 2021
Accepted 30 AUG 2021

Author Contributions:

Conceptualization: Manuela I. Brunner
Data curation: Manuela I. Brunner
Formal analysis: Manuela I. Brunner
Funding acquisition: Manuela I. Brunner
Methodology: Manuela I. Brunner, Eric Gilleland
Supervision: Eric Gilleland
Visualization: Manuela I. Brunner

Abstract Hydrologic extremes such as floods and droughts are often spatially related, which increases management challenges and potential impacts. However, these spatial relationships in high and low flows are often overlooked in risk assessments and we know little about their differences and origins. Here, we ask how spatial relationships of both types of hydrologic extremes and their potential hydro-meteorological drivers differ and vary by season. We propose lagged upper- and lower-tail correlation as a measure of extremal dependence for temporally ordered events to build complex networks of high and low flows. We compare complex networks of overall, low and high flows, determine hydro-meteorological drivers of these networks, and map past changes in spatial relationships using a large-sample data set in Central Europe. Our network comparison shows that low flows are correlated more strongly and over longer distances than high flows and high- and low-flow networks are strongest in spring and weakest in summer. Our driver analysis shows that high-flow dependence is most strongly governed by precipitation in winter and evapotranspiration in summer while low-flow dependence is most strongly governed by snowmelt in winter and evapotranspiration in fall. Finally, our change analysis shows that changes in connectedness (i.e., the number of catchments a catchments shows strong flow correlations with) vary spatially and are mostly positive for high flows. We conclude that spatial flow correlations are considerable for both high and particularly low flows as a result of a combination of spatially related hydro-meteorological drivers whose importance varies by extreme type and season.

Plain Language Summary Droughts and floods can happen in multiple locations at once with important implications for flood and drought risk. Still, the spatial relationships between events and the reasons for them are not well studied. Here, we therefore ask how spatial relationships of both types of extremes and their meteorological drivers differ and vary by season. We compare networks of overall, low and high flows, determine hydro-meteorological drivers of these networks, and map past changes in flow dependence using a large-sample data set in Central Europe. Our network comparison shows that low flows are correlated more strongly and over longer distances than high flows and both high- and low-flow networks are strongest in spring and weakest in summer. Our driver analysis shows that high-flow dependence is governed by precipitation dependence in winter and evapotranspiration dependence in summer and fall while low flow dependence is most strongly governed by snowmelt in winter and evapotranspiration and snowmelt in fall. Finally, our change analysis shows that changes in connectedness (i.e., the number of catchments a catchments shows strong flow correlations with) vary spatially and are mostly positive for high flows. We conclude that spatial flow correlation is considerable for both high and particularly low flows highlighting the need to consider it in risk assessments.

1. Introduction

Hydrologic extremes such as droughts and floods can be spatially related, that is they can co-occur in multiple catchments simultaneously. Such co-occurrence increases management challenges and potential economic and societal impacts because emergency measures need to be coordinated across catchments and damages may be widespread. A recent example of a drought event with widespread impacts is the 2018 drought in Europe, which affected Central and Northern Europe (Bakke et al., 2020; Brunner et al., 2019) and a variety of sectors including forestry, agriculture and water supply (Stephan et al., 2021). An example for the other side of the extreme spectrum is the widespread 2019 Mississippi flood, which affected the Missouri and Mississippi river basins and caused major damages to infrastructure (Pal et al., 2020). Despite the potential importance of spatial correlations with respect to impacts, we often overlook these relationships

Writing – original draft: Manuela I. Brunner
Writing – review & editing: Eric Gilleland

and study extremes from a local univariate perspective, which can lead to a misestimation of hazard and impacts (Thieken et al., 2015).

A few studies have tried to consider the regional dimension of hydrologic extremes in hazard assessments and to statistically model the spatial dependencies in flood and drought occurrence. For floods, spatial dependencies have been considered in modeling and hazard assessment efforts, for example by Keef et al. (2013) and Quinn et al. (2019) who have adopted the Heffernan and Tawn model (Heffernan & Tawn, 2004) to simulate widespread synthetic flood events and to calculate the economic loss of large flood events or by Brunner, Papalexiou, et al. (2020) who have used a spatio-temporal stochastic model to simulate spatially consistent flood events and quantify regional flood hazard over the United States. For droughts, the spatial dimension has mostly been considered in terms of drought extent (Brunner, Swain, et al., 2021; Rudd et al., 2019), by computing regional drought indices (Rossi et al., 1992), performing frequency analysis using severity-area-frequency curves (Andreadis et al., 2005; Henriques & Santos, 1999; Hisdal & Tallaksen, 2003), and in stochastic modeling through the use of max-stable models (Oesting & Stein, 2018). While it is increasingly common to consider spatial dependencies in modeling efforts, less attention has been paid to understanding the variations and drivers of these spatial dependencies. Only recently, Brunner, Gilleland, et al. (2020) have shown that spatial flood dependencies vary in time and space and are substantially modulated not just by precipitation but also land-surface processes, particularly in mountainous regions where snow accumulation and melt modulates the timing of flood occurrence. The drivers of spatial drought extents but not spatial dependencies have been studied by Brunner, Swain, et al. (2021), who have shown that soil moisture and temperature are important drivers of drought extent from winter to spring and in summer, respectively.

While these previous studies suggest that the spatial dependencies of hydrologic extremes vary seasonally and are governed by a variety of factors, they focus on a single type of extreme. Hence, we lack an understanding of how the dependencies in high and low flows differ from each other and from dependencies in the 'normal' state (daily streamflow). Here, we therefore compare the spatial dependencies of both types of hydrologic extremes and their potential hydro-meteorological drivers to each other and to spatial dependencies in the 'normal' state. Specifically, we ask: (a) How are hydrologic extremes (both high and low flows) spatially related across Central, Western, and Northern Europe and how do their correlations differ from general correlation patterns and vary per extreme type and season? (b) How are spatial correlations in hydrologic extremes related to correlations in potential hydro-meteorological drivers and how does driver importance vary by extreme type? (c) How has spatial connectedness in high and low flows (i.e., the number of catchments a catchment shows strong flow correlations with) changed over time and how do these changes differ for the two types of extremes?

To address these questions, we use complex network theory, which enables visualizing and describing structures and connections of large data sets (Kolaczyk & Csardi, 2020; Luke, 2015). Complex networks are networks that describe patterns of connections between their elements, in our case, between pairs of catchments. Formally, they consist of a set of vertices or network nodes, and a set of edges or links connecting the nodes. In our context, nodes represent catchment outlets and edges indicate similarities in streamflow patterns between a pair of catchments. Using complex network theory will allow us to map pairs of catchments with similar streamflow behavior according to a specific metric, which might not necessarily lie on the same river network.

A common method to assign edges to node pairs is to use a linear (Pearson) correlation coefficient exceeding a specific threshold. Such a correlation-based approach has also been frequently used in the geo- and climate sciences for example to study climate networks using atmospheric variables such as temperature (Donges et al., 2009; Tsonis & Roebber, 2004; Yamasaki et al., 2008), the spatial connections in rainfall networks (Guo et al., 2017; Han et al., 2020; Jha & Sivakumar, 2017; Jha et al., 2015; Naufan et al., 2018; Scarsoglio et al., 2013; Sivakumar & Woldemeskel, 2015; Tiwari et al., 2020), or the properties of streamflow networks. Such streamflow networks have been analyzed by mapping connections of annual (Han et al., 2018) or monthly streamflow (Sivakumar & Woldemeskel, 2014) and by identifying structures of similar catchments (Fang et al., 2017; Halverson & Fleming, 2015) or critical nodes in a measurement network (Sarker et al., 2019).

Much less attention than to studying the overall correlation behavior has been devoted to studying networks of extreme events. In the context of extremes, event synchronization has been proposed as a measure to define network edges. This measure counts the number of temporally coinciding events in two time series by considering potential time lags between events at two locations (Malik et al., 2012) and has been applied to study networks of extreme precipitation (Boers et al., 2014; Ozturk et al., 2018) and drought (Konapala & Mishra, 2017). While the event synchronization measure focuses on the number of co-occurring events, it does not quantify how strongly the co-occurring events are related, highlighting the need for additional suitable measures to study networks of extremes. In addition, the focus of past climate network studies has mainly been on mapping networks or identifying network communities and less on explaining network structure. One of the few studies also investigating potential drivers of relationships is the study by Jha et al. (2015) who tried to explain spatial correlations across a raingauge network with both topographic characteristics and rainfall properties.

Here, we expand complex network analysis in hydrology beyond studying overall correlation or temporal co-occurrence by proposing a measure of upper and lower tail correlation to define networks of hydrological extremes. Using complex networks as a graphical tool will allow us to identify spatial connections indicated through strong flow dependencies going beyond river network boundaries and using network descriptors will enable comparisons between networks of different types. In addition, we go beyond analyzing network structure by studying the potential drivers of correlation strength in such extreme networks. We map spatial high- and low-flow connections (i.e., connect pairs of catchments with strong flow correlations) in Central Europe using a tail dependence measure considering the temporal order of events and use multiple linear models to link spatial dependencies in hydrologic extremes to spatial dependencies in potential hydro-meteorological drivers including precipitation, evapotranspiration, and snowmelt. A better understanding of similarities and differences in networks of droughts and floods and their drivers will potentially help us to better understand potential future changes in spatial dependencies of extremes and related regional drought and flood risk.

2. Methods and Materials

Our analysis consists of three main steps: (a) mapping and comparing complex networks describing normal, high- and low-flow conditions in Central Europe, (b) explaining the different complex network patterns by dependencies in potential hydro-meteorological drivers, and (c) assessing past changes in complex network properties.

2.1. Data

For our analysis of complex networks under normal and extreme flow conditions, we use a large-sample data set of 937 catchments in Central Europe, which are partly nested (Figure 1), available through the Global Runoff Data Centre (GRDC; The Global Runoff Data Centre 56068 Koblenz Germany, 2019). These catchments have daily streamflow observations from 1969 through 2011, enabling the temporal analysis in Step 3, and are part of the global streamflow indices and metadata archive (GSIM; Do et al., 2018a), which provides the catchment boundaries needed to compute hydro-meteorological drivers at the catchment scale for the driver analysis in Step 2.

For each of the 937 catchments, we determine the membership to one of the large European river basins, for example the Rhine or Loire, and derive time series of hydro-meteorological variables at the catchment scale including mean temperature (2 m), precipitation, evapotranspiration, snow-water equivalent (SWE), and snowmelt, from the gridded ERA5-Land data set (ECMWF, 2019; Muñoz-Sabater et al., 2021). ERA5-Land relies on atmospheric forcing from the ERA5 reanalysis (Hersbach et al., 2020) and provides additional land-influenced reanalysis variables at a spatial resolution of 9 km for the period 1981 to present. We use the period 1981–2011 overlapping with the streamflow data set to compute time series of hydro-meteorological variables for the driver analysis.

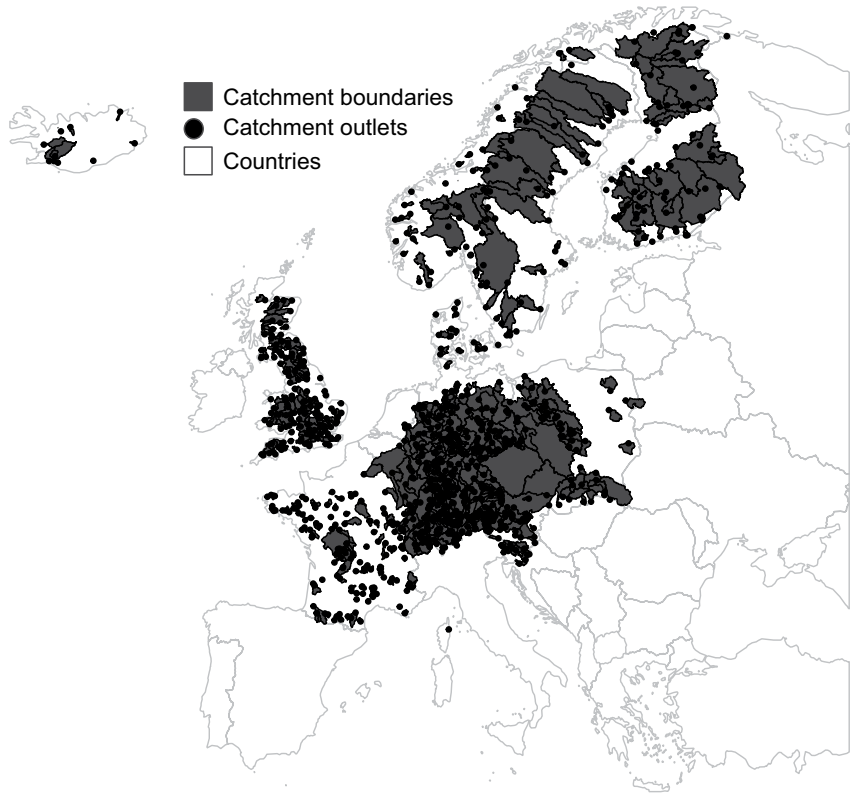

Figure 1. Map of 937 study catchments in Europe: catchment areas (gray) and gauge locations (black).

2.2. Networks

We map complex discharge networks, for normal and extreme conditions using Pearson's linear correlation coefficient over the whole distribution and in the upper and lower tail, respectively. Such complex discharge networks will highlight pairs of catchments with strong spatial dependencies in terms of daily flow or extremes, which do not necessarily need to correspond to pairs of catchments that are physically connected through the river network. Pearson's correlation quantifies the linear association between two variables. Alternatively, one could work with a nonlinear measure of dependence such as Kendall's tau, which does, however, not substantially change network structures (Figure A1 in Supporting Information S1). Upper tail correlation is computed for events jointly exceeding a quantile threshold at a pair of stations as

$$\rho_{\mathcal{A}}(x,y) = \frac{\sum_{i\in\mathcal{A}}(x_i - \overline{x}_{\mathcal{A}})(y_i - \overline{y}_{\mathcal{A}})}{\sqrt{\sum_{i\in\mathcal{A}}(x_i - \overline{x}_{\mathcal{A}})^2}\sqrt{\sum_{i\in\mathcal{A}}(y_i - \overline{y}_{\mathcal{A}})^2}},$$

where $\mathcal{A}$ is the set of points $\{x > q_\alpha(x), y > q_\alpha(y)\}$ with $q_\alpha(x)$ the α quantile of $x_1,\ldots,x_n$ and similarly for y. The averages $\overline{x}_{\mathcal{A}}$ and $\overline{y}_{\mathcal{A}}$ are similarly calculated over the set $\mathcal{A}$.

Lower tail correlation is computed in the same way but using a time series of negative values $-x$ and $-y$. That is, we take the negative daily streamflow time series, extract events where a threshold is jointly exceeded at a pair of stations, and compute the correlation of the threshold exceedances. When computing lower tail correlation, we excluded zeros for quantile computation but included them in correlation computations. We use this measure of upper and lower tail correlation as a measure for extremal dependence instead of classical tail dependence coefficients (Coles, 2001). This new measure allows us to consider the temporal order of events, which are jointly extreme at a pair of sites. This is not the case for classical tail dependence coefficients where events are rearranged according to magnitude (i.e., the largest event in catchment A is matched with the largest event in catchment B, etc.).

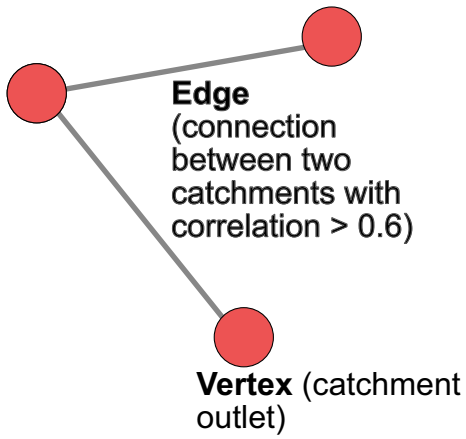

Figure 2. Illustration of simple network graph including vertices and edges.

Upper and lower tail correlation are computed for different exceedance thresholds of 0.8, 0.9, 0.95, and 0.99 to assess the sensitivity of network structure to threshold choice (Figure A2 in Supporting Information S1). A threshold of 0.9 is chosen for the final analysis to focus on extreme events but still guarantee a large enough sample size for reliable correlation computation. The median sample size of the joint upper tail of two catchments is 252 event pairs over all seasons, 62 in winter, 59 in spring, 62 in summer, and 66 in fall. The median sample size of the joint lower tail is 301 event pairs over all seasons, 63 in winter, 68 in spring, 86 in summer, and 63 in fall. To consider travel times, that is the phenomenon that event occurrence might be slightly time shifted in upstream and downstream locations, we consider a potential temporal lag of up to 2 days to compute both overall and tail correlation. We do so by computing overall and tail correlation for each pair of stations for lags of +/− 0 to +/− 2 days and choosing the maximum correlation within this window. We tested windows of 0, 2, and 5 days, and found that correlations/tail dependencies are slightly stronger for a lag of 2 instead of 0 days but do not substantially increase when moving to a window of 5 days (Figure A3 in Supporting Information S1). We therefore focus our analyses on a potential lag of 2 days.

Using the three dependence measures overall, upper tail and lower tail correlation, we map and compare complex networks of normal, high- and low-flows using the R-package *igraph* (Csárdi, 2020). A network graph $G = (V, E)$ is a mathematical structure consisting of a set V of vertices (also called nodes) and a set E of edges (also called links) (Figure 2). The number of vertices and the number of edges are called the order and size of a network graph, respectively (Kolaczyk & Csardi, 2020). More specifically, we construct weighted and undirected networks. Undirected networks are networks where there is no ordering in the vertices defining an edge and weighted networks are networks where edges are attributed different weights with heavier weights indicating stronger connections. Here, we focus on undirected networks because we are interested in identifying pairs of catchments where daily, high- or low-flows are strongly related independent of whether they are located on the same river network or not. We also focus on weighted networks instead of unweighted networks because we are interested in identifying pairs of catchments that are strongly related in terms of daily, high- or low-flows.

The weighted networks are constructed using different types of dependence measures as edge weights, which will inform us about how strongly pairs of catchments are related in terms of the different flow metrics considered. Edges representing correlations below a certain threshold are removed. We tested thresholds of 0.5, 0.6, 0.7, and 0.8 (Figure A4 in Supporting Information S1) and decided to use a threshold of 0.6 for all dependence measures to only include substantial correlations without excluding too many potential relationships. We perform the analysis for the whole year and for each season individually (winter: December–February, spring: March–May, summer: June–August, and fall: September–November) to map seasonal differences in network properties. To compare the network properties of the three different network types and the seasonal networks, we compute a measure of connectedness and a measure of connectedness length for each of the catchments. Connectedness is described by the network degree (also called centrality degree), which defines the number of edges incident on a certain node, that is how many catchments a certain catchment is correlated with. Connectedness length is defined for each existing edge as the Euclidean distance (Martinez & Chavez, 2019) between catchment outlets of flow-dependent but not necessarily physically connected catchments. We compare overall network degree and connectedness length for different network types by comparing their degree and connectedness length distributions and the medians of these distributions. Furthermore, we determine seasonal differences in dependence networks for high and low flows by comparing median degrees and connectedness lengths of networks representing different seasons.

Table 1
Networks Constructed for the Analysis of Streamflow Dependence Drivers

	Variable					
Complex network type	Discharge	T	ET	P	SWE	SMELT
			Lag			
Daily flow	2	0	0	0	0	0
Upper tail	2	0	0	0	0	0
Lower tail	2	0	0	0	0	0

Note. For each driver and network type, the temporal lag considered to compute the dependence measures is indicated.

2.3. Driver Analysis

Next, we want to understand whether spatial dependence in streamflow can be explained by spatial dependence in potential hydro-meteorological drivers. To do so, we look at networks of potential hydro-meteorological drivers including temperature (T), evapotranspiration (ET), precipitation (P), SWE, and snowmelt (SMELT). For each of these potential drivers, we also compute overall, lower- and upper-tail correlation but without considering a temporal lag as travel times do not apply. As for the streamflow networks, we only included connections that exceeded a dependence threshold of 0.6 in network construction. The different network types for the different variables are summarized together with the lag used to compute dependence measures in Table 1.

To identify potential links between driver and flow correlations, we fit multiple linear regression models to the three flow correlation networks and their corresponding driver correlation networks. These models are of the following form:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_p x_{ip} + \epsilon, \quad (1)$$

where the index i refers to the different pairs of catchments, y_i represents the dependent variable, that is the streamflow dependence measure, x_i represents the explanatory variables, that is the dependencies in the hydro-meteorological drivers, β_0 represents the intercept, β_p the slope coefficients for each explanatory variable, and ϵ the model's error term. For example, we model the upper-tail flow correlation network y_i as a function of the upper-tail correlation networks of the five hydro-meteorological drivers considered $x_{i1},\ldots,x_{i5}$.

We compute the variable inflation factor (VIF) to identify potentially redundant predictors. Including all five explanatory variables leads to relatively high VIF values (>10 in the case of the lower tail correlation model) with particularly high values for T and SWE (Table A1 in Supporting Information S1). This result indicates that T and SWE contain somewhat redundant information, which is already included in the other three variables P, ET, and SMELT. We therefore exclude T and SWE from the modeling efforts and fit a multiple regression model with only three explanatory variables, that is P, ET, and SMELT. The significant regression coefficients of this multiple regression model provide some insight on which drivers may be important for explaining spatial flow dependence patterns. We repeat the model fitting for each season to identify seasonally relevant explanatory variables for the three dependence measures.

2.4. Change Analysis

To assess temporal changes in complex discharge networks, we construct networks for two periods, an 'early' period from 1969–1990 and a 'late' period from 1991 to 2011, for the three dependence measures overall, upper- and lower-tail correlation. We determine connectedness for each catchment in the data set by determining its vertex degree for the networks of both periods. Then, we compute relative changes in vertex degrees per catchment to assess changes in spatial flow connectedness as: $(d_l - d_e)/d_e$, where d_e and d_l refer to a catchment's vertex degree in the early and late period, respectively.

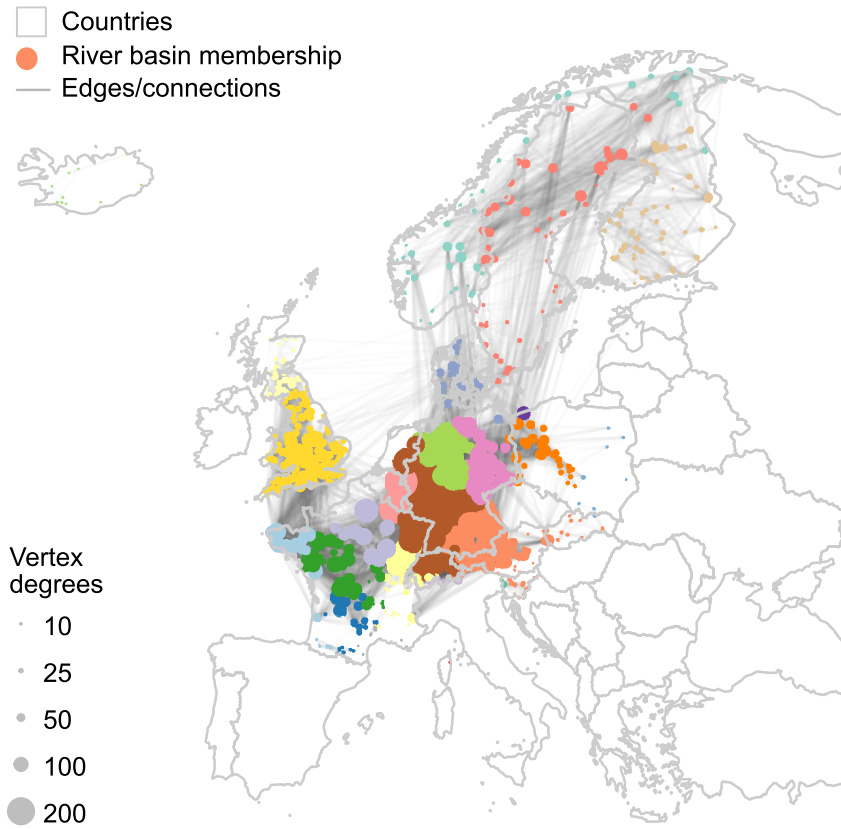

Figure 3. Correlation network of daily streamflow in Central Europe. Connections (edges) are plotted between pairs of catchments with dependence >0.6 and vertices are scaled by vertex degree and colored by basin membership.

3. Results

3.1. Networks of Overall and Tail Correlation

The spatial connections among pairs of stations showing flow correlations >0.6 are particularly evident in Central and Western Europe (Germany and France) (Figure 3) and weaker in Great Britain and Scandinavia. Connections are not limited to connections over land but also comprise connections between France and Great Britain or Central Europe and Scandinavia. Complex networks change as we turn to high and low flows (Figure 4). Overall, high-flow connections are less abundant than connections in daily streamflow, that is vertex degrees are generally lower in the high-than in the daily network. However, there are some long connections between Central Europe and Western and Northern Europe. Low-flow connections are particularly pronounced in Scandinavia, Great Britain, and the Alps where we see higher vertex degrees and seem to span long spatial scales with links reaching from Scandinavia and Iceland to Central Europe.

These visual differences in network properties are confirmed when we compare the normal, high-, and low-flow networks in terms of the distribution of vertex degree as a measure of connectedness and the distribution of edge length as a measure of connectedness length (Figure 5). Among the three flow types, daily flows are most strongly connected as indicated by the wide distribution and the high degree median (Figure 5a). That is, individual catchments show high streamflow correlations with many other catchments, while these connections are shortest as indicated by the low median connectedness length (Figure 5b), that is the dependent catchments are located close to each other. High flows are comparably weakly connected but over slightly longer distances while low flows are slightly more strongly connected and over the longest distances.

Connectedness (vertex degrees) and connectedness lengths (edge lengths) not only differ by flow type but also by season (Figure 6). Connectedness is strongest in spring and weakest in summer for both high and

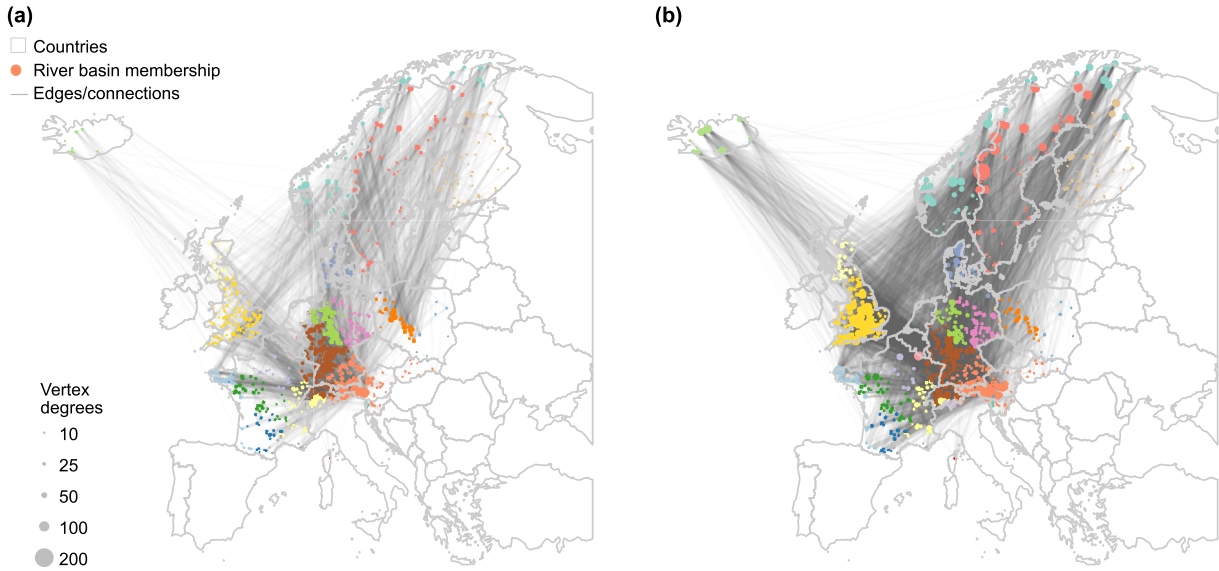

Figure 4. Comparison of dependence networks in Central Europe for (a) high flows (upper tail correlation) and (b) low flows (lower tail correlation). Connections (edges) are drawn between pairs of catchments with dependence >0.6 and vertices are scaled by vertex degree and colored by basin membership.

low flows. In spring, low flows in a particular catchment are, on average, related to low flows in roughly 80 other catchments while they are only related to roughly 40 catchments in summer. In contrast, high flows in a particular catchment are, on average, related to 40 catchments in spring and 20 in summer. Connectedness length is longest for low flows in winter with, on average, 1,000 km between correlated catchments, and for high flows in summer with, on average, 800 km between correlated catchments.

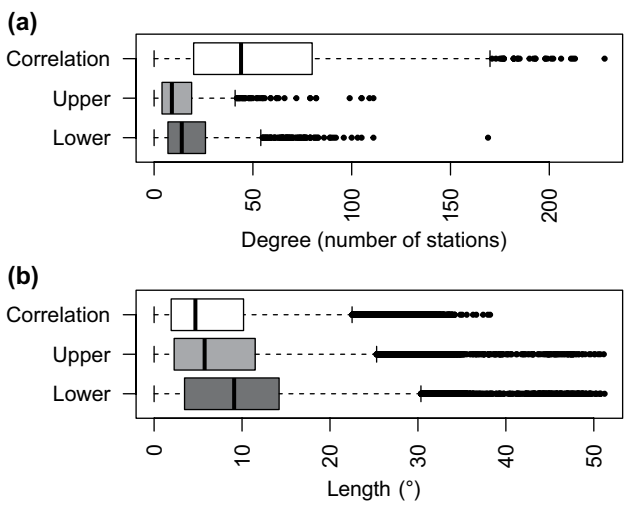

Figure 5. Comparison of dependence networks with respect to the distribution of (a) connectedness (degree) and (b) connectedness length (Euclidean distance between related catchment outlets) across catchments displayed by boxplots for different phenomena: daily streamflow (correlation), high flows (upper tail correlation), and low flows (lower tail correlation). Vertical lines indicate the medians over all catchments. 1° corresponds to roughly 111 km.

3.2. Driver Importance

Pairwise flow dependencies for the three flow characteristics (daily, high, and low flow) can be explained by dependencies in hydro-meteorological drivers (Figure 7) to a substantial degree (adjusted R^2 of fitted seasonal regression models ranges from 0.4 to 0.8). The importance of the different potential drivers P, ET, and SMELT varies by dependence type and season. Predictive power is generally highest for low-flow dependencies and lowest for overall dependencies. Daily streamflow correlations are best explained by precipitation correlations in most seasons except in spring and fall where snowmelt and ET correlations also play an important role, respectively (Figure 7a). That is, streamflow correlations between pairs of stations are highest for pairs where precipitation is either jointly high or jointly low, in spring for pairs with correlated snowmelt patterns, and in fall for pairs with correlated ET patterns. High-flow dependencies are also strongly related to precipitation and ET dependencies (Figure 7b). That is, high-flows tend to co-occur in pairs of catchments that show strong precipitation correlations and/or high ET correlations. High-flow dependencies are particularly positively related to precipitation in winter and to ET in summer and fall. That is, in winter, catchments with either jointly low precipitation or jointly high precipitation show either jointly low high-flows or jointly high high-flows, respectively. Similarly, in summer, jointly dry soils because of high ET or jointly wet soils because of low ET lead to high high-flow or low high-flow connectedness. In addition, snowmelt plays a supporting role from winter to summer, that is high flows tend to be spatially dependent if a pair of stations shows a strong relationship in snowmelt. Low-flow dependencies show different

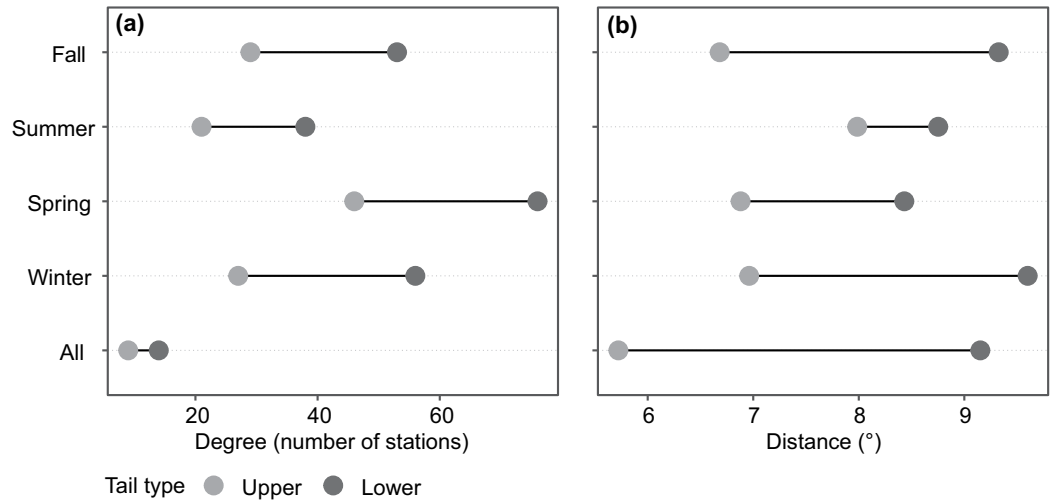

Figure 6. Comparison of seasonal dependence networks for high and low flows with respect to median (a) strength of connectedness (degree) and (b) length of connectedness (distance, 1° corresponds to roughly 111 km) across catchments.

driver dependence relationships than daily and high flows (Figure 7c). ET dependencies are relatively important to explain low-flow dependencies in spring and summer while snowmelt dependencies are relatively important to explain low-flow dependencies in summer and fall. That is, if a pair of stations shows jointly high/low ET values or jointly high/low snowmelt contributions, it is also likely that they jointly experience high/low low flows. Precipitation dependencies are comparably less relevant and mostly important in fall.

3.3. Connectedness Changes

Connectedness in terms of vertex degree has changed over time when comparing a past time period (1969–1990) to a more recent period (1991–2011) for daily, high, and low flows (Figure 8). Daily connectedness changes show spatial patterns with increases in Great Britain and along the North Sea coast and decreases in the greater Alpine region and northern Scandinavia (8a). Relative changes in high and low flow connectedness (i.e., vertex degrees of different catchments) are stronger than changes in daily flow connectedness but show less clear spatial patterns. High-flow connection strength has increased for most catchments with exceptions in southern Great Britain and southern Scandinavia. In contrast, low-flow connectedness has decreased in Great Britain, the northern part of Central Europe and southern Scandinavia but increased in the Alps and parts of Scandinavia.

4. Discussion

Our results highlight that overall and extreme complex discharge networks have different properties. Daily flows seem to be particularly related within regions with similar flow seasonality, which explains the abundant but rather short connections. For example, flows within Scandinavia are related because of relatively homogeneous snow-dominated regimes but relatively few connections exist to Central Europe where regimes are more rainfall dominated (Brunner & Tallaksen, 2019). In contrast, connections become less abundant but longer as we move to networks of extremes. Connections span from Scandinavia and Great Britain to Central Europe for high and particularly low flows. These connections are likely rather related to particular synoptic weather patterns than overall flow seasonality. For example, a frontal storm may lead to widespread precipitation which in combination with large-scale wet soil conditions may lead to widespread flooding. Similarly, a stable high-pressure system over Scandinavia and Central Europe (as e.g., observed during the 2018 drought; Bakke et al., 2020) can lead to synchronized soil drying and lack of precipitation. The drought connections tend to be longer than flood connections because the drought-triggering weather

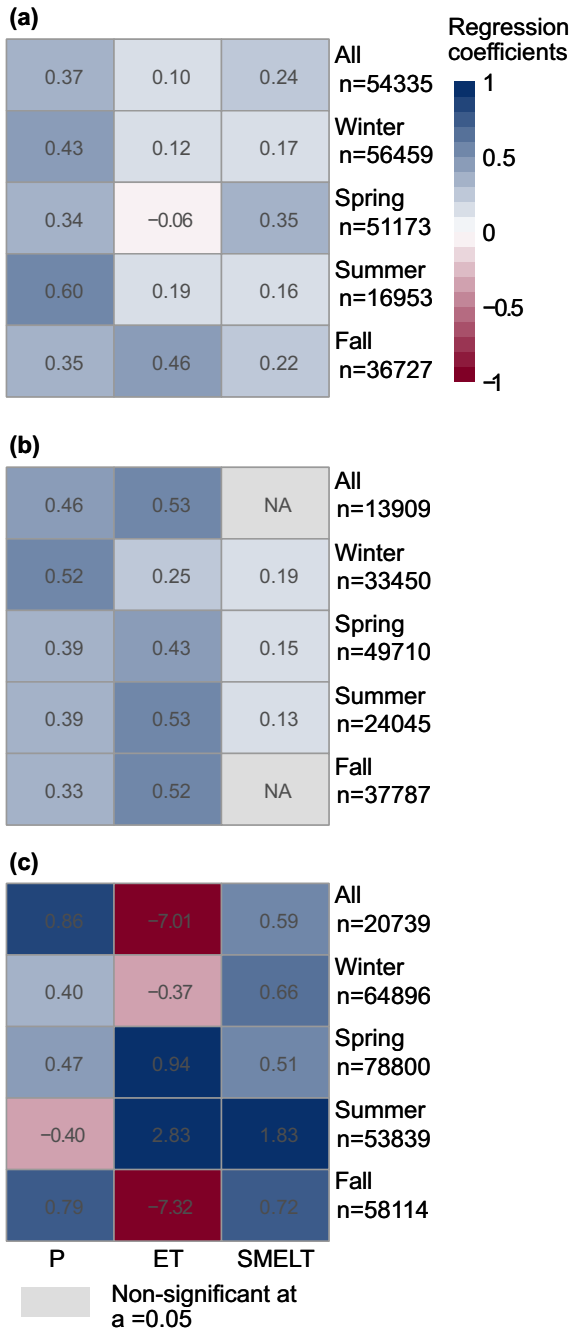

Figure 7. Predictor importance (regression coefficient) of precipitation (P), evapotranspiration (ET), and snowmelt (SMELT) dependencies in predicting positive daily streamflow dependence in the (a) whole distribution, (b) upper tail, and (c) lower tail. Non significant predictors (at $\alpha = 0.05$) are displayed in gray. The sample size n for each of the models is given on the right.

conditions are active at larger spatial scales than flood-triggering conditions. How exactly dominant spatial connections are linked to weather patterns remains to be investigated.

Our findings show that normal, high- and low-flow networks not only differ in their characteristics but also their hydro-meteorological drivers. While precipitation seems to be the most relevant driver for normal flow networks, ET and SMELT have important roles in governing high- and low-flow networks. This result corroborates findings by Brunner, Gilleland, et al. (2020), who have shown that spatial flood dependencies are, besides precipitation, substantially modulated by land-surface processes such as soil moisture and snowmelt, and by Brunner, Swain, et al. (2021), who highlighted the importance of temperature-related factors such as soil moisture and snowmelt in determining drought extents. In the case of high flows, ET is important in summer and fall because widespread dry soils might prevent widespread flood occurrence even in the case of widespread precipitation because a lot of water will be able to infiltrate instead of building direct runoff. Similarly, widespread wet soils might also lead to joint high flows in multiple catchments. In addition, spatially extensive snowmelt contributions may synchronize scattered extreme precipitation events, which otherwise might have led to localized flooding. In the case of low flows, spatial ET dependencies are also strongly related to flow dependencies because high ET rates can lead to dry soils and therefore reduced recharge at large spatial scales. In addition, missing snowmelt contributions in summer can lead to jointly dry conditions in several catchments at once.

In addition to network properties and drivers, different flow types differ in their network changes. These changes are strongest for high and low flow networks which mostly show increases and decreases in connectedness (vertex degrees), respectively. These changes are likely related to changes in the connectedness in individual hydro-meteorological drivers and their interplay. It remains to be investigated how future changes in driver connectedness as a potential result of climate change might translate into future changes in high- and low-flow connectedness and related to these to changes in flood and drought risk. Such an assessment requires setting up or using a chain of climate and hydrological models. However, uncertainties in the representation of spatial dependencies can be introduced at several points along this modeling chain.

The complex network construction depends on a few methodological choices including the tail threshold, time lag or network threshold, which affect network properties such as degree or length. For example, the choice of the tail threshold, that is the threshold above which flows are considered extreme, affects the number of edges (connections) drawn and therefore the node degrees and the weights of the edges (Figure A2 in Supporting Information S1). Similarly, correlation strength is influenced by the maximum time lag chosen to determine correlation (Figure A3 in Supporting Information S1) and the number of edges drawn by the cutoff threshold used to define when an edge is drawn or not drawn between a pair of catchments (Figure A4 in Supporting Information S1). While these methodological choices change the strength of the connection pattern, they do not change the overall spatial pattern indicating which regions are connected to which other regions. In addition to these methods choices, network properties may to some degree also depend on the spatial distribution of the catchments in the data set and on the strength of nestedness. Streamflow gauging stations are distributed irregularly

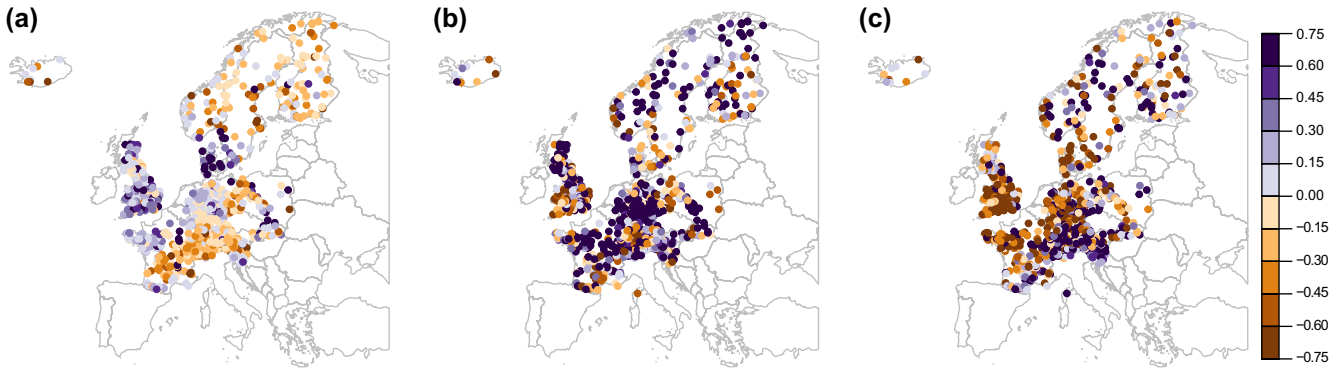

Figure 8. Relative changes (−) in connectedness (vertex degrees) from an early (1969–1990) to a late period (1991–2011) for (a) daily streamflow, (b) high flows, and (c) low flows. Purple and brown colors indicate increases and decreases in connectedness, respectively. The darker the color, the stronger is the change in connectedness.

in space, which cannot be avoided. Catchments with denser neighborhoods of catchments may have larger node degrees than those in less densely gauged regions. However, the observation that the catchments in Germany, where station density is highest, is not necessarily the region with the highest node degrees (see e.g., Figure 4) does not support this hypothesis. Instead, this observation suggests that other factors than spatial proximity are important to explain connectedness. Furthermore, catchments which are nested within a larger catchment may show larger node degrees than nonnested catchments as flow connections of pairs of catchments that are part of the same river network might be stronger than connections of pairs of catchments that are not connected through their river network. One potential way around the problem of varying station density would be to use a gridded data set of streamflow simulations. However, such a grid-based approach would require the use of a distributed hydrological model, which comes at the cost of introducing a few sources of uncertainty along the modeling chain (Clark et al., 2016). Such uncertainties have been shown to be particularly pronounced for high and low flows (Kempen et al., 2021; Pokorny et al., 2021). In addition, it has been shown that modeling spatial dependencies in hydrologic extremes is very challenging and often unsatisfactory using commonly used model calibration strategies (Brunner, Melsen, et al., 2021; Prudhomme et al., 2011). Therefore, regularly spaced simulated data might not be more suitable to study spatial flow connections than the irregularly spaced observations used in this study.

This work mainly focused on understanding the properties of complex discharge networks and their hydro-meteorological drivers. Future work could try to establish a relationship between large scale spatial coherence and synoptic weather patterns or climate oscillation indices, such as the Atlantic Multidecadal Oscillation or the Mediterranean Oscillation Index, among others (Han et al., 2020). The existence of such relationships may help to improve the prediction of spatial extreme events such as widespread floods and drought and to project how flow connectedness may change in future.

The network approach based on upper and lower tail correlation can be generalized to other variables and regions. It could for example be applied to visualize spatial connections of heatwaves or to map spatial dependencies in groundwater levels. Applied to a larger region, it could also be used to detect potential teleconnection signals.

5. Conclusions

We propose upper- and lower-tail correlation as a measure for extremal dependence to map complex networks of high- and low-flows and assess their differences, potential drivers, and past changes. Our analyses show that complex networks of hydrologic extremes substantially differ from complex networks of daily streamflow, low flows are connected (in terms of correlations exceeding 0.6) more strongly and over longer distances than high flows, and connections are strongest in spring and weakest in summer for both high and low flows. In addition, we show that hydro-meteorological driver importance varies by dependence type and season. Evapotranspiration dependence plays a larger role overall in explaining high-flow dependence than precipitation dependence except for winter where precipitation is the main driver. Snowmelt

dependence plays a supporting role in explaining high-flow dependence from winter through summer. Low-flow dependence is partly governed by snowmelt in all seasons, by ET in spring and summer, and by precipitation in fall. Our change assessment shows that changes in high- and low-flow connectedness (i.e., the number of catchments a catchment shows high correlations with) are stronger than for daily flow in general and that the direction of change varies spatially. High-flow connectedness generally increased while low-flow connectedness decreased in the North Sea region but increased in the Alps and parts of Scandinavia. We conclude that spatial flow dependencies are considerable for both high and particularly low flows as a result of a combination of spatially correlated hydro-meteorological drivers whose importance varies by extreme type, season, and in time. These temporal changes suggest that (climate) change assessments should in addition to looking at changes in extreme flows in individual catchments also look at changes in the spatial dependence of extreme flows and their drivers. Such consideration will enable more reliable regional drought and flood risk assessments as the impacts of extreme events affecting multiple locations at once likely exceed impacts of local events.

Data Availability Statement

The raw data used for the analysis are available through the Global Runoff Data Center (streamflow, The Global Runoff Data Centre 56068 Koblenz Germany, 2019), the global streamflow indices and metadata archive (catchment boundaries, Do et al., 2018b), and the Copernicus climate data store (ERA5-land hydro-climatic variables, ECMWF (2019)). The processed data set including a shapefile of the 937 catchments, their streamflow time series from 1969-2011 and the corresponding time series of the hydro-meteorological variables temperature, evapotranspiration, precipitation, SWE, and snowmelt can be downloaded from the HydroShare repository: https://www.hydroshare.org/resource/7392b4c3471a4f5e81da099aed230b80/.

Acknowledgments
This work was supported by the Swiss National Science Foundation via a PostDoc.Mobility grant (Number: P400P2_183844, granted to MIB). Open access funding enabled and organized by Projekt DEAL.

References

Andreadis, K. M., Clark, E. A., Wood, A. W., Hamlet, A. F., & Lettenmaier, D. P. (2005). Twentieth-century drought in the conterminous United States. *Journal of Hydrometeorology*, *6*(6), 985–1001. https://doi.org/10.1175/JHM450.1

Bakke, S., Ionita, M., & Tallaksen, L. (2020). The 2018 northern European hydrological drought and its drivers in a historical perspective. *Hydrology and Earth System Sciences*, *24*, 5621–5653. https://doi.org/10.5194/hess-24-5621-2020

Boers, N., Bookhagen, B., Barbosa, H. M., Marwan, N., Kurths, J., & Marengo, J. A. (2014). Prediction of extreme floods in the eastern Central Andes based on a complex networks approach. *Nature Communications*, *5*, 1–7. https://doi.org/10.1038/ncomms6199

Brunner, M. I., Gilleland, E., Wood, A., Swain, D. L., & Clark, M. (2020). Spatial dependence of floods shaped by spatiotemporal variations in meteorological and land-surface processes. *Geophysical Research Letters*, *47*, e2020GL088000. https://doi.org/10.1029/2020GL088000

Brunner, M. I., Liechti, K., & Zappa, M. (2019). Extremeness of recent drought events in Switzerland: Dependence on variable and return period choice. *Natural Hazards and Earth System Sciences*, *19*(10), 2311–2323. https://doi.org/10.5194/nhess-19-2311-2019

Brunner, M. I., Melsen, L. A., Wood, A. W., Rakovec, O., Mizukami, N., Knoben, W. J. M., & Clark, M. P. (2021). Flood spatial coherence, triggers and performance in hydrological simulations: Large-sample evaluation of four streamflow-calibrated models. *Hydrology and Earth System Sciences*, *25*, 105–119. https://doi.org/10.5194/hess-25-105-2021

Brunner, M. I., Papalexiou, S., Clark, M. P., & Gilleland, E. (2020). How probable is widespread flooding in the United States? *Water Resources Research*, *56*, e2020WR028096. https://doi.org/10.1029/2020WR028096

Brunner, M. I., Swain, D. L., Gilleland, E., & Wood, A. (2021). Increasing importance of temperature as a driver of streamflow drought spatial extent. *Environmental Research Letters*, *16*, 024038. https://doi.org/10.1088/1748-9326/abd2f0

Brunner, M. I., & Tallaksen, L. M. (2019). Proneness of European catchments to multiyear streamflow droughts. *Water Resources Research*, *55*, 8881–8894. https://doi.org/10.1029/2019WR025903

Clark, M. P., Wilby, R. L., Gutmann, E. D., Vano, J. A., Gangopadhyay, S., Wood, A. W., et al. (2016). Characterizing uncertainty of the hydrologic impacts of climate change. *Current Climate Change Reports*, *2*(2), 55–64. https://doi.org/10.1007/s40641-016-0034-x

Coles, S. (2001). *An introduction to statistical modeling of extreme values*. London. Springer London. https://doi.org/10.1007/978-1-4471-3675-0

Csárdi, G. (2020). *Igraph: Network analysis and visualization*. CRAN. Retrieved from https://cran.r-project.org/web/packages/igraph/index.html

Do, H. X., Gudmundsson, L., Leonard, M., & Westra, S. (2018a). The Global Streamflow Indices and Metadata Archive (GSIM)-Part 1: The production of a daily streamflow archive and metadata. *Earth System Science Data*, *10*(2), 765–785. https://doi.org/10.5194/essd-10-765-2018

Do, H. X., Gudmundsson, L., Leonard, M., & Westra, S. (2018b). *The global streamflow indices and metadata archive - Part 1: Station catalog and catchment boundary*. PANGAEA. https://doi.org/10.1594/PANGAEA.887477

Donges, J. F., Zou, Y., Marwan, N., & Kurths, J. (2009). Complex networks in climate dynamics: Comparing linear and nonlinear network construction methods. *The European Physical Journal - Special Topics*, *174*(1), 157–179. https://doi.org/10.1140/epjst/e2009-01098-2

ECMWF. (2019). ERA5-Land hourly data from 1981 to present. *Copernicus*. https://doi.org/10.24381/cds.e2161bac

Fang, K., Sivakumar, B., & Woldemeskel, F. M. (2017). Complex networks, community structure, and catchment classification in a large-scale river basin. *Journal of Hydrology*, *545*, 478–493. https://doi.org/10.1016/j.jhydrol.2016.11.056

Guo, H., Ramos, A. M., Macau, E. E., Zou, Y., & Guan, S. (2017). Constructing regional climate networks in the Amazonia during recent drought events. *PloS One*, *12*(10), 1–16. https://doi.org/10.1371/journal.pone.0186145

Halverson, M. J., & Fleming, S. W. (2015). Complex network theory, streamflow, and hydrometric monitoring system design. *Hydrology and Earth System Sciences*, *19*(7), 3301–3318. https://doi.org/10.5194/hess-19-3301-2015

Han, X., Mehrotra, R., & Sharma, A. (2020). Measuring the spatial connectivity of extreme rainfall. *Journal of Hydrology*, *590*, 125510. https://doi.org/10.1016/j.jhydrol.2020.125510

Han, X., Sivakumar, B., Woldmeskel, F. M., & Guerra de Aguilar, M. (2018). Temporal dynamics of streamflow: Application of complex networks. *Geoscience Letters*, *5*(1). https://doi.org/10.1186/s40562-018-0109-8

Heffernan, J. E., & Tawn, J. (2004). A conditional approach to modelling multivariate extreme values. *Journal of the Royal Statistical Society - Series B: Statistical Methodology*, *66*(3), 497–546. https://doi.org/10.1111/j.1467-9868.2004.02050.x

Henriques, A. G., & Santos, M. J. J. (1999). Regional drought distribution model. *Physics and Chemistry of the Earth - Part B: Hydrology, Oceans and Atmosphere*, *24*(1–2), 19–22. https://doi.org/10.1016/S1464-1909(98)00005-7

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., et al. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*, 1999–2049. https://doi.org/10.1002/qj.3803

Hisdal, H., & Tallaksen, L. M. (2003). Estimation of regional meteorological and hydrological drought characteristics: A case study for Denmark. *Journal of Hydrology*, *281*(3), 230–247. https://doi.org/10.1016/S0022-1694(03)00233-6

Jha, S. K., & Sivakumar, B. (2017). Complex networks for rainfall modeling: Spatial connections, temporal scale, and network size. *Journal of Hydrology*, *554*, 482–489. https://doi.org/10.1016/j.jhydrol.2017.09.030

Jha, S. K., Zhao, H., Woldemeskel, F. M., & Sivakumar, B. (2015). Network theory and spatial rainfall connections: An interpretation. *Journal of Hydrology*, *527*, 13–19. https://doi.org/10.1016/j.jhydrol.2015.04.035

Keef, C., Tawn, J. A., & Lamb, R. (2013). Estimating the probability of widespread flood events. *Environmetrics*, *24*(1), 13–21. https://doi.org/10.1002/env.2190

Kempen, G. V., Wiel, K. V. D., & Melsen, L. A. (2021). The impact of hydrological model structure on the simulation of extreme runoff events. *Natural Hazards and Earth System Sciences*, *21*, 961–976. https://doi.org/10.5194/nhess-21-961-2021

Kolaczyk, E. D., & Csardi, G. (2020). *Statistical analysis of network data with R*. In R. Gentleman, K. Hornik, & G. Parmigiani (Eds.), (2nd ed.). Cham. Springer. https://doi.org/10.18637/jss.v066.b01

Konapala, G., & Mishra, A. (2017). Review of complex networks application in hydroclimatic extremes with an implementation to characterize spatio-temporal drought propagation in continental USA. *Journal of Hydrology*, *555*, 600–620. https://doi.org/10.1016/j.jhydrol.2017.10.033

Luke, D. A. (2015). *A user's guide to network analysis in R*. In G. P. Robert Gentleman Kurt Hornik (Ed.), (1st ed.). Cham. Springer. https://doi.org/10.18637/jss.v072.b03

Malik, N., Bookhagen, B., Marwan, N., & Kurths, J. (2012). Analysis of spatial and temporal extreme monsoonal rainfall over South Asia using complex networks. *Climate Dynamics*, *39*(3), 971–987. https://doi.org/10.1007/s00382-011-1156-4

Martinez, J. H., & Chavez, M. (2019). Comparing complex networks: In defence of the simple. *New Journal of Physics*, *21*(1), 13033. https://doi.org/10.1088/1367-2630/ab0065

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., et al. (2021). ERA5-Land: A state-of-the-art global reanalysis dataset for land applications. *Earth System Science Data*, *13*, 4349–4383. https://doi.org/10.5194/essd-13-4349-2021

Naufan, I., Sivakumar, B., Woldemeskel, F. M., Raghavan, S. V., Vu, M. T., & Liong, S.-Y. (2018). Spatial connections in regional climate model rainfall outputs at different temporal scales: Application of network theory. *Journal of Hydrology*, *556*, 1232–1243. https://doi.org/10.1016/j.jhydrol.2017.05.029

Oesting, M., & Stein, A. (2018). Spatial modeling of drought events using max-stable processes. *Stochastic Environmental Research and Risk Assessment*, *32*(1), 63–81. https://doi.org/10.1007/s00477-017-1406-z

Ozturk, U., Marwan, N., Korup, O., Saito, H., Agarwal, A., Grossman, M. J., et al. (2018). Complex networks for tracking extreme rainfall during typhoons. *Chaos*, *28*, 075301. https://doi.org/10.1063/1.5004480

Pal, S., Lee, T. R., & Clark, N. (2020). The 2019 Mississippi and Missouri River flooding and its impact on atmospheric boundary layer dynamics. *Geophysical Research Letters*, *47*(6). https://doi.org/10.1029/2019GL086933

Pokorny, S., Stadnyk, T. A., Ali, G., Lilhare, R., Dery, S. J., & Koenig, K. (2021). Cumulative effects of uncertainty on simulated streamflow in a hydrologic modeling environment. *Elementa. Science of the Anthropocene*, *9*, 1–20. https://doi.org/10.1525/elementa.431

Prudhomme, C., Parry, S., Hannaford, J., Clark, D. B., Hagemann, S., & Voss, F. (2011). How well do large-scale models reproduce regional hydrological extremes: In Europe? *Journal of Hydrometeorology*, *12*(6), 1181–1204. https://doi.org/10.1175/2011JHM1387.1

Quinn, N., Bates, P. D., Neal, J., Smith, A., Wing, O., Sampson, C., et al. (2019). The spatial dependence of flood hazard and risk in the United States. *Water Resources Research*, *55*, 1890–1911. https://doi.org/10.1029/2018WR024205

Rossi, G., Benedini, M., Tsakiris, G., & Giakoumakis, S. (1992). On regional drought estimation and analysis. *Water Resources Management*, *6*(4), 249–277. https://doi.org/10.1007/BF00872280

Rudd, A. C., Kay, A. L., & Bell, V. A. (2019). National-scale analysis of future river flow and soil moisture droughts: Potential changes in drought characteristics. *Climatic Change*, *156*(3), 323–340. https://doi.org/10.1007/s10584-019-02528-0

Sarker, S., Veremyev, A., Boginski, V., & Singh, A. (2019). Critical nodes in river networks. *Scientific Reports*, *9*(1), 11178. https://doi.org/10.1038/s41598-019-47292-4

Scarsoglio, S., Laio, F., & Ridolfi, L. (2013). Climate dynamics: A network-based approach for the analysis of global precipitation. *PloS One*, *8*(8), e71129. https://doi.org/10.1371/journal.pone.0071129

Sivakumar, B., & Woldemeskel, F. M. (2014). Complex networks for streamflow dynamics. *Hydrology and Earth System Sciences*, *18*(11), 4565–4578. https://doi.org/10.5194/hess-18-4565-2014

Sivakumar, B., & Woldemeskel, F. M. (2015). A network-based analysis of spatial rainfall connections. *Environmental Modelling & Software*, *69*, 55–62. https://doi.org/10.1016/j.envsoft.2015.02.020

Stephan, R., Erfurt, M., Terzi, S., Žun, M., Kristan, B., Haslinger, K., & Stahl, K. (2021). An inventory of Alpine drought impact reports to explore past droughts in a mountain region. *Natural Hazards and Earth System Sciences*, *21*(8), 2485–2501. https://doi.org/10.5194/nhess-21-2485-2021

The Global Runoff Data Centre 56068 Koblenz Germany. (2019). *Global runoff data centre*. Retrieved from https://www.bafg.de/GRDC/EN/02_srvcs/21_tmsrs/riverdischarge_node.html

Thieken, A. H., Apel, H., & Merz, B. (2015). Assessing the probability of large-scale flood loss events: A case study for the river Rhine, Germany. *Journal of Flood Risk Management*, *8*(3), 247–262. https://doi.org/10.1111/jfr3.12091

Tiwari, S., Jha, S. K., & Singh, A. (2020). Quantification of node importance in rain gauge network: Influence of temporal resolution and rain gauge density. *Scientific Reports*, *10*(1), 1–17. https://doi.org/10.1038/s41598-020-66363-5

Tsonis, A. A., & Roebber, P. J. (2004). The architecture of the climate network. *Physica A: Statistical Mechanics and its Applications*, *333*(1–4), 497–504. https://doi.org/10.1016/j.physa.2003.10.045

Yamasaki, K., Gozolchiani, A., & Havlin, S. (2008). Climate networks around the globe are significantly affected by El Niño. *Physical Review Letters*, *100*(22), 1–4. https://doi.org/10.1103/PhysRevLett.100.228501