

## Low-frequency variability of European runoff

L. Gudmundsson<sup>1</sup>, L. M. Tallaksen<sup>1</sup>, K. Stahl<sup>1,2</sup>, and A. K. Fleig<sup>1,\*</sup>

<sup>1</sup>Department of Geosciences, University of Oslo, Oslo, Norway

<sup>2</sup>Institute of Hydrology, University of Freiburg, Freiburg, Germany

\* current address: Norwegian Water Resources and Energy Directorate, Oslo, Norway

Received: 1 February 2011 – Published in Hydrol. Earth Syst. Sci. Discuss.: 7 February 2011

Revised: 2 September 2011 – Accepted: 6 September 2011 – Published: 13 September 2011

**Abstract.** This study investigates the low-frequency components of observed monthly river flow from a large number of small catchments in Europe. The low-frequency components, defined as fluctuations on time scales longer than one year, were analysed both with respect to their dominant space-time patterns as well as their contribution to the variance of monthly runoff.

The analysis of observed streamflow and corresponding time series of precipitation and temperature, showed that the fraction of low-frequency variance of runoff is on average larger than, and not correlated to, the fraction of low-frequency variance of precipitation and temperature. However, it is correlated with mean climatic conditions and is on average lowest in catchments with significant influence of snow. Furthermore, it increases (decreases) under drier (wetter) conditions – indicating that the average degree of catchment wetness may be a primary control of low-frequency runoff dynamics. The fraction of low-frequency variance of runoff is consistently lower in responsive catchments, with a high variability of daily runoff.

The dominant space-time patterns of low-frequency runoff in Europe, identified using nonlinear dimension reduction, revealed that low-frequency runoff can be described with three modes, explaining together 80.6% of the variance. The dominant mode has opposing centres of simultaneous variations in northern and southern Europe. The secondary mode features a west-east pattern and the third mode has its centre of influence in central Europe. All modes are closely related to the space-time patterns extracted from time series of precipitation and temperature.

In summary, it is shown that the dynamics of low-frequency runoff follows well known continental-scale atmospheric features, whereas the proportion of variance attributed to low-frequency fluctuations is controlled by catchment processes and varies with mean climatic conditions. The results may have implications for interpreting the impact of changes in temperature and precipitation on river-flow dynamics.

### 1 Introduction

Catchment runoff depends on atmospheric water input (precipitation) and loss (evapotranspiration) as well as on catchment processes, which determine how atmospheric fluctuations are translated into runoff. On short time scales (days, months) a multitude of processes is known to influence runoff generation. In a changing climate, fluctuations on longer time scales (years, decades) gain increasing importance. On these time scales runoff variability is known to have systematic continental scale space-time patterns (e.g. Lins, 1997; Gudmundsson et al., 2011) that are related to large-scale atmospheric drivers (e.g. Shorthouse and Arnell, 1997; Barlow et al., 2001; Tootle and Piechota, 2006). However, the variance of annual runoff varies largely among catchments (for a global analysis see McMahan et al., 2007) and is on average larger than the variance induced by annual precipitation and evapotranspiration alone (as demonstrated for US streamflow by Sankarasubramanian and Vogel, 2002). Several studies have shown that the difference between the variance of annual runoff and of precipitation is dependent on the climatic water balance, expressed by the aridity index



Correspondence to: L. Gudmundsson  
(lukas.gudmundsson@geo.uio.no)

(e.g. Dooge, 1992; Koster and Suarez, 1999; Sankarasubramanian et al., 2000; Milly and Dunne, 2002).

It is also of interest to look at the relative strength of the variability on long time scales as compared to sub-annual variability. In the following, low-frequency variability is defined as variability on time-scales longer than one year. Studies that are concerned with low-frequency variability of monthly runoff (e.g. Shun and Duffy, 1999; Hanson et al., 2004; Kumar and Duffy, 2009), focus primarily on its temporal evolution and less emphasis is given to its strength. When the strength of the low-frequency variability of runoff (or defined sub-signals e.g. quasiperiodic oscillations) is reported, it is often measured as the fraction of the total variance of monthly runoff time series, which can be attributed to fluctuations on time-scales longer than one year. Generally, it is found that the fraction of low-frequency variance of runoff is usually larger than the fraction of low-frequency variance in the forcing.

A high fraction of low-frequency variance in time series arising from hydrological systems is closely related to the presence of “long-term memory” or persistence, which is often quantified by the Hurst coefficient (e.g. Mandelbrot and Wallis, 1968, 1969; Klemeš, 1974; Vogel et al., 1998; Koutsoyiannis, 2002). However, studies addressing persistence in annual runoff series have several common features that distinguish them from those analysing low-frequency variance of monthly time series. For example, it is common only to consider annual time series (e.g. Tallaksen et al., 1997; Vogel et al., 1998; Koutsoyiannis, 2002, 2003, 2010), and the results are thus not directly comparable to results obtained from analyses of monthly time series. Only a few “Hurst” studies consider daily or monthly time series (e.g. Montanari et al., 1997; Mudelsee, 2007), in which case the series are de-seasonalised prior to further analysis. However, the seasonal cycle is one of the most important phenomena in hydrology and contributes significantly to the variance of monthly runoff. Therefore, the present paper does not intend to study the relevance of the Hurst phenomenon of long-range dependence to explain its findings.

Alongside stochastic models (e.g. Montanari et al., 1997; Koutsoyiannis, 2002), various physical mechanisms have been proposed to explain the relatively high fraction of low-frequency variance (i.e. persistence) in hydrological time series. Amongst these are climate instationarities (Potter, 1976) – possibly due to land-atmosphere interactions (Bierkens and van den Hurk, 2007), storage mechanisms (Klemeš, 1974), groundwater upwelling (Shun and Duffy, 1999), and channel routing (Mudelsee, 2007).

This study aims at providing further insight into the low-frequency variability of monthly runoff, both with respect to its spatio-temporal patterns as well as with respect to its strength, measured by the fraction of low-frequency variance. The formulation of the working hypothesis, the design of the data analysis as well as the interpretation of the results, are based on the assumptions that hydrological systems

can be sufficiently characterised by a simple water-balance model

$$\frac{dS}{dt} = P - E - Q \quad (1)$$

where  $dS/dt$  denotes changes in the terrestrial water storage  $S$  (e.g. snow, soil moisture, groundwater, lakes),  $P$  precipitation,  $E$  actual evapotranspiration and  $Q$  runoff. Evapotranspiration is usually not directly observed, but estimated as a function of the atmospheric water demand (i.e. potential evapotranspiration) and water availability (i.e. soil moisture). Here, temperature ( $T$ ) will be used as a surrogate for potential evapotranspiration driving  $E$ , which is a common approach (e.g. Kingston et al., 2009). Runoff is assumed to be a function of storage

$$Q = h(S) \quad (2)$$

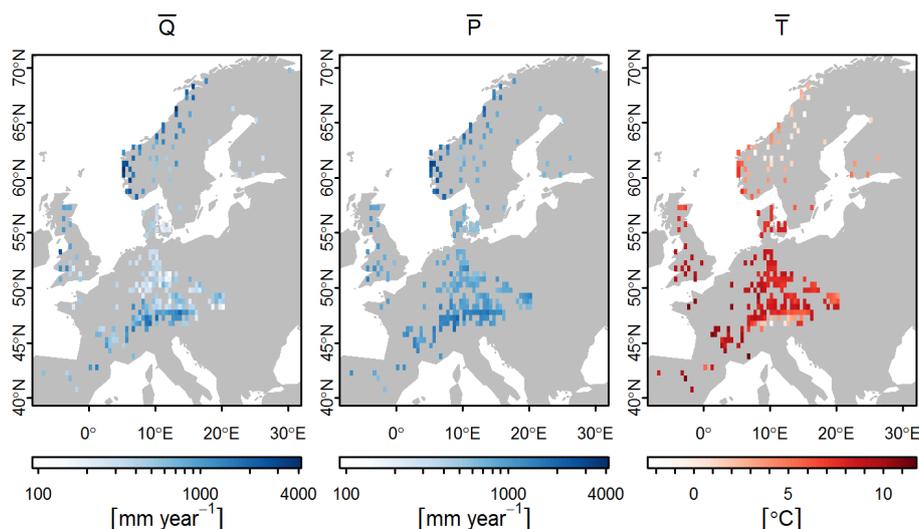
where  $h$  summarises various storage response functions in a catchment (see Clark et al., 2008, 2011 for different plausible formulations). Here, no specific assumptions on the form of  $h$  are made, instead focus is on detecting empirical relationships based on observations. However, it shall be noted that  $h$  can be characterised as a spectral filter (e.g. Milly and Wetherald, 2002) that reduces the high-frequency variance of the precipitation input, for example due to the retention of water in soils. Similarly,  $h$  can also be thought of as an amplifier of low-frequency fluctuations in the atmospheric forcing, for example if multi-year storage of groundwater is considered.

To better understand the different roles of meteorological forcing and terrestrial processes on low-frequency runoff variability, the following questions were formulated:

1. Can the space-time patterns of low-frequency runoff be directly related to the equivalent patterns of precipitation ( $P$ ) and temperature ( $T$ ) – or do catchment processes alter these patterns?
2. Is the fraction of low-frequency variance in runoff related to the fraction of low-frequency variance in the forcing – or do catchment processes have a major impact?
3. Can the fraction of low-frequency variance in runoff be related to catchment processes – here represented by a few, easily accessible, catchment characteristics?

In order to approach these questions we adopt the idea that hydrological systems can be considered as low-pass filters and that the strength of this filtering depends on catchment properties such as topography, hydrogeology, land-cover and climatic conditions. Any low-pass filtering implies a spectral representation of the time series. Accordingly, a runoff series  $Q$  can be decomposed into a set of additive sub-series,

$$Q = \sum_{f \in F} Q_f \quad (3)$$



**Fig. 1.** Mean annual runoff ( $\bar{Q}$ ), precipitation ( $\bar{P}$ ) and temperature ( $\bar{T}$ ) for the period 1962–2000.

where  $F$  denotes a set of frequency bands  $f$ , covering  $f_{\max} > f \geq f_{\min}$  ( $f_{\max}$  and  $f_{\min}$  are the upper and the lower frequency bounds, respectively). Each sub-series  $Q_f$ , can additionally be characterised by its variance  $\sigma_f^2$ . In this study, focus is on runoff variability on time scales longer than 12 months ( $1/12 > f > 0 \text{ months}^{-1}$ ). The corresponding sub-series is denoted as  $Q_{\text{Long}}$  and the fraction of low-frequency variance of runoff is

$$\Phi_Q = \frac{\sigma_{Q_{\text{Long}}}^2}{\sigma_Q^2} \quad (4)$$

where  $\sigma_{Q_{\text{Long}}}^2$  is the variance of  $Q_{\text{Long}}$  and  $\sigma_Q^2$  is the total variance of the runoff-series  $Q$ . Any reduction in the high frequency variance (variance on time scales smaller than one year) will increase the fraction of the low-frequency variance  $\Phi_Q$ , whereas the temporal evolution of the corresponding low-frequency component  $Q_{\text{Long}}$ , will not change.

This framework is then used to assess the low-frequency variability of monthly runoff. After isolating  $Q_{\text{Long}}$  and quantifying  $\Phi_Q$  these quantities will be analysed separately in order to provide answers to the questions formulated above. Finally, the results are discussed with respect to possible mechanistic explanations.

## 2 Data

The study was based on 358 time series of daily runoff, which cover the period 1963 to 2000 and were aggregated to monthly values. This unique data set consists of small, near-natural catchments that are not nested (median catchment size:  $300 \text{ km}^2$ ). Most records, including catchment area and coordinates of the gauging station, originate from the European Water Archive (EWA), a database assembled by the

Euro-FRIEND<sup>1</sup> program. The EWA is held by the Global Runoff Data Centre<sup>2</sup> (GRDC) which also manages data requests. The EWA was recently updated and complemented by national data from partners in the WATCH project<sup>3</sup>. For a detailed overview on data availability see Stahl et al. (2008, 2010). Mean catchment elevation and slope, estimated from a high resolution digital elevation model, were obtained from the pan-European river and catchment database CCM2 (Catchment Characterisation and Modelling 2; Vogt et al., 2007).

Observed temperature and precipitation series were not available for the catchments, so instead the WATCH forcing data (WFD; Weedon et al., 2010, 2011) were used. The WFD provide bias corrected variables, based on the ERA-40 reanalysis (Uppala et al., 2005), on a  $0.5^\circ$  grid. In the centre of the spatial domain of this study, at  $55^\circ \text{ N}$ , grid-cell cover  $1994 \text{ km}^2$ . Only grid-cells with one or more runoff stations were used, resulting in a total of 246 grid-cells. In case of more than one runoff station per grid-cell, the area-weighted average of the runoff values was used.

Mean climatic conditions of all grid-cells were characterised by the mean annual temperature ( $\bar{T}$ ), precipitation ( $\bar{P}$ ) and runoff ( $\bar{Q}$ ) (Fig. 1). The fraction of low-frequency variance of runoff ( $\Phi_Q$ ), precipitation ( $\Phi_P$ ) and temperature ( $\Phi_T$ ) was used to characterise the strength of fluctuations on long time scales.

The ratio of the 25 (low-flows) to the 75 (high-flows) percentile of daily runoff ( $Q_{25}/Q_{75}$ ) was used to characterise catchment response to precipitation. The notion of percentiles follows the statistical convention (representing

<sup>1</sup><http://ne-friend.bafg.de/servlet/is/7413/>, last access: 15 May 2011

<sup>2</sup><http://grdc.bafg.de>, last access: 15 May 2011

<sup>3</sup><http://eu-watch.org/>, last access: 15 May 2011

cumulative frequencies) and not the hydrological one (representing exceedance frequencies). The ratio  $Q_{25}/Q_{75}$  is related to the shape of the flow duration curve (FDC; Vogel and Fennessey, 1994). High values indicate a flat FDC, reflecting the relatively low variability of flows around the median and thus a dampened response. A low value corresponds to a steep FDC, and is an indicator of a fast responding flow regime with a high variance of daily runoff (Gustard et al., 1992).

### 3 Methods

#### 3.1 Extracting low-frequency components

The low-frequency components of runoff ( $Q_{\text{Long}}$ ), temperature ( $T_{\text{Long}}$ ) and precipitation ( $P_{\text{Long}}$ ) were obtained using the “Seasonal-Trend Decomposition Procedure Based on Loess” (STL; Cleveland et al., 1990). The STL-algorithm is one of many time series decomposition techniques that are available and is limited to the decomposition of time series into a low-frequency, a seasonal and a residual component. STL has been applied in several hydrological studies, for example, for deseasonalisation of runoff time series (Montanari et al., 1997) or for the detection of nonlinear trends in groundwater levels (Shamsudduha et al., 2009). In addition, other, more flexible, techniques were considered for the analysis, including Singular System Analysis (SSA; e.g. Ghil et al., 2002), Wavelet filtering (e.g. Torrence and Compo, 1998) and Empirical Mode Decomposition (EMD; Huang et al., 1998). The choice of the STL algorithm was motivated by its suitability for isolating low-frequency components from time series while analytically controlling the spectral leakage of high frequency variability into the low-frequency component.

STL is based on locally weighted scatter-plot smoothing, LOESS (Cleveland and Devlin, 1988). Assuming a dependent variable  $x_i$  and an independent variable  $t_i$  (for  $i = 1$  to  $n$ ), the LOESS estimates smoothed values of the dependent variable  $\hat{g}(t_i)$  for any value  $t_i$ . First, a positive integer  $\lambda$  is chosen. Then the subset of the  $\lambda$  nearest  $t_j$  to  $t_i$  are selected, where  $j$  is the index of the subset. For this selection weights are computed as

$$w_j = \left[ 1 - \left( \frac{|t_j - t_i|}{\delta_\lambda} \right)^3 \right]_+^3 \quad (5)$$

where  $[\ ]_+$  denotes the positive part and  $\delta_\lambda$  is the distance from the  $\lambda$ -th farthest  $t_j$  from  $t_i$ . Finally, the locally smoothed value of  $x_i$  at  $t_i$  are computed as a polynomial fit to the selected subset, weighted with  $w_j$ , to obtain the smoothed value  $\hat{g}(t_i)$ . Here a polynomial of degree one (i.e. a locally linear fit) was used.

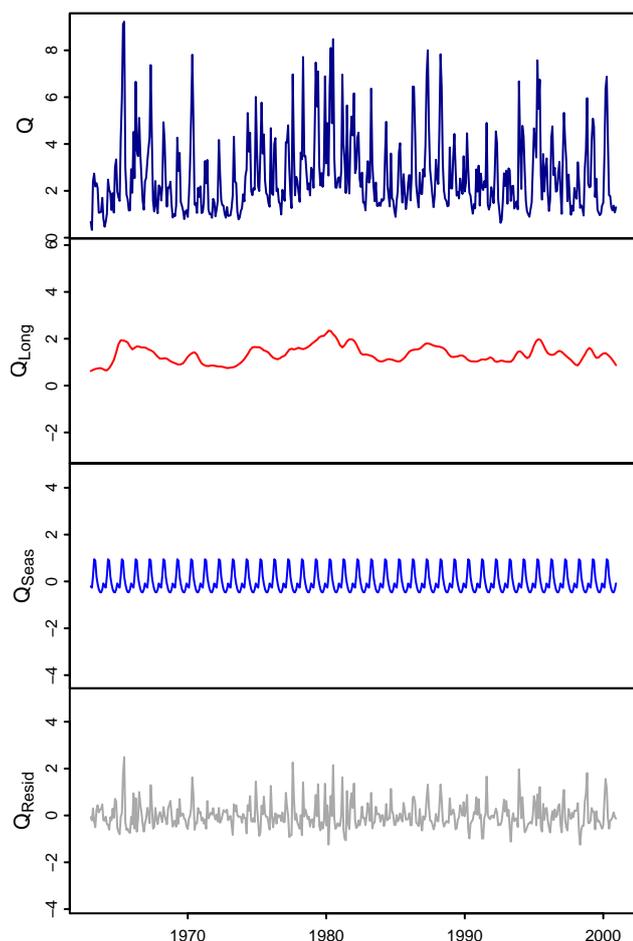
STL decomposes a time series  $X$  into

$$X = X_{\text{Long}} + X_{\text{Seas}} + X_{\text{Resid}} \quad (6)$$

where the seasonal ( $X_{\text{Seas}}$ ) and the low-frequency ( $X_{\text{Long}}$ ) components are separated from the residual ( $X_{\text{Resid}}$ ). STL is an iterative procedure involving an inner and an outer loop, each applying a sequence of LOESS. In the inner loop,  $X$  is decomposed into the three sub-series (Eq. 6). The seasonal component is identified by first smoothing the seasonal sub-series (i.e. the series of all Januaries, Februaries, ...) with LOESS (parameter  $\lambda_{\text{Seas}}$ ), which are then low-pass filtered by a sequence of moving averages and an additional application of LOESS. After removing the seasonal component,  $X_{\text{Long}}$  is separated from  $X_{\text{Resid}}$  using LOESS (parameter  $\lambda_{\text{Long}}$ ). In the outer loop, robustness weights are calculated that are used in the next iteration of the inner loop to reduce the influence of outliers. The full STL algorithm has six free parameters, determining the value of  $\lambda$  for the different applications of LOESS as well as the number of iterations. In this study only  $\lambda_{\text{Seas}}$  and  $\lambda_{\text{Long}}$  need to be controlled and all other parameters were set to default values. (See Cleveland et al., 1990, for recommendations and the documentation of the function “stl” in the R-software; R Development Core Team, 2011). Both  $\lambda_{\text{Seas}}$  and  $\lambda_{\text{Long}}$  are set in such a manner that an optimal identification of  $X_{\text{Long}}$ , the target variable, is guaranteed. For the identification of  $X_{\text{Seas}}$ , the LOESS parameter was set to  $\lambda_{\text{Seas}} = 10n + 1$ , where  $n = 444$  is the length of the monthly observations. This effectively replaces the smoothing by the mean of each month and guarantees that all low-frequency variability is captured by  $X_{\text{Long}}$ . For the identification of  $X_{\text{Long}}$ , the LOESS parameter was set to  $\lambda_{\text{Long}} = 19$ . This choice follows the recommendations of Cleveland et al. (1990), who showed analytically that setting  $\lambda_{\text{Long}}$  equal to the smallest integer satisfying  $\lambda_{\text{Long}} \geq 1.5p/(1 - (1.5/\lambda_{\text{Seas}}))$ , is optimal with respect to a minimal spectral leakage of high frequency components into the low-frequency components ( $p = 12$  is the periodicity of the seasonal cycle). Note that the choice of  $\lambda_{\text{Long}} = 19$  is closely related to separating the variance of a power-spectrum at a frequency of  $f = 1/19$  and thus a small part of the low-frequency variance is not captured by  $X_{\text{Long}}$ .

Figure 2 illustrates such a decomposition and reflects the chosen value of the parameter  $\lambda_{\text{Seas}}$ , that enforces annual cycles that do not change over time. It should be noted that the amplitudes of the three different components are actually of the same order of magnitude and thus these three sub signals are of similar importance. However, only  $X_{\text{Long}}$  is further analysed both with respect to its fractional variance (Eq. 4) and its space time pattern.

The fraction of low-frequency variance  $\Phi_X$ , is a parameter characterising the variance distribution of the power-spectrum of the time series  $X$ . However, the STL-algorithm used to estimate  $\Phi_X$  is primarily designed for time series decomposition and may thus lead to biased estimates of spectral properties. Appendix A summarises the results of a supplementary analysis comparing  $\Phi_X$  derived from runoff, precipitation and temperature to an alternative (spectral) estimate, based on the multi taper method (Thomson, 1982). Further,



**Fig. 2.** Decomposition of a monthly runoff series  $Q$ : [ $\text{mm d}^{-1}$ ] into low-frequency ( $Q_{\text{Long}}$ ), seasonal ( $Q_{\text{Seas}}$ ) and residual ( $Q_{\text{Resid}}$ ) components.

due to sampling errors and finite size effects, the estimated values of  $\Phi_X$  are uncertain and thus error-bars for  $\Phi_X$  are shown. This uncertainty was quantified using a bootstrapping technique and the resulting confidence intervals of  $\Phi_X$  are also presented in Appendix A. The results show that both estimates are closely related and that  $\Phi_X$  estimated using the STL algorithm is reliable.

## 3.2 Factors influencing $\Phi_Q$

### 3.2.1 Hydroclimatic regime

Differences in the seasonality of runoff may have an influence on the magnitude of  $\Phi_Q$  and therefore it was chosen to group observed runoff into two classes of hydroclimatic regimes. Regime class one (RC1) is a snow-regime where snow accumulation and melt lead to an annual cycle with minimum discharge in winter and a maximum in spring. Regime class two (RC2) is an evapotranspiration-regime, where the mean annual cycle of runoff follows the seasonal

pattern of the potential evapotranspiration with maximum discharge in winter and a summer minimum. Although a two-class description is a strong simplification of the complexity of seasonal patterns of runoff found throughout Europe (e.g. Haines et al., 1988; Harris et al., 2000; Krasovskaia et al., 2003; Renner and Bernhofer, 2011), the separation into snow- and evaporation-regimes allows for an easy assessment of the influence of snow on  $\Phi_Q$ .

The two groups were separated based on the shape of the hydroclimatic regime, defined as the mean annual cycle of runoff, using cluster analysis. Such an approach is also referred to as classification by shape (Harris et al., 2000). At each grid-cell, the mean monthly flow was first standardised by removing the mean and then dividing by the standard deviation. The standardised hydroclimatic regimes were then grouped into two regime classes using Ward's hierarchical clustering algorithm with an euclidean distance measure (Ward, 1967). The difference in  $\Phi_Q$  between the two regime classes were tested using Wilcoxon's rank test (Wilcoxon, 1945). Wilcoxon's test is comparable to the t-test which assesses whether the mean of two groups are different. However, unlike the t-test, it does not rely on the assumption of normality.

### 3.2.2 Catchment characteristics and climatic conditions

Spearman's rank-correlation coefficient  $\rho$  (Spearman, 1987, reprint from Spearman 1904), was used to test the influence of catchment topography, mean climatic conditions and climatic variability on the fraction of low-frequency variance of runoff,  $\Phi_Q$ . Spearman's  $\rho$  is equivalent to the Pearson correlation coefficient between ranked variables. The advantage of Spearman's  $\rho$  over the Pearson correlation coefficient is that it does not rely on observations from a bivariate normal distribution. It is robust to outliers and can also detect and quantify the strength of any monotonic (nonlinear) relation.

The significance of Spearman's  $\rho$  between two variables  $x$  and  $y$  was assessed using a modified version of the t-test, which tests the null hypothesis that  $\rho = 0$  based on the assumption that  $\rho$  is approximately normally distributed. However, the standard t-test assumes that the two correlated variables are independent, which is not necessarily the case. To account for possible spatial dependence in both  $x$  and  $y$  a modified version of the t-test (Clifford et al., 1989) was used, which estimates the (reduced) effective sample size for spatially autocorrelated processes. The test of Clifford et al. (1989) was initially developed for Pearson's correlation coefficient, but has been shown also to be valid for Spearman's  $\rho$  (Haining, 1991). The effective sample size  $m_{\text{eff}}$  is estimated, following the notation of Dale and Fortin (2009), as

$$m_{\text{eff}} = 1 + \frac{m}{1 + \frac{2}{m} \sum_l m(l) r_x(l) r_y(l)} \quad (7)$$

where  $m$  is the number of points contributing to  $\rho$ .  $r_x(l)$  and  $r_y(l)$  are the spatial autocorrelations at lag  $l$  and  $m(l)$

is the number of pairs. The appropriate estimate for  $r(l)$  is still subject to discussion (Dale and Fortin, 2009; Dutilleul et al., 1993). Here, we follow Pepin et al. (2011), who applied this test for atmospheric variables, and estimate the degree of spatial autocorrelation using Moran's  $I$  statistic (Moran, 1950) to estimate  $r(l)$ . To evaluate  $r(l)$  on unevenly spaced spatial samples (such as grid-cells defined by a latitude-longitude lattice), the spatial correlation needs to be evaluated for discrete distance classes, (e.g. Bjørnstad and Falck, 2001). Here, a distance class of 100 km was used, i.e.  $r(l)$  is only evaluated at  $l=0, 100, 200, \dots$  km. In practice this means that the distances between grid-cells are rounded to multiples of 100 prior to further computations. For long lags the number of pairs  $m(l)$  declines and the resulting estimate of  $r(l)$  becomes highly uncertain. To reduce this uncertainty only lags up to a maximal distance equal to the length of the shorter geographical dimension (latitude or longitude) of the spatial domain were evaluated, following the suggestion of Peltonen et al. (2002). (Moran's  $I$  was estimated using R-package "ncf"; Bjørnstad, 2009). Once the effective sample size is determined, the modified t-statistic can be retrieved as

$$t_\rho = \frac{\rho \sqrt{m_{\text{eff}} - 2}}{\sqrt{1 - \rho^2}}. \quad (8)$$

The null hypothesis of zero correlation is rejected with the significance level  $\alpha$  if  $|t_\rho| < t_{m_{\text{eff}}-2, 1-\alpha/2}$ , where  $t_{m_{\text{eff}}-2, 1-\alpha/2}$  is the corresponding value of the t-distribution with  $m_{\text{eff}}-2$  degrees of freedom. Significance is reported for  $\alpha = 0.001$ .

### 3.3 Spatial patterns of simultaneous variations

Dominant spatial patterns of simultaneous variations of geophysical time series are commonly identified using so-called dimension reduction techniques. One of the most popular methods is principle component analysis (PCA), also known as empirical orthogonal function (EOF) analysis (e.g. von Storch and Zwiers, 1999). (For applications to continental scale river flow, see Lins, 1997, and Shorthouse and Arnell, 1999). The results of PCA are often depicted as a series of maps, where each map corresponds to one temporal signal and the mapped value quantifies how the time series at each location are correlated to this signal. However, PCA relies on the assumption of linearity, which is not necessarily appropriate for geophysical phenomena (e.g. Monahan, 2001; Gamez et al., 2004; Mahecha et al., 2010). The consequences of the linearity assumptions can be illustrated by considering distances on the earth surface. The linear (euclidean) distance between a location on the northern hemisphere (e.g. Oslo) and a location on the southern hemisphere (e.g. Cape Town) would go through the solid earth. Thus, it does not reflect their true distance, which is rather characterised by the geodesic distance that describes the shortest path between two locations on the curved earth surface.

However, distances between nearby locations on the earth-surface can be approximated using linear (euclidean) measures and the distance between locations far from each other can be approximated as the sum of those local linear distances. Similar arguments hold for the analysis of nonlinear physical processes and isometric feature mapping (ISOMAP; Tenenbaum et al., 2000), which takes advantage of these considerations, was used in this study to characterise the sets of spatially distributed time series of  $Q_{\text{Long}}$ ,  $P_{\text{Long}}$  and  $T_{\text{Long}}$ .

Let  $\mathbf{X}$  be a matrix containing spatially distributed time series with  $m$  rows representing the locations and  $n$  columns representing the time steps. First ISOMAP estimates a  $m \times m$  geodesic distance matrix  $\mathbf{G}$ , which is a symmetric matrix and each element  $g_{ij}$  quantifies the strength of common variability between pairs of time series.  $\mathbf{G}$  is estimated based on the assumption that the geodesic distances can be approximated as the shortest path on a neighbourhood graph. This neighbourhood graph is computed using the algorithm of Dijkstra (1959), which connects each point to its  $k$  nearest neighbours.

The geodesic distance matrix  $\mathbf{G}$  is then subject to classical multidimensional scaling (e.g. Torgerson, 1952; Borg and Groenen, 2005), which seeks an euclidean space  $\mathbf{Y}$  that sufficiently describes  $\mathbf{G}$  with few dimensions.  $\mathbf{Y}$  is found as the first columns of

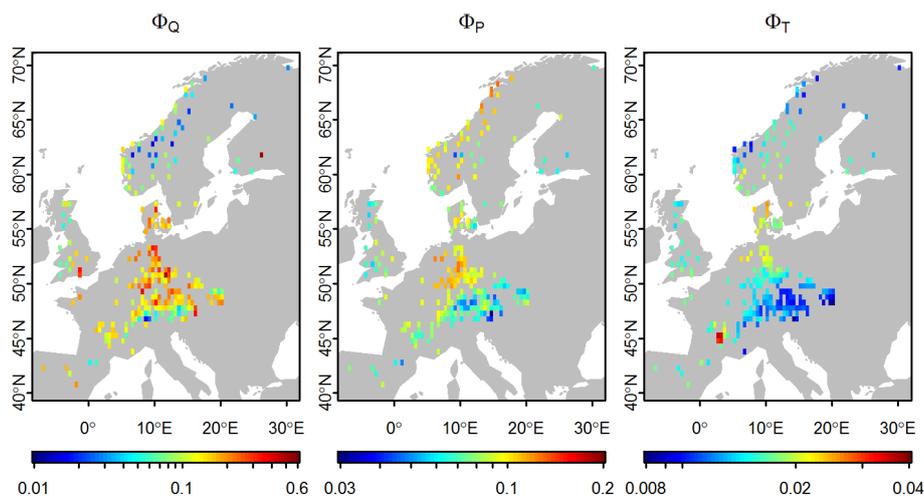
$$\mathbf{Y} = \mathbf{E} \Lambda^{1/2} \quad (9)$$

where  $\Lambda^{1/2}$  is a diagonal matrix containing the square-root of the eigenvectors of  $\tau(\mathbf{G})$  in a decreasing order and  $\mathbf{E}$  is a column matrix containing the corresponding eigenvectors. ( $\tau(\mathbf{G}) = \frac{1}{2} \mathbf{H} \mathbf{G}^{(2)} \mathbf{H}$  is a double centering operator, where  $\mathbf{G}^{(2)}$  is a matrix of squared distances and  $\mathbf{H}_{ij} = \delta_{ij} - \frac{1}{N}$  and  $\delta_{ij}$  is the Kronecker delta).

Each dimension (column) of the space  $\mathbf{Y}$  has  $m$  entries, which represent the different spatial locations and similar values indicate simultaneous temporal evolution. Thus, maps of the leading dimensions are used to show spatial patterns of simultaneous variations. Only the first columns of  $\mathbf{Y}$  that explain more than one percent of the variance of  $\mathbf{G}$ , will be considered as they capture the most dominant part of the signal.

ISOMAP has one free parameter,  $k$ , the number of nearest neighbours used to estimate the geodesic distance. Following the procedure of Mahecha et al. (2007),  $k=20$  was found to provide the best choice for representing the spatial structure of  $Q_{\text{Long}}$ ,  $T_{\text{Long}}$  and  $P_{\text{Long}}$  with a limited number of dominant modes. Geodesic distances were estimated based on euclidean distance matrices of the standardised versions of the low-frequency components to emphasise common temporal evolution. (ISOMAP was performed using the R-package "vegan"; Oksanen et al., 2010).

The similarity of the leading ISOMAP dimensions of  $Q_{\text{Long}}$ ,  $P_{\text{Long}}$  and  $T_{\text{Long}}$ , was quantified using linear Procrustes analysis (PA; e.g. Borg and Groenen, 2005). In principle, Procrustes analysis assesses whether a space  $\mathbf{B}$  can be



**Fig. 3.** Spatial patterns of the fraction of low-frequency variance of runoff ( $\Phi_Q$ ), precipitation ( $\Phi_P$ ) and temperature ( $\Phi_T$ ). (Note that each panel has its own logarithmic colour scale).

transformed into a target space  $\mathbf{A}$  by a limited set of operations such as scaling, rotation and reflection. Here, the matrix  $\mathbf{A}$  contains the standardised leading ISOMAP dimensions of  $Q_{\text{Long}}$  and  $\mathbf{B}$  contains the standardised leading ISOMAP dimensions of either  $P_{\text{Long}}$  or  $T_{\text{Long}}$ . In principle, the resulting parameters can be analysed in detail, however, in this study, only the goodness of fit will be considered, which together with a test of statistical significance, will allow to quantify the relation between the leading ISOMAP components of two variables. This is achieved using the procedure introduced by Peres-Neto and Jackson (2001), which solves  $\mathbf{A} = \mathbf{TB}$ , where the rotation matrix  $\mathbf{T}$  is found to minimise the sum of squared differences (i.e. linear PA). The strength of this fit can be quantified by

$$r^2 = 1 - \frac{2(1 - \text{trace}(\mathbf{W}))}{\text{trace}(\mathbf{W})^2} \quad (10)$$

where  $\mathbf{W}$  is a diagonal matrix containing the singular values of  $\mathbf{A}^T\mathbf{B}$ .  $r^2$  scales between 0 and 1 and its interpretation is equal to Pearson's correlation coefficient. The significance of  $r^2$  is tested using a resampling test (Peres-Neto and Jackson, 2001). (Procrustes analysis was conducted using the R-package “vegan”; Oksanen et al., 2010).

## 4 Results

### 4.1 Fraction of low-frequency variance

On average, the fraction of low-frequency variance was found to be highest for runoff ( $0.01 \leq \Phi_Q \leq 0.53$ ), followed by precipitation ( $0.03 \leq \Phi_P \leq 0.13$ ) and lowest for temperature ( $0.01 \leq \Phi_T \leq 0.04$ ). The large spread in the fraction of low-frequency variance of runoff, shows that catchments vary significantly in their sensitivity to long-term climatic

fluctuations. Note that the distribution of  $\Phi_Q$  is extremely skewed. All, but nine out of 246 grid-cells, have  $\Phi_Q \leq 0.3$  indicating that  $Q_{\text{Long}}$  accounts, in most of the cases, for less than one third of the total runoff variance. The fractions of low-frequency variance of runoff ( $\Phi_Q$ ), precipitation ( $\Phi_P$ ) and temperature ( $\Phi_T$ ) have large spatial variability (Fig. 3).  $\Phi_Q$  is on average largest in central Europe and lowest in the Alps and the inland parts of Scandinavia. The fraction of low-frequency variance of runoff is generally largest in regions where the mean annual runoff and precipitation are low and the mean annual temperature is high (Fig. 1). The spatial patterns of  $\Phi_P$  and  $\Phi_T$  differ from the spatial distribution of  $\Phi_Q$ .  $\Phi_P$  has the highest values in central Europe and Scandinavia, whereas  $\Phi_T$  has its largest values in southern France and Denmark.

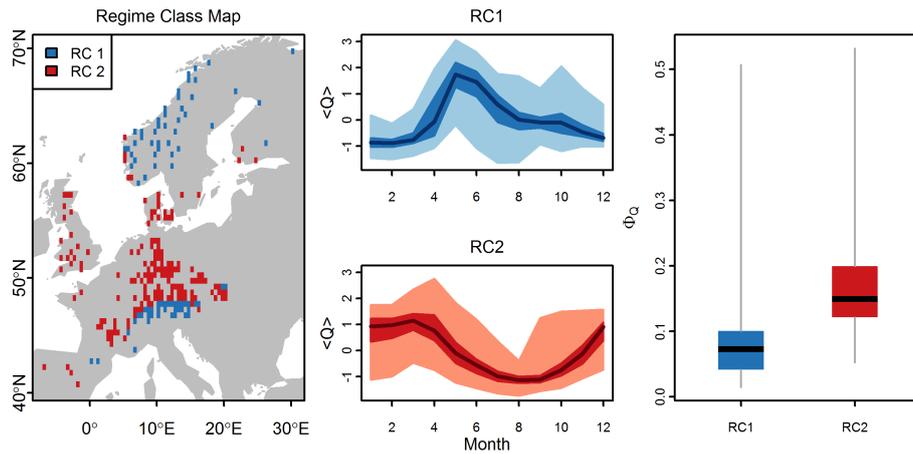
#### 4.1.1 Influence of snow on $\Phi_Q$

Figure 4 shows the results of the cluster analysis used to group grid-cells according to their hydroclimatic regime. Both the snow dominated (RC1, 81 grid-cells) and the evapotranspiration dominated regime class (RC2, 165 grid-cells) have distinct seasonal and regional patterns. RC1 has a winter minimum and a spring maximum and is separated into two regions, one in Scandinavia and one in the Alps. RC2 has a summer minimum and a winter maximum. Most grid-cells of this regime class are located in the centre of the spatial domain, between the two regions of RC1.

The median  $\Phi_Q$  of the two regime classes is significantly different ( $p < 0.001$ , Wilcoxon test). On average  $\Phi_Q$  is significantly lower in RC1 (median:  $\Phi_{Q,RC1} = 0.07$ ) than in RC2 (median:  $\Phi_{Q,RC2} = 0.15$ ) and the inter-quartile ranges of  $\Phi_Q$  do not overlap (Fig. 4).

**Table 1.** Spearman's rank correlation coefficients ( $\rho$ ) between the fraction of low-frequency variance of runoff ( $\Phi_Q$ ) and the following set of variables: Fraction of low-frequency variance of precipitation ( $\Phi_P$ ) and temperature ( $\Phi_T$ ); mean annual runoff ( $\bar{Q}$ ), precipitation ( $\bar{P}$ ) and temperature ( $\bar{T}$ ); the ratio of 25 and 75 percentile of daily runoff  $Q_{25}/Q_{75}$ ; as well as mean catchment elevation (Elev.), slope and area.  $m_{\text{eff}}$ : effective sample size.  $t_\rho$ : t-statistics for  $m_{\text{eff}} - 2$  degrees of freedom. Significance (signif.) of the t-test is reported for  $p \leq 0.001$  (\*\*\*).

	All grid-cells				Grid-cells in RC 1				Grid-cells in RC 2			
	$\rho$	$m_{\text{eff}}$	$t_\rho$	signif.	$\rho$	$m_{\text{eff}}$	$t_\rho$	signif.	$\rho$	$m_{\text{eff}}$	$t_\rho$	signif.
$\Phi_P$	0.11	50.52	0.77		-0.17	57.32	-1.27		0.25	40.66	1.60	
$\Phi_T$	0.11	37.17	0.65		-0.17	62.45	-1.33		0.01	44.94	0.09	
$\bar{Q}$	-0.69	60.17	-7.22	***	-0.02	76.01	-0.21		-0.62	78.66	-6.90	***
$\bar{P}$	-0.39	59.28	-3.25		0.33	75.41	3.02		-0.43	76.04	-4.11	***
$\bar{T}$	0.59	66.56	5.86	***	0.63	65.47	6.39	***	0.20	94.43	1.95	
Elev.	-0.41	47.25	-3.04		-0.19	56.31	-1.45		-0.22	48.30	-1.53	
Slope	-0.46	56.25	-3.77	***	0.07	59.02	0.51		-0.30	66.07	-2.56	
Area	-0.02	116.55	-0.20		-0.09	75.92	-0.80		0.03	97.25	0.28	
$Q_{25}/Q_{75}$	0.58	53.36	5.04	***	0.61	58.07	5.81	***	0.56	64.61	5.38	***

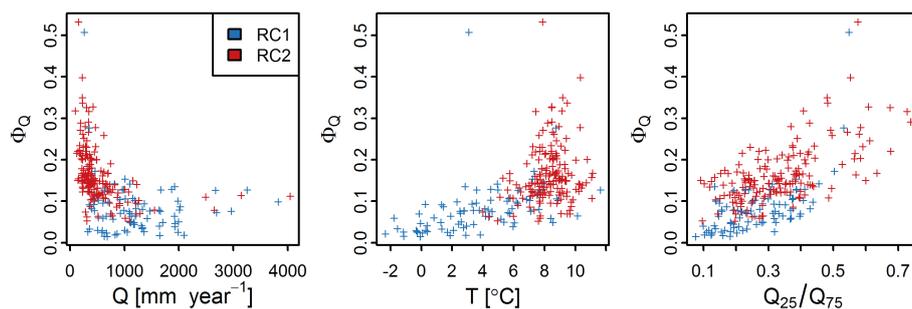


**Fig. 4.** Effect of the hydroclimatic regime on  $\Phi_Q$ , the fraction of low-frequency variance of runoff. RC1: Snow dominated regime (winter minimum–summer maximum). RC2: evapotranspiration dominated regime (summer minimum–winter maximum). The map shows the geographical domain of each regime class. The two central panels show the average shape of each hydroclimatic regime; dark line: median, dark shaded area: inter-quartile range, light shaded area: range;  $\langle Q \rangle$  denotes standardisation of runoff). The boxplots show the difference of  $\Phi_Q$  between the two hydroclimatic regime classes (RC1 and RC2). Black horizontal line: median, the box covers the interquartile, the grey whiskers range from the minimum to the maximum value.

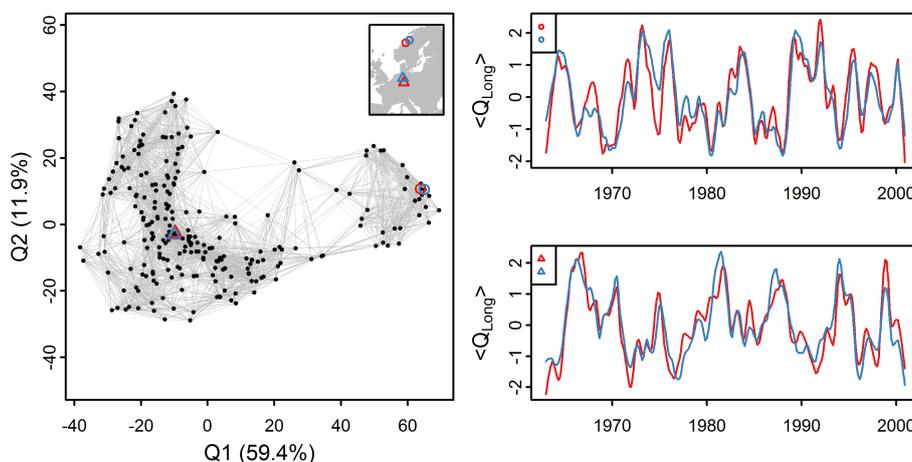
#### 4.1.2 Influence of catchment properties and climatic conditions on $\Phi_Q$

Table 1 shows the results of the correlation analysis relating  $\Phi_Q$  to climate and catchment characteristics, both for the entire study domain as well as for the two regime classes separately. Among all variables only mean annual runoff ( $\bar{Q}$ ), precipitation ( $\bar{P}$ ) and temperature ( $\bar{T}$ ), the mean catchment slope and  $Q_{25}/Q_{75}$  have significant correlations with  $\Phi_Q$ . The weakest correlations are found for the fraction of low-frequency variance of precipitation ( $\Phi_P$ ) and temperature ( $\Phi_Q$ ) as well as for catchment area.

A negative correlation is found between  $\Phi_Q$  and  $\bar{Q}$  for all grid-cells as well as the grid-cells in the evapotranspiration dominated regime class (RC2), indicating that the fraction of low-frequency variance of runoff decreases under wetter conditions. In RC2,  $\Phi_Q$  and  $\bar{P}$  are also correlated, which further supports this. However,  $\Phi_Q$  and  $\bar{Q}$  are not correlated in the snow-regime (RC1). As illustrated in Fig. 5, the  $\Phi_Q$  values of RC1 are almost constant over the entire flow-range. The fraction of low-frequency variance of runoff and mean annual temperature are positively correlated for all grid-cells and grid-cells in RC1. Thus  $\Phi_Q$  is larger under warmer conditions, as also seen in Fig. 5, where the  $\Phi_Q$  values for RC1 increase almost linearly with temperature. However, this is



**Fig. 5.** Scatter plots illustrating the influence of selected variables (mean annual runoff  $\bar{Q}$  and temperature  $\bar{T}$  and  $Q_{25}/Q_{75}$ ) on the fraction of low-frequency variance of runoff ( $\Phi_Q$ ).



**Fig. 6.** Left panel: the space spanned by the two leading ISOMAP components derived from  $Q_{Long}$ . The grey lines are the neighbourhood graph used to estimate the geodesic distances. Two pairs of neighbouring points are marked (triangles and circles) and their spatial location is indicated on the map. Right panel: time series of  $Q_{Long}$  for the two pairs of points marked in the left panel ( $\langle Q_{Long} \rangle$  denotes standardised low-frequency components of runoff).

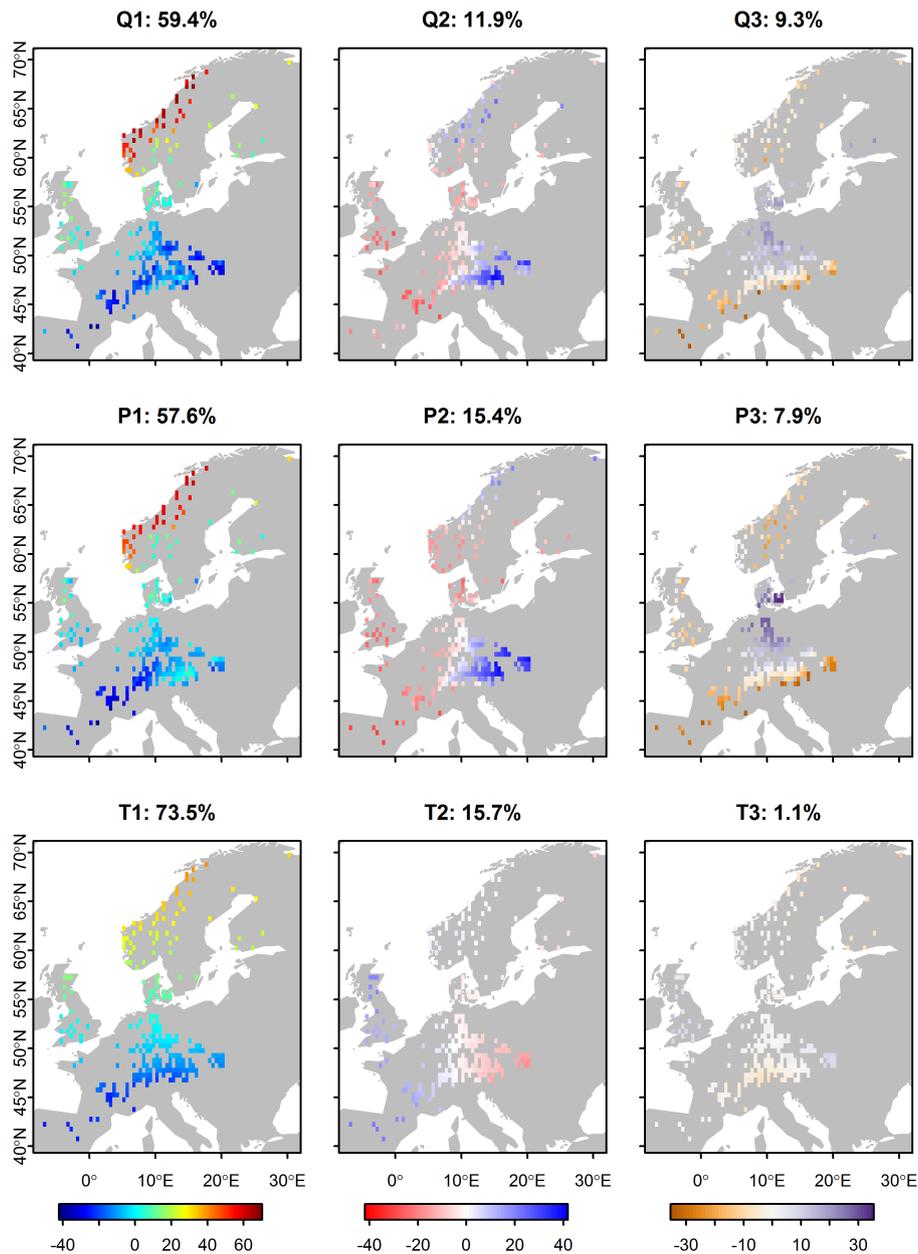
not the case for the evapotranspiration-regime (RC2). The negative correlation between mean catchment slope and the fraction of low-frequency variance is only found if all grid-cells are considered. Only  $Q_{25}/Q_{75}$  has a positive correlation with  $\Phi_Q$  in all cases. Figure 5 also shows that the rate of increase in  $\Phi_Q$  with increasing  $Q_{25}/Q_{75}$  are comparable in both regime classes, although overall higher  $\Phi_Q$  values are found for RC2.

The significance of the correlations reported above is directly influenced by the degree of spatial dependence of the variables (Sect. 3.2.2). The spatial dependence of all variables considered are reported in Appendix B, both for all data points as well as for the two regime classes.

#### 4.2 Space-time patterns of low-frequency runoff

Figure 6 is used to illustrate how the results of ISOMAP can be interpreted in the context of spatially distributed time series, using the ISOMAP of  $Q_{Long}$ . The left panel shows the space spanned by the two leading components (i.e. the

first two columns of the space  $\mathbf{Y}$ , Eq. 9). Each point represents a spatial location (i.e. grid-cell). The grey lines are the neighbourhood graph used to estimate the geodesic distances. Neighbouring points in this plot indicate that the underlying time series have a high agreement. The further apart the points are, the larger the difference in the underlying time series. This is illustrated using two pairs of points that are marked with circles and triangles. The corresponding time series are shown in the two right panels. The time series of neighbouring points are very similar, whereas time series of the more distant points do not share common patterns. These gradual changes of temporal evolution with spatial location can be due to various processes. In the framework of ISOMAP, the first component represents the most dominant signal and other signals are captured by subsequent components. The percentage of variance explained by each component indicates how strong these signals are. Maps of the components show the relation of the time series to each other, and how this relation changes in space.



**Fig. 7.** Spatial patterns of the three dominant ISOMAP components of the low-frequency variability of runoff ( $Q$ ), precipitation ( $P$ ) and temperature ( $T$ ). The numbers in the figure headings are the percentage of explained variance.

The three leading ISOMAP components of  $Q_{\text{Long}}$ ,  $P_{\text{Long}}$  and  $T_{\text{Long}}$  (Fig. 7) explain a large amount of the variance of the input matrices (runoff: 80.6%, precipitation: 80.8%, temperature: 90.3%). The three runoff components are strongly related to the corresponding precipitation components ( $r^2=0.93$ ,  $p \leq 0.001$ , Procrustes analysis) and moderately related to the temperature components ( $r^2=0.79$ ,  $p \leq 0.001$ ). The remaining components account for less than 1% of the variance. The first components of runoff ( $Q1$ : 59.4%), precipitation ( $P1$ : 57.6%) and temperature ( $T1$ : 73.5%) have very similar spatial patterns. Their

common feature is a pronounced north south gradient across Europe. For runoff and precipitation the gradient is slightly tilted northwest in Scandinavia. The main feature of the second components of runoff ( $Q2$ : 11.9%), precipitation ( $P2$ : 15.4%) and temperature ( $T2$ : 15.7%) is a west east gradient. This pattern is also common to all three variables with some departures of the runoff component ( $Q2$ ) in Scandinavia. Furthermore, the third components of runoff ( $Q3$ : 9.3%) and precipitation ( $P3$ : 7.9%) have similar patterns. Their common feature is a gradient from the centre of the spatial domain (around Denmark) toward the north and

south. The spatial pattern of the third temperature component ( $T_3$  1.1 %) is different and does not appear to be related to runoff.

## 5 Discussion

### 5.1 Fraction of low-frequency variance

The fact that  $\Phi_Q$  is on average larger than and not correlated to  $\Phi_P$  and  $\Phi_T$ , shows that the fraction of low-frequency variance of runoff is not proportional to the fraction of low-frequency variance of precipitation and temperature. Thus, it is likely that the magnitude of  $\Phi_Q$  is primarily controlled by catchment processes. In principle,  $\Phi_Q$  can be altered by either a change in the variance of high frequency components (i.e. the variance of  $Q_{\text{Seas}}$  and  $Q_{\text{Resid}}$ ) or by a change in the variance of the low-frequency component ( $Q_{\text{Long}}$ ).

The fact that  $\Phi_Q$  is significantly lower in hydroclimatic regimes where the mean annual cycle of runoff is dominated by snow (Fig. 4), suggests that snow accumulation and melt lead to a pronounced annual cycle causing a low  $\Phi_Q$ . The relatively strong positive correlation of  $\Phi_Q$  with the mean annual temperature in the snow dominated RC1 ( $\rho=0.63$ ) suggests that this effect gets more pronounced in colder environments, where the influence of snow is more important.

The large differences in  $\Phi_Q$  within both the snow dominated and the evapotranspiration dominated regime class, however, indicate that the impact of snow on the shape of the mean annual cycle is not sufficient to fully explain the variations. The strong negative correlation of  $\Phi_Q$  with mean annual runoff in the evaporation dominated regime class (RC2,  $\rho=-0.62$ ) shows that the fraction of low-frequency variance decreases under wetter conditions. In humid environments, catchment storages are often close to being saturated and runoff is more likely to respond quickly to precipitation, maintaining more of its high-frequency variance (resulting in a low  $\Phi_Q$ ). In this context it appears noteworthy that mean annual precipitation is only correlated to  $\Phi_Q$  in RC2. This may be related to comparing precipitation estimates from a global reanalysis product with large grid-cells to runoff from small catchments. Precipitation amounts are known to have a large spatial variability, depending on orographic effects (e.g. Barstad et al., 2007), which are neither resolved in the WFD nor in the underlying ERA40 data (Weedon et al., 2011). A dynamical down-scaling of the re-analysis data as suggested by Barstad et al. (2009) would be a possible solution for further studies.

The ratio  $Q_{25}/Q_{75}$  is the only variable found to have a strong positive correlation with the fraction of low-frequency variance of runoff for all grid-cells as well as both regime classes, indicating that  $\Phi_Q$  is larger for rivers with a dampened runoff response. The ratio  $Q_{25}/Q_{75}$  characterises the spread of the distribution of daily runoff, which is often described using the flow duration curve (FDC). The FDC can

empirically be related to storage properties (Gustard et al., 1992), but both theoretical (Botter et al., 2008, 2009) and empirical (Castellarin et al., 2004) investigations suggest that the average degree of catchment saturation, which in turn is related to the climatic water balance, also play a role.

The only physical catchment property to be correlated to  $\Phi_Q$ , the fraction of low-frequency variance of runoff, is mean catchment slope. However, this correlation is only significant if all grid-cells are considered and is neither in RC1 nor in RC2 significant. The mean catchment slope differs significantly between the two regime classes ( $p < 0.001$ , Wilcoxon test). This suggests that the high correlation between catchment slope and  $\Phi_Q$  found for all grid-cells follows the influence of snow and may not represent an additional mechanism. None of the other physical catchment properties considered (elevation and area) were found to be correlated to  $\Phi_Q$ , suggesting that rainfall-runoff processes controlled by these properties have little influence. However, it is likely that storage properties of the catchments have an effect and it would be desirable to include information on the hydrogeology and lake percentage of each catchment into the analysis. Unfortunately, no suitable data with the necessary precision are available at the pan-European scale.

Further investigations are needed to fully resolve the different roles of catchment properties and mean climatic conditions on the fraction of low-frequency variance of runoff. However, the current lack of large-scale high quality data on catchment properties hinders further investigations, and supplementary approaches based on hydrological modelling would go beyond the scope of this study. In this context, it would be interesting to test whether strictly formulated model assumptions are able to capture not only the spatial variability of  $\Phi_Q$ , but also of other signatures of hydrological variability. Examples could include spatial patterns of selected statistical moments of runoff variability, measures of streamflow recession and the Hurst coefficient as an indicator of long range dependence.

### 5.2 Space-time patterns

The observation that European low-frequency runoff exhibits strong spatial patterns of simultaneous variations (Fig. 7) confirms the results of several studies that have reported a high degree of coherence in the inter-annual variability of streamflow in relatively small regions in Europe (e.g. Denmark, Hisdal and Tallaksen, 2003; southern Germany, Stahl and Demuth, 1999; Lange and Bernhardt, 2004; the Iberian peninsula, Lorenzo-Lacruz et al., 2011; Iceland, Jónsdóttir and Uvo, 2009; and Turkey, Kalayci and Kahya, 2006). However, the limited spatial extent of previous studies did not allow for a full description of the emerging continental scale patterns.

The simultaneous variations of European low-frequency runoff can be efficiently described by three components, each having a distinct spatial pattern. The large amount of

variance explained by the first ISOMAP component shows that the low-frequency variations of European runoff are dominated by opposing centres of simultaneous variations in the north and the south. A similar north south pattern has previously been identified for the direction of annual and seasonal streamflow trends, revealing increasing streamflow in northern and decreasing streamflow in southern Europe throughout the last decades (Stahl et al., 2010). The results of the present study suggest that the previously identified spatial patterns in streamflow trends are an element of a more general pattern of inter-annual streamflow dynamics. This north south pattern may be related to the zonal structure of atmospheric circulation where the south of Europe is influenced by subtropical and the north by arctic weather systems. Similar patterns of precipitation (van Loon and Rogers, 1978; Hurrell, 1995; Hurrell and van Loon, 1997; Wibig, 1999; López-Moreno and Vicente-Serrano, 2008), temperature (van Loon and Rogers, 1978), runoff (Shorthouse and Arnell, 1997, 1999) and peak discharges (Bouwer et al., 2008) have previously been related to the North Atlantic Oscillation. The west east gradient of the second component may be related to a shift in influences between Atlantic and continental weather systems. Located in the zone of westerly winds, climate in western Europe is strongly influenced by the Atlantic Ocean. This influence diminishes with increasing distance from the coast. The third components of runoff and precipitation also have similar spatial patterns and explain a considerable amount of the variance. The location of the pattern in the centre of the spatial domain, indicates that it may be related to the spatial distribution of the observations and thus may be an artifact of the analysis. However, comparable structures have been identified as the leading principal component of the standardised precipitation index on time-scales of 24 months (Bordi et al., 2009).

In the context of this study, the pronounced similarity to the equivalent patterns found in precipitation and temperature is a strong evidence that the space-time patterns of low-frequency runoff follow closely the atmospheric drivers. A full interpretation of the patterns found for precipitation and temperature falls in the domain of atmospheric sciences and is not within the scope of this study.

## 6 Conclusions

This study aimed at analysing the low-frequency variability of monthly European runoff and to provide insights into the controlling factors. It was shown that the space-time patterns of low-frequency runoff ( $Q_{\text{Long}}$ ) can be described by a few modes of oscillation that have their direct counterparts in precipitation and temperature. This demonstrated that continental-scale patterns of low-frequency runoff dynamics are directly driven by large-scale climatic variability.

The fraction of low-frequency variance of runoff ( $\Phi_Q$ ), however, was found to be on average larger than, and not

correlated to, the fraction of low-frequency variance of precipitation ( $\Phi_P$ ) and temperature ( $\Phi_T$ ), suggesting that catchment processes amplify low-frequency fluctuations in the forcing. The large spread of  $\Phi_Q$  across Europe indicates large differences in the sensitivity of runoff to low-frequency variability in the forcing.  $\Phi_Q$  is on average lowest in regions with a significant influence of snow on the hydro-climatic regime and here a decrease with temperature is found. In evapotranspiration dominated regimes,  $\Phi_Q$  is generally larger in catchments with lower annual runoff. Overall  $\Phi_Q$  was found to be largest for catchments with a dampened rainfall-runoff response. Further, the fraction of low-frequency variance of runoff, increases under drier and warmer conditions where catchments respond less directly to precipitation input.

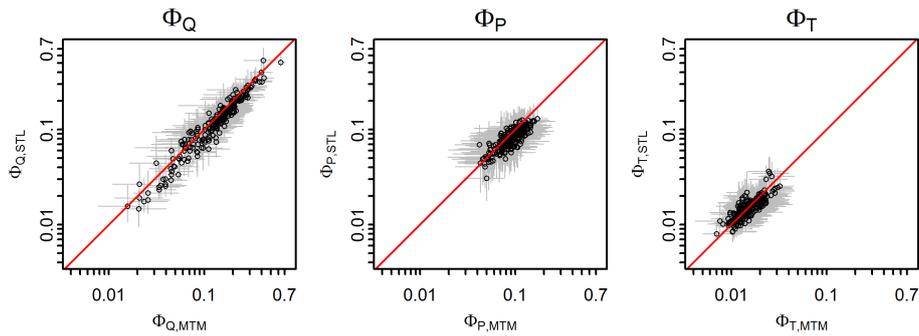
The mechanisms underlying these observations need to be explored in more detail. Nonetheless, these findings may be of interest for studies assessing the impact of climate variability and change on terrestrial water balance components. The influence of any climate signal may vary largely between rivers, depending on the long-term water budget. Climatic change, however, may influence the mean water budget, eventually changing  $\Phi_Q$ . In the case of increasingly wetter conditions, low-frequency runoff variability is likely to decline, simplifying water management on a year to year basis. In the case of increasingly drier conditions, low-frequency runoff variability is likely to increase, eventually decreasing predictability and challenge water management.

## Appendix A

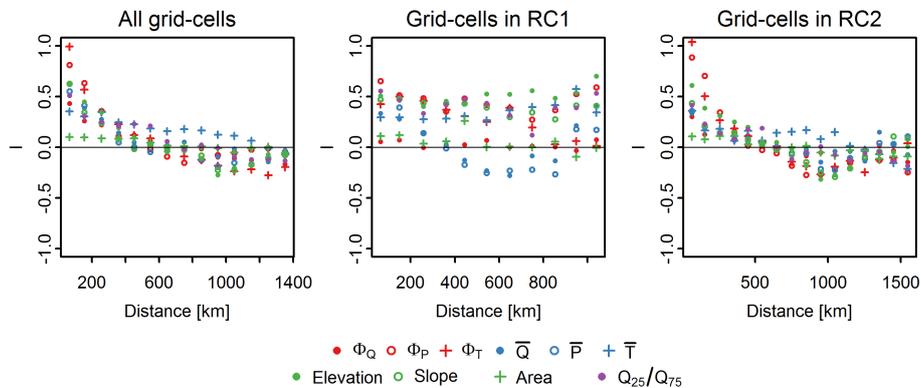
### Stability of $\Phi_X$

This study aimed at a parallel analysis of the space-time patterns of the low-frequency components of runoff as well as an analysis of its spectral properties. To achieve this  $\Phi_X$ , the fraction of low-frequency variance of the series  $X$ , characterising the shape of the power spectrum, was introduced. The time series analysis was based on the STL-algorithm (Cleveland et al., 1990), which is designed for time series decomposition. To assess whether the STL-algorithm produces biased estimates, a supplementary analysis was conducted, assessing the stability of the estimated values of  $\Phi_X$ . This test was performed by comparing  $\Phi_{X,\text{STL}}$ , the estimate used in the paper, to an alternative estimate,  $\Phi_{X,\text{MTM}}$ , which relies on the multi taper method (MTM; Thomson, 1982) to estimate spectral properties. The MTM is a reliable estimator of the spectral density and has been recommended for the analysis of climate variables (Ghil et al., 2002).

$\Phi_{X,\text{MTM}}$  was estimated by first computing the power-spectrum of  $X$  using MTM. Following the recommendations of Ghil et al. (2002) “discrete prolate spheroidal sequences” (DPSS) – tapers were used. In a second step,  $\Phi_{X,\text{MTM}}$  was determined as the fraction of variance explained by



**Fig. A1.** Comparison of  $\Phi_X$  estimated with the STL algorithm ( $\Phi_{X,STL}$ ) and the multi taper method ( $\Phi_{X,MTL}$ ) for runoff ( $\Phi_Q$ ), precipitation ( $\Phi_P$ ) and temperature ( $\Phi_T$ ). Grey bars are 95 % bootstrap confidence intervals.



**Fig. B1.** Spatial correlation functions for all variables of the correlation analysis, estimated using Moran’s  $I$ .

frequencies  $<1/19$  months (being consistent with the STL parameter  $\lambda_{Long}$ ). (MTM spectra were estimated using the R-package “sapa”; Constantine and Percival, 2010).

Confidence intervals of both  $\Phi_{X,STL}$  and  $\Phi_{X,MTM}$  were obtained using a bootstrapping procedure. To account for the strong seasonality and high serial correlation in the runoff time series, only the residuals of  $X$  with the mean annual cycle and the linear trend being removed were resampled. As the residuals may exhibit some serial correlation they were resampled using a moving block bootstrap, where blocks of a time series with fixed length are resampled instead of single data points (e.g. Efron and Tibshirani, 1993; Mudelsee, 2010). For each series the optimal block length  $n_b$  is estimated as a function of the lag one autocorrelation coefficient using the technique of Carlstein (1986) that has been adapted by Sherman et al. (1998) for moving block bootstrap. If the lag one autocorrelation is close to zero the, estimated block length of Carlstein (1986) may be zero. In such cases the block lengths were set to  $n_b = 1$ . The bootstrapping was based on 2000 replications and the bootstrap sample was used to construct bias-corrected and accelerated (BCa) 95 % confidence intervals (e.g. Efron and Tibshirani, 1993; Mudelsee, 2010). (Bootstrapping was performed using the R-packages “boot”, Canty and Ripley, 2011; “bootstrap”, Leisch, 2007, and “tseries”, Trapletti and Hornik, 2009).

Figure A1 summarises the results. The estimates of  $\Phi_{Q,STL}$  and  $\Phi_{Q,MTM}$  are closely related ( $R^2 = 0.93$ ), indicating that the two approaches provide quantitatively comparable results. The grey horizontal and vertical bars are the bootstrap confidence intervals of each data point and are measures of stability. For both measures, the confidence intervals have comparable magnitudes. This shows that both estimates of  $\Phi_Q$  yield quantitatively comparable and equally robust results. Both  $\Phi_{P,STL}$  and  $\Phi_{P,MTM}$  as well as  $\Phi_{T,STL}$  ( $R^2 = 0.66$ ) and  $\Phi_{T,MTM}$  ( $R^2 = 0.67$ ) have a lower correlation, which points toward some inconsistencies between the two estimates. Note however, that both  $\Phi_P$  and  $\Phi_T$  scatter over a much smaller range and that the confidence intervals of most data points do overlap. This suggests that the spatial variability of both  $\Phi_P$  and  $\Phi_T$  may not be significant, which is consistent with the finding that neither was found to be significantly correlated with  $\Phi_Q$  (see Table 1).

## Appendix B

### Moran’s $I$ for all variables

The significance of Spearman’s rank correlation coefficient was tested in this study taking spatial dependence into

account (Sect. 3.2.2). Spatial dependence was quantified using Moran's  $I$  (Moran, 1950) and Fig. B1 shows Moran's  $I$  for all the variables considered in the correlation analysis. Spatial correlation is generally highest for neighbouring points and declines with distance. Most variables have comparable degrees of spatial correlation if all grid-cells are considered. One exception is the mean catchment area, which exhibits almost zero correlations independent of distance. The spatial correlations in RC2 have a similar pattern, whereas grid-cells in RC1 have on average lower correlations and exhibit more scatter among variables. Likely, this can be related to the complex topography in the Alps and in Scandinavia.

**Acknowledgements.** This research contributes to the European Union (FP6) funded Integrated Project WATCH (Contract No. 036946). The provision of streamflow data by all agencies that contributed data to the EWA and to the WATCH project is gratefully acknowledged. The thorough review of M. Mudelsee, the comments of one anonymous reviewer and the remarks from the editor D. Koutsoyiannis helped increasing the clarity of the paper.

Edited by: D. Koutsoyiannis

## References

- Barlow, M., Nigam, S., and Berbery, E. H.: ENSO, Pacific Decadal Variability, and U.S. Summertime Precipitation, Drought, and Stream Flow, *J. Climate*, 14, 2105–2128, doi:10.1175/1520-0442(2001)014<2105:EPDVAU>2.0.CO;2, 2001.
- Barstad, I., Grabowski, W. W., and Smolarkiewicz, P. K.: Characteristics of large-scale orographic precipitation: Evaluation of linear model in idealized problems, *J. Hydrol.*, 340, 78–90, doi:10.1016/j.jhydrol.2007.04.005, 2007.
- Barstad, I., Sorteberg, A., Flatø, F., and Déqué, M.: Precipitation, temperature and wind in Norway: dynamical downscaling of ERA40, *Clim. Dynam.*, 33, 769–776, doi:10.1007/s00382-008-0476-5, 2009.
- Bierkens, M. F. P. and van den Hurk, B. J. J. M.: Groundwater convergence as a possible mechanism for multi-year persistence in rainfall, *Geophys. Res. Lett.*, 34, L02402, doi:10.1029/2006GL028396, 2007.
- Bjørnstad, O. N.: ncf: spatial nonparametric covariance functions, <http://CRAN.R-project.org/package=ncf>, last access: 9 September 2011, r package version 1.1-3, 2009.
- Bjørnstad, O. N. and Falck, W.: Nonparametric spatial covariance functions: Estimation and testing, *Environ. Ecol. Stat.*, 8, 53–70, doi:10.1023/A:1009601932481, 2001.
- Bordi, I., Fraedrich, K., and Sutera, A.: Observed drought and wetness trends in Europe: an update, *Hydrol. Earth Syst. Sci.*, 13, 1519–1530, doi:10.5194/hess-13-1519-2009, 2009.
- Borg, I. and Groenen, P. J. F.: *Modern Multidimensional Scaling – Theory and Applications*, Springer Series in Statistics, 2 Edn., Springer, 2005.
- Botter, G., Zanardo, S., Porporato, A., Rodriguez-Iturbe, I., and Rinaldo, A.: Ecohydrological model of flow duration curves and annual minima, *Water Resour. Res.*, 44, W08418, doi:10.1029/2008WR006814, 2008.
- Botter, G., Porporato, A., Rodriguez-Iturbe, I., and Rinaldo, A.: Nonlinear storage-discharge relations and catchment streamflow regimes, *Water Resour. Res.*, 45, W10427, doi:10.1029/2008WR007658, 2009.
- Bouwer, L. M., Vermaat, J. E., and Aerts, J. C. J. H.: Regional sensitivities of mean and peak river discharge to climate variability in Europe, *J. Geophys. Res.*, 113, D19103, doi:10.1029/2008JD010301, 2008.
- Canty, A. and Ripley, B. D.: boot: Bootstrap R (S-Plus) Functions, <http://CRAN.R-project.org/package=boot>, last access: 9 September 2011, r package version 1.3-2, 2011.
- Carlstein, E.: The Use of Subseries Values for Estimating the Variance of a General Statistic from a Stationary Sequence, *Ann. Stat.*, 14, 1171–1179, 1986.
- Castellari, A., Galeati, G., Brandimarte, L., Montanari, A., and Brath, A.: Regional flow-duration curves: reliability for ungauged basins, *Adv. Water Resour.*, 27, 953–965, doi:10.1016/j.advwatres.2004.08.005, 2004.
- Clark, M. P., Slater, A. G., Rupp, D. E., Woods, R. A., Vrugt, J. A., Gupta, H. V., Wagener, T., and Hay, L. E.: Framework for Understanding Structural Errors (FUSE): A modular framework to diagnose differences between hydrological models, *Water Resour. Res.*, 44, W00B02, doi:10.1029/2007WR006735, 2008.
- Clark, M. P., McMillan, H. K., Collins, D. B. G., Kavetski, D., and Woods, R. A.: Hydrological field data from a modeller's perspective: Part 2: process-based evaluation of model hypotheses, *Hydrol. Process.*, 25, 523–543, doi:10.1002/hyp.7902, 2011.
- Cleveland, R. B., Cleveland, W. S., Mcrae, J. E., and Terpenning, I.: STL: A Seasonal-Trend Decomposition Procedure Based on Loess, *J. Official Stat.*, 6, 3–73, 1990.
- Cleveland, W. S. and Devlin, S. J.: Locally Weighted Regression: An Approach to Regression Analysis by Local Fitting, *J. Am. Stat. Assoc.*, 83, 596–610, 1988.
- Clifford, P., Richardson, S., and Hemon, D.: Assessing the Significance of the Correlation between Two Spatial Processes, *Biometrics*, 45, 123–134, 1989.
- Constantine, W. and Percival, D.: sapa: Insightful Spectral Analysis for Physical Applications, REvolution Computing and Applied Physics Laboratory and University of Washington, <http://CRAN.R-project.org/package=sapa>, last access: 9 September 2011, r package version 1.0-3, 2010.
- Dale, M. and Fortin, M.-J.: Spatial autocorrelation and statistical tests: Some solutions, *J. Agr. Biol. Environ. St.*, 14, 188–206, doi:10.1198/jabes.2009.0012, 2009.
- Dijkstra, E. W.: A note on two problems in connexion with graphs, *Numer. Math.*, 1, 269–271, doi:10.1007/BF01386390, 1959.
- Dooge, J. C. I.: Sensitivity of Runoff to Climate Change: A Hortonian Approach, *B. Am. Meteorol. Soc.*, 73, 2013–2024, doi:10.1175/1520-0477(1992)073<2013:SORTCC>2.0.CO;2, 1992.
- Dutilleul, P., Clifford, P., Richardson, S., and Hemon, D.: Modifying the  $t$  Test for Assessing the Correlation Between Two Spatial Processes, *Biometrics*, 49, 305–314, 1993.

- Efron, B. and Tibshirani, R. J.: An Introduction to the bootstrap, no. 57 in *Monographs on Statistics and Applied Probability*, Chapman & Hall/CRC, Boca Raton, Fla., 1993.
- Gámez, A. J., Zhou, C. S., Timmermann, A., and Kurths, J.: Nonlinear dimensionality reduction in climate data, *Nonlin. Processes Geophys.*, 11, 393–398, doi:10.5194/npg-11-393-2004, 2004.
- Ghil, M., Allen, M. R., Dettinger, M. D., Ide, K., Kondrashov, D., Mann, M. E., Robertson, A. W., Saunders, A., Tian, Y., Varadi, F., and Yiou, P.: Advanced spectral methods for climatic time series, *Rev. Geophys.*, 40, 1003, doi:10.1029/2000RG000092, 2002.
- Gudmundsson, L., Tallaksen, L. M., and Stahl, K.: Spatial cross-correlation patterns of European low, mean and high flows, *Hydrol. Process.*, 25, 1034–1045, doi:10.1002/hyp.7807, 2011.
- Gustard, A., Bullock, A., and Dixon, J. M.: Low flow estimation in the United Kingdom, Tech. Rep. 108, Institute of Hydrology, <http://www.ceh.ac.uk/products/publications/lowflowestimationintheunitedkingdom.html> (last access: 15 May 2011), 1992.
- Haines, A. T., Finlayson, B. L., and McMahon, T. A.: A global classification of river regimes, *Appl. Geogr.*, 8, 255–272, doi:10.1016/0143-6228(88)90035-5, 1988.
- Haining, R.: Bivariate Correlation with Spatial Data, *Geogr. Anal.*, 23, 210–227, doi:10.1111/j.1538-4632.1991.tb00235.x, 1991.
- Hanson, R. T., Newhouse, M. W., and Dettinger, M. D.: A methodology to assess relations between climatic variability and variations in hydrologic time series in the southwestern United States, *J. Hydrol.*, 287, 252–269, 2004.
- Harris, N. M., Gurnell, A. M., Hannah, D. M., and Petts, G. E.: Classification of river regimes: a context for hydroecology, *Hydrol. Process.*, 14, 2831–2848, doi:10.1002/1099-1085(200011/12)14:16/17<2831::AID-HYP122>3.0.CO;2-O, 2000.
- Hisdal, H. and Tallaksen, L. M.: Estimation of regional meteorological and hydrological drought characteristics: a case study for Denmark, *J. Hydrol.*, 281, 230–247, doi:10.1016/S0022-1694(03)00233-6, 2003.
- Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N.-C., Tung, C. C., and Liu, H. H.: The Empirical Mode Decomposition and the Hilbert Spectrum for Nonlinear and Non-Stationary Time Series Analysis, *Proceedings: Mathematical, Phys. Eng. Sci.*, 454, 903–995, 1998.
- Hurrell, J. and van Loon, H.: Decadal variations in climate associated with the North Atlantic Oscillation, *Climatic Change*, 36, 301–326, doi:10.1023/A:1005314315270, 1997.
- Hurrell, J. W.: Decadal Trends in the North Atlantic Oscillation: Regional Temperatures and Precipitation, *Science*, 269, 676–679, doi:10.1126/science.269.5224.676, 1995.
- Jónsdóttir, J. F. and Uvo, C. B.: Long-term variability in precipitation and streamflow in Iceland and relations to atmospheric circulation, *Int. J. Climatol.*, 29, 1369–1380, doi:10.1002/joc.1781, 2009.
- Kalayci, S. and Kahya, E.: Assessment of streamflow variability modes in Turkey: 1964–1994, *J. Hydrol.*, 324, 163–177, doi:10.1016/j.jhydrol.2005.10.002, 2006.
- Kingston, D. G., Todd, M. C., Taylor, R. G., Thompson, J. R., and Arnell, N. W.: Uncertainty in the estimation of potential evapotranspiration under climate change, *Geophys. Res. Lett.*, 36, L20403, doi:10.1029/2009GL040267, 2009.
- Klemeš, V.: The Hurst Phenomenon: A puzzle?, *Water Resour. Res.*, 10, 675–688, doi:10.1029/WR010i004p00675, 1974.
- Koster, R. D. and Suarez, M. J.: A Simple Framework for Examining the Interannual Variability of Land Surface Moisture Fluxes, *J. Climate*, 12, 1911–1917, doi:10.1175/1520-0442(1999)012<1911:ASFFET>2.0.CO;2, 1999.
- Koutsoyiannis, D.: The Hurst phenomenon and fractional Gaussian noise made easy, *Hydrolog. Sci. J.*, 47, 573–595, 2002.
- Koutsoyiannis, D.: Climate change, the Hurst phenomenon, and hydrological statistics, *Hydrolog. Sci. J.*, 48, 3–24, 2003.
- Koutsoyiannis, D.: *HESS Opinions* “A random walk on water”, *Hydrol. Earth Syst. Sci.*, 14, 585–601, doi:10.5194/hess-14-585-2010, 2010.
- Krasovskaia, I., Gottschalk, L., Leblois, E., and Sauquet, E.: Dynamics of River Flow Regimes Viewed through Attractors, *Nord. Hydrol.*, 34, 461–476, 2003.
- Kumar, M. and Duffy, C. J.: Detecting Hydroclimatic Change Using Spatio-Temporal Analysis of Time Series in Colorado River Basin, *J. Hydrol.*, 374, 1–15, doi:10.1016/j.jhydrol.2009.03.039, 2009.
- Lange, H. and Bernhardt, K.: Long-term components and regional synchronization of river runoffs, in: *Hydrology: Science and Practice for the 21st Century*, edited by: Butler, A., British Hydrological Society, London, 165–170, 2004.
- Leisch, F.: bootstrap: Functions for the Book “An Introduction to the Bootstrap”, <http://CRAN.R-project.org/package=bootstrap>, last access: 9 September 2011, r package version 1.0-22, 2007.
- Lins, H. F.: Regional streamflow regimes and hydroclimatology of the United States, *Water Resour. Res.*, 33, 1655–1667, doi:10.1029/97WR00615, 1997.
- López-Moreno, J. I. and Vicente-Serrano, S. M.: Positive and Negative Phases of the Wintertime North Atlantic Oscillation and Drought Occurrence over Europe: A Multitemporal-Scale Approach, *J. Climate*, 21, 1220–1243, doi:10.1175/2007JCLI1739.1, 2008.
- Lorenzo-Lacruz, J., Vicente-Serrano, S. M., López-Moreno, J. I., González-Hidalgo, J. C., and Morán-Tejada, E.: The response of Iberian rivers to the North Atlantic Oscillation, *Hydrol. Earth Syst. Sci.*, 15, 2581–2597, doi:10.5194/hess-15-2581-2011, 2011.
- Mahecha, M. D., Martinez, A., Lischeid, G., and Beck, E.: Nonlinear dimensionality reduction: Alternative ordination approaches for extracting and visualizing biodiversity patterns in tropical montane forest vegetation data, *Ecol. Inform.*, 2, 138–149, doi:10.1016/j.ecoinf.2007.05.002, 2007.
- Mahecha, M. D., Fürst, L. M., Gobron, N., and Lange, H.: Identifying multiple spatiotemporal patterns: A refined view on terrestrial photosynthetic activity, *Pattern Recogn. Lett.*, 31, 2309–2317, doi:10.1016/j.patrec.2010.06.021, 2010.
- Mandelbrot, B. B. and Wallis, J. R.: Noah, Joseph, and Operational Hydrology, *Water Resour. Res.*, 4, 909–918, doi:10.1029/WR004i005p00909, 1968.
- Mandelbrot, B. B. and Wallis, J. R.: Some Long-Run Properties of Geophysical Records, *Water Resour. Res.*, 5, 321–340, doi:10.1029/WR005i002p00321, 1969.
- McMahon, T. A., Vogel, R. M., Peel, M. C., and Pegram, G. G.: Global streamflows – Part 1: Characteristics of annual streamflows, *J. Hydrol.*, 347, 243–259, doi:10.1016/j.jhydrol.2007.09.002, 2007.

- Milly, P. C. D. and Dunne, K. A.: Macroscale water fluxes 2. Water and energy supply control of their interannual variability, *Water Resour. Res.*, 38, 1206, doi:10.1029/2001WR000760, 2002.
- Milly, P. C. D. and Wetherald, R. T.: Macroscale water fluxes 3. Effects of land processes on variability of monthly river discharge, *Water Resour. Res.*, 38, 1235, doi:10.1029/2001WR000761, 2002.
- Monahan, A. H.: Nonlinear Principal Component Analysis: Tropical Indo-Pacific Sea Surface Temperature and Sea Level Pressure, *J. Climate*, 14, 219–233, doi:10.1175/1520-0442(2001)013<0219:NPACAT>2.0.CO;2, 2001.
- Montanari, A., Rosso, R., and Taquq, M. S.: Fractionally Differenced ARIMA Models Applied to Hydrologic Time Series: Identification, Estimation, and Simulation, *Water Resour. Res.*, 33, 1035–1044, doi:10.1029/97WR00043, 1997.
- Moran, P. A. P.: Notes on Continuous Stochastic Phenomena, *Biometrika*, 37, 17–23, 1950.
- Mudelsee, M.: Long memory of rivers from spatial aggregation, *Water Resour. Res.*, 43, W01202, doi:10.1029/2006WR005721, 2007.
- Mudelsee, M.: Climate Time Series Analysis: Classical Statistical and Bootstrap Methods, vol. 42 of Atmospheric and Oceanographic Sciences Library, Springer, Dordrecht, doi:10.1007/978-90-481-9482-7, 2010.
- Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. H., and Wagner, H.: *vegan: Community Ecology Package*, <http://CRAN.R-project.org/package=vegan>, last access: 9 September 2011, r package version 1.17-4, 2010.
- Peltonen, M., Liebhold, A. M., Bjornstad, O. N., and Williams, D. W.: Spatial synchrony in forest insect outbreaks: roles of regional stochasticity and dispersal, *Ecology*, 83, 3120–3129, doi:10.1890/0012-9658(2002)083[3120:SSIFIO]2.0.CO;2, 2002.
- Pepin, N. C., Daly, C., and Lundquist, J.: The influence of surface versus free-air decoupling on temperature trend patterns in the western United States, *J. Geophys. Res.*, 116, D10109, doi:10.1029/2010JD014769, 2011.
- Peres-Neto, P. and Jackson, D.: How well do multivariate data sets match? The advantages of a Procrustean superimposition approach over the Mantel test, *Oecologia*, 129, 169–178, doi:10.1007/s004420100720, 2001.
- Potter, K. W.: Evidence for nonstationarity as a physical explanation of the Hurst Phenomenon, *Water Resour. Res.*, 12, 1047–1052, doi:10.1029/WR012i005p01047, 1976.
- R Development Core Team: *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, <http://www.R-project.org>, last access: 15 May 2011.
- Renner, M. and Bernhofer, C.: Long term variability of the annual hydrological regime and sensitivity to temperature phase shifts in Saxony/Germany, *Hydrol. Earth Syst. Sci.*, 15, 1819–1833, doi:10.5194/hess-15-1819-2011, 2011.
- Sankarasubramanian, A. and Vogel, R. M.: Annual hydroclimatology of the United States, *Water Resour. Res.*, 38, 1083, doi:10.1029/2001WR000619, 2002.
- Sankarasubramanian, A., Vogel, R. M., and Limbrunner, J. F.: Climate Elasticity of Streamflow in the United States, *Water Resour. Res.*, 37, 1771–1781, doi:10.1029/2000WR900330, 2000.
- Shamsudduha, M., Chandler, R. E., Taylor, R. G., and Ahmed, K. M.: Recent trends in groundwater levels in a highly seasonal hydrological system: the Ganges-Brahmaputra-Meghna Delta, *Hydrol. Earth Syst. Sci.*, 13, 2373–2385, doi:10.5194/hess-13-2373-2009, 2009.
- Sherman, M., Speed, F. M., and Speed, F. M.: Analysis of tidal data via the blockwise bootstrap, *J. Appl. Stat.*, 25, 333–340, 1998.
- Shorthouse, C. and Arnell, N.: FRIEND'97 – Regional Hydrology: Concepts and Models for Sustainable Water Resource Management, vol. 246, chap. Spatial and temporal variability in European river flows and the North Atlantic oscillation, The International Association of Hydrological Sciences (IAHS), 77–85, 1997.
- Shorthouse, C. and Arnell, N.: The effects of climatic variability on spatial characteristics of European river flows, *Phys. Chem. Earth B*, 24, 7–13, doi:10.1016/S1464-1909(98)00003-3, 1999.
- Shun, T. and Duffy, C. J.: Low-frequency oscillations in precipitation, temperature, and runoff on a west facing mountain front: A hydrogeologic interpretation, *Water Resour. Res.*, 35, 191–201, doi:10.1029/98WR02818, 1999.
- Spearman, C.: The Proof and Measurement of Association between Two Things, *Am. J. Psychol.*, 100, 441–471, 1987.
- Stahl, K. and Demuth, S.: Linking streamflow drought to the occurrence of atmospheric circulation patterns, *Hydrolog. Sci. J. – Journal Des Sciences Hydrologiques*, 44, 467–482, 1999.
- Stahl, K., Hisdal, H., Tallaksen, L., van Lanen, H., Hannaford, J., and Sauquet, E.: Trends in low flows and streamflow droughts across Europe, Tech. rep., UNESCO, Paris, 2008.
- Stahl, K., Hisdal, H., Hannaford, J., Tallaksen, L. M., van Lanen, H. A. J., Sauquet, E., Demuth, S., Fendekova, M., and Jódar, J.: Streamflow trends in Europe: evidence from a dataset of near-natural catchments, *Hydrol. Earth Syst. Sci.*, 14, 2367–2382, doi:10.5194/hess-14-2367-2010, 2010.
- Tallaksen, L. M., Madsen, H., and Clausen, B.: On the definition and modelling of streamflow drought duration and deficit volume, *Hydrolog. Sci. J. – Journal Des Sciences Hydrologiques*, 42, 15–33, 1997.
- Tenenbaum, J. B., de Silva, V., and Langford, J. C.: A global geometric framework for nonlinear dimensionality reduction, *Science*, 290, 2319–2323, 2000.
- Thomson, D.: Spectrum estimation and harmonic analysis, *Proc. IEEE*, 70, 1055–1096, 1982.
- Tootle, G. A. and Piechota, T. C.: Relationships between Pacific and Atlantic ocean sea surface temperatures and U.S. streamflow variability, *Water Resour. Res.*, 42, W07411, doi:10.1029/2005WR004184, 2006.
- Torgerson, W. S.: Multidimensional scaling: I. Theory and method, *Psychometrika*, 17, 401–419, 1952.
- Torrence, C. and Compo, G. P.: A Practical Guide to Wavelet Analysis, *B. Am. Meteorol. Soc.*, 79, 61–78, doi:10.1175/1520-0477(1998)079<0061:APGTWA>2.0.CO;2, 1998.
- Trapletti, A. and Hornik, K.: *tseries: Time Series Analysis and Computational Finance*, <http://CRAN.R-project.org/package=tseries>, last access: 9 September 2011, r package version 0.10-22., 2009.
- Uppala, S. M., Kållberg, P. W., Simmons, A. J., Andrae, U., Bechtold, V. D. C., Fiorino, M., Gibson, J. K., Haseler, J., Hernandez, A., Kelly, G. A., Li, X., Onogi, K., Saarinen, S., Sokka, N., Allan, R. P., Andersson, E., Arpe, K., Balmaseda, M. A., Beljaars,

- A. C. M., Berg, L. V. D., Bidlot, J., Bormann, N., Caires, S., Chevallier, F., Dethof, A., Dragosavac, M., Fisher, M., Fuentes, M., Hagemann, S., Hólm, E., Hoskins, B. J., Isaksen, L., Janssen, P. A. E. M., Jenne, R., McNally, A. P., Mahfouf, J.-F., Morcrette, J.-J., Rayner, N. A., Saunders, R. W., Simon, P., Sterl, A., Trenberth, K. E., Untch, A., Vasiljevic, D., Viterbo, P., and Woollen, J.: The ERA-40 re-analysis, *Q. J. Roy. Meteorol. Soc.*, 131, 2961–3012, doi:10.1256/qj.04.176, 2005.
- van Loon, H. and Rogers, J. C.: The Seesaw in Winter Temperatures between Greenland and Northern Europe, Part I: General Description, *Mon. Weather Rev.*, 106, 296–310, doi:10.1175/1520-0493(1978)106<0296:TSIWTB>2.0.CO;2, 1978.
- Vogel, R. M. and Fennessey, N. M.: Flow-duration, I: New interpretation and confidence intervals, *J. Water Resour. Plan. Manage.*, 120, 458–504, 1994.
- Vogel, R. M., Tsai, Y., and Limbrunner, J. F.: The Regional Persistence and Variability of Annual Streamflow in the United States, *Water Resour. Res.*, 34, 3445–3459, doi:10.1029/98WR02523, 1998.
- Vogt, J., Soille, P., de Jager, A., Rimaviciute, E., and Mehl, W.: A pan-European River and Catchment Database, Tech. Rep. EUR 22920 EN, EC-JRC, Luxembourg, [http://ccm.jrc.ec.europa.eu/documents/CCM2-Report\\_EUR-22920-EN\\_2007\\_STD.pdf](http://ccm.jrc.ec.europa.eu/documents/CCM2-Report_EUR-22920-EN_2007_STD.pdf) (last access: 7 August 2011), 2007.
- von Storch, H. and Zwiers, F. W.: Statistical analysis in climate research, Cambridge University Press, Cambridge, UK, 1999.
- Ward, J.: Hierarchical Clustering to Optimize an Objective Function, *J. Am. Stat. Assoc.*, 58, 236–244, 1967.
- Weedon, G., Gomes, S., Viterbo, P., Österle, H., Adam, J., Bellouin, N., Boucher, O., and Best, M.: The WATCH forcing data 1958–2001: a meteorological forcing dataset for land surface- and hydrological-models, WATCH Technical Report 22, <http://www.eu-watch.org> (last access: 15 May 2011), 2010.
- Weedon, G. P., Gomes, S., Viterbo, P., Shuttleworth, W. J., Blyth, E., Österle, H., Adam, J. C., Bellouin, N., Boucher, O., and Best, M.: Creation of the WATCH Forcing Data and its use to assess global and regional reference crop evaporation over land during the twentieth century, *J. Hydrometeorol.*, in press, doi:10.1175/2011JHM1369.1, 2011.
- Wibig, J.: Precipitation in Europe in relation to circulation patterns at the 500 hPa level, *Int. J. Climatol.*, 19, 253–269, doi:10.1002/(SICI)1097-0088(19990315)19:3<253::AID-JOC366>3.0.CO;2-0, 1999.
- Wilcoxon, F.: Individual Comparisons by Ranking Methods, *Biometrics Bull.*, 1, 80–83, 1945.