# Large-Scale Hydrology in Europe

## Observed Patterns and Model Performance

Dissertation for the degree of
Philosophiae Doctor (Ph.D.)

Lukas Gudmundsson

Department of Geosciences
University of Oslo

June 2011

*Series of dissertations submitted to the*
*Faculty of Mathematics and Natural Sciences, University of Oslo*
*No. 1109*

ISSN 1501-7710



Cover: Inger Sandved Anfinsen.
Printed in Norway: AIT Oslo AS.

Produced in co-operation with Unipub.
The thesis is produced by Unipub merely in connection with the
thesis defence. Kindly direct all inquiries regarding the thesis to the copyright
holder or the unit which grants the doctorate.

# Synopsis

In a changing climate, terrestrial water storages are of great interest as water availability impacts key aspects of ecosystem functioning. Thus, a better understanding of the variations of wet and dry periods will contribute to fully grasp processes of the earth system such as nutrient cycling and vegetation dynamics. Currently, river runoff from small, nearly natural, catchments is one of the few variables of the terrestrial water balance that is regularly monitored with detailed spatial and temporal coverage on large scales. River runoff, therefore, provides a foundation to approach European hydrology with respect to observed patterns on large scales, with regard to the ability of models to capture these.

The analysis of observed river flow from small catchments, focused on the identification and description of spatial patterns of simultaneous temporal variations of runoff. These are dominated by large-scale variations of climatic variables but also altered by catchment processes. It was shown that time series of annual low, mean and high flows follow the same atmospheric drivers. The observation that high flows are more closely coupled to large scale atmospheric drivers than low flows, indicates the increasing influence of catchment properties on runoff under dry conditions. Further, it was shown that the low-frequency variability of European runoff is dominated by two opposing centres of simultaneous variations, such that dry years in the north are accompanied by wet years in the south.

Large-scale hydrological models are simplified representations of our current perception of the terrestrial water balance on large scales. Quantification of the models strengths and weaknesses is the prerequisite for a reliable interpretation of simulation results. Model evaluations may also enable to detect shortcomings with model assumptions and thus enable a refinement of the current perception of hydrological systems. The ability of a multi model ensemble of nine large-scale hydrological models and the land surface scheme of a high resolution regional climate model to capture various aspects of runoff were assessed. In general high and mean flows were better captured than low flows, pointing toward deficiencies with the models representation of storage processes. The analysis of the multi model ensemble also showed that the mean of all model simulations generally provided a more robust and accurate estimator of continental scale variations of catchment runoff than any individual model.

# Acknowledgements

I want to express my gratitude to Prof. Dr. Lena M. Tallaksen for the insightful guidance through the many stages of this thesis. I am grateful for her patience, for keeping me on track and for the large amount of liberty that I had to develop my scientific ideas. I am also indebted to Dr. Kerstin Stahl from the University of Freiburg for keeping me motivated and for her refreshing ideas. Her countless contributions have always been a great help. I also want to thank Dr. Kolbjørn Engeland from SINTEF Energy Research, for the many accurate suggestions, often pointing me with only few words toward the right direction.

I appreciated very much the stimulating discussions and the hospitality of Prof. Dr. Thorsten Wagener from the Pennsylvania State University who hosted me one summer at his institution.

This dissertation would possibly never have been written without Prof. Dr. Holger Lange from the Norwegian Forest and Landscape Institute who supported my move to Norway. His creativity and inquisitiveness have always been a source of inspiration.

The journal articles constituting the main body of this thesis would not be the same without the help of my many co-authors. I want to thank all of them for the fruitful discussions and contributions.

I am in debit to all my friends and colleagues at the Department of Geosciences, for the many inspiring conversations. Special thanks go to Anne Fleig and Stian Solvang Johansen for the "hydrological" discussions and also for buffering some of the scientific frustrations that arose throughout this research endeavour. Many thanks to Kirsten Fristad for proof reading the introductory chapters of this dissertation.

This work would not have been possible without the excellent resources provided by the Water and Global Change project (WATCH, Contract No. 036946), funded by the European Union. Many of my travels to conferences and the stay at the Pennsylvania State University, would not have been possible without the financial support of the Industrial Liaison (IL) fund.

I want to thank all my friends in Oslo, for introducing me to the many things this city has to offer. I have not only shared home with my flatmates Anne Gerd and Magnus, they also have become dear friends. Indre, Petr and Marianne, with whom I have organised the *moloch filmklubb* on an almost weekly basis, where always good in distracting me from the sometimes tiring academic business.

Finally, I want to thank my parents – for the alls support throughout the years – and my tow brothers – for regularly reminding me on the limits of natural sciences.

Lukas Gudmundsson

# Contents

# Part A.

# Introduction and Synthesis

# 1. Introduction

Terrestrial water stores and fluxes are key variables controlling many aspects of ecosystem- and climate dynamics. Examples include a wide range of phenomena ranging from micro-scale processes related to nutrient turnover in soils to regional scale influences of precipitation patterns. Many of the processes underlying these phenomena happen on very small scales and their physical, chemical and biological principles are relatively well understood. On large (e.g. ecosystem and landscape) scales, however, these small scale principles are often not sufficient to explain the full range of natural variability. The reasons for this are manifold and rooted in the complexity of environmental systems. However, in many instances the large-scale environmental settings such as the geological framework and average climatic conditions define a framework that restricts the observed phenomena to a limited number of possibilities [e.g. *Raich and Potter*, 1995; *Clark et al.*, 2001]. The rationale underlying this is based on the fact that many small scale processes are controlled by the availability of water, energy or nutrients. Vegetation, for example, is often regarded as a function of environmental conditions [e.g. *Haxeltine et al.*, 1996; *Knapp et al.*, 2002; *Caylor et al.*, 2006] and this relation has been used to predict global vegetation patterns as a function of mean water availability [e.g. *Stephenson*, 1990; *Gerten et al.*, 2004]. Other interesting aspects are nutrient and carbon cycling, which largely depend on average water availability [e.g. *Davidson et al.*, 2000; *Reichstein et al.*, 2003; *Lohse et al.*, 2009; *Ju et al.*, 2010]. In recent years, the influence of episodic changes in terrestrial water availability on large-scale carbon cycling, including droughts [e.g. *Saigusa et al.*, 2010; *Ju et al.*, 2010], has received increasing attention. Theoretical [e.g. *Porporato et al.*, 2004] and experimental [e.g. *Knapp et al.*, 2002; *Miller et al.*, 2005; *Knorr and Blodau*, 2009] investigations have shown that changes in the dynamical properties of terrestrial water availability may lead to large changes in carbon and nutrient dynamics.

Terrestrial water storage does not only impact ecosystem processes but also influences atmospheric dynamics [*Seneviratne et al.*, 2010, and references therein]. Commonly this influence is described by two positive feedback loops, describing the coupling of soil-moisture, precipitation and temperature. The soil-moisture precipitation feedback loop is based on the fact that evaporation rates are higher for wet soils, leading to increasing precipitation, which in turn causes an increase in soil moisture [e.g. *Findell and Eltahir*, 1997; *Koster and Suarez*, 2001; *Koster et al.*, 2004, 2003]. The soil-moisture temperature feedback loop is linked to evaporation processes. Evaporation requires energy and thus leads to a reduction in air temperature. On the other hand, if soil moisture is low, less water is being evaporated and thus the cooling effect of evaporation is reduced, leading to increasing temperature. This soil-moisture temperature feedback has been suggested as a key factor for the generation of heat-waves and droughts [e.g. *Fischer et al.*, 2007; *Zampieri et al.*, 2009].

As the influence of terrestrial water storages and fluxes on various aspects of the earth system are increasingly recognised, the need for a sound understanding of their fluctuations on large, continental, scales is of a vital interest. Both for the scientific community, trying to disentangle the complexity of the earth system, as well as for the many practitioners dealing with the challenges of resources management on a daily basis. This dissertation aims to contribute to this task with a special emphasis on the large-scale hydrology in Europe. Focus is put both on the characterisation of observed patterns of runoff on continental scales, and on the skill of comprehensive large-scale hydrological models to important features of runoff variability in Europe.

This introduction is continued with a summary of the physical framework underlying the terrestrial water balance (Section 1.1) and a review that covers an overview on relevant observations (Section 1.2), a summary of observed runoff patterns on large scales, with emphasis on Europe (Section 1.3), and a recapitulation of the principals of large-scale hydrological modelling. (Section 1.4). Finally the scope of the thesis is specified in more detail in Section 2. The main body of this dissertation is a collection of five Articles (Articles I, II, III, IV and V), which are summarised in Chapter 3, where general conclusions are also drawn.

# 1.1. A Physical Framework

## 1.1.1. The Terrestrial Water Balance

The fundamental principle underlying all hydrological phenomena is the terrestrial water balance

$$\frac{dS}{dt} = P - E - Q, \tag{1.1}$$

where $dS/dt$ describes changes in the total amount of water stored in a land unit. The terrestrial water storage $S$, includes amongst others, water stored in vegetation, snow, soil-moisture, groundwater and surface water bodies. Precipitation, $P$, denotes atmospheric water input, either as rain or snowfall, and $E$ denotes Evapotranspiration. The runoff term $Q$ stands for all water flowing out of the land unit, both at the surface as streamflow and below the surface as groundwater runoff. Runoff is usually assumed to be a function of storage

$$Q = h(S), \tag{1.2}$$

where the function $h$ summarises all relevant terrestrial drainage of water. The function $h$ is subject to hydrological research and a many possible formulations are known, each attempting to describe the most relevant processes. Principally, all water movements are based on small-scale physical principles such as the flow through porous media (soils) or channels (macro-pores, river beds) and plant water uptake. However, due to the complexity of natural systems and the fact that large parts of the subsurface are unobservable, physical descriptions of hydrological systems incorporate simplifications. These simplifications are often tailored for distinct spatial or temporal resolutions. Due to the need for simplification, however, the form of $h$ has many degrees of freedom, resulting in a variety of many plausible forms. In general $h$ is assumed to be dependent on land properties such as topography, vegetation cover, storage capacity and hydraulic conductivity of soils and aquifers. (See *Clark et al.* [2008, 2011] for a comprehensive overview of so called lumped hydrological models that have a large degree of spatial aggregation and *Kampf and Burges* [2007] for so called distributed hydrological models that describe the spatial distribution of water flows explicitly). Evapotranspiration $E$, the other flux of water leaving the system, couples the terrestrial water balance to the surface energy balance

$$\frac{dH}{dt} = R_n - \lambda E - SH - G, \tag{1.3}$$

where $dH/dt$ describes changes in the surface energy balance, $R_n$ is net radiation, $\lambda E$ the latent heat flux, $SH$ the sensible heat flux and $G$ the ground heat flux. The latent heat of vaporisation $\lambda$ is the energy required to evaporate a given quantity of water. Evapotranspiration is, similar to runoff, a function of terrestrial water storage but also dependent on the amount of energy that is available to evaporate water, such that

$$E = \beta(S) \times E_p, \tag{1.4}$$

where the function $\beta$ describes the release of water from soils and accounts for other influencing factors such as air humidity, wind speed and plant water uptake [e.g. *Allen et al.*, 1998]. $E_p$ is the potential evaporation and under idealised conditions, when all energy is converted into evapotranspiration and none into heat, potential evaporation is defined as $E_p = R_n/\lambda$ [*Arora*, 2002; *Gerrits et al.*, 2009].

The key variables of the terrestrial water cycle that are of interest for many aspects of ecosystem and climate dynamics are storage variables, especially soil moisture and ground water. As these variables can often not be directly monitored with the desired spatial and temporal resolution, it is useful to recall that runoff is a function of storage (Eq. (1.2)). Thus runoff rates can be used to infer information on terrestrial water storage by inverting the storage-discharge relation (Eq. (1.2)) such that

$$S = h^{-1}(Q), \tag{1.5}$$

[e.g. *Kirchner*, 2009]. Even if the form of the function $h(S)$, respectively $h^{-1}(Q)$ is not known, it implies that any fluctuations in (observed) $Q$ are proportional to fluctuations in the terrestrial water storage.

### 1.1.2. Temporal and Spatial Scales

Many environmental phenomena can be characterised by their spatial and temporal scales. Although only vaguely defined in the context of hydrology [e.g. *Blöschl and Sivapalan*, 1995; *Klemeš*, 1983] and atmospheric sciences [e.g. *von Storch and Zwiers*, 1999], the following qualitative definitions of temporal and spatial scale, following *von Storch and Zwiers* [1999, chapter 3.0.3], are sufficient for many purposes: The *timescale* is the characteristic duration of the phenomenon/process of interest. The *spatial scale* is the characteristic length that is representative for the spatial variations relevant to the phenomenon/process of interest.

From a quantitative perspective, temporal and spatial scales are often characterised by means of a de-correlation time or length [e.g. *von Storch and Zwiers*, 1999; *Skøien et al.*, 2003], that is the distance in time or space at which two observations no longer share common variations. De-correlation times or lengths can be quantified using statistical techniques such as the autocorrelation function or the semi variogram. For periodical phenomena, such as the seasonal cycle (temporal) or ripples in the sediment of a river bed (spatial), the characteristic scale is often defined by means of the period (temporal scale) or the wavelength (spatial scale). Usually scales are quantified not by exact numbers but by orders of magnitude. For spatial scales often multiples of ten of a length measure are used e.g. 1km, 10km and 100km. Similarly, temporal scales can be characterised by "days", "months" and "years".

Historically, hydrology has mostly been focusing on temporal phenomena, often with the aim of predicting river flow. As the temporal evolution of hydrological variables is directly related the atmospheric forcing, the temporal scales of hydrological phenomena is often characterised in conjunction with the timescale of meteorological variables [*Blöschl and Sivapalan*, 1995]. The term "event scale" is usually used to characterise the duration of runoff peaks, that follow rainfall events. Depending on the hydrological system, the event scale is usually considered to have an order of magnitude of days. The "seasonal scale" is determined by the periodic annual cycle, which is driven by the revolution of the earth around the sun and impacts hydrological phenomena by changing meteorological conditions. Further, timescales longer than one year, often referred to as "inter-annual", "long term" or "climatic" timescales, are usually assumed to reflect the response of hydrological systems to climatic variability.

Spatial scales of hydrological phenomena have often been associated with the spatial properties of landscapes, including topographic features and sub surface properties [*Blöschl and Sivapalan*, 1995]. At the "local" or "plot scale" (order of magnitude: 1 m) the hydraulic properties of the soils including macro pores (e.g. root channels) influences the capacity of soils to act as a medium for water storage and transport. At the "catchment scale" (order of magnitude: 10 km) both topographic features and the spatial distribution of soil types determine how precipitation is translated to runoff. Finally, at the "regional" (order of magnitude 1000 km) scale geology determines predominant soil types and also large-scale topographic features, including stream network density. However, the spatial scale of hydrological phenomena does not only depend on the properties of landscape, but also on the spatial scales of the atmospheric forcing, i.e. precipitation and the atmospheric water demand driving evapotranspiration. Convective precipitation and thunderstorms have, for example, often relatively short durations (a few hours) and a small spatial extent (some kilometres). This is contrasted by the large areas covered by low and high pressure systems, which last from several days to possibly several weeks. Finally, the mean climatic conditions, including the annual cycle of precipitation and temperature, cover even larger regions. The spatial scale of terrestrial hydrological phenomena is thus a result of the combined effects of land properties and atmospheric forcing.

To assess the different roles of land properties and atmospheric forcing on terrestrial water storages, in particular soil moisture, *Vinnikov et al.* [1996] suggested that soil moisture has two distinct spatial scales. A short, *land property scale* that is related to small scale variations of land properties such as soil texture and micro topography and a longer *atmospheric scale* that reflects the spatial scale of the atmospheric forcing (Figure 1.1). *Vinnikov et al.* [1996] hypothesised that both the land property scale and the atmospheric scale of soil moisture can be characterised using two spatial (and temporal) autocorrelation functions with an exponential decay law. At short distances the spatial autocorrelation decreases rapidly due to the heterogeneity of land properties. At long distances and long time intervals, however, the autocorrelation decreases slower and follows atmospheric variability.

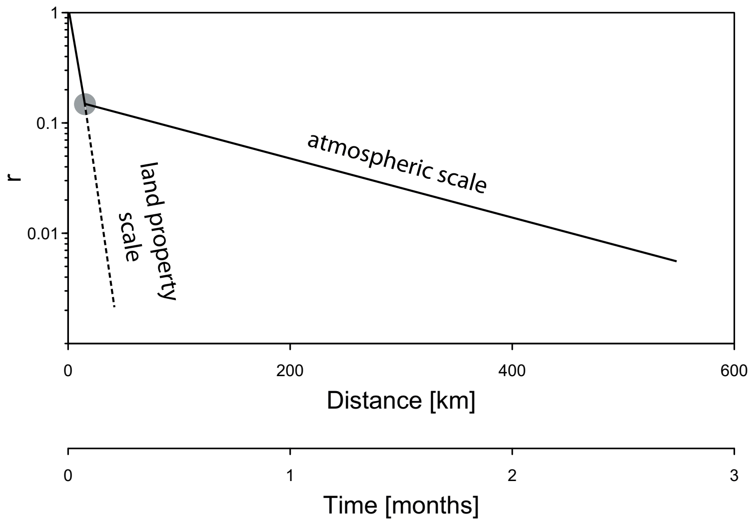


**Figure 1.1.:** Schematic diagram of hydrological and meteorological scales of soil moisture variations. $r$ is the autocorrelation function. The scales are determined by the slopes of the curves: for time scales, $r(t) = e^{-t/T}$, $t$ is time and $T$ is the timescale; for spatial scales, $r(l) = e^{-l/L}$, $l$ is distance and $T$ is the length scale. Adapted from *Robock et al.* [1998], modified annotation.

Empirical investigations have indeed suggested that soil moisture has two scales of variability, where the land property scale has an order of magnitude of tens of meters and the atmospheric scale has an order of magnitude of hundreds of kilometres [*Entin et al.*, 2000; *Robock et al.*, 1998; *Vinnikov et al.*, 1996]. This is in agreement with the common notion that soil moisture has a huge amount of spatial variability in the field [e.g. *Famiglietti et al.*, 2008; *Wilson et al.*, 2004; *Daly and Porporato*, 2005; *Rodríguez-Iturbe et al.*, 2006; *Western et al.*, 2002] and the observation that satellite based, continental scale, estimates of soil-moisture are highly correlated to the equivalent precipitation and evapotranspiration patterns [*Liu et al.*, 2009].

River flow, the most commonly observed hydrological variable, is in contrast to many other variables of the terrestrial water balance (including soil moisture, ground water and evaporation), not continuous in space. River flow is constrained to the drainage network, which extends over a multitude of spatial scales, from micro topography to large continental river basins. Thus, the spatial scale of a continental river system may be predominantly determined and constrained by the drainage network and the temporal scale is likely to be primarily controlled by hydrodynamic routing. However, regions of simultaneous variation of seasonal river flow from large number of stations from all over the United States were not in correspondence with large-scale drainage basins [*Lins*, 1997] but rather related to features of large-scale atmospheric variability [*Tootle and Piechota*, 2006; *Barlow et al.*, 2001]. Similarly, process controls of the seasonality of runoff (i.e. evapotranspiration and snow processes) do not correspond to large-scale drainage basins but rather vary with the mean climate. This suggests that the spatial scale of river flow is determined by the atmospheric drivers – at least if the scale of considered catchment is smaller than the scale of the atmospheric forcing variables. As catchments differ substantially in their characteristics, the "land property scale" of river flow is likely to be larger than the one of soil moisture suggested by *Vinnikov et al.* [1996]. *Skøien et al.* [2003] estimated the spatial scale of runoff from a large number of catchments in Austria as the $e$-folding distance, which is the distance at which the spatial variogram reaches $1 - 1/e$ of its maximum. They found that the spatial scale of river flow ranged between 18 and 59 km, depending on estimation technique and catchment size. As the spatial extent of this study was limited to Austria it is likely that the atmospheric scale of streamflow is not fully resolved. Thus, these results suggest that the "land property scale" of catchment runoff may be an order of magnitude of tenths of kilometres in regions with climates and landscapes that are comparable to Austria.

## 1.2. Observational Data

### 1.2.1. Terrestrial Variables

To better understand terrestrial water storage and flux and their influence on climate and ecosystem processes, observations are essential. Ideally observations would cover information on the temporal evolution of the amount of water stored in a landscape unit (state variables) as well as information on the amount of water leaving the system (flux variables). To date, however, the only variable of the terrestrial water cycle being reliably monitored with relatively high spatial and temporal coverage is *river discharge*, also referred to as (*catchment*) *runoff*, *river flow* and *streamflow*. River discharge provides information on the spatially integrated outflow from water stores for a given catchment. The temporal and spatial coverage of gauging stations varies between regions, and observations from small catchments and large river basins cannot be easily compared. *Hannah et al.* [2010, 2011] provide a comprehensive overview on various large-scale river flow archives. Examples include the large amount of discharge series available from the U.S. Geological Survey[1] (USGS) which provide also the quality assured observations collected in the Hydro-Climatic Data Network (HCDN) [*Slack and Landwehr*, 1992] and the Global Runoff Data Center[2] (GRDC) which hosts several data sources, such as discharge series from the worlds largest rivers (e.g. Rhine, Danube, Congo) and some of their tributaries as well as the European Water Archive (EWA). The EWA is a collection of a large number of gauging stations in Europe, gathered by the European partners of the Flow Regime from International Experimental and Network Data[3] (FRIEND) program, and is to date the data source with the most exhaustive coverage of streamflow in Europe.

Observational networks of water stores that are suited for studying large-scale hydrology are scarce. Albeit groundwater is monitored regularly in the context of local resource management, data availability varies with national authorities and no harmonised transnational data products exist. The situation is slightly better for soil moisture where the Global Soil Moisture Data Bank [*Robock et al.*, 2000] collects a large number of observations based on gravimetric measurements. The observations are concentrated on grassland in mid-latitude regions of the Eurasian continent. A variety of other in-situ observation techniques are increasingly used (see *Vereecken et al.* [2008] for a review) and their systematic application may eventually augment available data sources.

The principal limitation of any observation network is, besides the observed time window, the distribution of stations. Stations are usually clustered in regions with high population density and high levels of industrialisation. However, the use of

[1] http://www.usgs.gov/, last accessed: 4 June 2011
[2] www.grdc.bafg.de, last accessed: 4 June 2011
[3] http://ne-friend.bafg.de, last accessed: 4 June 2011

modern remote sensing techniques are promising. In recent years, soil moisture estimates based on data from passive and active microwave sensors are becoming increasingly available [e.g. *de Jeu et al.*, 2008; *Wagner et al.*, 2007] and some data products combining the information of multiple sensors extend as far back as the late 1970s [e.g. *Owe et al.*, 2008]. However, microwave based soil moisture estimates have a relatively large uncertainty in regions with high vegetation density [e.g. *de Jeu et al.*, 2008]. The twin-satellites of the Gravity Recovery and Climate Experiment (GRACE) measure variations in the gravity field of the earth with approximately monthly resolution [*Ramillien et al.*, 2008; *Tapley et al.*, 2004; *Rodell and Famiglietti*, 1999]. Regional accumulations of water (e.g. ground water, soil moisture and snow cover) change the gravity field, which in turn can be used to retrieve estimates of fluctuations in the total terrestrial water storage. In recent years several studies have demonstrated the reliability of GRACE products for regional investigation of terrestrial water stores [e.g. *Chen et al.*, 2010; *Schmidt et al.*, 2006] and some advances have been made regarding the interpretation of fluctuations in the gravity field with respect to small basins [*Longuevergne et al.*, 2010]. To date, however, remote sensing based estimates of terrestrial water storage are not yet fully developed and research on both sensor technology and data processing algorithms is still ongoing.

### 1.2.2. Atmospheric Variables

Precipitation and evaporation, the dominant water fluxes between land and atmosphere, have great influence on terrestrial hydrology. Precipitation is regularly monitored by national meteorological authorities and a large collection of station data (and derived data products) is collected by the Global Precipitation Climatology Centre[4] [GPCC *Schneider et al.*, 2010; *Fuchs*, 2009; *Rudolf and Schneider*, 2005] on behalf of the World Meteorological Organisation[5] (WMO). Evapotranspiration is, in contrast to precipitation, not regularly monitored and usually estimated by means of models that take variables related to the surface energy budget (e.g. temperature, net short and long wave radiation) and land properties

[4]http://gpcc.dwd.de, last accessed: 4 June 2011
[5]www.wmo.int, last accessed: 25 June 2011

(e.g. vegetation types, soil types or lake cover) into account. (See *Kingston et al.* [2009b]; *Kalma et al.* [2008]; *Allen et al.* [1998] for reviews and comments on different approaches). Only recently the FLUXNET[6] initiative [*Baldocchi et al.*, 2001] established a global network to monitor ecosystem scale evapotranspiration (and $CO_2$ fluxes). However, there is still a limited number of monitoring sites (mainly in Europe and north America) and the observational window is relatively short. Pan evaporation, the amount of water being evaporated from a dish on the earth surface is another estimator for evaporation. Many countries have over the past built relatively large monitoring networks and the observational time window is long [*Roderick et al.*, 2009a]. However, due to different standards in instrumentation and dependency on surrounding vegetation as well as other factors that are often difficult to control, interpretation of the resulting time series, albeit possible, has to be done with caution [*Roderick et al.*, 2009a,b].

For most atmospheric variables that are monitored on a regular basis (e.g. precipitation, temperature, air pressure), various data products exist that provide estimates of these variables on a regular spatial grid. These so-called gridded data are either available for a given region of interest or for the entire globe. Values at grid-cells without observations are estimated using sophisticated statistical interpolation techniques. The Climate Research Unit[7] (CRU) of the University of University of East Anglia, for example, provides global estimates of a wide range of atmospheric variables with a monthly resolution [*Mitchell and Jones*, 2005]. Another example are the high resolution estimates of gridded daily surface temperature and precipitation data introduced by *Haylock et al.* [2008].

Another source of gridded data are reanalysis products, where dynamical weather forecast models are used to assimilate *in situ* observations and remote sensing data to provide consistent estimates of atmospheric variables. These products have the advantage that they combine observed values while also providing estimates of variables that are not being monitored based on physical principles. Widely used examples of comprehensive reanalysis products are the ERA40 reanalysis[8] [*Uppala et al.*, 2005], which covers the period 1957 - 2002 and the

[6] http://www.fluxnet.ornl.gov/fluxnet, last accessed: 4 June 2011
[7] http://www.cru.uea.ac.uk, last accessed: 4 June 2011
[8] http://www.ecmwf.int/research/era/do/get/era-40, last accessed: 4 June 2011

NCEP/NCAR[9] reanalysis [*Kalnay et al.*, 1996]. However, reanalysis products are known to have significant biases (and other inaccuracies) in variables related to the terrestrial water budget such as near surface temperature [e.g. *Pitman and Perkins*, 2009; *Simmons et al.*, 2004; *Uppala et al.*, 2005] and precipitation [e.g *Serreze and Hurst*, 2000; *Sheffield et al.*, 2004; *Uppala et al.*, 2005]. In recent years, several authors [e.g *Sheffield et al.*, 2006; *Weedon et al.*, 2011] have suggested the creation of so called forcing data. These forcing data are based on reanalysis products, which are modified by a series of statistical corrections that remove systematic differences between the reanalysis and the observed values. Such forcing data are currently assumed to be one of the most accurate data sources available to provide near-surface estimates of atmospheric variables relevant to the hydrological cycle. One of the most comprehensive data sets are the so called WATCH forcing data [*Weedon et al.*, 2011, 2010], which have been developed within the European Framework Project "Water and Global Change"[10] (WATCH). The WFD provide bias corrected estimates of a wide range of atmospheric variables based on the ERA40 reanalysis.

One issue common to most data sets with gridded atmospheric variables, is related to the effects that topography can have on the variables, such as changes in temperature and air pressure with elevation. Especially, the orographic effect on precipitation [*Barstad et al.*, 2007] is often not resolved, which is known to lead to systematic biases in regions with complex topography [*Adam and Lettenmaier*, 2003]. This issue can in principle be tackled using downscaling techniques that estimate local variability of atmospheric variables based on coarse gridded data (e.g. from a 1° grid), taking landscape features (e.g. topography) into account. Currently two downscaling approaches are distinguished. In statistical downscaling [e.g. *Wilby et al.*, 1998], statistical models are built that aim at reproducing the observed values based on the grid-cell value and additional variables such as elevation. Once such a model is built it can be used to estimate the variable value at any desired location. In dynamical downscaling [e.g. *Barstad et al.*, 2009], the grid-cell value (and possible values of neighbouring grid-cells) are used as boundary conditions for dynamical atmospheric models that provide estimates of the desired

[9] http://www.esrl.noaa.gov/psd/data/reanalysis/reanalysis.shtml, last accessed: 4 June 2011
[10] www.eu-watch.org, last accessed: 25 June 2011

variables on a finer spatial grid. To date both approaches are used and seem in general to have comparable a precision [e.g. *Haylock et al.*, 2006; *Schmidli et al.*, 2007].

# 1.3. Large-scale Hydrology - A Review of Observed Patterns

The following section aims at providing a summary on observed patterns of large-scale hydrology, focusing on findings from Europe. In some instances, additional results from the north American continent are reported to provide a more complete picture of hydrological phenomena on large scales in a region comparable to Europe. As mentioned in Section 1.2.1, river flow is the most commonly monitored variable of the terrestrial water balance. Consequently the work referred to in the following sections is based on river flow. Nevertheless, the summarised results are likely to be relevant for analogue patterns in variables such as soil moisture or ground water.

## 1.3.1. The Global Water Cycle

Regional fluctuations of terrestrial water storage and the corresponding fluxes of water (Eq. (1.1)) contribute to the global water cycle (Figure 1.2). Driven by solar radiation, water evaporates from oceans and land. The water vapour is first transported in the atmosphere and than precipitated nearby or at a distant location. Globally, about 60% of the precipitation over land evaporates, whereas 40% is discharged through rivers and groundwater to the oceans [*Oki and Kanae*, 2006]. The precise figures vary in space and time, depending on factors such as the mean climatic conditions and geographic location. Especially transport of water vapour from the oceans to land is highly dependent on the atmospheric circulation and the fraction of precipitation that is evaporated again will vary largely, being much higher in warm than in cold regions.

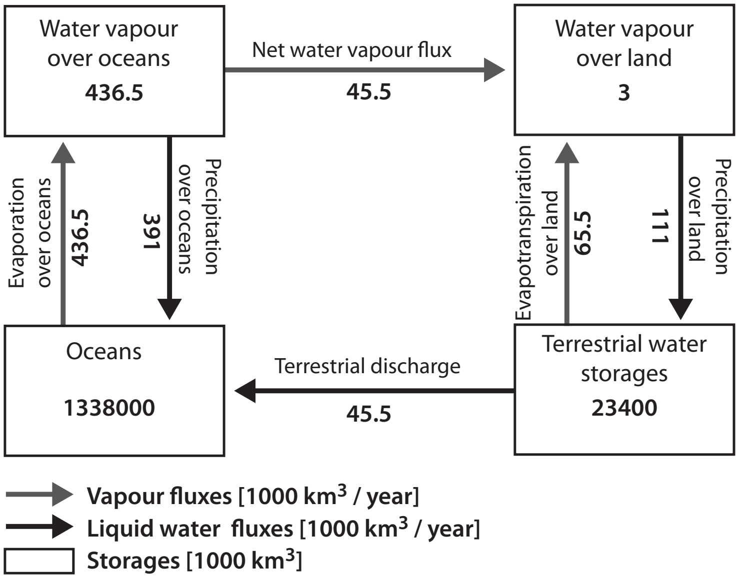


**Figure 1.2.:** Global water fluxes between oceans, atmosphere and land as well as the corresponding storage volumes. Numbers are taken from [*Oki and Kanae*, 2006].

## 1.3.2. Longterm Summaries of Hydrological Variables

### The Coupled Water Energy Balance

The terrestrial water balance (Eq. (1.1)) is to a wide extend determined by the trade off between water input from precipitation, $P$, and losses driven by the potential evaporation $E_p$. The relation between atmospheric water input and demand can be characterised by the ratio $E_p/P$, which is also referred to as the aridity index. The aridity index is usually computed on the basis of long term averages and allows for an easy classification of arid i.e. water limited ($E_p/P > 1$) and humid i.e. energy limited ($E_p/P < 1$) environments. However, land processes involved in separating the incoming precipitation into evaporation and runoff do also play a considerable role. The observable counterpart to the aridity index is the evapotranspiration ratio $E/P$. Unlike the aridity index, which grows toward infinity for large $E_p$ and small $P$, the evaporation ratio has an upper limit of $E/P = 1$, which occurs if all precipitation is evaporated and none is discharged. In some cases the longterm evapotranspiration ratio ($E/P$) is replaced with the longterm runoff ratio $Q/P$, which provides essentially the same information under the assumption that changes in the terrestrial water balance (Eq. (1.1)) are negligible such that $dS/dt \approx 0$. The aridity index and the evaporation ratio can be put into relation such that

$$\frac{E}{P} = F\left(\frac{E_p}{P}\right), \tag{1.6}$$

where $F$ is a parametrisation of those processes that influence the partition of precipitation into evapotranspiration and runoff as a function of water and energy availability and many different forms have been suggested [see e.g. *Szilagyi and Jozsa*, 2009, for an comprehensive overview]. Equation(1.6) was initially introduced by *Budyko* [1974] and is also known as Budykos hypothesis.

Various authors have used Budykos framework to characterise the hydroclimatic conditions [e.g. *Koster and Suarez*, 1999; *Sankarasubramanian et al.*, 2000; *Arora*, 2002; *Sankarasubramanian and Vogel*, 2003; *Donohue et al.*, 2007; *Yang et al.*, 2007; *Oudin et al.*, 2008; *Yang et al.*, 2008; *Szilagyi and Jozsa*, 2009; *Yang et al.*, 2009], but only relatively few assessed its variability across large spatial scales and along hydroclimatic gradients. For the eastern part of the continental United States (US) *Milly* [1994] was one of the first to use a physically based parametrisation of $F$ to estimate mean annual runoff. Budykos framework has also been used to characterise the average water balance of the entire US [*Sankarasubramanian and Vogel*, 2003, 2002]. Throughout the continent, the aridity index was found to vary from values close to zero (in the northwest of the country) to values of $E_p/P > 3$ (in the southwest of the country). *Sankarasubramanian and Vogel* [2003] pointed out that the predictive potential of $E/P$ for water balance components became increasingly limited in dry regions, where the water holding capacity of soils gained importance. In Europe, the aridity index has only been used regionally to characterise hydroclimatic conditions. In Austria, the majority of 459 catchments were found to be in humid (energy limited) environments and the aridity index was found to be a good predictor for the evapotranspiration ratio [*Merz and Blöschl*, 2009]. Similar results were found for a large number of catchments in France and

Great Britain [*Oudin et al.*, 2008]. Due to the regional focus of these studies they only provide an incomplete picture of the coupled energy-water balance in Europe, focusing on regions with predominantly humid climates. However, large parts of southern Europe have arid climates [e.g. *Kottek et al.*, 2006; *Peel et al.*, 2007].

### Mean Annual Runoff

Most large-scale hydroclimatic analyses do not focus on the spatial variability of the coupled water-energy balance, but rather on directly observable variables, especially river runoff per unit area. Global estimates of average runoff per unit area range from 339 mm year$^{-1}$ [*McMahon et al.*, 2007b] to 404 mm year$^{-1}$ [*Dettinger and Diaz*, 2000]. The mean annual runoff rate varies amongst regions depending on the climatic conditions [e.g. *McMahon et al.*, 2007a]. In Europe, runnoff rates have been reported to range from $<$ 150 mm year$^{-1}$ to $>$ 2900 mm year$^{-1}$ [*McMahon et al.*, 2007a; *Oudin et al.*, 2008]. Figure 1.3 provides an overview on observed and modelled mean annual runoff rates in Europe. Observed runoff rates range from $<$ 40 mm year$^{-1}$ in central and southern Europe to $>$ 4000 mm year$^{-1}$ on the Scandinavian west coast. The modelled runoff rates are lower ranging from $<$ 10 mm year$^{-1}$ in the Mediterranean to $>$ 2700 mm year$^{-1}$ at the Scandinavian west coast. The mean observed runoff rate is 730; mm year$^{-1}$ which is well above the mean of the modelled runoff rates (373; mm year$^{-1}$). Note however, that the observation network is geographically biased and does thus not provide a consistent estimate of mean runoff in Europe. Comparable maps based on a combination of observations and empirical water balance estimates [*Rees et al.*, 1997] are available from the European Environmental Agency[11].

### Seasonality of Runoff

To understand hydrological variability on large scales, not only the annual average values of a runoff are of interest, but also its seasonality, that is, the average evolution throughout a year. The seasonality of runoff in mid latitudes is in general determined by the interplay of two different processes, which are both related to

[11] http://www.eea.europa.eu/data-and-maps/figures/average-annual-runoff, , last accessed: 21 June 2011

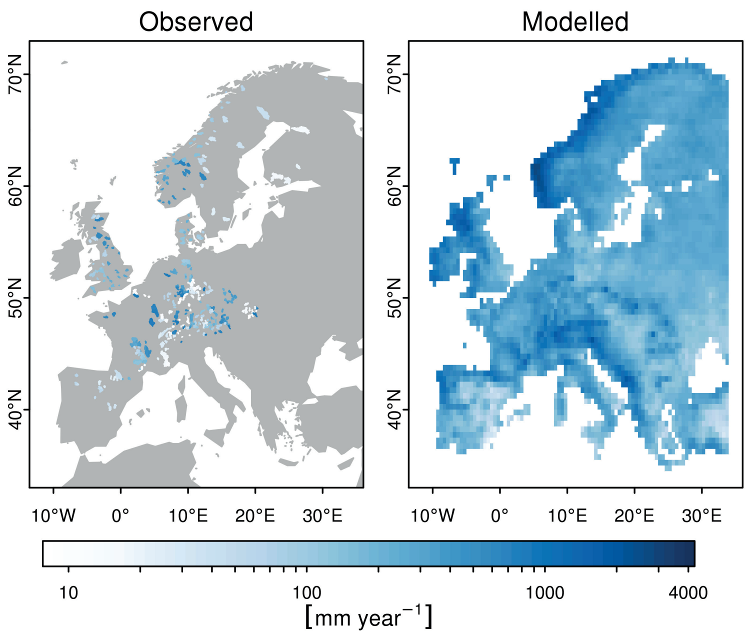


**Figure 1.3.:** Observed (left) and modelled (right) mean annual runoff rates in Europe. The observed values originate form the European Water Archive. The modelled values are the mean of a multi model ensemble (See *Gudmundsson et al.* [2011d] (Article III) and *Gudmundsson et al.* [2011c] (Article IV) for a description of the model ensemble and a verdict on the quality of the simulations.)

the seasonal cycle of incoming solar radiation. In warm climates, the seasonality of runoff is closely related to the mean annual cycle of evapotranspiration, which is typically highest in summer. Consequently, the mean annual cycle of runoff has a summer minimum (when most of the incoming precipitation is evaporated) and an autumn or winter maximum. In cold climates, the mean annual cycle of runoff is predominantly influenced by snow processes. Precipitation falling as snow is stored in a snow pack, often leading to a winter minimum in runoff, and later discharged during snow melt, typically causing pronounced spring floods. In many regions of Europe both evapotranspiration and snow processes influence the mean annual cycle of runoff causing gradually changing spatial patterns of runoff seasonality [*Dettinger and Diaz*, 2000].

One approach to summarise hydroclimatic patterns is to group hydrological units with respect to their mean annual cycle, often leading to the identification of geographical regions with common behaviour. Regions with a common seasonality in runoff are often referred to as flow regimes. In western Europe a broad distinction can be drawn between *Atlantic*, *Mediterranean* and *Continental* regimes based on the timing of the minimum and the maximum of the mean annual cycle of monthly runoff [*Arnell*, 2002; *Krasovskaia et al.*, 1994; *Gottschalk et al.*, 1979]. The Atlantic regimes are most abundant in western Europe and are characterised by a winter maximum and summer minimum. These regimes are predominantly influenced by evaporation, as precipitation rates are almost constant throughout the year. The Mediterranean regimes have seasonal patterns comparable to the Atlantic regimes, but have a larger amplitude because most precipitation falls during the winter months. The continental regimes are strongly influenced by snow melt, leading to peak flows in spring. However, they differ in the timing of low flows (either in summer due to evaporation or in winter as all precipitation is stored as snow) and some have a secondary peak due to the seasonality of rainfall. Continental regimes occur in the Nordic region, east of central Germany and in European mountain ranges.

Figure 1.4 shows the month of the minimum and the maximum of the mean annual cycle of observed and modelled runoff. Observed values originate from the EWA (see Section 1.2.1). The modelled runoff is derived as the mean of a multi model ensemble which is described and evaluated in detail in *Gudmundsson et al.* [2011d] (Article III) and *Gudmundsson et al.* [2011c] (Article IV). In western and central Europe, along the Scandinavian west coast and in the Mediterranean region, the mean annual cycle follows the atmospheric water demand and has its minimum in summer and a maximum in winter. In the inland regions of Scandinavia, the Alps and in eastern Europe, the mean annual cycle of runoff is mostly influenced by snow accumulation and melt with winter minima and spring floods.

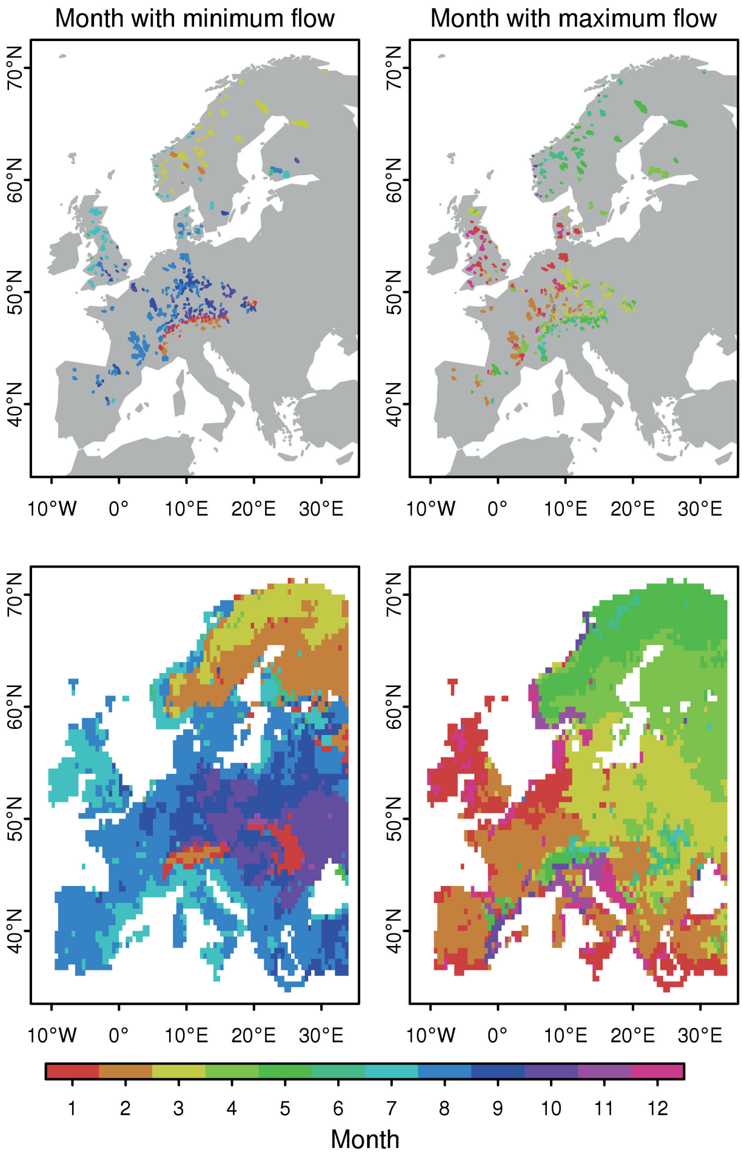


**Figure 1.4.:** Month of the minimum and the maximum of the mean annual cycle of observed (top) and modelled (bottom) runoff rates.

### Other Longterm Summaries

In addition to the mean annual flow and the seasonal cycle, a large variety of other statistical summaries of streamflow are used to characterise the hydroclimatic conditions. *Clausen and Biggs* [2000] and *Olden and Poff* [2003] provide comprehensive comparisons of up to 171 different indices. Many of these are based on the empirical cumulative distribution of runoff, also referred to as the Flow Duration Curve [FDC; e.g. *Vogel and Fennessey*, 1994] and include the probability of exceeding different flow levels as well as measures of the difference between high and low flows which are suitable for the characterisation of extremes. Drought indices or low flow statistics include the probability of not exceeding a low flow level, and the average duration of spells below a flow level and its deficit volume [e.g. *Zelenhasić and Salvai*, 1987; *Tallaksen et al.*, 1997; *Smakhtin*, 2001; *Tallaksen and van Lanen*, 2004; *Fleig et al.*, 2006]. Flood characteristics are often quantified using the probability of exceeding certain flow levels and statistical moments related to the variability of runoff such as the coefficient of variation or the skewness [e.g. *Hosking and Wallis*, 1997; *Merz and Blöschl*, 2009, 2008a,b]. For both drought [e.g. *Smakhtin*, 2001; *Laaha and Blöschl*, 2005; *Engeland et al.*, 2006; *Laaha and Blöschl*, 2007] and flood [e.g. *Blöschl and Sivapalan*, 1997; *Merz and Blöschl*, 2009, 2008a,b; *Sivapalan et al.*, 2005] characteristics a large range of statistical and process oriented studies have assessed influencing factors. In summary most of these studies find that both the average climatic conditions as well as land properties are influencing regional patterns.

## 1.3.3. Decadal and Inter-Annual Variability

As the terrestrial water balance is directly coupled to the dynamic atmosphere, runoff climatologies such as mean annual runoff, the seasonal cycle or flood and drought statistics are likely to change systematically from year to year. In the following, all variability of runoff on timescales longer than one year will be referred to as low-frequency variability. Low-frequency variability of runoff has often been analysed with special emphasis on phenomena such as trends or (quasi) periodic oscillations with timescales ranging from several years to several decades.

First studies that have analysed runoff with respect to trends will be summarised, followed by a review of investigations focusing on (quasi-)periodic oscillations on timescales longer than one year. The low-frequency variability of runoff from neighbouring rivers usually have a coherent temporal evolution and the spatial extent as well as emerging regions of coherent low-frequency runoff variability will be commented upon. Finally, links between the low-frequency variability of runoff and large-scale atmospheric drivers will be summarised.

#### Trends

A variety of studies have analysed streamflow trends throughout Europe, either based on data sets covering large parts of the continent [*Stahl et al.*, 2010; *Hisdal et al.*, 2001] or for specific regions of interest [e.g. *Hannaford and Marsh*, 2006; *Mudelsee et al.*, 2003; *Kingston et al.*, 2011]. Most of these studies analysed trends of miscellaneous annual summary statistics such as mean annual runoff, seasonal trends and trends in flood and drought characteristics. Trends of mean annual runoff in Europe have coherent regional patterns. In southern Europe streamflow has a tendency to decrease throughout the last decades, indicating drying conditions. The drying trends in southern Europe are contrasted by wetting trends in annual runoff north of the alps [*Stahl et al.*, 2010]. Trends in seasonal runoff (e.g. trends in mean January, February, ... flow) have coherent spatial patterns. However, there are pronounced shifts between the seasons. In the winter months, runoff increases quite consistently throughout Europe. In the spring and summer months, however, monthly runoff generally decreases, especially in central and southern Europe, indicating increasingly dry summers [*Stahl et al.*, 2010]. Trends in annual low flow [e.g. *Hannaford and Marsh*, 2006; *Hisdal et al.*, 2001; *Stahl et al.*, 2010] and high flow [e.g. *Kundzewicz et al.*, 2005; *Hannaford and Marsh*, 2008] statistics often follow the patterns of the corresponding seasonal trends. However, the spatial patterns of trends in low and high flows are at times less coherent than trends in mean annual or seasonal runoff. Comparable patterns of regional coherence have been found for trends in various aspects of North American streamflow, where the availability of continental scale observation networks has triggered a rich body of literature of continental scale trend analysis [e.g. *Lettenmaier et al.*, 1994; *Rose*, 2009; *Lins and Slack*, 1999; *Small et al.*, 2006; *Regonda et al.*, 2005; *Khaliq et al.*, 2008].

### Low-frequency Oscillations

Studies that analyse the temporal variability of runoff with respect to trends assume monotonic (often linear) changes over time. However, streamflow exhibits, like many other variables of the earth system, other systematic fluctuations on timescales longer than one year. Analysis of river discharge from several stations around the globe [*Labat*, 2008; *Labat et al.*, 2005; *Pekarova et al.*, 2003] showed that river discharge has quasi-periodic oscillation with periods ranging from a few years to several decades. Examples in Europe include the Seine River where two periodic modes (17 and 5-9 years) have been identified [*Massei et al.*, 2010] and the analysis of river flow in England and Wales, which was found to have systematic variations of decadal timescales [*Sen*, 2009]. In the Elbe and the Oder, which are the two largest rivers in central Europe, systematic (decadal) fluctuations in the extreme floods were found [*Mudelsee et al.*, 2004, 2003]. The existence of systematic variations on long timescales is consistent with the previous notion that high and low flow years in Europe have a tendency to cluster in time [*Arnell*, 1994]. The presence of systematic fluctuations in hydrological variables on timescales longer than one year has also been demonstrated for various locations across the North American continent [e.g. *Hanson et al.*, 2004; *Kumar and Duffy*, 2009; *Shun and Duffy*, 1999].

### Spatial Coherence

An interesting feature of fluctuations of runoff on inter-annual and decadal timescales is that these fluctuations are coherent in space. In relatively small regions such as southern Germany [*Stahl and Demuth*, 1999; *Lange and Bernhardt*, 2004], Denmark [*Hisdal and Tallaksen*, 2003] and Iceland [*Jónsdóttir and Uvo*, 2009], inter-annual fluctuations of various streamflow characteristics (including significant low-frequency oscillations and drought statistics) have been shown to have a high degree of coherence and their temporal evolution differs only marginally between locations. In larger regions (e.g. Iberian peninsula [*Lorenzo-Lacruz et al.*, 2011] or Turkey [*Kalayci and Kahya*, 2006]) gradual changes in the degree of simultaneous variations become visible and are associated with varying influences of dominant temporal signals. Only few studies have assessed the coherence of the inter-annual and low-frequency variability of runoff in Europe on continental scales. Overall, large parts of Europe experience similar streamflow variations at the same time and only the Nordic countries have been reported to differ from this pattern [*Arnell*, 1994; *Shorthouse and Arnell*, 1999]. Comparable results were found for the analysis of streamflow droughts across Europe which exhibit spatial patterns of highly coherent temporal variations [*Zaidman et al.*, 2002]. The analysis of streamflow anomalies in North America revealed large regions with highly intercorrelated streams. These regions corresponded only marginally with major physiographic divisions of the US [*Bartlein*, 1982; *Lins*, 1997]. For both the US [*McCabe and Wolock*, 2002] and Canada [*Ehsanzadeh et al.*, 2011] regionally coherent changes in annual streamflow statistics (including low and high flows) have been reported.

#### Atmospheric Drivers

The inter-annual and decadal variability or runoff is related to equivalent fluctuations in precipitation and the atmospheric water demand, driving evapotranspiration [e.g. *Shun and Duffy*, 1999; *Kumar and Duffy*, 2009]. These variables are in turn driven by large-scale patterns of atmospheric circulation and thus, river flow will respond to changes in atmospheric dynamics. Atmospheric circulation is commonly characterised using spatial pattens of atmospheric pressure (indicating air flow direction), which can be classified into regions with coherent temporal variability [*Barnston and Livezey*, 1987]. In these regions the temporal variability is often summarised by so called atmospheric oscillation indices.

European climate is closely related to the North Atlantic Oscillation (NAO), which is characterised by fluctuations in the pressure gradient between the Azores high and the Icelandic low [e.g. *Hurrell and van Loon*, 1997; *Hurrell*, 1995; *van Loon and Rogers*, 1978]. Large-scale precipitation [e.g. *Hurrell and van Loon*, 1997; *Hurrell*, 1995; *Wibig*, 1999] and temperature [e.g. *van Loon and Rogers*, 1978] anomalies in Europe can be related to the NAO as well as to comparable patterns of atmospheric circulation. These large-scale precipitation and temperature patterns over Europe are often characterised by a seesaw, where anomalies in southern Europe are paired with opposing anomalies in the north. Unsurprisingly fluctuations in statistical summaries of river flow observations in Europe have been successfully related to the NAO and equivalent indicators of atmospheric circulation [e.g. *Shorthouse and Arnell*, 1997; *Rîmbu et al.*, 2002; *Jacobeit et al.*, 2003; *Bouwer et al.*, 2006; *Kalayci and Kahya*, 2006; *Kingston et al.*, 2006a,b; *Bouwer et al.*, 2008; *Hannaford and Marsh*, 2008; *Holman et al.*, 2009; *Jónsdóttir and Uvo*, 2009; *Kingston et al.*, 2009a; *Massei et al.*, 2010; *Lorenzo-Lacruz et al.*, 2011; *Morán-Tejeda et al.*, 2011]. Dependencies are usually found to be most pronounced in northwestern and southwestern Europe and the relation is generally reported to be stronger in winter. The NAO is not the only atmospheric circulation pattern that influences European weather, and influences of the Arctic Oscillation [*Schaefer et al.*, 2004; *Kingston et al.*, 2006a; *Ionita et al.*, 2011] on hydrological variables in northern Europe have also been identified. As river flow is not only controlled by the climatic water balance, but also depends on terrestrial processes, it is plausible that low, mean and high river flows may respond differently to large-scale circulation patterns. This is supported by fact that annual peak discharges are more strongly correlated to the NAO than mean discharges [*Bouwer et al.*, 2008].

Another approach to relate changes in hydrological variables to atmospheric circulation is based on weather types, also referred to as circulation patterns. Weather types are recurrent weather patterns, that are often characterised by dominant air flow directions [e.g. *Gerstengarbe et al.*, 1999; *James*, 2007]. The frequency of certain weather types have been related to mean annual runoff [*Bouwer et al.*, 2008, 2006] as well as to flood [*Mudelsee et al.*, 2004; *Prudhomme and Genevier*, 2011] and drought [*Fleig et al.*, 2011; *Stahl and Demuth*, 1999] statistics.

Common to many European studies attributing long term changes in river flow to large-scale atmospheric variability, is the concentration on relatively small regions of interest – or – the analysis of just one river flow series. Although such a regional focus enables detailed insights to local variations, it hampers the identification of clear patterns on large scales. For the North American continent the availability of appropriate data has enabled the demonstration that large-scale patterns of observed river flow can be directly attributed to fluctuations in atmospheric water balance variables [e.g. *Krakauer and Fung*, 2008; *Small et al.*, 2006; *Tootle and Piechota*, 2006] and large scale atmospheric circulation patterns [e.g. *Barlow et al.*, 2001; *Rajagopalan et al.*, 2000; *Tootle and Piechota*, 2006; *Tootle et al.*, 2005].

## 1.4. Principles of Large-scale Hydrological Modelling

Modelling of large-scale hydrology is, like any other hydrological modelling, based on solving the water balance equation (Eq. (1.1)), sometimes paired with explicit solutions of the energy balance (Eq. (1.3)). The crucial step in model building is to find appropriate descriptions of the terrestrial hydrological processes, which determine how much water is stored and how easily water is released as runoff or as evapotranspiration (Eq. (1.2) and (1.4)). The task is to find a compromise between a realistic mathematical description of the influencing factors such as vegetation, soil properties and topography, while keeping the solution of the resulting set of equations feasible. Thus, model building is often a trade-off between the level of detail (inclusion of distinct processes; spatial and temporal resolution) and the availability of data and computational resources. Throughout the years a multitude of plausible models that provide descriptions of the terrestrial water balance have been suggested [e.g. *Manabe*, 1969; *Deardorff*, 1978; *Haxeltine et al.*, 1996; *Meigh et al.*, 1999; *Hagemann and Dümenil Gates*, 2003; *Hanasaki et al.*, 2008; *Balsamo et al.*, 2009; *Best et al.*, 2011]. Although differing in the processes included and the level of detail, all large-scale hydrological models share the same principle architecture. In order to represent the spatial variability, the area of interest is subdivided into many small areas, commonly referred as grid cells. These grid-cells are the basic unit of large-scale hydrological models and are usually built to represent a vertical water balance, describing the exchange of water between the land and the atmosphere and the vertical drainage of water through soils.

The mathematical structures of model grid-cells are usually based on partitioning the terrestrial water storage ($S$) into smaller units within each grid-cell and the description of the fluxes of water between these. Figure 1.5 illustrates such a structure, summarising different storages and fluxes. Typically considered storages are the canopy or interception storage ($S_{can}$) that accounts for water stored on plant surfaces, the snow storage ($S_{snow}$), surface water bodies such as lakes or streams ($S_{surf}$, not included in Figure 1.5), several soil layers $S_{soil}$ and ground water ($S_{gw}$). The fluxes between these storages can be separated into upward fluxes feeding evapotranspiration and downward fluxes that eventually contribute to runoff. Amongst the upward fluxes are evaporation from soils ($E_{soil}$), transpiration of plants ($E_{can}$), sometimes paired with descriptions of root water uptake ($E_{root}$) and sublimation of snow ($E_{sub}$). The downward fluxes include throughfall ($P_t$), i.e. water that reaches the ground, snow melt ($Q_m$), infiltration into soils and recharge to groundwater ($Q_d$). Excess water leaves the grid-cells either as surface ($Q_s$) or sub-surface ($Q_{sb}$) runoff. The sum of surface and sub-surface runoff of each grid-cell is then used as input to the river network. It shall be noted that the mathematical structure of the individual grid-cells is in many cases closely related to the form of so-called "lumped catchment models", which are used to describe rainfall runoff processes of individual catchments.

Grid-cells of large-scale hydrological models usually cover large areas (hundreds to thousands of square kilometres). Consequently, land properties that influence water movements may vary largely within a grid-cell and several strategies to account for this sub-grid variability have been suggested. Popular approaches include incorporation of stochastic properties of topography [e.g. *Todini*, 1996] or soils [e.g. *Moore*, 1985; *Hagemann and Dümenil Gates*, 2003] and the sub-division of the area covered by a grid-cell depending on dominant vegetation types [e.g. *Essery et al.*, 2003; *Best et al.*, 2011].

Once a set of equations describing the water movements within a grid-cell has been established, their parameters, representing the land properties (e.g. the hydraulic conductivity of soils) need to be determined. In catchment modelling these parameters are usually determined through an extensive calibration process. This practise is contrasted by large-scale hydrological models that usually undergo no (or only a limited) calibration procedure. Large-scale hydrological models rather rely on mapped land properties such as the Harmonized World Soil Database [*FAO et al.*, 2009]. However, both the data products used to estimate the model parameters as well as the interpretation of the mapped values varies largely between different models, which results in considerable differences in simulated system behaviour [e.g. *Teuling et al.*, 2009]. Consequently, large-scale hydrological models

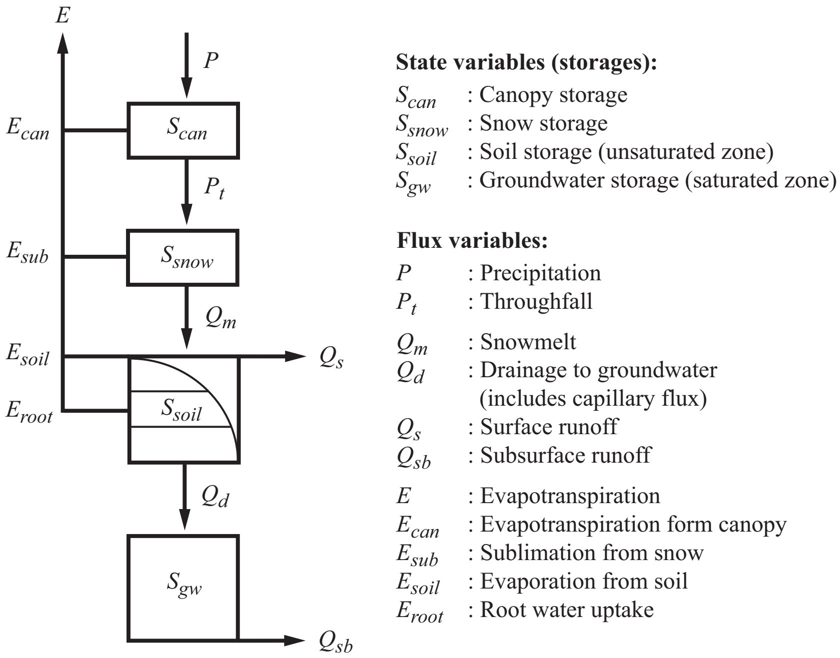


**Figure 1.5.:** Simplified conceptualisations of state (storage) and flux variables involved in runoff generation. Adapted from *Gudmundsson et al.* [2011d] (Article III)

are not only defined by their mathematical structure but also by the data sources used for parameter identification.

# 2. Scope of the Thesis

The superordinate goal of this thesis is to contribute to a better understanding of the terrestrial water storages and fluxes. Special emphasis is put on those patterns in runoff that become apparent on large, continental scales. The main part of this dissertation is a compilation of five journal articles, each focusing on different aspects of runoff European variability on large scales. The first two articles concentrate on the identification, description and interpretation of the spatial patterns associated with the inter-annual or low-frequency variability of observed runoff in Europe. The remaining three articles are concerned with the performance of large-scale hydrological models, as providing insights into their strengths and weaknesses may help to refine our understanding of the functioning of the terrestrial water balance on large, continental, scales.

## 2.1. Observed Patterns

Observations are the primary source of informations enabling insights to the functioning of hydrological systems. However, most available studies that assess the inter-annual or low-frequency variability of river runoff in Europe have a distinct regional focus. Therefore the knowledge about spatial patterns associated with the inter-annual or low-frequency variability of runoff in Europe on large, continental, scales is fare from exhaustive (Section 1.3). The two articles that focus on observed patterns of large-scale hydrology in Europe aim at bridging this gap.

**Article I: Spatial cross-correlation patterns of European low, mean and high flows** [*Gudmundsson et al.*, 2011a], assesses the spatial patterns associated with annual low, mean and high flow statistics. Article I also addresses the question whether the inter-annual variability of low, mean and high flows is controlled by different atmospheric drivers.

**Article II: Low-frequency variability of European runoff** [*Gudmundsson et al.*, 2011b], assesses the spatial patterns of simultaneous variability of the low-frequency components of monthly runoff and compares them to the spatial patterns of low-frequency precipitation and temperature. This study also searches for factors that determine the strength of low-frequency variability of individual runoff series.

## 2.2. Model Performance

Large-scale hydrological models implement our current understanding of those processes that dominate the terrestrial water balance on large scales. The inclusion of a multitude of small-scale processes, however, does not necessarily imply the emergence of large-scale patterns. Thus, models can be regarded as hypotheses on the functioning of hydrological systems which can be tested with respect to their plausibility [*Savenije*, 2009; *Sivapalan*, 2005]. The model evaluation studies, focus on the models ability to capture large-scale patterns of runoff dynamics. In a hypothesis testing framework this implies that specific processes are not tested, but rather whether a set of interacting processes does in fact lead to the emergence of macroscopic patterns. The complexity of the considered models may render a strict verification or falsification impossible [*Harte*, 2002]. However, the ability of the models to capture well defined macroscopic patterns as well as possible regularities in the model error may give hints on the reasons for specific virtues and shortcomings.

**Article III: Comparing Large-scale Hydrological Models to Observed Runoff Percentiles in Europe** [*Gudmundsson et al.*, 2011d], introduces a multi model ensemble of nine large-scale hydrological models and assesses its ability of capture the dominant features of the inter-annual variability of low, mean and high river flows in Europe.

**Article IV: Seasonal Evaluation of Nine Large-Scale Hydrological Models Across Europe** [*Gudmundsson et al.*, 2011c], assesses the ability of the multi model ensemble introduced in Article III to simulate the mean annual cycle of monthly runoff. Article IV also explores possibilities of inferring specific issues with the model formulation based on systematic structures in the model error.

**Article V: Streamflow data from small basins: a challenging test to high resolution regional climate modelling** [*Stahl et al.*, 2011], assesses the ability of the land surface scheme of a high resolution regional climate model to capture the dynamical properties of streamflow anomalies with special emphasis on low and high flows.

# 3. Summary and Synthesis

## 3.1. Data Sources

### Runoff Observations

All articles contributing to this dissertation are based on daily runoff observations from the European Water Archive (EWA, see Section 1.2.1), which were augmented with data held by national authorities [see *Stahl et al.*, 2010, for a detailed description of the various data sources]. Figure 3.1 gives an overview on the spatial distribution, the size (median catchment area 258 $km^2$) and the elevation (median catchment elevation 582 m) of the considered catchments. Catchment boundaries as well as mean catchment elevation and slope where taken from the pan-European river and catchment database CCM2 [Catchment Characterisation and Modelling 2 *Vogt et al.*, 2007]. In most instances runoff rates per unit area (e.g. mm $day^{-1}$) and not flow volumes (e.g. $m^3$ $sec^{-1}$) were analysed. This collection is to date the most comprehensive source of discharge observations in Europe from small, nearly undisturbed, catchments that are not nested. The time window covered by the observations ranges in some instance from 1932 to 2004, with large regional differences. In this dissertation, however, only observations covering the period 1962 to 2004 are considered as this is the most complete part of the data set. The specific selection of catchments as well as the chosen time window vary for the different articles contributing to this thesis, depending on different data quality requirements as well as on the spatial and temporal coverage of supplementary data sources.

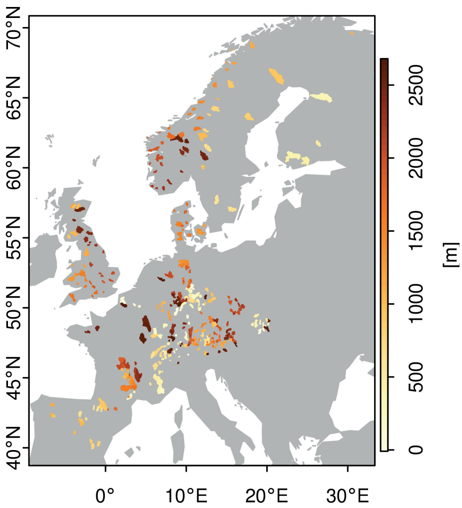


**Figure 3.1.:** Spatial distribution and mean elevation of the catchments with runoff observations considered in this dissertation.

### Atmospheric Variables

Atmospheric variables, matching the runoff observations from the EWA are in most instances taken from the WATCH Forcing Data (WFD, see Section 1.2.2). The WFD provide global estimates of near surface meteorology based on the ERA40 reanalysis which have been bias corrected [*Weedon et al.*, 2011]. The WFD can be assumed to be currently one of the most reliable global estimators for atmospheric near surface variables. The WFD are provided on the 0.5° grid, defined by the CRU (Climate Research Unit of the University of East Anglia) global land mask. The grid-cell size varies depending on the latitude and ranges from the northernmost to the southernmost gauging station of the runoff observations from 1065 $km^2$ (at 70°N) to 2387 $km^2$ (at 39.5°N). Thus the size of the grid-cells is about one order of magnitude larger than the median catchment size. The WFD are either used directly to provide estimates of precipitation and temperature (Article II) or indirectly as they are used to force models (Article III and IV) which in turn are evaluated using the observations from the EWA. To match the large (0.5°) grid-cells with runoff observations from small catchments, only those grid-cells with gauging stations were selected. The observed runoff series were assigned to the corresponding grid-cells and if several stations where located in one grid-cell the average of the runoff series (weighted by catchment area) was used.

### Model Simulations

The articles focusing on model performance rely on model simulations conducted by other work groups. Two articles (Article III and IV) evaluate the performance of a multi model ensemble comprising nine large-scale hydrological models. This multi model ensemble is closely related to the ensemble described by *Haddeland et al.* [2011] and has been developed as a joint effort within the WATCH project. All models are forced using the WFD but no effort has been made to harmonise model parameters. The simulation setup and the models contributing to the ensemble are described in detail in Article III. Observations were matched with the model simulations in the same manner as for the WFD. Only those 0.5° grid-cells, containing gauging stations were considered and the runoff observations were assigned to these grid-cells. Thus observed runoff from small catchments was compared to simulated runoff from relatively large grid-cells. This approach appears valid as the mathematical structure underlying the individual grid-cells is typically closely related to the mathematical structures of lumped catchment models which are commonly applied to assess runoff from small catchments.

Article V compares the runoff observations to simulation results of a high resolution regional climate model (0.12°). In this case simulated runoff from those model grid-cells that are located in a particular catchment have been selected and the sum of the grid-cell runoff was compared to the observations.

## 3.2. Features of Runoff Variability

A common approach for the analysis of environmental time series is to preprocess these in order to emphasis specific features of variability. Examples are the analysis of time series of annual mean values or monthly anomalies, i.e. monthly time series with the longterm mean of each month being removed. In all articles contributing to this dissertation data preprocessing is used to isolate different features of runoff variability. Although differing in their scope as well as in complexity, the preprocessing steps used in the different articles have in common that they operate on time series from individual catchments (or grid-cells) isolating selected temporal features of runoff variability.

### Percentile Series

Both Article I and III analyse time series of annual runoff percentiles in order get insights to the inter-annual variability of low, mean and high flows. Runoff percentiles are derived from daily time series and are defined as time series of the annual $5\%, 10\%, \ldots, 95\%$ percentile, where the 5% percentile denotes low and the 95% percentile denotes high flows. (Alternatively, the notation $5^{th}, 10^{th}, \ldots, 95^{th}$ has also been used.) Note that the definition of percentiles follows the statistical convention (representing cumulative frequencies) which is commonly used in North America and not the hydrological convention (representing exceedance frequencies) which is more often used in Europe.

### Low-frequency Components

Article II makes use of low-frequency components of monthly runoff. Within this dissertation the term low-frequency is used to address all variability on timescales longer than one year. Note, however, that the term "low-frequency" is used differently in other context and may for example be used to address the variability with timescales of centuries. By construction, the information content of low-frequency components of monthly runoff is closely related to the information content of annual mean or median series.

### Nonparametric Anomalies

Article V is based on nonparametric anomaly series of daily data, smoothed by a seven day moving average. These series were deseasonalised and standardised by replacing the actual values with the percentiles of the empirical frequency distribution of each calender day. This effectively removes the annual cycle and scales the time series between 0 and 100. Values of 0 are given for the day of the year with the lowest value in the observational window and values of 100 are given to the maximum value.

### Mean Annual Cycle

Both Article II and IV make use of the mean annual cycle of monthly runoff, which in this dissertation is also referred to as hydroclimatic regime. The mean annual cycle is defined be the longterm mean of each month and it's analysis also has implications for understanding time series or runoff percentiles, as the lowest and the highest flows values usually occur in distinct season (e.g. spring flood in snow dominated climates).

## 3.3. Observed Patterns

The two articles focusing on observed patterns of runoff variability focus on the inter-annual variability of runoff percentiles and the low-frequency components of monthly runoff. The main attention is not on the temporal evolution (i.e. the analysis of the sequence of wet and dry years) but on the associated spatial patterns. In other words, the question is asked: Which of the small rivers across Europe share common temporal variations and how does the relation of the temporal evolution of different rivers change in space? Once identified, these spatial patterns of common temporal variability derived from different variables (e.g. different percentile series or the low-frequency components of runoff and precipitation ) can be compared and further analysed. Similarity in these patterns are interpreted as an indication for physical coupling.

### Article I

#### Spatial cross-correlation patterns of European low, mean and high flows

The main question underlying this study was whether annual low, mean and high flows respond in a similar manner to inter-annual variations of the atmospheric forcing – or – whether catchment processes emphasise different modes of climatic variability in low and high flows. The rational behind this question is that high flows occur in direct response to precipitation or snow melt, whereas low flows occur in periods with now direct water input and thus depend on the amount of water stored in the catchment. Further, low and high flows usually occur in distinct seasons, which may also lead to different inter-annual dynamics, depending on the inter-annual variability of the corresponding weather patterns.

The study was based on analysing time series of annual streamflow percentiles ($5^{th}$, $10^{th}$, ..., $95^{th}$), covering the full range of flow dynamics. For each percentile level, cross-correlation matrices, summarising the interrelation of the inter-annual variability of all streams were then derived. Each of these cross-correlation matrices summarises to what degree all pairs of streams share common temporal variability, depending on the percentile level. Finally, the similarity among the cross-correlation matrices was quantified. In other words, it was tested whether streams with correlated low flows also have correlated high flows. The reason for comparing spatial correlation patterns instead of correlating time series of annual runoff percentiles directly are manifold. Spatial correlation patterns of annual runoff percentiles are, for example, likely to reflect equivalent patterns of atmospheric variables. Furthermore, the average strength of correlations between neighbouring stations may change if the influence of catchment processes gets more pronounced. This may provide insights on the importance of differing catchment properties.

The spatial correlation of all runoff percentiles are dominated by one common cross-correlation pattern. Further analysis of this pattern revealed six regions of coherent runoff-variability in Europe. This regional coherence suggests that time series of annual low mean and high flows follow essentially the same large-scale atmospheric drivers. However, both low and high flows deviate systematically from this pattern – especially with respect to the strength in the correlation of nearby catchments. On average, high flows exhibit a higher degree of correlation in space than low flows. This indicates that time series of annual high flow statistics do closely follow the corresponding atmospheric drivers. Low flows on the other hand, exhibit on average a much smaller degree of correlation between nearby stations, which points toward the influence of locally varying catchment properties including topography and aquifers.

## Article II

### Low-frequency variability of European runoff

The aim of this study was to get a comprehensive overview on the low-frequency variability of monthly runoff in Europe as well as to gather empirical evidence on influencing factors such as the atmospheric drivers and catchment properties.

To guide the analysis it was hypothesised that catchments can be characterised as spectral filters, dampening or amplifying fluctuations depending on the timescale. To characterise such effects the fraction of variance of monthly runoff, that can be attributed to the low-frequency components of monthly runoff (i.e. the fraction of low-frequency variance) was determined and compared to catchment properties and climatic conditions. Further, it was assumed that the temporal evolution of low-frequency runoff (and their associated spatial patterns) is primarily dependent on the atmospheric drivers.

The fraction of low-frequency variance was neither correlated to the fraction of low-frequency variance of precipitation and temperature nor to land properties such as catchment area, elevation or slope. It was, however, found to be significantly lower in those regions where snow accumulation and melt influence the mean annual cycle. Furthermore, the fraction of low-frequency variance of runoff increases (decreases) under drier (wetter) conditions and is lowest in catchments with a high variability of daily runoff. These results suggest that the mean climatic conditions have a significant influence on the strength of the low-frequency variability of monthly runoff.

The dominant space-time patterns of low-frequency runoff in Europe can efficiently be characterised by three dominant modes, each associated with a distinct spatial pattern. The dominant pattern has two opposing centres of simultaneous variations, suggesting that dry years in northern Europe are accompanied by wet years in the south (and *vice versa*). The secondary pattern features an west-east gradient, suggesting a gradual change in the temporal evolution of low-frequency runoff from maritime to continental climates. The third pattern features a gradient from the centre of the spatial domain toward the north and south. These spatial pattern of simultaneous variations of low-frequency runoff in Europe can be directly related to analogue patterns extracted from low-frequency precipitation and temperature. The precipitation pattern, however, was found to have a stronger influence.

## 3.4. Model Performance

The three articles focusing on model performance focus, like the studies concentrating on observed patterns, on specific features of runoff variability on large scales. The key assumption underlying the analysis is that different features of runoff variability may correspond to different flow processes and that a model may have appropriate representation of some of them but not necessarily for the others. One simple example is that simulated catchment runoff may be highly correlated to the observations (i.e. the temporal evolution of precipitation is correctly translated into runoff), while being biased (i.e. too much or too little water is discharged). Similar arguments have previously been used advocate “signature indices”, that measure specific properties of the runoff dynamics for model evaluation [*Gupta et al.*, 2008; *Yilmaz et al.*, 2008].

Some similarities in the text of the two articles that analyse the multi model ensemble of nine large-scale hydrological models (Articles III and IV), with respect to background literature as well as the description of models, were unavoidable. These articles do also make use of almost the same performance metrics, although different features of runoff variability are analysed. The first performance metric is the relative difference in the mean of observed and modelled runoff, also referred to as the relative bias. The second performance metric the relative difference in standard deviation. This performance metric provides insights on the model ability to capture the amplitude of the analysed features of runoff variability. The third performance metric is differs slightly for the two articles. Article III uses the squared Pearson correlation coefficient, to quantify the degree of similarity between observed and simulated runoff percentiles. Article IV also uses the unsquared Pearson correlation coefficient to compare the phasing of observed and simulated mean annual cycles.

Article V, however, differs as it is based on the evaluation of one single model with a relatively high spatial resolution and the assessment of model performance is based on the evaluation of nonparametric anomalies.

## Article III

### Comparing Large-scale Hydrological Models to Observed Runoff Percentiles in Europe

This study aimed at assessing how well a comprehensive ensemble of nine large-scale hydrological models captures the inter-annual variability of runoff in Europe. Special attention was given to aspects of runoff that are important for the simulation of large-scale floods and droughts.

For each grid-cell with runoff observations, observed and simulated daily series were aggregated to five time series of annual runoff percentiles (5%, 25%, 50%, 75% and 95%) representing low, mean and high flows. Finally, the resulting annual percentile series of the individual grid-cells were spatially aggregated to obtain five time series of average percentile values for every model as well as for the observations. Limiting the model evaluation to spatially averaged percentile series enabled to focus on the models' ability to capture the dominant features of the inter-annual variability of runoff.

Overall, the models capture the temporal evolution of annual low, mean and high flows reasonably well. The relatively small differences among the models suggests that this is likely related to the fact that all annual runoff percentiles closely follow the atmospheric drivers. However, errors in the mean and the standard deviations are getting increasingly pronounced for the low runoff percentiles, where differences among the models are also larger. As low flows are more dependent on catchment properties (especially storages), this emphasises the large uncertainty in the appropriate parametrisation of sub-surface processes.

The large differences among the models were contrasted by the overall good performance of the ensemble mean, which was treated as an additional model. The ensemble mean can thus be used as a robust estimator of spatially aggregated runoff percentiles.

## Article IV

### Seasonal Evaluation of Nine Large-Scale Hydrological Models Across Europe

This study focused on assessing the ability of an ensemble of nine large-scale hydrological models to simulate the mean annual cycle of runoff at the grid-cell scale. Besides a robust quantification of model skill, the study also attempted also to do a diagnostic model evaluation, with the aim of getting insights into the structural deficiencies leading to poor model performance. In order to achieve robust quantifications of model skill, the performance metrics derived for each grid-cell were averaged for regions with comparable mean annual cycles, also referred to as hydroclimatic regime classes.

The performance of all nine models were found to have a large spatial variability, indicating pronounced differences in model performance among the grid-cells. At this local scale, errors can at times be large and grid cells with biases exceeding 100% are often found close to nearly unbiased grid-cells. Thus, the interpretation of individual grid-cells is likely to result in biased conclusions. Instead, regime class wide averages of model performance were used as these provide a more robust picture. Model performance was found to vary systematically among regime classes and was best in regions with little influence of snow. In cold regions, many models were found to have shortcomings related to the timing of the mean annual cycle, which is likely related to issues in the modelling of snow accumulation and melt. Further, some of the observed errors are likely to be related to biases in the forcing data, as indicated by the long term mean of the observed runoff rates being in some instances, larger than the precipitation of the input data. Despite some considerable effort differences, in model performance could not be related to differences in model structure. This documents the limits of attributing model errors to particular model components in the case of complex large-scale hydrological models. These results thus also emphasise the need for transparent documentation of the mathematical structure of the models, the data sources used for parameter identification, and last but not least, the technical choices made throughout model implementation as a prerequisite for further diagnostic model evaluation.

### Article V

### Streamflow data from small basins: a challenging test to high resolution regional climate modelling

This study compared simulated runoff from the land-surface scheme of a high-resolution (0.12°) regional climate model to observed runoff in Europe, in order to assess the models' ability to capture the dynamical properties of streamflow anomalies, with special emphasis on low and high flows.

The data analysis was based on daily series which were deseasonalised and standardised by replacing the actual values with the percentiles of the empirical frequency distribution of each calender day. The resulting time series are referred to as nonparametric anomaly series. These series were then stratified into 19 equally spaced anomaly levels, in order to distinguish low, mean and high flows. For each anomaly level, four performance metrics were applied. The first two operate on daily timescales and focuses on the temporal evolution of wet and dry events as well as on their occurrence frequency. The second two are defined on the annual timescale and compare the inter-annual variability as well as trends of annual anomalies.

All four indices exhibited similar patterns of model performance among the different anomaly levels. Simulated runoff had the least agreement with the observations for the lowest anomaly levels. Model performance was found to be best for moderately high anomalies, decreasing again for the extremely wet conditions. For all anomaly levels, model performance was found to have a large spatial variability with a tendency of decreasing model performance from the west to the east and from low to high altitudes.

## 3.5. Conclusions

The overall aim of this dissertation was to contribute to a better understanding of the terrestrial water balance with special emphasis on runoff patterns emerging at large, continental, scales. Large-scale patterns of different aspects of runoff in space and time have been assessed focusing both on the identification of observed structures as well as the ability of large-scale hydrological models to capture these. Although differing in approach, the combination of studies focusing on observed patterns, and studies concentrating on model performance allowed for some new insights into features of European runoff variability and raised questions that can be used to guide further research.

### Spatial Patterns of Low-frequency Variability

Both the analysis of time series of annual runoff percentiles (Article I) and the examination of low-frequency components of monthly runoff (Article II) documented large similarity in the temporal patterns of nearby rivers. These investigations also showed how the similarity in the temporal evolution of runoff on timescales longer than one year changes over long spatial distances, leading to the identification of clear regional patterns of simultaneous runoff variability. The strength of these patterns as well as the similarity to analogue patterns found in low-frequency precipitation and temperature, suggests that runoff variability on timescales longer than one year, predominantly follows the atmospheric drivers. This is also supported by the fact that large-scale hydrological models were found to capture the temporal evolution of time series of annual runoff statistics reasonably well (Articles III and V). Topics that could not be addressed within the framework of this thesis are related to the characterisation of the actual sequence of wet and dry years that underly these patterns. Questions of specific interest may be directed to the identification of (quasi) periodic phenomena and the influence of large-scale atmospheric circulation patterns. In recent years, the feedback between terrestrial water storage and atmospheric variability has received increasing attention [*Bierkens and van den Hurk*, 2007; *Seneviratne et al.*, 2010]. In this context, it may be of interest to assess whether runoff observations from a large number of small catchments can provide added information that may assist to fully understand the influence of land-atmosphere interactions on climate variability.

### Differences Between Low and High Flows

The analysis of observed runoff percentiles (Article I) revealed that time series of annual high flow statistics have a larger degree of spatial dependence than time series of annual mean and low flow statistics. This finding suggests that hydrological systems are more closely linked to the atmospheric drivers in periods of high flows. In periods of low flows, runoff is fed by terrestrial water storage (e.g. ground water, soil moisture and lakes) and their properties vary largely in space, which may explain the observed pattern. This result is to some extent unexpected, as the inter-annual variability of low flows follows predominantly slow fluctuations in the climatic water balance which have a high degrees of regional coherence. The influence of terrestrial water storage on the inter-annual variability of low runoff percentiles is also supported by the systematic decrease in model performance from high to low flows (Articles III and V). The relatively poor model performance as well as decreasing discrepancies among large-scale hydrological models for low flows, illustrate the large uncertainties in the appropriate description of processes involved in the storage and release of water, which is one of the main topics within hydrological research.

#### Influence of Climatic Conditions on Runoff Dynamics

Several measures of model performance (Articles IV and V) as well as observed characteristics of runoff variability (Article II), change systematically in space, often along climatic gradients. Among the relevant climatic characteristics identified were the seasonality of runoff as well as water balance components, such as mean annual runoff. This suggests that the mean annual conditions may act as a primary control of runoff dynamics. Although several of the underlying mechanisms, such as the impact of snow on runoff in cold regions, or the fact that catchments respond quicker to precipitation in wet conditions, are relatively well understood, the role of the mean climatic conditions on runoff dynamics has received relatively little attention. Therefore, studies that clearly document the influence of mean climatic conditions on the dynamic properties of runoff appearer to be of great interest. Further research on this topic could for example follow previous work that focuses on the role of soil moisture on the longterm water-energy balance [e.g. *Milly*, 1994] or be motivated by the findings of studies that relate both climatic conditions and catchment properties to statistical flood moments [e.g. *Merz and Blöschl*, 2009].

### Implications From Model Evaluation

The runoff simulations of the multi model ensemble (Article III and IV) were found, on average, to underestimate runoff, despite the fact that they have been forced with bias corrected atmospheric variables. The spatial patterns of underestimation of runoff, inconsistencies between forcing data and runoff observations, and available literature [e.g. *Adam et al.*, 2006; *Barstad et al.*, 2009; *Weedon et al.*, 2011] suggest that orographic effects underly these biases. The incapability of forcing data to fully resolve small-scale variations in precipitation, paired with the uncertainty of the mapped land properties that are used to derive model parameters, are likely to underly the large spatial variability in model performance (Articles IV and V). This rapid succession of good and poor model performance between neighbouring grid-cells highlights the limits of large-scale hydrological modelling. These models are built to capture the dominant features of the terrestrial water balance on large scales and interpreting their predictions locally (e.g. individual grid-cells) will lead to inconsistent findings.

The evaluation of the multi model ensemble also revealed large differences in the runoff simulations among the models. These differences highlight the dissension on the appropriate description of hydrological systems and no unambiguous criterion for the acceptability of a certain model could be identified. This implies that the application of any single model is associated with a high risk of biased conclusions. However, the mean of the simulations of a large number of models, also referred to as the “ensemble mean”, has been shown to be a reliable predictor for several features of runoff variability on large scales. Thus the ensemble mean is recommended to make runoff predictions of large-scale hydrological models more robust.

The complexity of large-scale hydrological models did render an unambiguous allocation of model errors to specific parameterisations impossible. However, the fact that regular structures in model error could be clearly related to distinct hydrological phenomena (e.g. differences between low and high flows, differences among hydroclimatic regimes) shows that much can be gained from using large-scale patterns for model validation. These patterns can, for example, be used for “process-based evaluation of model hypotheses” [*Clark et al.*, 2011], where the plausibility of distinct model components is judged upon their relevance for simulating observed features of runoff variability on large scales.

# Bibliography


Adam, J. C., and D. P. Lettenmaier (2003), Adjustment of global gridded precipitation for systematic bias, *J. Geophys. Res.*, *108(D9)*, 4257, doi:10.1029/2002JD002499.

Adam, J. C., E. A. Clark, D. P. Lettenmaier, and E. F. Wood (2006), Correction of global precipitation products for orographic effects, *Journal of Climate*, *19*(1), 15–38, doi:10.1175/JCLI3604.1.

Allen, R., L. Pereira, D. Raes, and M. Smith (1998), Crop evapotranspiration guidelines for computing crop water requirements, *FAO Irrigation and Drainage Paper 56*, FAO.

Arnell, N. (2002), *Hydrology and global environmental change*, 346 pp., Prentice Hall, Harlow.

Arnell, N. W. (1994), *FRIEND: Flow Regimes from International Experimental and Network Data*, vol. 221, chap. Variations over time in European hydrological behaviour: a spatial perspective, pp. 179 – 184, The International Association of Hydrological Sciences (IAHS).

Arora, V. K. (2002), The use of the aridity index to assess climate change effect on annual runoff, *Journal of Hydrology*, *265*(1-4), 164 – 177, doi:10.1016/S0022-1694(02)00101-4.

Baldocchi, D., et al. (2001), Fluxnet: A new tool to study the temporal and spatial variability of ecosystemscale carbon dioxide, water vapor, and energy flux densities, *Bulletin of the American Meteorological Society*, *82*(11), 2415–2434, doi:10.1175/1520-0477(2001)082⟨2415:FANTTS⟩2.3.CO;2.

Balsamo, G., A. Beljaars, K. Scipal, P. Viterbo, B. van den Hurk, M. Hirschi, and A. K. Betts (2009), A revised hydrology for the ecmwf model: Verification from field site to terrestrial water storage and impact in the integrated forecast system, *Journal of Hydrometeorology*, *10*(3), 623–643, doi:10.1175/2008JHM1068.1.

Barlow, M., S. Nigam, and E. H. Berbery (2001), Enso, pacific decadal variability, and u.s. summertime precipitation, drought, and stream flow, *Journal of Climate*, *14*(9), 2105–2128, doi:10.1175/1520-0442(2001)014⟨2105:EPDVAU⟩2.0.CO;2.

Barnston, A. G., and R. E. Livezey (1987), Classification, seasonality and persistence of low-frequency atmospheric circulation patterns, *Monthly Weather Review*, *115*(6), 1083–1126, doi:10.1175/1520-0493(1987)115⟨1083:CSAPOL⟩2.0.CO;2.

Barstad, I., W. W. Grabowski, and P. K. Smolarkiewicz (2007), Characteristics of large-scale orographic precipitation: Evaluation of linear model in idealized problems, *Journal of Hydrology*, *340*(1-2), 78 – 90, doi:10.1016/j.jhydrol.2007.04.005.

Barstad, I., A. Sorteberg, F. Flatø, and M. Déqué (2009), Precipitation, temperature and wind in norway: dynamical downscaling of era40, *Climate Dynamics*, *33*, 769–776, doi:10.1007/s00382-008-0476-5, 10.1007/s00382-008-0476-5.

Bartlein, P. J. (1982), Streamflow anomaly patterns in the u.s.a. and southern canada – 1951-1970, *Journal of Hydrology*, *57*(1-2), 49 – 63, doi:10.1016/0022-1694(82)90102-0.

Best, M. J., et al. (2011), The joint uk land environment simulator (jules), model description - part 1: Energy and water fluxes, *Geoscientific Model Development Discussions*, *4*(1), 595–640, doi:10.5194/gmdd-4-595-2011.

Bierkens, M. F. P., and B. J. J. M. van den Hurk (2007), Groundwater convergence as a possible mechanism for multi-year persistence in rainfall, *Geophys. Res. Lett.*, *34*, L02,402, doi:10.1029/2006GL028396.

Blöschl, G., and M. Sivapalan (1995), Scale issues in hydrological modelling: A review, *Hydrological Processes*, *9*(3-4), 251–290, doi:10.1002/hyp.3360090305.

Blöschl, G., and M. Sivapalan (1997), Process controls on regional flood frequency: Coefficient of variation and basin scale, *Water Resour. Res.*, *33*, 2967 – 2980, doi:10.1029/97WR00568.

Bouwer, L. M., J. E. Vermaat, and J. C. J. H. Aerts (2006), Winter atmospheric circulation and river discharge in northwest europe, *Geophys. Res. Lett.*, *33*(6), L06,403, doi:10.1029/2005GL025548.

Bouwer, L. M., J. E. Vermaat, and J. C. J. H. Aerts (2008), Regional sensitivities of mean and peak river discharge to climate variability in europe, *J. Geophys. Res.*, *113*, D19,103, doi:10.1029/2008JD010301.

Budyko, M. I. (1974), *Climate and life*, 508 pp., Academic Press, New York.

Caylor, K. K., P. D'Odorico, and I. Rodriguez-Iturbe (2006), On the ecohydrology of structurally heterogeneous semiarid landscapes, *Water Resources Research*, *42*(7), W07,424.

Chen, J. L., C. R. Wilson, and B. D. Tapley (2010), The 2009 exceptional amazon flood and interannual terrestrial water storage change observed by grace, *Water Resour. Res.*, *46*(12), W12,526, doi:10.1029/2010WR009383.

Clark, J. S., et al. (2001), Ecological forecasts: An emerging imperative, *Science*, *293*(5530), 657–660, doi:10.1126/science.293.5530.657.

Clark, M. P., A. G. Slater, D. E. Rupp, R. A. Woods, J. A. Vrugt, H. V. Gupta, T. Wagener, and L. E. Hay (2008), Framework for understanding structural errors (fuse): A modular framework to diagnose differences between hydrological models, *Water Resour. Res.*, *44*, W00B02, doi:10.1029/2007WR006735.

Clark, M. P., H. K. McMillan, D. B. G. Collins, D. Kavetski, and R. A. Woods (2011), 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.

Clausen, B., and B. J. F. Biggs (2000), Flow variables for ecological studies in temperate streams: groupings based on covariance, *Journal of Hydrology*, *237*(3-4), 184–197, doi:10.1016/S0022-1694(00)00306-1.

Daly, E., and A. Porporato (2005), A review of soil moisture dynamics: From rainfall infiltration to ecosystem response, *Environmental Engineering Science*, *22*(1), 9–24, doi:10.1089/ees.2005.22.9.

Davidson, E. A., L. V. Verchot, J. H. Cattônio, I. L. Ackerman, and J. Carvalho (2000), Effects of soil water content on soil respiration in forests and cattle pastures of eastern amazonia, *Biogeochemistry*, *48*(1), 53–69, doi:10.1023/A:1006204113917.

de Jeu, R., W. Wagner, T. Holmes, A. Dolman, N. van de Giesen, and J. Friesen (2008), Global soil moisture patterns observed by space borne microwave radiometers and scatterometers, *Surveys in Geophysics*, *29*, 399–420, doi:10.1007/

s10712-008-9044-0, 10.1007/s10712-008-9044-0.

Deardorff, J. W. (1978), Efficient prediction of ground surface temperature and moisture, with inclusion of a layer of vegetation, *J. Geophys. Res.*, *83*(C4), 1889–1903, doi:10.1029/JC083iC04p01889.

Dettinger, M., and H. Diaz (2000), Global characteristics of stream flow seasonality and variability, *Journal of Hydrometeorology*, *1*(4), 289–310, doi: 10.1175/1525-7541(2000)001⟨0289:GCOSFS⟩2.0.CO;2.

Donohue, R. J., M. L. Roderick, and T. R. McVicar (2007), On the importance of including vegetation dynamics in budyko's hydrological model, *Hydrology and Earth System Sciences*, *11*(2), 983–995.

Ehsanzadeh, E., T. B. M. J. Ouarda, and H. M. Saley (2011), A simultaneous analysis of gradual and abrupt changes in canadian low streamflows, *Hydrological Processes*, *25*(5), 727–739, doi:10.1002/hyp.7861.

Engeland, K., H. Hisdal, and S. Beldring (2006), Predicting low flows in ungauged catchments, in *Climate Variability and Change - Hydrological Impacts (Proceedings of the Fifth FRIEND World Conference held at Havana, Cuba, November 2006)*, 308, pp. 163–168, IAHS publ.

Entin, J. K., A. Robock, K. Y. Vinnikov, S. E. Hollinger, S. Liu, and A. Namkhai (2000), Temporal and spatial scales of observed soil moisture variations in the extratropics, *J. Geophys. Res.*, *105*, 865 – 877, doi:10.1029/2000JD900051.

Essery, R. L. H., M. J. Best, R. A. Betts, P. M. Cox, and C. M. Taylor (2003), Explicit representation of subgrid heterogeneity in a gcm land surface scheme, *Journal of Hydrometeorology*, *4*(3), 530–543, doi:10.1175/1525-7541(2003)004⟨0530:EROSHI⟩2.0.CO;2.

Famiglietti, J. S., D. Ryu, A. A. Berg, M. Rodell, and T. J. Jackson (2008), Field observations of soil moisture variability across scales, *Water Resour. Res.*, *44*(1), W01,423, doi:10.1029/2006WR005804.

FAO, IIASA, ISRIC, ISSCAS, and JRC (2009), Harmonized world soil database (version 1.1), *Tech. rep.*, FAO, Rome, Italy and IIASA, Laxenburg, Austria.

Findell, K. L., and E. A. B. Eltahir (1997), An analysis of the soil moisture-rainfall feedback, based on direct observations from illinois, *Water Resour. Res.*, *33*, 725 – 735.

Fischer, E. M., S. I. Seneviratne, D. Lthi, and C. Schr (2007), Contribution of land-atmosphere coupling to recent european summer heat waves, *Geophys. Res. Lett.*, *34*(6), L06,707, doi:10.1029/2006GL029068.

Fleig, A. K., L. M. Tallaksen, H. Hisdal, and S. Demuth (2006), A global evaluation of streamflow drought characteristics, *Hydrology and Earth System Sciences*, *10*(4), 535–552.

Fleig, A. K., L. M. Tallaksen, H. Hisdal, and D. M. Hannah (2011), Regional hydrological drought in north-western europe: linking a new regional drought area index with weather types, *Hydrological Processes*, *25*(7), 1163–1179, doi: 10.1002/hyp.7644.

Fuchs, T. (2009), Gpcc annual report for year 2008, *Tech. rep.*, Global Precipitation Climatology Centre (GPCC), DWD (available from reports and publications section of the GPCC homepage - gpcc.dwd.de).

Gerrits, A. M. J., H. H. G. Savenije, E. J. M. Veling, and L. Pfister (2009), Analytical derivation of the budyko curve based on rainfall characteristics and a simple evaporation model, *Water Resour. Res.*, *45*, W04,403.

Gerstengarbe, F.-W., P. C. Werner, and K. Fraedrich (1999), Applying non-hierarchical cluster analysis algorithms to climate classification: Some problems and their solution, *Theoretical and Applied Climatology*, *64*(3), 143–150.

Gerten, D., S. Schaphoff, U. Haberlandt, W. Lucht, and S. Sitch (2004), Terrestrial vegetation and water balance–hydrological evaluation of a dynamic global vegetation model, *Journal of Hydrology*, *286*(1-4), 249 – 270, doi: 10.1016/j.jhydrol.2003.09.029.

Gottschalk, L., J. L. Jensen, D. Lundquist, R. Solantie, and A. Tollan (1979), Hydrologic regions in the nordic countries, *Nordic Hydrology*, *10*(5), 273 – 286, doi:10.2166/nh.1979.019.

Gudmundsson, L., L. M. Tallaksen, and K. Stahl (2011a), Spatial cross-correlation patterns of european low, mean and high flows, *Hydrological Processes*, *25*(7), 1034–1045, doi:10.1002/hyp.7807.

Gudmundsson, L., L. M. Tallaksen, K. Stahl, and A. K. Fleig (2011b), Low-frequency variability of european runoff, *Hydrology and Earth System Sciences Discussions*, *8*(1), 1705–1727, doi:10.5194/hessd-8-1705-2011.

Gudmundsson, L., T. Wagener, L. Tallaksen, and K. Engeland (2011c), Seasonal evaluation of nine large-scale hydrological models across europe, *submitted to Water Resour. Res.*

Gudmundsson, L., et al. (2011d), Comparing large-scale hydrological models to observed runoff percentiles in europe, *in preparation for: Journal of Hydrometeorology*.

Gupta, H. V., T. Wagener, and Y. Liu (2008), Reconciling theory with observations: elements of a diagnostic approach to model evaluation, *Hydrological Processes*, *22*(18), 3802–3813.

Haddeland, I., et al. (2011), Multi-model estimate of the global terrestrial water balance: Setup and first results, *Journal of Hydrometeorology*, *preprint*, doi: 10.1175/2011JHM1324.1.

Hagemann, S., and L. Dümenil Gates (2003), Improving a subgrid runoff parameterization scheme for climate models by the use of high resolution data derived from satellite observations, *Climate Dynamics*, *21*(3), 349–359, doi: 10.1007/s00382-003-0349-x.

Hanasaki, N., S. Kanae, T. Oki, K. Masuda, K. Motoya, N. Shirakawa, Y. Shen, and K. Tanaka (2008), An integrated model for the assessment of global water resources part 1: Model description and input meteorological forcing, *Hydrology and Earth System Sciences*, *12*(4), 1007–1025.

Hannaford, J., and T. Marsh (2006), An assessment of trends in uk runoff and low flows using a network of undisturbed catchments, *International Journal of Climatology*, *26*(9), 1237–1253, doi:10.1002/joc.1303.

Hannaford, J., and T. J. Marsh (2008), High-flow and flood trends in a network of undisturbed catchments in the uk, *International Journal of Climatology*, *28*(10), 1325–1338, doi:10.1002/joc.1643.

Hannah, D. M., S. Demuth, H. A. J. van Lanen, U. Looser, C. Prudhomme, G. Rees, K. Stahl, and L. M. Tallaksen (2010), A review of the status, research opportunities and future of large-scale river flow archives, in *Global Change: Facing Risks and Threats to Water Resources*, *IAHS Publication*, vol. 340, pp. 584 – 590.

Hannah, D. M., S. Demuth, H. A. J. van Lanen, U. Looser, C. Prudhomme, G. Rees, K. Stahl, and L. M. Tallaksen (2011), Large-scale river flow archives:

importance, current status and future needs, *Hydrological Processes*, *25*(7), 1191–1200, doi:10.1002/hyp.7794.

Hanson, R. T., M. W. Newhouse, and M. D. Dettinger (2004), A methodology to asess relations between climatic variability and variations in hydrologic time series in the southwestern united states, *Journal of Hydrology*, *287*, 252–269.

Harte, J. (2002), Toward a synthesis of the newtonian and darwinian worldviews, *Physics Today*, *55*(10), 29–34, doi:10.1063/1.1522164.

Haxeltine, A., I. C. Prentice, and I. D. Creswell (1996), A coupled carbon and water flux model to predict vegetation structure, *Journal of Vegetation Science*, *7*(5), 651–666, doi:10.2307/3236377.

Haylock, M. R., G. C. Cawley, C. Harpham, R. L. Wilby, and C. M. Goodess (2006), Downscaling heavy precipitation over the united kingdom: a comparison of dynamical and statistical methods and their future scenarios, *International Journal of Climatology*, *26*(10), 1397–1415, doi:10.1002/joc.1318.

Haylock, M. R., N. Hofstra, A. M. G. Klein Tank, E. J. Klok, P. D. Jones, and M. New (2008), A european daily high-resolution gridded data set of surface temperature and precipitation for 1950 - 2006, *J. Geophys. Res.*, *113*, D20,119, doi:10.1029/2008JD010201.

Hisdal, H., and L. M. Tallaksen (2003), Estimation of regional meteorological and hydrological drought characteristics: a case study for denmark, *Journal of Hydrology*, *281*(3), 230–247.

Hisdal, H., K. Stahl, L. M. Tallaksen, and S. Demuth (2001), Have streamflow droughts in europe become more severe or frequent?, *International Journal of Climatology*, *21*, 317–333, doi:10.1002/joc.619.

Holman, I. P., M. Rivas-Casado, N. J. K. Howden, J. P. Bloomfield, and A. T. Williams (2009), Linking north atlantic ocean-atmosphere teleconnection patterns and hydrogeological responses in temperate groundwater systems, *Hydrological Processes*, *23*(21), 3123–3126, doi:10.1002/hyp.7466.

Hosking, J. R. M., and J. R. Wallis (1997), *Regional frequency analysis: an approach based on L-moments*, Cambridge University Press, Cambridge, U.K.

Hurrell, J., and H. van Loon (1997), Decadal variations in climate associated with the north atlantic oscillation, *Climatic Change*, *36*(3-4), 301–326, doi:10.1023/A:1005314315270.

Hurrell, J. W. (1995), Decadal trends in the north atlantic oscillation: Regional temperatures and precipitation, *Science*, *269*(5224), 676–679, doi:10.1126/science.269.5224.676.

Ionita, M., N. Rimbu, and G. Lohmann (2011), Decadal variability of the elbe river streamflow, *Int. J. Climatol.*, *31*(1), 22–30, doi:10.1002/joc.2054.

Jacobeit, J., R. Glaser, J. Luterbacher, and H. Wanner (2003), Links between flood events in central europe since ad 1500 and large-scale atmospheric circulation modes, *Geophys. Res. Lett.*, *30*(4), 1172.

James, P. M. (2007), An objective classification method for hess and brezowsky grosswetterlagen over europe, *Theoretical and Applied Climatology*, *88*(1), 17–42.

Jónsdóttir, J. F., and C. B. Uvo (2009), Long-term variability in precipitation and streamflow in iceland and relations to atmospheric circulation, *International Journal of Climatology*, *29*(10), 1369–1380, doi:10.1002/joc.1781.

Ju, W., S. Wang, G. Yu, Y. Zhou, and H. Wang (2010), Modeling the impact of drought on canopy carbon and water fluxes for a subtropical evergreen coniferous plantation in southern china through parameter optimization using an ensemble kalman filter, *Biogeosciences*, *7*(3), 845–857.

Kalayci, S., and E. Kahya (2006), Assessment of streamflow variability modes in turkey: 1964-1994, *Journal of Hydrology*, *324*(1-4), 163 – 177, doi:10.1016/j.jhydrol.2005.10.002.

Kalma, J., T. McVicar, and M. McCabe (2008), Estimating land surface evaporation: A review of methods using remotely sensed surface temperature data, *Surveys in Geophysics*, *29*, 421–469, doi:10.1007/s10712-008-9037-z.

Kalnay, E., et al. (1996), The ncep/ncar 40-year reanalysis project, *Bulletin of the American Meteorological Society*, *77*(3), 437–471, doi:10.1175/1520-0477(1996)077⟨0437:TNYRP⟩2.0.CO;2.

Kampf, S. K., and S. J. Burges (2007), A framework for classifying and comparing distributed hillslope and catchment hydrologic models, *Water Resour. Res.*, *43*, W05,423, doi:10.1029/2006WR005370.

Khaliq, M. N., T. B. M. J. Ouarda, P. Gachon, and L. Sushama (2008), Temporal evolution of low-flow regimes in canadian rivers, *Water Resource Research*, *44*, W08,436.

Kingston, D. G., D. M. Lawler, and G. R. McGregor (2006a), Linkages between atmospheric circulation, climate and streamflow in the northern north atlantic: research prospects, *Progress in Physical Geography*, *30*(2), 143–174, doi:10.1191/0309133306pp471ra.

Kingston, D. G., G. R. McGregor, D. M. Hannah, and D. M. Lawler (2006b), River flow teleconnections across the northern north atlantic region, *Geophys. Res. Lett.*, *33*, L14,705, doi:10.1029/2006GL026574.

Kingston, D. G., D. M. Hannah, D. M. Lawler, and G. R. McGregor (2009a), Climate-river flow relationships across montane and lowland environments in northern europe, *Hydrological Processes*, *23*(7), 985–996, doi:10.1002/hyp.7202.

Kingston, D. G., M. C. Todd, R. G. Taylor, J. R. Thompson, and N. W. Arnell (2009b), Uncertainty in the estimation of potential evapotranspiration under climate change, *Geophys. Res. Lett.*, *36*(20), L20,403.

Kingston, D. G., D. M. Hannah, D. M. Lawler, and G. R. McGregor (2011), Regional classification, variability, and trends of northern north atlantic river flow, *Hydrological Processes*, *25*(7), 1021–1033, doi:10.1002/hyp.7655.

Kirchner, J. W. (2009), Catchments as simple dynamical systems: Catchment characterization, rainfall-runoff modeling, and doing hydrology backward, *Water Resour. Res.*, *45*, W02,429.

Klemeš, V. (1983), Conceptualization and scale in hydrology, *Journal of Hydrology*, *65*(1-3), 1 – 23, doi:DOI:10.1016/0022-1694(83)90208-1.

Knapp, A. K., et al. (2002), Rainfall variability, carbon cycling, and plant species diversity in a mesic grassland, *Science*, *298*(5601), 2202–2205, doi:10.1126/science.1076347.

Knorr, K.-H., and C. Blodau (2009), Impact of experimental drought and rewetting on redox transformations and methanogenesis in mesocosms of a northern fen soil, *Soil Biology and Biochemistry*, *41*(6), 1187 – 1198, doi:10.1016/j.soilbio.2009.02.030.

Koster, R. D., and M. J. Suarez (1999), A simple framework for examining the interannual variability of land surface moisture fluxes, *Journal of Climate*, *12*(7), 1911–1917, doi:10.1175/1520-0442(1999)012⟨1911:ASFFET⟩2.0.CO;2.

Koster, R. D., and M. J. Suarez (2001), Soil moisture memory in climate models, *Journal of Hydrometeorology*, *2*(6), 558–570, doi:10.1175/1525-7541(2001)

002⟨0558:SMMICM⟩2.0.CO;2.

Koster, R. D., M. J. Suarez, R. W. Higgins, and H. M. Van den Dool (2003), Observational evidence that soil moisture variations affect precipitation, *Geophys. Res. Lett.*, *30*(5), 1241, doi:10.1029/2002GL016571.

Koster, R. D., et al. (2004), Regions of strong coupling between soil moisture and precipitation, *Science*, *305*(5687), 1138–1140, doi:10.1126/science.1100217.

Kottek, M., J. Grieser, C. Beck, B. Rudolf, and F. Rubel (2006), World map of the köppen-geiger climate classification updated, *Meteorologische Zeitschrift*, *15*(3), 259 – 263, doi:10.1127/0941-2948/2006/0130.

Krakauer, N. Y., and I. Fung (2008), Mapping and attribution of change in streamflow in the coterminous united states, *Hydrology and Earth System Sciences*, *12*(4), 1111–1120.

Krasovskaia, L., N. W. Arnell, and L. Gottschalk (1994), Flow regimes in northern and western europe: development and application of procedures for classifying flow regimes, in *FRIEND: Flow Regimes from International Experimental and Network Data*, IAHS pubs No 221, pp. 185 – 192.

Kumar, M., and C. J. Duffy (2009), Detecting hydroclimatic change using spatiotemporal analysis of time series in colorado river basin, *Journal of Hydrology*, *374*(1 – 2), 1 – 15, doi:10.1016/j.jhydrol.2009.03.039.

Kundzewicz, Z. W., D. Graczyk, T. Maurer, I. Pińskwar, M. Radziejewski, C. Svensson, and M. Szwed (2005), Trend detection in river flow series: 1. annual maximum flow / détection de tendance dans des séries de débit fluvial: 1. dbit maximum annuel, *Hydrological Sciences Journal - Journal Des Sciences Hydrologiques*, *50*(5), 797 – 810, doi:10.1623/hysj.2005.50.5.797.

Laaha, G., and G. Blöschl (2005), Low flow estimates from short stream flow records–a comparison of methods, *Journal of Hydrology*, *306*(1-4), 264 – 286, doi:10.1016/j.jhydrol.2004.09.012.

Laaha, G., and G. Blöschl (2007), A national low flow estimation procedure for austria, *Hydrological Sciences Journal - Journal Des Sciences Hydrologiques*, *52*(4), 625–644, doi:10.1623/hysj.52.4.625.

Labat, D. (2008), Wavelet analysis of the annual discharge records of the world's largest rivers, *Advances in Water Resources*, *31*(1), 109 – 117, doi:10.1016/j.advwatres.2007.07.004.

Labat, D., J. Ronchail, and J. L. Guyot (2005), Recent advances in wavelet analyses: Part 2–amazon, parana, orinoco and congo discharges time scale variability, *Journal of Hydrology*, *314* (1-4), 289 – 311, doi:10.1016/j.jhydrol.2005.04.004.

Lange, H., and K. Bernhardt (2004), Long-term components and regional synchronization of river runoffs, in *Hydrology: Science and Practice for the 21st Century*, edited by A. Butler, pp. 165 – 170, British Hydrological Society, London.

Lettenmaier, D. P., E. F. Wood, and J. R. Wallis (1994), Hydro-climatological trends in the continental united states, 1948-88, *Journal of Climate*, *7*(4), 586–607, doi:10.1175/1520-0442(1994)007⟨0586:HCTITC⟩2.0.CO;2.

Lins, H. F. (1997), Regional streamflow regimes and hydroclimatology of the united states, *Water Resources Research*, *33*, 1655–1667, doi:10.1029/97WR00615.

Lins, H. F., and J. R. Slack (1999), Streamflow trends in the united states, *Geophysical Research Letters*, *26* (2), 227–230, doi:10.1029/1998GL900291.

Liu, Y. Y., A. I. J. M. van Dijk, R. A. M. de Jeu, and T. R. H. Holmes (2009), An analysis of spatiotemporal variations of soil and vegetation moisture from a 29-year satellite-derived data set over mainland australia, *Water Resour. Res.*, *45*, W07,405.

Lohse, K. A., P. D. Brooks, J. C. McIntosh, T. Meixner, and T. E. Huxman (2009), Interactions between biogeochemistry and hydrologic systems, *Annual Review of Environment and Resources*, *34* (1), 65–96, doi:10.1146/annurev.environ.33.031207.111141.

Longuevergne, L., B. R. Scanlon, and C. R. Wilson (2010), Grace hydrological estimates for small basins: Evaluating processing approaches on the high plains aquifer, usa, *Water Resour. Res.*, *46* (11), W11,517, doi:10.1029/2009WR008564.

Lorenzo-Lacruz, J., S. M. Vicente-Serrano, J. I. López-Moreno, J. C. González-Hidalgo, and E. Morán-Tejeda (2011), The response of iberian rivers to the north atlantic oscillation, *Hydrology and Earth System Sciences Discussions*, *8* (3), 4459–4493, doi:10.5194/hessd-8-4459-2011.

Manabe, S. (1969), Climate and the ocean circulation - 1: The atmospheric circulation and the hydrology of the earths surface, *Monthly Weather Review*, *97* (11), 739–774, doi:10.1175/1520-0493(1969)097⟨0739:CATOC⟩2.3.CO;2.

Massei, N., B. Laignel, J. Deloffre, J. Mesquita, A. Motelay, R. Lafite, and A. Durand (2010), Long-term hydrological changes of the seine river flow (france) and their relation to the north atlantic oscillation over the period 1950 – 2008, *International Journal of Climatology*, *30*(14), 2146–2154, doi:10.1002/joc.2022.

McCabe, G. J., and D. M. Wolock (2002), A step increase in streamflow in the conterminous united states, *Geophys. Res. Lett.*, *29*, 2185, doi:10.1029/2002GL015999.

McMahon, T. A., M. C. Peel, R. M. Vogel, and G. G. Pegram (2007a), Global streamflows - part 3: Country and climate zone characteristics, *Journal of Hydrology*, *347*(3-4), 272 – 291, doi:DOI:10.1016/j.jhydrol.2007.09.013.

McMahon, T. A., R. M. Vogel, M. C. Peel, and G. G. Pegram (2007b), Global streamflows - part 1: Characteristics of annual streamflows, *Journal of Hydrology*, *347*(3-4), 243 – 259, doi:10.1016/j.jhydrol.2007.09.002.

Meigh, J. R., A. A. McKenzie, and K. J. Sene (1999), A grid-based approach to water scarcity estimates for eastern and southern africa, *Water Resources Management*, *13*(2), 85–115, doi:10.1023/A:1008025703712.

Merz, R., and G. Blöschl (2008a), Flood frequency hydrology: 1. temporal, spatial, and causal expansion of information, *Water Resour. Res.*, *44*, W08,432.

Merz, R., and G. Blöschl (2008b), Flood frequency hydrology: 2. combining data evidence, *Water Resour. Res.*, *44*, W08,433, doi:10.1029/2007WR006745.

Merz, R., and G. Blöschl (2009), Process controls on the statistical flood moments - a data based analysis, *Hydrological Processes*, *23*(5), 675–696, doi:10.1002/hyp.7168.

Miller, A. E., J. P. Schimel, T. Meixner, J. O. Sickman, and J. M. Melack (2005), Episodic rewetting enhances carbon and nitrogen release from chaparral soils, *Soil Biology and Biochemistry*, *37*(12), 2195 – 2204, doi:10.1016/j.soilbio.2005.03.021.

Milly, P. C. D. (1994), Climate, soil water storage, and the average annual water balance, *Water Resour. Res.*, *30*(7), 2143–2156, doi:10.1029/94WR00586.

Mitchell, T. D., and P. D. Jones (2005), An improved method of constructing a database of monthly climate observations and associated high-resolution grids, *International Journal of Climatology*, *25*(6), 693–712, doi:10.1002/joc.1181.

Moore, R. J. (1985), The probability-distributed principle and runoff production at point and basin scales, *Hydrological sciences journal*, *30*(2), 273 – 297.

Morán-Tejeda, E., L.-M. Ignacio, C.-B. Antonio, and V.-S. Sergio M. (2011), Evaluating duero's basin (spain) response to the nao phases: spatial and seasonal variability, *Hydrological Processes*, *25*(8), 1313–1326, doi:10.1002/hyp.7907.

Mudelsee, M., M. Borngen, G. Tetzlaff, and U. Grunewald (2003), No upward trends in the occurrence of extreme floods in central europe, *Nature*, *425*(6954), 166–169, doi:10.1038/nature01928.

Mudelsee, M., M. Börngen, G. Tetzlaff, and U. Grünewald (2004), Extreme floods in central europe over the past 500 years: Role of cyclone pathway "zugstrasse vb", *J. Geophys. Res.*, *109*, D23,101, doi:10.1029/2004JD005034.

Oki, T., and S. Kanae (2006), Global Hydrological Cycles and World Water Resources, *Science*, *313*(5790), 1068–1072, doi:10.1126/science.1128845.

Olden, J. D., and N. L. Poff (2003), Redundancy and the choice of hydrologic indices for characterizing streamflow regimes, *River Research and Applications*, *19*(2), 101–121, doi:10.1002/rra.700.

Oudin, L., V. Andrassian, J. Lerat, and C. Michel (2008), Has land cover a significant impact on mean annual streamflow? an international assessment using 1508 catchments, *Journal of Hydrology*, *357*(3-4), 303 – 316, doi: 10.1016/j.jhydrol.2008.05.021.

Owe, M., R. de Jeu, and T. Holmes (2008), Multisensor historical climatology of satellite-derived global land surface moisture, *Journal of Geophysical Research (Earth Surface)*, *113*(F12), 1002, doi:10.1029/2007JF000769.

Peel, M. C., B. L. Finlayson, and T. A. McMahon (2007), Updated world map of the köppen-geiger climate classification, *Hydrology and Earth System Sciences*, *11*(5), 1633–1644.

Pekarova, P., P. Miklanek, and J. Pekar (2003), Spatial and temporal runoff oscillation analysis of the main rivers of the world during the 19th-20th centuries, *Journal of Hydrology*, *274*(1-4), 62–79.

Pitman, A. J., and S. E. Perkins (2009), Global and regional comparison of daily 2-m and 1000-hpa maximum and minimum temperatures in three global reanalyses, *Journal of Climate*, *22*(17), 4667–4681, doi:10.1175/2009JCLI2799.1.

Porporato, A., E. Daly, and I. Rodriguez-Iturbe (2004), Soil water balance and ecosystem response to climate change., *The American Naturalist*, *164*(5), 625–632, doi:10.1086/424970.

Prudhomme, C., and M. Genevier (2011), Can atmospheric circulation be linked to flooding in europe?, *Hydrological Processes*, *25*(7), 1180–1190, doi:10.1002/hyp.7879.

Raich, J. W., and C. S. Potter (1995), Global patterns of carbon dioxide emissions from soils, *Global Biogeochem. Cycles*, *9*(1), 23–36, doi:10.1029/94GB02723.

Rajagopalan, B., E. Cook, U. Lall, and B. K. Ray (2000), Spatiotemporal variability of enso and sst teleconnections to summer drought over the united states during the twentieth century, *Journal of Climate*, *13*(24), 4244–4255.

Ramillien, G., J. Famiglietti, and J. Wahr (2008), Detection of continental hydrology and glaciology signals from grace: A review, *Surveys in Geophysics*, *29*, 361–374, doi:10.1007/s10712-008-9048-9, 10.1007/s10712-008-9048-9.

Rees, H. G., K. M. Croker, N. S. Reynard, and A. Gustard (1997), Estimation of renewable water resources in the european union, in *FRIEND97Regional Hydrology: Concepts and Models for Sustainable Water Resource Managemen*.

Regonda, S. K., B. Rajagopalan, M. Clark, and J. Pitlick (2005), Seasonal cycle shifts in hydroclimatology over the western united states, *Journal of Climate*, *18*(2), 372–384, doi:10.1175/JCLI-3272.1.

Reichstein, M., et al. (2003), Modeling temporal and large-scale spatial variability of soil respiration from soil water availability, temperature and vegetation productivity indices, *Global Biogeochem. Cycles*, *17*(4), 1104, doi:10.1029/2003GB002035.

Rîmbu, N., C. Boroneant, C. Buta, and M. Dima (2002), Decadal variability of the danube river flow in the lower basin and its relation with the north atlantic oscillation, *International Journal of Climatology*, *22*(10), 1169–1179, doi:10.1002/joc.788.

Robock, A., C. A. Schlosser, K. Y. Vinnikov, N. A. Speranskaya, J. K. Entin, and S. Qiu (1998), Evaluation of the amip soil moisture simulations, *Global and Planetary Change*, *19*(1-4), 181 – 208, doi:10.1016/S0921-8181(98)00047-2.

Robock, A., K. Y. Vinnikov, G. Srinivasan, J. K. Entin, S. E. Hollinger, N. A. Speranskaya, S. Liu, and A. Namkhai (2000), The global soil moisture data bank, *Bulletin of the American Meteorological Society*, *81*(6), 1281–1299, doi: 10.1175/1520-0477(2000)081⟨1281:TGSMDB⟩2.3.CO;2.

Rodell, M., and J. S. Famiglietti (1999), Detectability of variations in continental water storage from satellite observations of the time dependent gravity field, *Water Resour. Res.*, *35*(9), 2705–2723, doi:10.1029/1999WR900141.

Roderick, M. L., M. T. Hobbins, and G. D. Farquhar (2009a), Pan evaporation trends and the terrestrial water balance. i. principles and observations, *Geography Compass*, *3*(2), 746–760, doi:10.1111/j.1749-8198.2008.00213.x.

Roderick, M. L., M. T. Hobbins, and G. D. Farquhar (2009b), Pan evaporation trends and the terrestrial water balance. ii. energy balance and interpretation, *Geography Compass*, *3*(2), 761–780, doi:10.1111/j.1749-8198.2008.00214.x.

Rodríguez-Iturbe, I., V. Isham, D. R. Cox, S. Manfreda, and A. Porporato (2006), Space-time modeling of soil moisture: Stochastic rainfall forcing with heterogeneous vegetation, *Water Resour. Res.*, *42*, W06D05.

Rose, S. (2009), Rainfall-runoff trends in the south-eastern usa: 1938-2005, *Hydrological Processes*, *23*(8), 1105–1118.

Rudolf, B., and U. Schneider (2005), Calculation of gridded precipitation data for the global land-surface using in-situ gauge observations, in *2nd Workshop of the International Precipitation Working Group (available at: reports and publications section of the GPCC homepage - gpcc.dwd.de).*

Saigusa, N., et al. (2010), Impact of meteorological anomalies in the 2003 summer on gross primary productivity in east asia, *Biogeosciences*, *7*(2), 641–655.

Sankarasubramanian, A., and R. M. Vogel (2002), Annual hydroclimatology of the united states, *Water Resour. Res.*, *38*(6), 1083, doi:10.1029/2001WR000619.

Sankarasubramanian, A., and R. M. Vogel (2003), Hydroclimatology of the continental united states, *Geophys. Res. Lett.*, *30*(7), 1363, doi:10.1029/2002GL015937.

Sankarasubramanian, A., R. M. Vogel, and J. F. Limbrunner (2000), Climate elasticity of streamflow in the united states, *Water Resour. Res.*, *37*, 1771 – 1781.

Savenije, H. H. G. (2009), Hess opinions "the art of hydrology", *Hydrology and Earth System Sciences*, *13*(2), 157–161.

Schaefer, K., A. S. Denning, and O. Leonard (2004), The winter arctic oscillation and the timing of snowmelt in europe, *Geophys. Res. Lett.*, *31*, L22,205, doi: 10.1029/2004GL021035.

Schmidli, J., C. M. Goodess, C. Frei, M. R. Haylock, Y. Hundecha, J. Ribalaygua, and T. Schmith (2007), Statistical and dynamical downscaling of precipitation: An evaluation and comparison of scenarios for the european alps, *J. Geophys. Res.*, *112*(D4), D04,105, doi:10.1029/2005JD007026.

Schmidt, R., et al. (2006), Grace observations of changes in continental water storage, *Global and Planetary Change*, *50*(1-2), 112 – 126, doi:10.1016/j.gloplacha.2004.11.018.

Schneider, U., A. Becker, A. Meyer-Christoffer, M. Ziese, and B. Rudolf (2010), Global precipitation analysis products of the gpcc, *Tech. rep.*, Global Precipitation Climatology Centre (GPCC), DWD (available from reports and publications section of the GPCC homepage - gpcc.dwd.de).

Sen, A. K. (2009), Spectral-temporal characterization of riverflow variability in england and wales for the period 1865 – 2002, *Hydrological Processes*, *23*(8), 1147–1157, doi:10.1002/hyp.7224.

Seneviratne, S. I., T. Corti, E. L. Davin, M. Hirschi, E. B. Jaeger, I. Lehner, B. Orlowsky, and A. J. Teuling (2010), Investigating soil moisture-climate interactions in a changing climate: A review, *Earth-Science Reviews*, *99*(3-4), 125 – 161, doi:10.1016/j.earscirev.2010.02.004.

Serreze, M. C., and C. M. Hurst (2000), Representation of mean arctic precipitation from ncep/ncar and era reanalyses, *Journal of Climate*, *13*(1), 182–201, doi: 10.1175/1520-0442(2000)013⟨0182:ROMAPF⟩2.0.CO;2.

Sheffield, J., A. D. Ziegler, E. F. Wood, and Y. Chen (2004), Correction of the high-latitude rain day anomaly in the ncep/ncar reanalysis for land surface hydrological modeling, *Journal of Climate*, *17*(19), 3814–3828, doi: 10.1175/1520-0442(2004)017⟨3814:COTHRD⟩2.0.CO;2.

Sheffield, J., G. Goteti, and E. F. Wood (2006), Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling, *Journal of Climate*, *19*(13), 3088–3111, doi:10.1175/JCLI3790.1.

Shorthouse, C., and N. Arnell (1997), *FRIEND97 - 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, pp. 77 – 85, The International Association of Hydrological Sciences (IAHS).

Shorthouse, C., and N. Arnell (1999), The effects of climatic variability on spatial characteristics of european river flows, *Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere*, *24*(1-2), 7 – 13, doi:10.1016/S1464-1909(98)00003-3, european Water Resources and Climate Changes Processes.

Shun, T., and C. J. Duffy (1999), Low-frequency oscillations in precipitation, temperature, and runoff on a west facing mountain front: A hydrogeologic interpretation, *Water Resour. Res.*, *35*(1), 191–201, doi:10.1029/98WR02818.

Simmons, A. J., P. D. Jones, V. da Costa Bechtold, A. C. M. Beljaars, P. W. Kllberg, S. Saarinen, S. M. Uppala, P. Viterbo, and N. Wedi (2004), Comparison of trends and low-frequency variability in cru, era-40, and ncep/ncar analyses of surface air temperature, *J. Geophys. Res.*, *109*(D24), D24,115, doi:10.1029/2004JD005306.

Sivapalan, M. (2005), *Encyclopedia of Hydrological Sciences*, chap. 13: Pattern, Process and Function: Elements of a Unified Theory of Hydrology at the Catchment Scale, pp. 193 – 219, John Wiley & Sons, Ltd.

Sivapalan, M., G. Blöschl, R. Merz, and D. Gutknecht (2005), Linking flood frequency to long-term water balance: Incorporating effects of seasonality, *Water Resour. Res.*, *41*, W06,012.

Skøien, J. O., G. Blöschl, and A. W. Western (2003), Characteristic space scales and timescales in hydrology, *Water Resour. Res.*, *39*(10), 1304, doi:10.1029/2002WR001736.

Slack, J. R., and J. M. Landwehr (1992), Hydroclimatic data network (hcdn): A usgs streamflow data set for the united states for the study of climate variations, 1874-1988, *USGS Open-File Report*, *92*, 129–193.

Smakhtin, V. U. (2001), Low flow hydrology: a review, *Journal of Hydrology*, *240*(3-4), 147–186, doi:10.1016/S0022-1694(00)00340-1.

Small, D., S. Islam, and R. M. Vogel (2006), Trends in precipitation and streamflow in the eastern us: Paradox or perception?, *Geophysical Research Letters*, *33*(3), L03,403.

Stahl, K., and S. Demuth (1999), Linking streamflow drought to the occurrence of atmospheric circulation patterns, *Hydrological Sciences Journal - Journal Des Sciences Hydrologiques*, *44*, 467 – 482.

Stahl, K., H. Hisdal, J. Hannaford, L. M. Tallaksen, H. A. J. van Lanen, E. Sauquet, S. Demuth, M. Fendekova, and J. Jódar (2010), Streamflow trends in europe: evidence from a dataset of near-natural catchments, *Hydrology and Earth System Sciences*, *14*(12), 2367–2382, doi:10.5194/hess-14-2367-2010.

Stahl, K., L. M. Tallaksen, L. Gudmundsson, and J. H. Christensen (2011), Streamflow data from small basins: a challenging test to high resolution regional climate modeling, *Journal of Hydrometeorology*, *preprint*, doi:10.1175/2011JHM1356.1.

Stephenson, N. L. (1990), Climatic control of vegetation distribution: The role of the water balance, *The American Naturalist*, *135*(5), pp. 649–670.

Szilagyi, J., and J. Jozsa (2009), Complementary relationship of evaporation and the mean annual water-energy balance, *Water Resour. Res.*, *45*, W09,201, doi:10.1029/2009WR008129.

Tallaksen, L., and H. van Lanen (Eds.) (2004), *Hydrological Drought: Processes and Estimation Methods for Streamflow and Groundwater*, Developements in Water Science, Elsevier.

Tallaksen, L. M., H. Madsen, and B. Clausen (1997), On the definition and modelling of streamflow drought duration and deficit volume, *Hydrological Sciences Journal-Journal Des Sciences Hydrologiques*, *42*(1), 15–33.

Tapley, B. D., S. Bettadpur, M. Watkins, and C. Reigber (2004), The gravity recovery and climate experiment: Mission overview and early results, *Geophys. Res. Lett.*, *31*(9), L09,607, doi:10.1029/2004GL019920.

Teuling, A. J., R. Uijlenhoet, B. van den Hurk, and S. I. Seneviratne (2009), Parameter sensitivity in lsms: An analysis using stochastic soil moisture models and eldas soil parameters, *Journal of Hydrometeorology*, *10*(3), 751–765, doi:10.1175/2008JHM1033.1.

Todini, E. (1996), The arno rainfall–runoff model, *Journal of Hydrology*, *175*(1-4), 339 – 382, doi:10.1016/S0022-1694(96)80016-3.

Tootle, G. A., and T. C. Piechota (2006), Relationships between pacific and atlantic ocean sea surface temperatures and u.s. streamflow variability, *Water Resour. Res.*, *42*, W07,411, doi:10.1029/2005WR004184.

Tootle, G. A., T. C. Piechota, and A. Singh (2005), Coupled oceanic-atmospheric variability and u.s. streamflow, *Water Resour. Res.*, *41*, W12,408, doi:10.1029/2005WR004381.

Uppala, S. M., et al. (2005), The era-40 re-analysis, *Quarterly Journal of the Royal Meteorological Society*, *131*(612), 2961–3012, doi:10.1256/qj.04.176.

van Loon, H., and J. C. Rogers (1978), The seesaw in winter temperatures between greenland and northern europe. part i: General description, *Monthly Weather Review*, *106*(3), 296–310, doi:10.1175/1520-0493(1978)106⟨0296:TSIWTB⟩2.0.CO;2.

Vereecken, H., J. A. Huisman, H. Bogena, J. Vanderborght, J. A. Vrugt, and J. W. Hopmans (2008), On the value of soil moisture measurements in vadose zone hydrology: A review, *Water Resour. Res.*, *44*, W00D06, doi:10.1029/2008WR006829.

Vinnikov, K. Y., N. A. S. A. Robock, and C. A. Schlosser (1996), Scales of temporal and spatial variability of midlatitude soil moisture, *Journal of Geophysical Research*, *101(D3)*, 71637174.

Vogel, R. M., and N. M. Fennessey (1994), Flow-duration. i: New interpretation and confidence intervals, *J. Water Resour. Plann. Manage.*, *120*(4), 458 – 504.

Vogt, J., P. Soille, A. de Jager, E. Rimaviciute, and W. Mehl (2007), A pan-european river and catchment database, *Tech. Rep. EUR 22920 EN*, EC-JRC, Luxembourg.

von Storch, H., and F. W. Zwiers (1999), *Statistical analysis in climate research*, 484 pp., Cambridge University Press, Cambridge, UK.

Wagner, W., V. Naeimi, K. Scipal, R. de Jeu, and J. Martinez-Fernandez (2007), Soil moisture from operational meteorological satellites, *Hydrogeology Journal*, *15*(1), 121–131.

Weedon, G., S. Gomes, P. Viterbo, H. sterle, J. Adam, N. Bellouin, O. Boucher, and M. Best (2010), The watch forcing data 1958-2001: a meteorological forcing dataset for land surface- and hydrological-models, *WATCH Technical Report 22*, (available at www.eu-watch.org).

Weedon, G. P., et al. (2011), Creation of the watch forcing data and its use to assess global and regional reference crop evaporation over land during the twentieth century., *Journal of Hydrometeorology*, *early view*(0), doi:

10.1175/2011JHM1369.1.

Western, A. W., R. B. Grayson, and G. Blöschl (2002), Scaling of soil moisture: A hydrologic perspective, *Annual Review of Earth and Planetary Sciences*, *30*(1), 149–180, doi:10.1146/annurev.earth.30.091201.140434.

Wibig, J. (1999), Precipitation in europe in relation to circulation patterns at the 500 hpa level, *International Journal of Climatology*, *19*(3), 253–269, doi: 10.1002/(SICI)1097-0088(19990315)19:3⟨253::AID-JOC366⟩3.0.CO;2-0.

Wilby, R. L., T. M. L. Wigley, D. Conway, P. D. Jones, B. C. Hewitson, J. Main, and D. S. Wilks (1998), Statistical downscaling of general circulation model output: A comparison of methods, *Water Resour. Res.*, *34*(11), 2995–3008, doi:10.1029/98WR02577.

Wilson, D. J., A. W. Western, and R. B. Grayson (2004), Identifying and quantifying sources of variability in temporal and spatial soil moisture observations, *Water Resour. Res.*, *40*, W02,507, doi:10.1029/2003WR002306.

Yang, D., F. Sun, Z. Liu, Z. Cong, G. Ni, and Z. Lei (2007), Analyzing spatial and temporal variability of annual water-energy balance in nonhumid regions of china using the budyko hypothesis, *Water Resour. Res.*, *43*, W04,426.

Yang, D., W. Shao, P. J.-F. Yeh, H. Yang, S. Kanae, and T. Oki (2009), Impact of vegetation coverage on regional water balance in the nonhumid regions of china, *Water Resour. Res.*, *45*, W00A14.

Yang, H., D. Yang, Z. Lei, and F. Sun (2008), New analytical derivation of the mean annual water-energy balance equation, *Water Resour. Res.*, *44*, W03,410.

Yilmaz, K. K., H. V. Gupta, and T. Wagener (2008), A process-based diagnostic approach to model evaluation: Application to the nws distributed hydrologic model, *Water Resour. Res.*, *44*(9), W09,417, doi:10.1029/2007WR006716.

Zaidman, M. D., H. G. Rees, and A. R. Young (2002), Spatio-temporal development of streamflow droughts in north-west europe, *Hydrology and Earth System Sciences*, *6*(4), 733–751.

Zampieri, M., F. D' Andrea, R. Vautard, P. Ciais, N. de Noblet-Ducoudre, and P. Yiou (2009), Hot european summers and the role of soil moisture in the propagation of mediterranean drought, *Journal of Climate*, *22*(18), 4747–4758, doi:10.1175%2F2009JCLI2568.1.

Zelenhasić, E., and A. Salvai (1987), A method of streamflow drought analysis, *Water Resour. Res.*, *23*(1), 156–168, doi:10.1029/WR023i001p00156.

# Part B.

# Journal Articles

# Article I

---

## Spatial cross-correlation patterns of European low, mean and high flows

Lukas Gudmundsson, Lena M. Tallaksen and Kerstin Stahl

*Published in:*
Hydrological Processes — 2011
Volume: 25. Pages: 1034 – 1045.
DOI: 10.1002/hyp.7807

# Article II

## Low-frequency variability of European runoff

Lukas Gudmundsson, Lena M. Tallaksen, Kerstin Stahl
and Anne K. Fleig

*Revised version of the discussion paper published in:*
Hydrology and Earth System Sciences Discussions — 2011
Volume: 8. Pages: 1705 – 1727.
DOI: 10.5194/hessd-8-1705-2011

# Article II

## Low-frequency variability of European runoff[1]

## 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*

[1] This is a revised version of the discussion paper published online as: Gudmundsson L, Tallaksen L M, Stahl K, Fleig A K. 2011. *Low-frequency variability of European runoff.* Hydrology and Earth System Sciences Discussions, 8, 1705-1727. doi:10.5194/hessd-8-1705-2011

*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 centers 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 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 space-time patterns [e.g. *Lins*, 1997; *Gudmundsson et al.*, 2011] that are related to large-scale atmospheric drivers [e.g. *Barlow et al.*, 2001; *Tootle and Piechota*, 2006]. However, the variance of annual runoff varies largely among catchments [for a global analysis see *McMahon 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 [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. Further, 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 form hydrological systems is closely related to the presence of "longterm-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 analyzing 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 deseasonalised prior to further analysis. However, the seasonal cycle is one of the most important phenomenas in hydrology and contributes significantly to the variance of monthly runoff.

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

$$\frac{dS}{dt} = P - E - Q \tag{1}$$

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

$$Q = h(S) \tag{2}$$

where the function $h$ summarizes various storage response functions in a catchment [see *Clark et al.*, 2011, 2008, 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 characterized 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 storages of groundwater are considered.

To better understand the different roles of meteorological forcing and terrestrial processes on low-frequency runoff variability, the following questions guiding the analysis 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 data – 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, easy 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 and accordingly, a runoff series $Q$ can be decomposed into a set of additive sub-series,

$$Q = \sum_{f \in F} Q_f \tag{3}$$

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

$$\Phi_{\mathrm{Q}} = \frac{\sigma^2_{\mathrm{Q_{Long}}}}{\sigma^2_{\mathrm{Q}}} \tag{4}$$

where $\sigma^2_{\mathrm{Q_{Long}}}$ is the variance of $Q_{\mathrm{Long}}$ and $\sigma^2_{\mathrm{Q}}$ 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 $\Phi_{\mathrm{Q}}$, whereas the temporal evolution of the corresponding low-frequency component $Q_{\mathrm{Long}}$, will not change.

This framework is then used to assess the low-frequency variability of monthly runoff. After isolating $Q_{\mathrm{Long}}$ and quantifying $\Phi_{\mathrm{Q}}$ 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 a set of 358 time series of daily runoff covering the period 1963 to 2000 and where aggregated to monthly values. This unique data set consists of small, near-natural catchments that are not nested (median catchment size: $300\,\mathrm{km}^2$). Most records originate from the European Water Archive (EWA), a database assembled by the Euro-FRIEND[1] program. The EWA is held by the Global Runoff Data Centre[2] (GRDC) which also manages data requests. The EWA was recently updated and complemented by national data from partners in the WATCH project[3]. For a detailed overview on data availability see *Stahl et al.* [2010, 2008]. 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 *Weedon et al.* [WFD; 2010, 2011] were used. The WFD provide bias corrected variables, based on the ERA-40 reanalysis [*Uppala et al.*, 2005], on a 0.5 degree grid. In the centre of the spatial domain of this study, at 55°N, grid-cell cover $1994\,\mathrm{km}^2$. 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 characterized by the mean annual temperature ($\bar{T}$), precipitation ($\bar{P}$) and runoff ($\bar{Q}$) (Fig. 1). The fraction of low-

[1] http://ne-friend.bafg.de/servlet/is/7413/, last accessed: 15 May 2011
[2] http://grdc.bafg.de, last accessed: 15 May 2011
[3] http://eu-watch.org/, last accessed: 15 May 2011

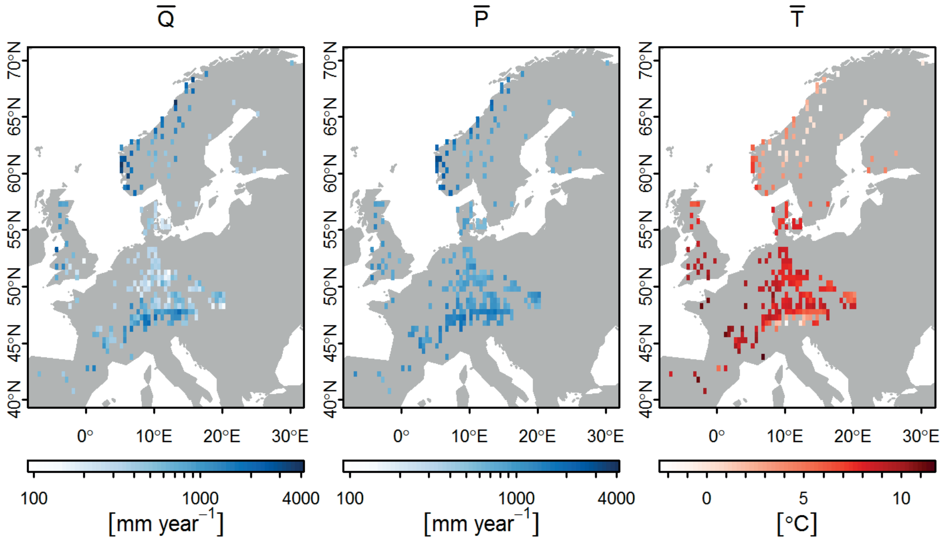


**Figure 1.:** Mean annual runoff ($\bar{Q}$), precipitation ($\bar{P}$) and temperature ($\bar{T}$) for the period 1962–2000.

frequency variance of runoff ($\Phi_{\mathrm{Q}}$), precipitation ($\Phi_{\mathrm{P}}$) and temperature ($\Phi_{\mathrm{T}}$) was used to characterize 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 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 thus 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 ($Q_{\mathrm{Long}}$), temperature ($T_{\mathrm{Long}}$) and precipitation ($P_{\mathrm{Long}}$) 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 $x_{\mathrm{i}}$ and an independent variable $t_{\mathrm{i}}$ (for $i = 1$ to $n$), the LOESS estimates smoothed values of the dependent variable $\hat{g}(t_{\mathrm{i}})$ for any value $t_{\mathrm{i}}$. First, a positive integer $\lambda$ is chosen. Then the subset of the $\lambda$ nearest $t_{\mathrm{j}}$ to $t_{\mathrm{i}}$ are selected, where $j$ is the index of the subset. For this selection weights are computed as

$$w_{\mathrm{j}} = \left[1 - \left(\frac{|t_{\mathrm{j}} - t_{\mathrm{i}}|}{\delta_{\lambda}}\right)^3\right]_+^3 \tag{5}$$

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

STL decomposes a time series $X$ into

$$X = X_{\text{Long}} + X_{\text{Seas}} + X_{\text{Resid}} \tag{6}$$

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 $\lambda_{Long}$ equal to the smallest integer satisfying $\lambda_{Long} \geq 1.5p/(1 - (1.5/\lambda_{Seas})$, 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 $\lambda_{Long} = 19$ 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 $X_{\text{Long}}$.

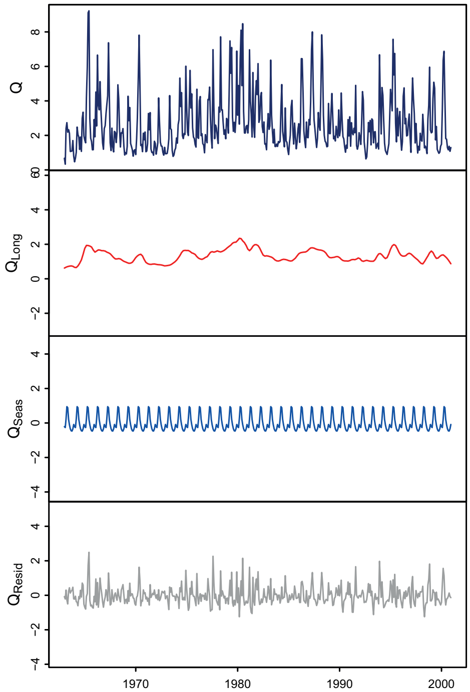


**Figure 2.:** Decomposition of a monthly runoff series $Q$ [mm $d^{-1}$] into low-frequency ($Q_{\text{Long}}$), seasonal ($Q_{\text{Seas}}$) and residual ($Q_{\text{Resid}}$) components.

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 amplitude of the three different components are actually of the same order of magnitude and thus these three sub signals of similar importance. However, only $X_{\text{Long}}$ is further analyzed both with respect to its fractional variance (Eq. (4)) and its space time pattern.

The fraction of low-frequency variance $\Phi_{\text{X}}$, is a parameter characterizing the variance distribution of the power-spectrum of the time series $X$. However, the STL-algorithm used to estimate $\Phi_{\text{X}}$ is primarily designed for time-series decomposition and may thus lead to biased estimates of spectral properties. Appendix 6 summarizes the results of a supplementary analysis comparing $\Phi_{\text{Q}}$ to an alternative (spectral) estimate, based on the multi taper method [e.g. *Ghil et al.*, 2002]. The results show that both estimates are closely related and that $\Phi_X$ estimated using the STL algorithm is reliable. Further, due to sampling errors and finite size effects, the estimated values of $\Phi_X$ are uncertain and thus error-bars for $\Phi_Q$. This uncertainty was quantified using a bootstrapping technique and the resulting confidence intervals of $\Phi_Q$ are also presented in Appendix 6.

## 3.2. Factors influencing $\Phi_Q$

### 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. One is a snow-regime where snow accumulation and melt lead to an annual cycle with minimum discharge in winter and a maximum in spring. The other 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 standardized by removing the mean and then dividing by the standard deviation. These standardized hydroclimatic regimes were than 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 will be 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.

#### 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, $\Phi_{\mathrm{Q}}$. 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.

A statistical testing of the significance of $\rho$ is hampered by the fact that all variables considered are likely to exhibit some spatial dependence, which may lead to spuriously high correlations. Therefore, no formal statistical testing is applied and instead only values satisfying $|\rho| \geq \sqrt{0.25} = 0.5$, explaining more than 25% of the variance of the ranks, will be considered.

### 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 characterized 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 thus isometric feature mapping [ISOMAP; *Tenenbaum et al.*, 2000], which takes advantage of these consideration, was used to characterize the sets of spatially distributed time series of $Q_{\mathrm{Long}}$, $P_{\mathrm{Long}}$ and $T_{\mathrm{Long}}$, separately.

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 $g_{\mathrm{ij}}$ 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

$$\mathbf{Y} = \mathbf{E}\Lambda^{1/2} \tag{7}$$

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, $\mathbf{H}_{ij} = \delta_{ij} - \frac{1}{N}$ 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 $Q_{\mathrm{Long}}$, $T_{\mathrm{Long}}$ and $P_{\mathrm{Long}}$ 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 emphasize common temporal evolution.

The similarity in the spatial patterns of simulations variation, found as the leading ISOMAP dimensions of $Q_{\mathrm{Long}}$, $P_{\mathrm{Long}}$ and $T_{\mathrm{Long}}$, 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 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 standardized leading ISOMAP dimensions of $Q_{\mathrm{Long}}$ and $\mathbf{B}$ contains the standardised leading ISOMAP dimensions of either $P_{\mathrm{Long}}$ or $T_{\mathrm{Long}}$. In principle, the resulting parameters can be analysed in detail, however, in this study, only the goodness of the 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 minimize the sum of squared differences (i.e. linear PA). The strength of this fit can be quantified by

$$r^2 = 1 - \frac{2(1 - \mathrm{trace}(\mathbf{W}))}{\mathrm{trace}(\mathbf{W})^2}, \tag{8}$$

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.

# 4. Results

## 4.1. Fraction of low-frequency variance

On average, the fraction of low-frequency variance was found to be highest for runoff ($0.01 \leq \Phi_{\mathrm{Q}} \leq 0.53$), followed by precipitation ($0.03 \leq \Phi_{\mathrm{P}} \leq 0.13$) and lowest for temperature ($0.01 \leq \Phi_{\mathrm{T}} \leq 0.04$). The large spread in the fraction

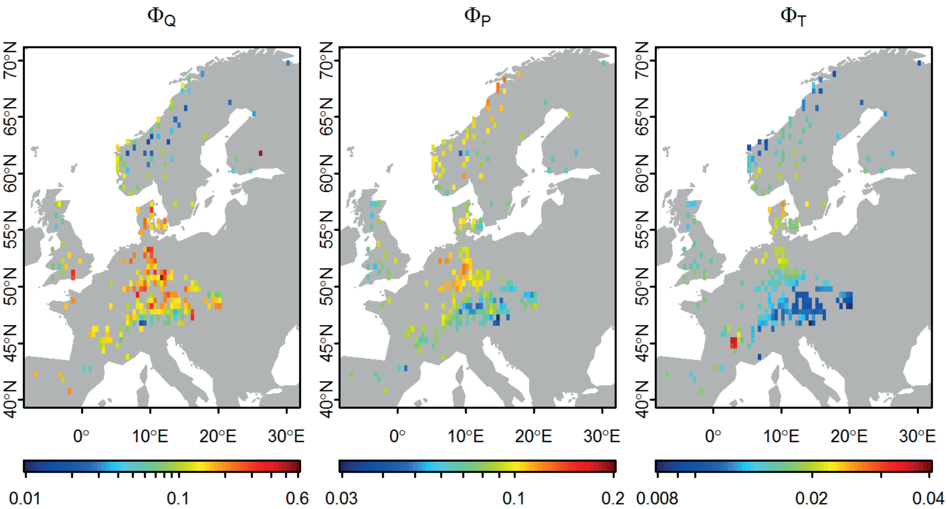


**Figure 3.:** Spatial patterns of the fraction of low-frequency variance of runoff ($\Phi_Q$), precipitation ($\Phi_P$) and temperature ($\Phi_P$). (Note that each panel has its own logarithmic color scale).

of low-frequency variance of runoff, suggests 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 $Q_{\mathrm{Long}}$ 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. Comparison with the spatial distributions of mean annual runoff, precipitation and temperature (Fig. 1) shows that 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. 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.

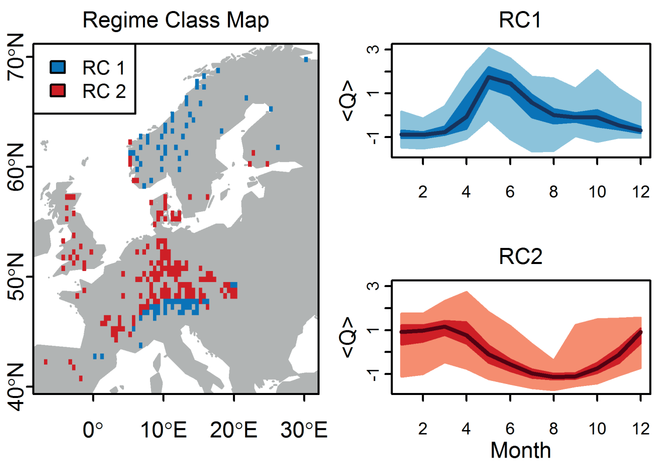


**Figure 4.:** Classification of grid-cells into regime classes. RC1: Winter minimum - summer maximum regime. RC2: Summer minimum - winter maximum regime. The map shows the geographical domain of each regime class. The two small panels show the average shape of each hydroclimatic regime; dark line: median, dark shaded area: inter-quartile range, light shaded area: range; (< > denotes standardisation).

### 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 center of the spatial domain, between the two regions of RC1.

The median $\Phi_Q$ of the two regime classes is significantly different ($p \leq 0.01$, 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. 5a).

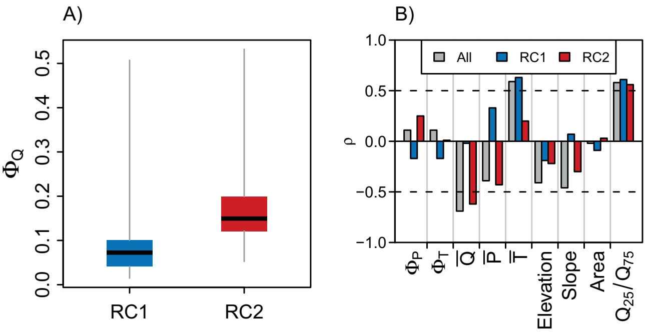


**Figure 5.:** A) Boxplots of $\Phi_Q$ for the two hydroclimatic regime classes (RC1 and RC2). The black horizontal line is the median, the box covers the interquartile range and the gray whiskers range from the minimum to the maximum value. B) Spearman's rank correlation coefficients ($\rho$) between the fraction of low-frequency variance of runoff ($\Phi_\mathrm{Q}$) and the following set of variables: fraction of low-frequency variance of precipitation ($\Phi_\mathrm{P}$) and temperature ($\Phi_\mathrm{T}$); mean annual runoff ($\bar{Q}$), mean annual precipitation ($\bar{P}$) and temperature $\bar{T}$; mean catchment elevation, slope and area as well as the ratio of 25% and 75% percentile of daily runoff $Q_{25}/Q_{75}$.

### Correlating $\Phi_\mathrm{Q}$ with catchment properties and climatic conditions

Figure 5b summarizes the results of the correlation analysis relating $\Phi_\mathrm{Q}$ 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 ($\bar{Q}$), mean annual temperature ($\bar{T}$) and $Q_{25}/Q_{25}$ have correlations satisfying $|\rho| \geq \sqrt{0.25} = 0.50$. (The correlation among these three variables never exceeds $|\rho| \geq 0.50$). The weakest correlations are found for the fraction of low-frequency variance of precipitation ($\Phi_\mathrm{P}$) and temperature ($\Phi_\mathrm{Q}$) as well as for catchment area.

A negative correlation if found between $\Phi_\mathrm{Q}$ and $\bar{Q}$ for all grid-cells as well as the grid-cells in RC2, indicating that the fraction of low-frequency variance of runoff decreases under wetter conditions. However, in the snow-regime (RC1) $\Phi_\mathrm{Q}$ and $\bar{Q}$ have almost zero correlation. As illustrated in Fig. 6a, the $\Phi_\mathrm{Q}$ 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 those in RC1. Thus $\Phi_\mathrm{Q}$ is larger under warmer conditions, as also seen in Fig. 6b, were the $\Phi_\mathrm{Q}$ values for RC1

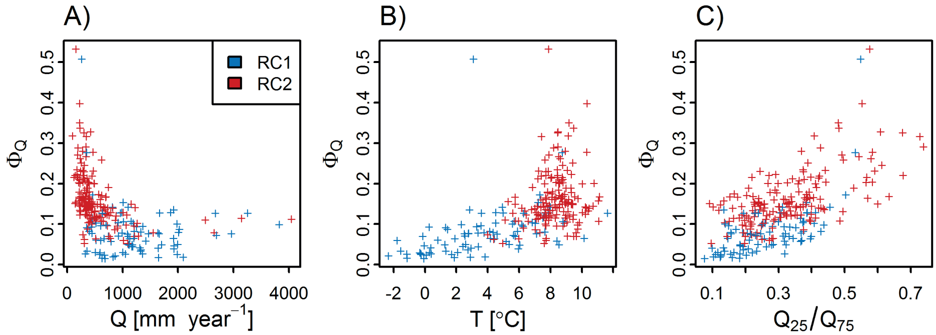


**Figure 6.:** Scatter plots illustrating the influence of selected variables (mean annual runoff $\bar{Q}$ and temperature $\bar{T}$; the ratio of 25% and 75% percentile of daily runoff $Q_{25}/Q_{75}$) on the fraction of low-frequency variance of runoff ($\Phi_{\mathrm{Q}}$).

increase almost linearly with temperature. However, this is not the case for the evapotranspiration-regime (RC2).

Only $Q_{25}/Q_{75}$ has approximately equally strong positive correlations with $\Phi_{\mathrm{Q}}$ in all cases. Figure 6b shows that the rate of increase in $\Phi_{\mathrm{Q}}$ with increasing $Q_{25}/Q_{25}$ are comparable in both regime classes, although overall higher $\Phi_{\mathrm{Q}}$ values are found for RC2.

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

Figure 7 is used to illustrate how the results of ISOMAP can be interpreted in the context of spatially distributed time series, using the ISOMAP of $Q_{\mathrm{Long}}$. The left panel shows the space spanned by the two leading components (i.e. the first two columns of the space $\mathbf{Y}$, Eq. (7)). Each point represents a spatial location (i.e. grid-cell). The gray lines are the neighbourhood graph used to estimate the geodesic distances. Neighboring points in this plot indicate that the underlying time series have a large agreement. The further the points are apart, the larger the difference of 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. This change in temporal evolution occurs gradually along each component (i.e. the axis of the

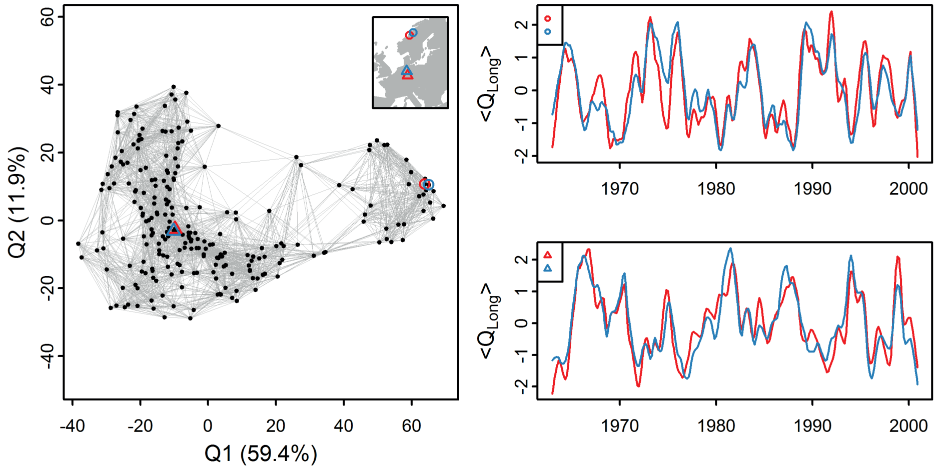


**Figure 7.:** Left: the space spanned by the two leading ISOMAP components derived from $Q_{\mathrm{Long}}$. The gray lines are the neighborhood graph used to estimate the geodesic distances. Two pairs of neighbouring points are marked (triangles and circles) and their spatial location is indicated on the map. Right: time series of $Q_{\mathrm{Long}}$ for the two pairs of points marked in the left panel ($< >$ denotes standardisation).

left panel). These changes of temporal evolution with spatial location can be due to various processes, each of them generating a signal. 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 thus the relation of the time series to each other, and how this relation changes in space.

The three leading ISOMAP components of $Q_{\mathrm{Long}}$, $P_{\mathrm{Long}}$ and $T_{\mathrm{Long}}$ (Fig. 8) 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 ($r^2 = 0.93$, $p \leq 0.001$, see Eq. (8)) 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 center of the spatial domain (around Denmark) toward the north and south. The spatial pattern of the third temperature component ($T3$ : 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 $\Phi_{\mathrm{Q}}$ is on average larger than and not correlated to $\Phi_{\mathrm{P}}$ and $\Phi_{\mathrm{T}}$ (Fig. 5b), 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 $\Phi_{\mathrm{Q}}$ is primarily controlled by catchment processes. In principle, $\Phi_{\mathrm{Q}}$ can be altered by either a change in the variance of high frequency components (i.e. the variance of $Q_{\mathrm{Seas}}$ and $Q_{\mathrm{Resid}}$) or by a change in the variance of the low-frequency component ($Q_{\mathrm{Long}}$). The fact that $\Phi_Q$ is significantly lower in hydroclimatic regimes where the mean annual cycle of runoff is dominated by snow (Fig. 5a) suggests that snow accumulation and melt lead to a pronounced annual cycle, which implies a low $\Phi_{\mathrm{Q}}$. The relatively strong positive correlation of $\Phi_{\mathrm{Q}}$ with the mean annual temperature in 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 $\Phi_{\mathrm{Q}}$ 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 all the variations. The strong negative correlation of $\Phi_{\mathrm{Q}}$ with mean annual runoff in the evaporation dominated regime class (RC2, $\rho = -0.62$) shows that the fraction of

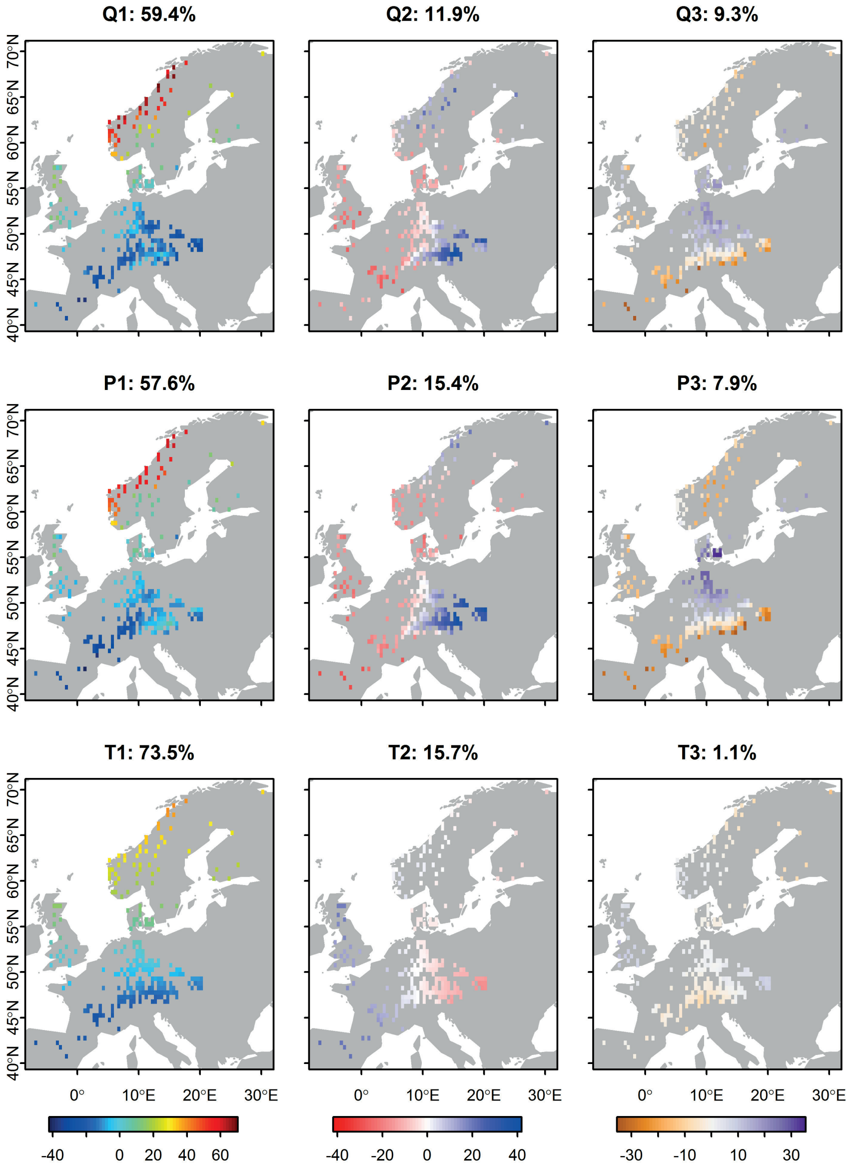


**Figure 8.:** Spatial patterns of the three dominant ISOMAP components of low-frequency variability of runoff ($Q$), precipitation ($P$) and temperature ($T$). The numbers in the figure headings are the percentage of explained variance.

low-frequency variance decreases under wetter conditions. In humid environments, catchment storages are often close to being saturated and thus 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 $\Phi_Q$ is not correlated to mean annual precipitation. 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 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.

Interestingly, none of the considered catchment properties (elevation, slope and area) were found to be correlated to the fraction of low-frequency variance of runoff, suggesting that rainfall-runoff processes controlled by these properties have little influence on $\Phi_Q$. 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.

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 both regime classes, indicating that $\Phi_Q$ is larger for rivers with a dampened runoff response. The ratio $Q_{25}/Q_{75}$ is a parameter characterising the spread of the distribution of daily runoff, which is often characterised by 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 vital role.

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.

## 5.2. Space-time 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 centers of simultaneous variations in the north and the south. 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 [*López-Moreno and Vicente-Serrano*, 2008], 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. With increasing distance from the coast this influence diminishes. Also the third components of runoff and precipitation have similar spatial patterns. Both explain a considerable amount of the variance. The location of the pattern in the center 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 standardized precipitation index on time-scales of 24 months [*Bordi et al.*, 2009].

In principle, such large-scale patterns in runoff can be related to the general features of atmospheric oscillation [e.g. *Tootle and Piechota*, 2006; *Barlow et al.*, 2001; *Shorthouse and Arnell*, 1997]. In the context of this study, however, the pronounced similarity to the equivalent patterns found in precipitation and temperature is a sufficient proof 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, which focuses on terrestrial hydrology.

# 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 ($Q_{\mathrm{Long}}$) 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 and are unlikely to be altered by catchment processes.

The fraction of low-frequency variance of runoff ($\Phi_{\mathrm{Q}}$), however, was found to be on average larger than, and not correlated to, the fraction of low-frequency variance of precipitation ($\Phi_{\mathrm{P}}$) and temperature ($\Phi_{\mathrm{T}}$), suggesting that catchment processes amplify low-frequency fluctuations in the forcing. The large spread of $\Phi_{\mathrm{Q}}$ across Europe indicates large differences in the sensitivity of runoff to low-frequency variability in the forcing. $\Phi_{\mathrm{Q}}$ is on average lowest in regions with a significant influence of snow on the hydroclimatic regime and here a decrease with temperature is found. In evapotranspiration dominated regimes $\Phi_{\mathrm{Q}}$ is generally larger in catchments with lower annual runoff. Overall $\Phi_{\mathrm{Q}}$ was found to be largest for catchments with a dampened rainfall-runoff response. In general, 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. None the less, these findings may be of interest for studies of climate variability and change. 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 $\Phi_{\mathrm{Q}}$. 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: 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 $X$, as an albeit simplistic parameter characterizing the shape of the power spectrum was introduced and all time-series analysis was based on the STL-algorithm, 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 $\Phi_{\mathrm{Q}}$, the fraction of low-frequency variance of runoff. This test was performed by comparing $\Phi_{\mathrm{Q,STL}}$, the estimate used in the main section of the paper, to an alternative estimate, $\Phi_{\mathrm{Q,MTM}}$, which relies on the multi-taper method [MTM; e.g. *Ghil et al.*, 2002] to estimate spectral properties.

$\Phi_{\mathrm{Q,MTM}}$ was estimated by first computing the power-spectrum of $Q$ using MTM. (Following the recommendations of *Ghil et al.* [2002] “discrete prolate spheroidal sequences” (DPSS) - tapers were used). In a second step, $\Phi_{\mathrm{Q,MTM}}$ was determined as the fraction of variance explained by frequencies $> 1/19$ months (being consistent with the STL parameter $\lambda_{\mathrm{Long}}$).

Confidence intervals of both $\Phi_{\mathrm{Q,STL}}$ and $\Phi_{\mathrm{Q,MTM}}$ were obtained using a bootstrapping procedure. To account for the strong seasonality and high serial correlation in the runoff time series we used a block bootstrap, where blocks of a time series with fixed length are resampled instead of single data points [e.g. *Efron and Tibshirani*, 1993]. There is no standard recommendation on the choice of the block length $n_{\mathrm{b}}$, and often $n_{\mathrm{b}}$ is set in an ad-hoc fashion. In the present case, the seasonal pattern of the input series needs to be accounted for, requiring a block-length of at least $n_{\mathrm{b}} = 12$ (months). In order to be consistent with the parameter value chosen for the parameter $\lambda_{Long}$, the block length was set to $n_{\mathrm{b}} = \lambda_{Long} = 19$. The bootstrapping was based on 1000 replications and the 2.5% and the 97.2% percentiles of the bootstrap sample were used to construct 95% confidence intervals.

Figure 9 summarizes the results. The estimates of $\Phi_{\mathrm{Q,STL}}$ and $\Phi_{\mathrm{Q,MTM}}$ (circles) are closely related ($R^2 = 0.93$), indicating that the two approaches provide quantitatively comparable results. The gray horizontal and vertical bars are the

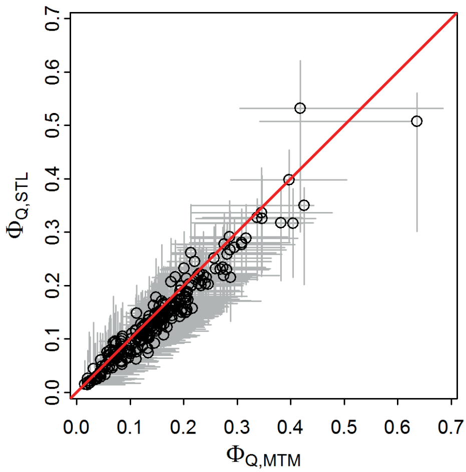


**Figure 9.:** Comparison of $\Phi_{Q,STL}$ and $\Phi_{Q,MTL}$. Gray bars are the 95% confidence intervals based on bootstrap.

boot-strap confidence intervals (CI) of each data point and are measures of stability. For both measures, the CIs have comparable magnitudes. Note that the estimates fall outside of the CI in a few cases (STL: 1.6% and MTM: 0.9 %). This may be related to (a) that the "model" introduced by the block bootstrap may not be fully appropriate, and (b) that 5% of the observations are expected to lie outside the 95% confidence-intervals by construction. Based on these results we conclude that the estimate $\Phi_{Q,STL}$ is equally robust and comparable to the alternative estimate $\Phi_{Q,MTM}$.


## 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 helpful comments of M. Mudelsee and one anonymous reviewer as well as from the editor D. Koutsoyiannis did increase the clarity of the paper.

# Bibliography


Barlow, M., S. Nigam, and E. H. Berbery (2001), Enso, pacific decadal variability, and u.s. summertime precipitation, drought, and stream flow, *Journal of Climate*, *14*(9), 2105–2128, doi:10.1175/1520-0442(2001)014⟨2105:EPDVAU⟩2.0.CO;2.

Barstad, I., W. W. Grabowski, and P. K. Smolarkiewicz (2007), Characteristics of large-scale orographic precipitation: Evaluation of linear model in idealized problems, *Journal of Hydrology*, *340*(1-2), 78 – 90, doi:10.1016/j.jhydrol.2007.04.005.

Barstad, I., A. Sorteberg, F. Flatø, and M. Déqué (2009), Precipitation, temperature and wind in norway: dynamical downscaling of era40, *Climate Dynamics*, *33*, 769–776, doi:10.1007/s00382-008-0476-5, 10.1007/s00382-008-0476-5.

Bierkens, M. F. P., and B. J. J. M. van den Hurk (2007), Groundwater convergence as a possible mechanism for multi-year persistence in rainfall, *Geophys. Res. Lett.*, *34*, L02,402, doi:10.1029/2006GL028396.

Bordi, I., K. Fraedrich, and A. Sutera (2009), Observed drought and wetness trends in europe: an update, *Hydrology and Earth System Sciences*, *13*(8), 1519–1530, doi:10.5194/hess-13-1519-2009.

Borg, I., and P. J. F. Groenen (2005), *Modern Multidimensional Scaling - Theory and Applications*, Springer Series in Statistics, 2 ed., Springer.

Botter, G., S. Zanardo, A. Porporato, I. Rodriguez-Iturbe, , and A. Rinaldo (2008), Ecohydrological model of flow duration curves and annual minima, *Water Resources Research*, *44*, W08,418, doi:10.1029/2008WR006814.

Botter, G., A. Porporato, I. Rodriguez-Iturbe, and A. Rinaldo (2009), Nonlinear storage-discharge relations and catchment streamflow regimes, *Water Resour. Res.*, *45*, W10,427, doi:10.1029/2008WR007658.

Bouwer, L. M., J. E. Vermaat, and J. C. J. H. Aerts (2008), Regional sensitivities of mean and peak river discharge to climate variability in europe, *J. Geophys. Res.*, *113*, D19,103, doi:10.1029/2008JD010301.

Castellarin, A., G. Galeati, L. Brandimarte, A. Montanari, and A. Brath (2004), Regional flow-duration curves: reliability for ungauged basins, *Advances in Water Resources*, *27*(10), 953–965, doi:10.1016/j.advwatres.2004.08.005.

Clark, M. P., A. G. Slater, D. E. Rupp, R. A. Woods, J. A. Vrugt, H. V. Gupta, T. Wagener, and L. E. Hay (2008), Framework for understanding structural errors (fuse): A modular framework to diagnose differences between hydrological models, *Water Resour. Res.*, *44*, W00B02, doi:10.1029/2007WR006735.

Clark, M. P., H. K. McMillan, D. B. G. Collins, D. Kavetski, and R. A. Woods (2011), 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.

Cleveland, R. B., W. S. Cleveland, J. E. Mcrae, and I. Terpenning (1990), Stl: A seasonal-trend decomposition procedure based on loess, *Journal of Official Statistics*, *6*(1), 3–73.

Cleveland, W. S., and S. J. Devlin (1988), Locally weighted regression: An approach to regression analysis by local fitting, *Journal of the American Statistical Association*, *83*(403), 596 – 610.

Dijkstra, E. W. (1959), A note on two problems in connexion with graphs, *Numerische Mathematik*, *1*, 269–271, doi:10.1007/BF01386390, 10.1007/BF01386390.

Dooge, J. C. I. (1992), Sensitivity of runoff to climate change: A hortonian approach, *Bulletin of the American Meteorological Society*, *73*(12), 2013–2024, doi:10.1175/1520-0477(1992)073⟨2013:SORTCC⟩2.0.CO;2.

Efron, B., and R. J. Tibshirani (1993), *An Introduction to the bootstrap*, no. 57 in Monographs on Statistics and Applied Probability, 436 pp., Chapman & Hall/CRC, Boca Raton, Fla.

Gamez, A. J., C. S. Zhou, A. Timmermann, and J. Kurths (2004), Nonlinear dimensionality reduction in climate data, *Nonlinear Processes In Geophysics*, *11*(3), 393–398.

Ghil, M., et al. (2002), Advanced spectral methods for climatic time series, *Rev. Geophys*, *40*(1), 1003, doi:10.1029/2000RG000092.

Gudmundsson, L., L. M. Tallaksen, and K. Stahl (2011), Spatial cross-correlation patterns of european low, mean and high flows, *Hydrological Processes*, *25*(7), 1034–1045, doi:10.1002/hyp.7807.

Gustard, A., A. Bullock, and J. M. Dixon (1992), Low flow estimation in the united kingdom, *Tech. Rep. 108*, Institute of Hydrology.

Haines, A. T., B. L. Finlayson, and T. A. McMahon (1988), A global classification of river regimes, *Applied Geography*, *8*(4), 255 – 272, doi:10.1016/0143-6228(88)90035-5.

Hanson, R. T., M. W. Newhouse, and M. D. Dettinger (2004), A methodology to asess relations between climatic variability and variations in hydrologic time series in the southwestern united states, *Journal of Hydrology*, *287*, 252–269.

Harris, N. M., A. M. Gurnell, D. M. Hannah, and G. E. Petts (2000), Classification of river regimes: a context for hydroecology, *Hydrol. Process.*, *14*(16-17), 2831–2848.

Huang, N. E., Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu (1998), The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis, *Proceedings: Mathematical, Physical and Engineering Sciences*, *454*(1971), 903–995.

Kingston, D. G., M. C. Todd, R. G. Taylor, J. R. Thompson, and N. W. Arnell (2009), Uncertainty in the estimation of potential evapotranspiration under climate change, *Geophys. Res. Lett.*, *36*(20), L20,403.

Klemeš, V. (1974), The hurst phenomenon: A puzzle?, *Water Resour. Res.*, *10*(4), 675–688, doi:10.1029/WR010i004p00675.

Koster, R. D., and M. J. Suarez (1999), A simple framework for examining the interannual variability of land surface moisture fluxes, *Journal of Climate*, *12*(7), 1911–1917, doi:10.1175/1520-0442(1999)012⟨1911:ASFFET⟩2.0.CO;2.

Koutsoyiannis, D. (2002), The hurst phenomenon and fractional gaussian noise made easy, *Hydrological Sciences Journal*, *47*(4), 573 – 595, doi:10.1080/02626660209492961.

Koutsoyiannis, D. (2003), Climate change, the hurst phenomenon, and hydrological statistics, *Hydrological Sciences Journal*, *48*(1), 3 – 24, doi:10.1623/hysj.48.1.3.43481.

Koutsoyiannis, D. (2010), Hess opinions "a random walk on water", *Hydrology and Earth System Sciences*, *14*(3), 585–601, doi:10.5194/hess-14-585-2010.

Krasovskaia, I., L. Gottschalk, E. Leblois, and E. Sauquet (2003), Dynamics of river flow regimes viewed through attractors, *Nordic Hydrology*, *34*, 461 – 476.

Kumar, M., and C. J. Duffy (2009), Detecting hydroclimatic change using spatio-temporal analysis of time series in colorado river basin, *Journal of Hydrology*, *374*(1 – 2), 1 – 15, doi:10.1016/j.jhydrol.2009.03.039.

Lins, H. F. (1997), Regional streamflow regimes and hydroclimatology of the united states, *Water Resources Research*, *33*, 1655–1667, doi:10.1029/97WR00615.

López-Moreno, J. I., and S. M. Vicente-Serrano (2008), Positive and negative phases of the wintertime north atlantic oscillation and drought occurrence over europe: A multitemporal-scale approach, *Journal of Climate*, *21*(6), 1220–1243, doi:10.1175/2007JCLI1739.1.

Mahecha, M. D., A. Martinez, G. Lischeid, and E. Beck (2007), Nonlinear dimensionality reduction: Alternative ordination approaches for extracting and visualizing biodiversity patterns in tropical montane forest vegetation data, *Ecological Informatics*, *2*(2), 138–149, doi:10.1016/j.ecoinf.2007.05.002.

Mahecha, M. D., L. M. Frst, N. Gobron, and H. Lange (2010), Identifying multiple spatiotemporal patterns: A refined view on terrestrial photosynthetic activity, *Pattern Recognition Letters*, *31*(14), 2309 – 2317, doi:10.1016/j.patrec.2010.06.021.

Mandelbrot, B. B., and J. R. Wallis (1968), Noah, joseph, and operational hydrology, *Water Resour. Res.*, *4*(5), 909–918, doi:10.1029/WR004i005p00909.

Mandelbrot, B. B., and J. R. Wallis (1969), Some long-run properties of geophysical records, *Water Resour. Res.*, *5*(2), 321 – 340, doi:10.1029/WR005i002p00321.

McMahon, T. A., R. M. Vogel, M. C. Peel, and G. G. Pegram (2007), Global streamflows - part 1: Characteristics of annual streamflows, *Journal of Hydrol-*

*ogy*, *347*(3-4), 243 – 259, doi:10.1016/j.jhydrol.2007.09.002.

Milly, P. C. D., and K. A. Dunne (2002), Macroscale water fluxes 2. water and energy supply control of their interannual variability, *Water Resour. Res.*, *38*(10), 1206, doi:10.1029/2001WR000760.

Milly, P. C. D., and R. T. Wetherald (2002), Macroscale water fluxes 3. effects of land processes on variability of monthly river discharge, *Water Resour. Res.*, *38*(11), 1235, doi:10.1029/2001WR000761.

Monahan, A. H. (2001), Nonlinear principal component analysis: Tropical indo-pacific sea surface temperature and sea level pressure, *Journal of Climate*, *14*(2), 219–233, doi:10.1175/1520-0442(2001)013⟨0219:NPCATI⟩2.0.CO;2.

Montanari, A., R. Rosso, and M. S. Taqqu (1997), Fractionally differenced arima models applied to hydrologic time series: Identification, estimation, and simulation, *Water Resour. Res.*, *33*(5), 1035–1044, doi:10.1029/97WR00043.

Mudelsee, M. (2007), Long memory of rivers from spatial aggregation, *Water Resour. Res.*, *43*(1), W01,202, doi:10.1029/2006WR005721.

Peres-Neto, P., and D. Jackson (2001), How well do multivariate data sets match? the advantages of a procrustean superimposition approach over the mantel test, *Oecologia*, *129*(2), 169–178, doi:10.1007/s004420100720.

Potter, K. W. (1976), Evidence for nonstationarity as a physical explanation of the hurst phenomenon, *Water Resour. Res.*, *12*(5), 1047–1052, doi:10.1029/WR012i005p01047.

R Development Core Team (2011), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, last accessed: 15 May 2011.

Renner, M., and C. Bernhofer (2011), Long term variability of the annual hydrological regime and sensitivity to temperature phase shifts in saxony/germany, *Hydrology and Earth System Sciences Discussions*, *8*(1), 811–853, doi:10.5194/hessd-8-811-2011.

Sankarasubramanian, A., and R. M. Vogel (2002), Annual hydroclimatology of the united states, *Water Resour. Res.*, *38*(6), 1083, doi:10.1029/2001WR000619.

Sankarasubramanian, A., R. M. Vogel, and J. F. Limbrunner (2000), Climate elasticity of streamflow in the united states, *Water Resour. Res.*, *37*, 1771 – 1781.

Shamsudduha, M., R. E. Chandler, R. G. Taylor, and K. M. Ahmed (2009), Recent trends in groundwater levels in a highly seasonal hydrological system: the ganges-brahmaputra-meghna delta, *Hydrology and Earth System Sciences*, *13*(12), 2373–2385, doi:10.5194/hess-13-2373-2009.

Shorthouse, C., and N. Arnell (1997), *FRIEND97 - 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, pp. 77 – 85, The International Association of Hydrological Sciences (IAHS).

Shorthouse, C., and N. Arnell (1999), The effects of climatic variability on spatial characteristics of european river flows, *Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere*, *24*(1-2), 7 – 13, doi:10.1016/S1464-1909(98)00003-3, european Water Resources and Climate Changes Processes.

Shun, T., and C. J. Duffy (1999), Low-frequency oscillations in precipitation, temperature, and runoff on a west facing mountain front: A hydrogeologic interpretation, *Water Resour. Res.*, *35*(1), 191–201, doi:10.1029/98WR02818.

Spearman, C. (1987), The proof and measurement of association between two things, *The American Journal of Psychology*, *100*(3/4), 441–471.

Stahl, K., H. Hisdal, L. Tallaksen, H. van Lanen, J. Hannaford, and E. Sauquet (2008), Trends in low flows and streamflow droughts across europe, *Tech. rep.*, UNESCO, Paris.

Stahl, K., H. Hisdal, J. Hannaford, L. M. Tallaksen, H. A. J. van Lanen, E. Sauquet, S. Demuth, M. Fendekova, and J. Jódar (2010), Streamflow trends in europe: evidence from a dataset of near-natural catchments, *Hydrology and Earth System Sciences*, *14*(12), 2367–2382, doi:10.5194/hess-14-2367-2010.

Tallaksen, L. M., H. Madsen, and B. Clausen (1997), On the definition and modelling of streamflow drought duration and deficit volume, *Hydrological Sciences Journal-Journal Des Sciences Hydrologiques*, *42*(1), 15–33.

Tenenbaum, J. B., V. de Silva, and J. C. Langford (2000), A global geometric framework for nonlinear dimensionality reduction, *Science*, *290*(5500), 2319 – 2323.

Tootle, G. A., and T. C. Piechota (2006), Relationships between pacific and atlantic ocean sea surface temperatures and u.s. streamflow variability, *Water Resour. Res.*, *42*, W07,411, doi:10.1029/2005WR004184.

Torgerson, W. S. (1952), Multidimensional scaling: I. theory and method, *Psychometrika*, *17*(4), 401–419.

Torrence, C., and G. P. Compo (1998), A practical guide to wavelet analysis, *Bulletin of the American Meteorological Society*, *79*(1), 61–78, doi:10.1175/1520-0477(1998)079⟨0061:APGTWA⟩2.0.CO;2.

Uppala, S. M., et al. (2005), The era-40 re-analysis, *Quarterly Journal of the Royal Meteorological Society*, *131*(612), 2961–3012, doi:10.1256/qj.04.176.

Vogel, R. M., and N. M. Fennessey (1994), Flow-duration. i: New interpretation and confidence intervals, *J. Water Resour. Plann. Manage.*, *120*(4), 458 – 504.

Vogel, R. M., Y. Tsai, and J. F. Limbrunner (1998), The regional persistence and variability of annual streamflow in the united states, *Water Resour. Res.*, *34*(12), 3445 – 3459, doi:10.1029/98WR02523.

Vogt, J., P. Soille, A. de Jager, E. Rimaviciute, and W. Mehl (2007), A pan-european river and catchment database, *Tech. Rep. EUR 22920 EN*, EC-JRC, Luxembourg.

von Storch, H., and F. W. Zwiers (1999), *Statistical analysis in climate research*, 484 pp., Cambridge University Press, Cambridge, UK.

Ward, J. (1967), Hierarchical clustering to optimize an objective function, *J. Am. Stat. Assoc.*, *58*, 236–244.

Weedon, G., S. Gomes, P. Viterbo, H. sterle, J. Adam, N. Bellouin, O. Boucher, and M. Best (2010), The watch forcing data 1958-2001: a meteorological forcing dataset for land surface- and hydrological-models, *WATCH Technical Report 22*, (available at www.eu-watch.org).

Weedon, G. P., et al. (2011), Creation of the watch forcing data and its use to assess global and regional reference crop evaporation over land during the twentieth century., *Journal of Hydrometeorology*, *early view*(0), doi:10.1175/2011JHM1369.1.

Wilcoxon, F. (1945), Individual comparisons by ranking methods, *Biometrics Bulletin*, *1*(6), 80 – 83.

# Article III

## Comparing Large-scale Hydrological Models to Observed Runoff Percentiles in Europe

Lukas Gudmundsson, Lena M. Tallaksen, Kerstin Stahl, Douglas B. Clark, Egon Dumont, Stefan Hagemann, Nathalie Bertrand, Dieter Gerten, Naota Hanasaki, Jens Heinke, Frank Voß and Sujan Koirala

*Prepared for submission to:*
Journal of Hydrometeorology

# Article IV

## Seasonal Evaluation of Nine Large-Scale Hydrological Models Across Europe

Lukas Gudmundsson, Thorsten Wagener, Lena M. Tallaksen and Kolbjørn Engeland

*Submitted to:*
Water Resources Research — 2011

# Article V

## Streamflow data from small basins: a challenging test to high resolution regional climate modeling

Kerstin Stahl, Lena M. Tallaksen, Lukas Gudmundsson
and Jens H. Christensen

*In press:*
Journal of Hydrometeorology — 2011
DOI: 10.1175/2011JHM1356.1