Doctoral Thesis

Hydrological climate-impact modelling in the Rhine catchment down to Cologne

Author(s):
Bosshard, Thomas

Publication Date:
2012

Permanent Link:
https://doi.org/10.3929/ethz-a-007316898

Rights / License:
In Copyright - Non-Commercial Use Permitted

This page was generated automatically upon download from the ETH Zurich Research Collection. For more information please consult the Terms of use.
Hydrological climate-impact modelling in the Rhine catchment down to Cologne

A dissertation submitted to
ETH ZURICH

for the degree of
Doctor of Sciences

presented by
THOMAS BOSSHARD
Dipl. Umwelt-Natw. ETH
citizen of Turbenthal - Switzerland

accepted on the recommendation of
Prof. Dr. C. Schär, examiner
Prof. Dr. R. Weingartner, co-examiner
Dr. S. Kotlarski, co-examiner
Dr. M. Zappa, co-examiner

2012
Summary

The impact of climate change on river streamflow has gained much attention from the public and in the scientific community during the last roughly 30 years. This high awareness of possible impacts is related to the close and manifold interactions between river systems and mankind as the former provide fresh water resources for multiple purposes but also have a damaging potential in cases of extreme high or low flows. Also, rivers have reacted sensitively to climate variability and past climate change, as demonstrated by numerous studies using hydrometeorological and discharge time-series. Thus, the prospect of climate change associated with unprecedented non-stationarities unavoidably leads to the question of what river streamflow will look like in a future climate. Changes in the seasonality of streamflow and runoff extremes might require measures to ensure water resource management related to water supply, navigation, hydropower production and flood protection.

Since the start of hydrological climate-impact studies, impact modelling chains consisting of prescribed greenhouse gas emission scenarios, general circulation models (GCMs) and hydrological models were applied. Considerable effort has been invested into developing methods to bridge the spatio-temporal gap at the interface between GCMs and hydrological models. Modern impact modelling chains use regional climate models (RCMs) to downscale the information of coarse-resolution GCMs to finer resolutions, which are closer to the scales of hydrological models. Due to climate model limitations, the dynamical downscaling approach does not provide a perfect representation of atmospheric variables, and the downscaled data thus need to be statistically post-processed before they can be used to force hydrological models.

In this thesis, several topics of hydrological climate-impact modelling are addressed in the context of the river Rhine upstream of Cologne, considering a time period from the recent past to the end of the current century. The thesis uses output of the ENSEMBLES set of global and regional climate simulations to drive the hydrological model PREVAH. The results are presented in the core of the thesis which consists of three publications.

In the first part, the statistical linkage between climate models and hydrological impact models is investigated. A common technique is the delta change method which scales observed time series according to an estimated future climate change signal and employs the scaled time series to drive impact models. To optimally represent the seasonal cycle, a spectral approach has been developed that allows to represent the annual cycle of the climate change signal of hydrometeorological variables with a daily resolution while, at the same time, fluctuations caused by natural variability are filtered out. Such fluctuations are random in their nature and should not be used to scale observed time series. The spectral approach has been applied to GCM-RCM data of the ENSEMBLES project to estimate annual cycles of the climate change signal of temperature and precipitation at measurement sites in Switzerland. These climate change signals are going to be used
in current climate-impact research projects such as “CCHydro” or “Klimaänderung und Wasserkraft”.

In the second part, the focus is put on the quantification of uncertainties in the hydrological climate-runoff projections arising from different elements of an impact modelling chain. The study region is the Alpine Rhine which is a challenging region for all the impact modelling chain elements. The experimental setup consists of nine GCM-RCMs provided by the ENSEMBLES project, two statistical post-processing methods and two hydrological models. The total uncertainty in the runoff projections is decomposed into contributions of different uncertainty sources using two statistical methods. Both methods identify the GCM-RCMs as the dominant source of uncertainty. However, the results also indicate that interactions between different uncertainty sources play an important role. This might be of major importance for the design of further impact-modelling systems.

In the third part, hydrological climate-impact modelling results for the Rhine river down to Cologne are presented for the two scenario periods 2021-2050 and 2070-2099 with respect to the control period 1979-2008. The study uses GCM-RCM data of the ENSEMBLES project, the improved version of the delta change methodology as developed in the first part of the thesis, and the semi-distributed conceptual hydrological model PREVAH. It is found that at the Cologne gauge, the projections agree on a decrease of runoff in summer and increase in winter for the period 2070-2099 relative to the period 1979-2008. Also, changes in extreme runoff indices are presented and discussed. Generally, it is found that the driving GCM has a strong influence on the runoff projections. Furthermore, the individual contributions of projected temperature and precipitation changes to the changes in runoff are investigated. In the Alpine region, the change in runoff is to a large extent determined by temperature changes. Further downstream, precipitation changes become more important.

It is important to note that projections of the climate impact on river streamflow should be interpreted with caution as they involve assumptions about future greenhouse gas emissions and about the representation of physical processes in the models. Such projections are best estimates based on our current knowledge. As scenario assumptions change, as models are likely to be improved and as more observations become available, one needs to reassess the climate impacts. In the future, climate models will be operated at spatial resolutions that allow limitations in the representation of physical processes to be overcome and the complex Alpine topography in the Rhine catchment to be considered in more detail. Thus, it is necessary to interpret climate-impact research as a continuous process of constant re-evaluation. Also, as more models and more uncertainty sources will be included into climate-impact studies, it is possible that the total uncertainty of climate-impact projections will increase in future studies. The quantification of uncertainties in climate-impact projections is of high value to water resource planners who should account explicitly for uncertainties arising from climate variability and climate change.
Zusammenfassung

Der Einfluss des Klimawandels auf Flusssysteme hat in den letzten 30 Jahren an Aufmerksamkeit in der öffentlichen Wahrnehmung und in der Wissenschaft gewonnen. Das Bewusstsein um die möglichen Auswirkungen gründet in der engen und vielfältigen Wechselwirkung zwischen Flusssystemen und der Anthroposphäre. So nutzt der Mensch die Flüsse für vielfältige Zwecke, ist aber auch gegen Schäden durch Dürreperioden oder Hochwassereignisse nicht gefeit. Untersuchungen von Messzeitreihen belegen, dass Flüsse sensitiv auf die Klimavariabilität und die schon messbare Klimaveränderung reagieren. Es erscheint daher nur logisch, dass im Zusammenhang mit dem prognostizierten Klimawandel die Frage gestellt wird, wie sich die Wasserführung unter zukünftigen Klimabedingungen verändern wird. Änderungen der Saisonaldigkeit oder der Abflussextreme machen möglicherweise Massnahmen nötig, um die Wasserversorgung, Schifffahrt und Energieproduktion weiterhin garantieren sowie vor Hochwasser schützen zu können.


Die vorliegende Arbeit präsentiert hydrologische Klimaprojektionen. Bei der Interpretation solcher Projektionen gilt es zu beachten, dass ihnen immer Annahmen über die Entwicklung der Treibhausgasemissionen und die Abbildung physikalischer Prozesse in den Modellen zu Grunde liegen. Solche Projektionen sind die zur Zeit bestmögliche Abschätzung der Klimazukunft basierend auf unserem heutigen Kenntnisstand. Sollten sich aber die
CONTENTS

4.6 Acknowledgments .......................................................... 80

5 Synthesis and outlook ....................................................... 83

A Supplementary material: Hydrological climate-impact projections for
the Rhine river .............................................................. 89
  A.1 Introduction ................................................................. 89
  A.2 Spatial pattern of changes of atmospheric variables ................. 90
  A.3 Changes in the annual cycle ............................................ 106

B Regional parameter allocation and predictive uncertainty estimation
of a RR model ............................................................... 109
  B.1 Introduction ................................................................. 110
  B.2 Methods ...................................................................... 112
  B.3 Results ......................................................................... 120
  B.4 Discussion ................................................................. 125
  B.5 Conclusion ................................................................. 127
  B.6 Acknowledgments ......................................................... 128

Bibliography ................................................................. 129

Special thanks .............................................................. 142
1 Introduction

This PhD thesis studies the impact of climate change on the hydrology of the Rhine river down to Cologne. It presents an improved method to link climate and hydrological models, discusses the relevance of several uncertainty sources for projections of changes in the runoff, and presents results of projected hydrological changes in the Rhine river induced by climate change. It is though not only now within the context of climate change that the Rhine river moves into the center of scientific interest. Rather, the Rhine river has always been closely connected to the culture, politics, economy and ecology of the riparian countries as they both benefit from its continuous streamflow and suffer from flood catastrophes (Blackbourn, 2006). Thus, given the importance of the Rhine river to society, the hydrology of the river Rhine has been subject to numerous studies.

In this introduction, the focus will be on a short description of the Rhine river and its importance nowadays, a brief history of hydrological modelling in the Rhine basin followed by an overview of the development of hydrological climate-impact research. At the end of the introduction, an overview of the different chapters of the PhD thesis is given.

1.1 The Rhine river

The Rhine river originates at the Lay da Tuma in the Swiss Alps, flows northwards and drains into the North Sea (see Fig. 1.1). It is the largest river in Western Europe with a length of about 1’230 km and a total catchment area of about 185’000 km$^2$. Based on its varying hydrological characteristics, the Rhine river is subdivided into different reaches. The Alpine and High Rhine are characterised by mountainous topography and a nival runoff regime. Both regions have a much higher specific discharge than the reaches further downstream and thus contribute a large portion to the total runoff along the Rhine river (Weingartner et al., 2007). The next reach, the Upper Rhine, includes the tributaries Neckar and Main which superimpose the snowmelt dominated regime by a rainfall dominated one. In the Middle Rhine, the Moselle river joins the Rhine. The Moselle does not alter the runoff regime substantially but often determines the flood peak during winter floods (Belz et al., 2007). After Bonn, the Lower Rhine reaches down to the German-Dutch border where the Rhine delta begins (not shown on the map). At Cologne, the largest city along the river, the mean runoff amounts to about 2’240 m$^3$/s.

The Rhine river provides water supply for domestic, industrial and agricultural purposes in a settlement area for about 58 million people (Linde et al., 2010). Its water is also used for hydropower production and navigation. It is Europe’s river with the highest traffic density. In 2007, for example, about 100-500 ships per day have transported 170 million
1 Introduction

Figure 1.1: Map of the Rhine catchment down to Cologne. The black dots denote the subdivision of the Rhine river into reaches with different characteristics. By courtesy of Nina Köplin.

freight tons on the lower Rhine (Heinz et al., 2009). The Rhine river is navigable all-year round from Basel down to the North Sea except for periods of pronounced low flows and high flows that can lead to reduced load capacities or even might halt the navigation. Apart from its benefits, the Rhine river also causes frequent large-scale floods. Uhlemann et al. (2010) counted 80 large-scale floods in Germany during the period from 1952 to 2002. They showed that most of the large-scale floods occurred in the hydrological winter and that the proportion of winter floods increased from 58% in the period 1952-1977 to 70.5% in 1978-2002. They also found that flood events have a tendency to cluster in time. Similarly, Schmocker-Fackel and Naef (2010) found periods with frequent floods alternating with quieter periods in an analysis of flood records of the last 150 years in Switzerland. Concerning the High Rhine catchment, the most severe recorded flood happened in August 2005. About one third of the gauging stations in the area recorded the highest discharges since the start of the measurements. Costs due to flood damages were estimated to be about 3 billion Swiss Francs (Bezzola and Hegg, 2007). However, the spatial extent of the 2005 flood was relatively small compared to the two recent large-scale floods in December 1993 and January 1995. Both events were caused by a superposition of flood waves
from the different tributaries. At Cologne, the gauges recorded in 1993 the second largest waterlevel in the 20th century and in 1995 an equal high waterlevel as during the largest flood in 1926 (Disse and Engel, 2001). The damage costs were estimated to be about 2.1 billion DM for the 1993 and about 3.3 billion DM for the 1995 flood event (Engel, 1999). These two floods induced the international Action Plan on Floods (APF) adopted by the International Commission for the Protection of the Rhine (ICPR) (ICPR, 2005). The goals of the APF are to reduce damage risks, reduce flood stages, increase awareness of flooding and improve the flood forecasting system. By 2005, a total flood retention volume of about 213 million m³ was available along the Rhine (ICPR, 2005). It became evident that an efficient operation of the retention volume, if it should have a decreasing effect on flood peak waterlevel at all, can only be guaranteed if the flood forecasting system’s reliability is increased for longer lead times.

1.2 Hydrological modelling in the Rhine basin

Early attempts to forecast the runoff based on statistical models were made in the 1950s (Lugiez et al., 1969). A cooperation of hydropower companies, shipping-firms and the Dutch water management authorities were the main drivers behind the first forecast modelling studies. They were interested in short- and long-term runoff forecasts. The snowmelt from the Alpine area with peak discharges in summer was noticed to have a balancing effect on the summer low flow conditions further downstream which has clear advantages for both all-year round hydropower production and navigation. Thus, the estimation of the snow storage in the Alpine area was used to estimate summer water levels further downstream. Later on, flood forecasting was added to the list of goals of runoff forecasts (Lang et al., 1987; Disse and Engel, 2001). In 1988, the International Commission for the Hydrology of the Rhine Basin (CHR) described the existing hydrological forecast system in the Rhine basin. The system consisted of individual models for different parts of the Rhine basin. The model types ranged from regression based models to models using unit hydrographs and routing schemes. The forecast lead times ranged from 6 hours to 4 days (Grebner et al., 1988). Nowadays, a variety of hydrological models that run on a daily or sub-daily time step are in use by different agencies of the riparian countries and coupling with numerical weather forecast models is done on an operational or semi-operational basis (see e.g. Cloke and Pappenberger, 2009, and references therein).
1 Introduction

1.3 Hydrological climate-impact research

The explanations above illustrate that changes in river streamflow along the river Rhine possibly induce strong impacts on society. Thus, it is important to investigate how projected climate changes might effect the hydrology of the Rhine. Evidence, physical considerations and climate modelling studies suggest an acceleration of the hydrological cycle on the global scale due to increased greenhouse gas concentrations in the atmosphere (Allen and Ingram, 2002; Wentz et al., 2007; Allan and Soden, 2008; Wild et al., 2008). Going from the global to the catchment scale, the overall acceleration of the hydrological cycle is superimposed by local catchment-specific effects such as regional climate changes of hydrometeorological variables and non-linear sensitivity of hydrological systems to changes in the forcing meteorology. Important processes involved are changes in seasonality of precipitation (Räisänen et al., 2004), changes of snow accumulation and melt dynamics caused by temperature changes (Barnett et al., 2005), changes in precipitation intensity and frequency distributions (Frei et al., 2006) and changes in evapotranspiration (Zappa and Kan, 2007). Also changes in land use might have an impact on the reaction of the hydrological system to climate change (Bronstert et al., 2002). In the Rhine river basin, changes of the runoff regimes have already been detected (Belz et al., 2007; Hänggi and Weingartner, 2011).

The basis for international hydrological climate-impact research was founded at the first World Climate Conference in 1979 (Schaake and Kaczmarek, 1979). Right from the start, it was thought that the combination of General Circulation Models (GCMs) and hydrological models is a promising strategy to assess the climate impact on the hydrological cycle. Seven years later, Gleick (1986) reviewed the research in the field of hydrological climate-impact research and came up with six criteria for evaluating the applicability of hydrologic models for climatic impact assessment:

- The inherent accuracy of the hydrological model
- The degree to which [hydrological] model accuracy depends upon the existing climatic conditions for which the model was initially developed and calibrated
- The availability of input data, including comparative historical climatic data
- The accuracy of the input data
- Model flexibility, ease of use, and adaptability to diverse climatic and hydrologic conditions
- Compatibility with existing general circulation models

The first five points in the list have general validity to hydrological modelling. The last
point addresses the coupling with GCMs which has proved to be a particularly challenging one to impact scientists. The different spatial and temporal scales of GCMs and hydrological models have been identified as the major problem (e.g. Xu, 1999; Maraun et al., 2010). While GCMs are designed to simulate the earth’s climate system on continental or global spatial and climatological temporal scales, hydrological models simulate individual catchments that range from the hillslope to the macroscale and require input data with hourly to monthly resolution.

In the first hydrological climate-impact studies, no GCM data were used at all. Instead, observed data were scaled according to hypothetical changes of precipitation and temperature and used to force a hydrological model (Stockton and Boggess, 1979; Némec and Schaake, 1982). This approach enabled the assessment of the sensitivity of a hydrological system to hypothetical changes in meteorological input data. Soon, attempts were made to derive the precipitation and temperature changes from GCM simulations (Cohen, 1986; Bultot et al., 1988; MacCabe and Ayers, 1989; Lettenmaier and Gan, 1990; Arnell, 1992). First, annual changes and later monthly changes were applied. These first studies proved that changes in meteorological input data have a potentially large impact on the runoff. In particular, changes in seasonality due to shifted snow accumulation and melt seasons were detected (Barnett et al., 2005). For the whole Rhine basin, the first climate-impact study on runoff dates back to 1995 (Kwadijk and Rotmans, 1995). See chapter 4 for a more detailed literature overview on climate-impact studies about runoff in the Rhine basin.

The scaling of observed data according to a climate change signal does only account for changes in the mean and does not alter the variability in the meteorological forcing data. It is well-known that changes in variability have an impact on hydrological systems with regard to extreme events such as droughts and floods (Shabalova et al., 2003; Lenderink et al., 2007; Fowler et al., 2007). The hydrological modelling tools to study such events are conceptual lumped/semi-distributed or physically based distributed models that use station site specific input data of at least temperature and precipitation with daily or sub-daily temporal resolution (Xu, 1999). Statistical downscaling models were developed to transfer GCM data to the spatial and temporal scales needed by the hydrological model (Wilby and Wigley, 1997; Fowler et al., 2007; Maraun et al., 2010). The downscaled data inherit the variability from the GCM and therefore enable the assessment of changes in the variability and its effects on hydrological systems. It has to be noted though, that advanced statistical downscaling models were developed and tested in numerous studies, but few of them have been used in hydrological impact studies (Maraun et al., 2010).

The statistical downscaling models are all based on the assumption that the derived statistical relationship between the large and local scales in a control period is transferable to future climates. An alternative to base the spatial disaggregation of GCM data on physical principles is to use dynamical downscaling. In dynamical downscaling, regional climate models (RCMs) are nested into the GCM and simulate in a limited area the climate system with a high spatial resolution. In principle, these approaches may represent
potential changes in the statistical relationships between large-scale and regional-scale climate variables, as well as potential changes in variability (Schär et al., 2004; Vidale et al., 2007).

In Europe, the PRUDENCE project (Christensen and Christensen, 2007) was one of the first internationally coordinated projects to provide a suite of dynamically downscaled GCM data. It employed 3 GCMs and 11 RCMs to provide regional climate scenarios for two 30-year time slices, one being the control period 1961-1990 and one the scenario period 2071-2100. The RCMs had a horizontal resolution of about 50 km. It was the first time that uncertainty arising from combining different GCMs and RCMs could be investigated in a systematic manner. One major lesson learned through PRUDENCE was that the driving GCM strongly influences the results of the regional models. The GCM provides the lateral boundaries for the RCM and, roughly speaking, broadly determines the range of atmospheric states that the RCM can take. For instance, if a GCM e.g. imposes a circulation pattern with a south-westerly flow, the RCM cannot simulate a flow that strongly deviates from the lateral flow pattern.

Hence, in the successive ENSEMBLES project (van der Linden and Mitchell, 2009), special attention was given to combine various GCMs with different RCMs. Up to now, the ENSEMBLES project provided a total of 27 regional climate scenarios for the European domain by combining 7 GCMs with RCMs from 15 research institutions. The climate scenarios either assume the IPCC SRES A1B or the A2 emission scenario. All of the ENSEMBLES simulations provide transient data for at least the period 1961-2050. 20 GCM-RCMs were integrated until 2099. Horizontal resolutions of 25 km and 50 km were applied. Compared to PRUDENCE, the ENSEMBLES project improved on state-of-the-art climate model formulations, the GCM-RCM combination scheme, the horizontal resolution and the use of a transient rather than a time-slice simulation strategy. The database of regional climate scenarios is the most extensive one currently available and allows for an unprecedented uncertainty analysis of regional climate scenarios.

Although current RCMs have a horizontal resolution that is similar to the density of stations which are usually used to drive hydrological models in well equipped river basins, the coupling of GCM-RCMs with hydrological models can still not be done directly. Hydrological models are calibrated on observed data. GCM-RCM data, when compared to observations, show substantial biases in virtually all statistical moments of hydrometeorological variables. If used directly, such biased GCM-RCM data would lead to significant biases in the output of hydrological models. Thus, some sort of statistical post-processing to correct for GCM-RCM biases is still necessary in order to fulfill the last point in Gleick’s list.

A variety of post-processing methods already exists (Fowler et al., 2007; Maraun et al., 2010). Among other characteristics, they vary in the degree they allow for changes in the variability and correct for GCM-RCM biases of different statistical moments. The simplest approach is the delta change method. It has been widely used in hydrological
climate-impact research () and is also the main statistical post-processing method in this thesis. Within the delta change method, observed time series of hydrometeorological variables are scaled according to the climate change signal extracted from GCM-RCM data. The method does not allow for changes in temporal variability or spatial covariance and is therefore arguably not necessarily suited for the analysis of changes in extreme events. However, as e.g. Lenderink et al. (2007) noted, methods that inherit changes in variability from GCM-RCMs are also prone to inherit biases in variability. In Alpine terrain, the complex topography and the pronounced spatial variability of climatic parameters poses a special challenge for climate models and can lead to substantial climate model biases in various statistical moments (see Frei et al., 2006, for an analysis of GCM-RCM precipitation). The key importance of spatial variations is particularly obvious for major river systems such as the Rhine, where floods often derive from the superposition of flood waves from different tributaries. As the delta change approach adopts the covariance from observed data, it neglects potential changes in spatial covariance, but it also ensures a high level of consistency under current climatic conditions. Therefore, conservative methods such as the delta change method provide robust climate scenarios, but some care is needed regarding the interpretation.

1.4 Outline of the thesis

The core of the thesis consists of three studies that investigate a few aspects of hydrological climate-impact research in more detail and apply the gained knowledge to modelling the climate impact in the Rhine basin down to Cologne (see Fig. 1.2).

Chapter 2 summarises the first study that deals with the representation of the annual cycle in the climate change signal. Early impact studies already pointed out the importance of changes in seasonality (i.e. in the mean annual cycle) of climatic parameters. In the first part, a focus is therefore put on developing a method that produces a detailed and credible representation of the annual cycle of the climate change signal. Subsequently, the improved methodology to extract the annual cycle of the climate change signal is used within the delta change approach to produce new climate scenarios of temperature and precipitation for impact studies within Switzerland that are based on GCM-RCM data of the ENSEMBLES project.

The second study is presented in chapter 3 and investigates the importance of different uncertainty sources in climate-impact modelling chains. Given the chaotic nature of the climate system, the imperfect knowledge of the governing processes and their simplified representation in climate models, all climate scenarios are subject to uncertainties. These uncertainties propagate through and may be amplified by the subsequent impact modelling chain. In principle, every modelling chain element contributes to the total uncertainty.
In this study, a multi-propagation experiment is conducted in the Alpine Rhine region - a challenging region for every impact modelling chain element. Two statistical methods are used to decompose the total uncertainty into contributions of different uncertainty sources. The knowledge about the importance of different uncertainty sources might be relevant for the design of future climate-impact modelling studies.

Chapter 4 contains the third study and presents hydrological climate-impact projections for the Rhine basin down to Cologne. 10 GCM-RCM chains of the ENSEMBLES project are linked via the improved delta change method developed in the first study to the hydrological model PREVAH. At Cologne, all the major hydrologically distinct reaches of the Rhine as explained in section 1.1 are captured within the upstream basin. Changes in
water balance quantities and the seasonality of runoff are investigated for all the sub-basins and the different reaches. Also, projected changes of extreme runoff indices are discussed. Finally, the importance of temperature and precipitation changes to the estimated changes in runoff is assessed.

Chapters 2-4 have been prepared as papers for the peer-reviewed scientific literature. In the same order as the chapters, the papers are


The core chapters of the thesis are followed by the chapter “Synthesis and outlook”. In this chapter, the relevance of the findings and some open scientific questions are discussed in the context of current hydrological climate-impact research. Also, an outlook on the future of transdisciplinary climate-impact research is presented.

In the appendix, chapter A shows additional analysis of results of the third paper while chapter B contains a paper about the probabilistic calibration of the model PREVAH in the Three Gorges Area (PR China). In general, probabilistic calibration of a hydrological model might be important for a complete quantification of uncertainties of hydrological climate-impact projections.
2 Spectral representation of the annual cycle in the climate change signal

Thomas Bosshard\textsuperscript{1}, Sven Kotlarski\textsuperscript{1}, Tracy Ewen\textsuperscript{1,*}, Christoph Schär\textsuperscript{1}

\textsuperscript{1}Institute for Atmospheric and Climate Science, ETH Zurich, Switzerland  
\textsuperscript{*}now at: Department of Geography, University of Zurich, Switzerland

(Published in Hydrology and Earth System Sciences, 15, 2777–2788, 2011)

Abstract

The annual cycle of temperature and precipitation changes as projected by climate models is of fundamental interest in climate impact studies. Its estimation, however, is impaired by natural variability. Using a simple form of the delta change method, we show that on regional scales relevant for hydrological impact models, the projected changes in the annual cycle are prone to sampling artefacts. For precipitation at station locations, these artefacts may have amplitudes that are comparable to the climate change signal itself. Therefore, the annual cycle of the climate change signal should be filtered when generating climate change scenarios. We test a spectral smoothing method to remove the artificial fluctuations. Comparison against moving monthly averages shows that sampling artefacts in the climate change signal can successfully be removed by spectral smoothing. The method is tested at Swiss climate stations and applied to regional climate model output of the ENSEMBLES project. The spectral method performs well, except in cases with a strong annual cycle and large relative precipitation changes.

2.1 Introduction

Impacts of climate change on the hydrological cycle are both of high scientific interest as well as of high relevance for society as a whole. The former is due to the intimate coupling of the hydrological cycle and the climate system (Allen and Ingram, 2002; Wentz et al., 2007; Wild et al., 2008) while the latter is based on the manifold interactions between the anthroposphere and the hydrosphere (Kundzewicz et al., 2007). Hydrological impact studies focussing on runoff often use statistically post-processed global climate model (GCM)
or regional climate model (RCM) data to drive a hydrological model (Hay et al., 2000; Leung et al., 2004; Wood et al., 2004; Kay et al., 2006; Buytaert et al., 2010). For this purpose, various statistical post-processing methods have been developed (see e.g. Fowler et al., 2007 or Maraun et al., 2010 for comprehensive reviews). All these methods are based on statistical relationships that bridge the spatial and temporal gaps between observations and modelled data, and attempt to correct for climate model biases. Most of the available methods focus on the hydrometeorological variables temperature and precipitation (abbreviated as $T$ and $P$ respectively in the remainder of this article) and usually include some representation of the annual cycle.

Natural variability, both on interannual as well as intraannual time scales, impairs parameter estimates of the statistical post-processing methods. The range of the natural variability can be assessed using e.g. resampling techniques. Prudhomme and Davies (2009b) and Wood et al. (2004), for example, resampled observed time series to estimate the range of natural variability of the climate change signal. Cross-validation has also been used to test the robustness of the parameter estimates to interannual variability (Terink et al., 2010; Schmidli et al., 2007; Widmann et al., 2003). However, only a few studies focussing on hydrological impacts have looked in detail at the intraannual variability of the parameters. Smoothing by averaging over seasons (see e.g. Schmidli et al., 2007) or months (see e.g. Middelkoop et al., 2001; Kleinn et al., 2005) is a common practise. An appropriate representation of the seasonal cycle, however, is not straightforward. On the one hand, the optimal choice of the averaging period is dependent on the magnitude of the natural variability, the spatial averaging, and the length of the data records. The stronger the natural variability, the smaller the spatial averaging area and the shorter the data record is, the wider the averaging window has to be chosen in order to reduce the effects of natural variability on the parameter estimates. On the other hand, hydrological impact modellers are interested in an accurate representation of the annual cycle and therefore prefer as narrow averaging windows as feasible. The optimal solution is thus not trivial to find and depends on the region and application under consideration. Despite its importance, the discussion of how to optimally represent the annual cycle in climate change scenarios is often neglected in recent impact modelling. In many cases, the averaging window width is mentioned without specific justification (e.g. Cameron et al., 2000; Jasper et al., 2004; Graham et al., 2007).

This paper elaborates on the representation of the annual cycle in the climate change signal within the delta change post-processing methodology. We chose the delta change method because of its simplicity, but the results appear relevant for more sophisticated methods as well. The delta change method has been used for hydrological impact studies ever since GCM data became available, and it is still used nowadays (Gleick, 1986; Hay et al., 2000; Prudhomme et al., 2002; Lenderink et al., 2007). More sophisticated combinations of the delta change approach and weather generators have also been developed (Kilsby et al., 2007). It is noteworthy that Gleick (1986) already stressed the importance of representing
Table 2.1: List of the employed climate model chains from the ENSEMBLES project.

<table>
<thead>
<tr>
<th>Institution</th>
<th>GCM</th>
<th>RCM</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMHI</td>
<td>ECHAM5</td>
<td>RCA</td>
</tr>
<tr>
<td>MPI</td>
<td>ECHAM5</td>
<td>REMO</td>
</tr>
<tr>
<td>KNMI</td>
<td>ECHAM5</td>
<td>RACMO</td>
</tr>
<tr>
<td>ICTP</td>
<td>ECHAM5</td>
<td>REGCM</td>
</tr>
<tr>
<td>DMI</td>
<td>ECHAM5</td>
<td>HIRHAM</td>
</tr>
<tr>
<td>ETHZ</td>
<td>HadCM3Q0</td>
<td>CLM</td>
</tr>
<tr>
<td>HC</td>
<td>HadCM3Q0</td>
<td>HadRM3Q0</td>
</tr>
<tr>
<td>SMHI</td>
<td>HadCM3Q3</td>
<td>RCA</td>
</tr>
<tr>
<td>CNRM</td>
<td>ARPEGE</td>
<td>ALADIN</td>
</tr>
<tr>
<td>SMHI</td>
<td>BCM</td>
<td>RCA</td>
</tr>
</tbody>
</table>

the climate change throughout the annual cycle since seasonal changes tend to cancel each other out in the annual average.

Here, we test the influence of sampling variability on the annual cycle of the climate change signal by using moving averages (MA) of different window widths. As an alternative to the MA, we present a spectral approach to estimate the climate change signal. Our analysis is carried out at observational station sites in Switzerland. The spectral estimation produces smoother annual cycles of the climate change signals than MAs.

The paper is structured as follows: in Sect. 2.2, we present the data used for the study. Section 2.3 introduces the delta change method and the estimation methods for the annual cycle. In Sect. 2.4, we study the effects of sampling variability on the estimation of the annual cycle using a stochastic rainfall generator. Section 2.5 presents the estimation of the annual cycle of the climate change signal using a spectral smoothing method and a comparison to estimates using MAs of 31 days window width. The methodology is then systematically explored at Swiss station sites. Section 3.5 summarises the findings and discusses their relevance for climate impact studies.

### 2.2 Data

We used daily near-surface $T$ and $P$ data from 10 GCM-RCM model chains provided by the ENSEMBLES project (van der Linden and Mitchell, 2009) to estimate the annual cycle of the climate change signal (see Table 2.1). All model chains use the A1B emission scenario, cover the period 1951–2099 and have a horizontal resolution of about 25 km. From the whole set of model chains available through the ENSEMBLES project, we excluded the HadCM3Q16 driven model chains as well as DMI-ARPEGE-ALADIN in our analysis, due
The geographical focus of our study is Switzerland. Throughout this paper, the estimation of the climate change signal is based on GCM-RCM data interpolated to station locations of MeteoSwiss (see Fig. 2.1). All the stations provide T and P data with at least daily resolution for the period 1980–2009. We used the four nearest gridpoints and the inverse distance weighting interpolation algorithm to spatially interpolate the GCM-RCM data to station locations. It should also be noted that any height correction is redundant since constant correction terms cancel each other in the delta change approach. Also, for simplicity, we neglect leap days in the data, unless stated otherwise.

For the stochastic rainfall generator experiments, two subsets of precipitation stations are used to estimate the rainfall generator parameters. These subsets are indicated by blue and red dots in Fig. 2.1 (see also Sect. 2.4). In addition, we used long-term data series from 26 stations with records going back to 1900 from the climate monitoring network of MeteoSwiss to constrain the harmonic smoothing model. Stars mark these stations in Fig. 2.1.
2.3 Methodology

2.3.1 The delta change method

In the delta change method, observational records are scaled according to a climate change signal. The climate change signal is derived from climate model data as the change between a scenario period (SCE) and a control period (CTL). As a result of the scaling, the spatio-temporal patterns as well as the correlations between the variables closely follow the observed records. Thus, the delta change method is considered a robust method to generate climate impact scenarios (Graham et al., 2007).

In this study, we applied the delta change method at station sites for the SCE periods 2021–2050 and 2070–2099, both relative to the CTL period 1980–2009. At each station \(i\), for each GCM-RCM \(j\) and for each day \(d\) in the year, we estimate the mean annual cycle of the variable of interest and denote it with \(X_{i,j}^{\text{CTL}}(d)\) for the control and \(X_{i,j}^{\text{SCE}}(d)\) for the scenario period where \(X\) stands for either \(T\) or \(P\). The delta change method then derives an additive (\(\Delta X_{i,j}^{\text{add}}(d)\)) and a multiplicative (\(\Delta X_{i,j}^{\text{mult}}(d)\)) climate change signal for \(T\) and \(P\), respectively, according to

\[
\Delta T_{i,j}^{\text{add}}(d) = T_{i,j}^{\text{SCE}}(d) - T_{i,j}^{\text{CTL}}(d) \quad (2.1)
\]

\[
\Delta P_{i,j}^{\text{mult}}(d) = \frac{P_{i,j}^{\text{SCE}}(d)}{P_{i,j}^{\text{CTL}}(d)} \quad (2.2)
\]

Let \(X_{i,\text{obs}}^{\text{CTL}}(y, d)\) denote the continuous observational time series at station sites in the CTL period 1980–2009. Here, \(y\) represents the years in the CTL period. In the delta change method, all observational time steps in the CTL period belonging to the same day \(d\) in the year are scaled with the corresponding climate change value. Again, one commonly uses an additive or multiplicative scaling for \(T\) and \(P\), respectively:

\[
T_{i,j}^{\text{SCE,add}}(y, d) = T_{i,j}^{\text{CTL}}(y, d) + \Delta T_{i,j}^{\text{add}}(d) \quad (2.3)
\]

\[
P_{i,j}^{\text{SCE,mult}}(y, d) = P_{i,j}^{\text{CTL}}(y, d) \times \Delta P_{i,j}^{\text{mult}}(d) \quad (2.4)
\]

Equations (2.1–2.4) reveal that a key issue in the delta change approach is the estimation of the climatological annual cycle in a predefined period. In fact, the delta change approach states nothing but how the climatological annual cycle changes in the transition of the climate state from CTL to SCE periods according to climate model simulations. If one fails to estimate a robust annual cycle in either the CTL or the SCE period, one will fail to project an accurate change of the annual cycle.
2.3.2 Estimation of the climatological annual cycle

It is not possible to derive the true climatological annual cycle of any variable but only an estimate thereof, due to the natural variability and the limited duration of observed or simulated data records. The uncertainty of the estimate might be represented by a stochastic component. Ideally, the estimated climatological annual cycle should be robust, not depend on the stochastic components in the time series, and preserve the amplitude of the annual cycle. Often, the optimisation of these criteria is a trade-off, and it is not trivial to choose an optimal method to estimate the climatological annual cycle.

In this study, we used MAs and a spectral approach as an alternative to the MA for the estimation of the climatological annual cycle of $T$ and $P$. In the MA approach, the terms $\bar{X}(d)_{\text{CTL}}$ and $\bar{X}(d)_{\text{SCE}}$ in Eqs. (2.1) and (2.2) become

$$\bar{X}_{i,j}(d) = \frac{1}{y_e - y_s + 1} \sum_{y=y_s}^{y_e} \left[ \frac{1}{2n + 1} \sum_{k=d-n}^{d+n} X_{i,j}(y,k) \right]$$

where $y_s$ and $y_e$ denote the start and end year of the chosen period and $n$ stands for the number of days before and after the day $d$ in each year $y$. We used MAs with window widths of 15 ($n = 7$), 31 ($n = 15$), 61 ($n = 30$) and 91 ($n = 45$) days. The larger the $n$, the smaller the effect of the natural variability on the estimate of the climatological annual cycle. However, the amplitude of the annual cycle is more strongly damped for larger $n$.

In the spectral approach, we investigated a spectral reconstruction of the climatological annual cycle by a superposition of harmonics with the base period $P_0 = 365$ days as

$$\bar{X}_{i,j}(d) = a_0^i + \sum_{k=1}^{H} \left[ a_k^i \cos(\omega_k d) + b_k^i \sin(\omega_k d) \right]$$

$$\omega_k = \frac{2k\pi}{P_0}.$$  

The superscript $k$ indicates the order of the harmonic components and $H$ is the maximum order retained. The coefficients $a_k^i$ and $b_k^i$ are estimated using the discrete Fourier transform (see e.g. von Storch and Zwiers, 1999) from the daily time series of the GCM-RCM $j$ at station site $i$ in the CTL and SCE period.

For GCM-RCMs using the Gregorian calendar, the base period $P_0$ is set to 365.25 days to account for leap years (Narapusetty et al., 2009). The HadCM3Q0 and HadCM3Q3 driven RCMs have a 360 days calendar. For these GCM-RCMs, we set $P_0$ to 360 days. Having estimated the coefficients of the harmonic smoothing model at each station site and for each GCM-RCM, we scale the different lengths of the annual cycle to fit 365 days by choosing $P_0 = 365$ days in the reconstruction of $\bar{X}_{i,j}(d)$ as in Eq. (2.6).

In the spectral framework of harmonics, the choice of the maximum order $H$ is the only free parameter. The larger $H$ is, the more the details of the annual cycle can be resolved,
2.4 Analysis of synthetically generated precipitation time series

Let’s assume we could sample two 30 yr long precipitation time series from a stationary climate and derive the annual cycle of the precipitation change between the two time series. Stationary here means that the mean climate state is the same in both samples, but the two realisations are modulated by natural variability. Since we know that the climate is stationary by assumption, the asymptotic solution of the precipitation change (expressed as a ratio) should equal one representing no precipitation change. Any deviation from one, e.g. the occurrence of an annual cycle, is solely caused by sampling variability, and does not contain any climate signal.

Here, we investigate the sampling artefacts in the annual cycle of the precipitation changes in an idealised stochastic setup using a rainfall generator with stationary parameters. The setup consists of four experiments which are representative for precipitation series of
individual station sites (STATION) and regions (REGION) both for the northern (CHN) and southern (CHS) parts of Switzerland, in order to study two distinct climates at station and regional scale. Following Wilks and Wilby (1999), we employed a first order Markov chain rainfall generator with the precipitation intensity being modelled by a two-parameter gamma distribution. Let $p_{dw}$ and $p_{ww}$ be the transition probabilities from a dry to wet and a wet to wet day, respectively. Given realisations of the uniform random number $r_1$ on the unit interval, the precipitation occurrence $Y(t)$ is modelled as

$$
Y(t) = \begin{cases} 
1 & \text{if } Y(t-1) = 0 \text{ and } r_1 \leq p_{dw} \\
1 & \text{if } Y(t-1) = 1 \text{ and } r_1 \leq p_{ww} \\
0 & \text{otherwise}
\end{cases}
$$

(2.8)

where 1 stands for a wet and 0 for a dry day. On wet days, the precipitation intensity $I(t)$ is sampled from a gamma probability density function according to

$$
f(I(t)) = \frac{(I(t)/\beta)^{\alpha-1}e^{-I(t)/\beta}}{\beta \Gamma(\alpha)}, \quad I(t), \alpha, \beta > 0
$$

(2.9)

where $\alpha$ and $\beta$ are parameters of the gamma distribution and $\Gamma$ is the gamma function. In this case, the synthetic precipitation time series $P_{synth}$ is derived as

$$
P_{synth}(t) = Y(t) \times I(t)
$$

(2.10)

For the single-station experiments CHN\text{STATION} and CHS\text{STATION}, we derived the parameters $p_{dw}$, $p_{ww}$, $\alpha$ and $\beta$ from the observed daily precipitation records in the period 1980–2009 at the stations Bern (BER) and Lugano (LUG) respectively. The parameters were estimated for each season separately. At the transition from one season to the other, the parameter set is changed but the wet/dry state from the last day of the previous season is taken for the continuation of the Markov chain.

For the experiments representing regional precipitation, we first selected all stations in a radius $r$ around BER ($r = 60$ km) and LUG ($r = 40$ km) that have less than 10% missing values in the period 1980-2009 and calculated the mean daily precipitation time series therefrom. Remaining missing data were ignored in the averaging. The selected stations are indicated by red and blue dots in Fig. 2.1. We are aware that this averaging does not
follow any spatial interpolation standards. The procedure suffices, though, to analyse the effect of spatial averaging on the fluctuations of the climate change signal. Table 2.2 lists the seasonal parameter settings for each of the four experiments.

For each experiment, we generated 100 realisations of a daily precipitation time series with a length of 30 yr. Subsequently, we randomly chose 500 pairs out of the 100 realisations and calculated the multiplicative precipitation change signal by MAs with window widths of 15, 31, 61 and 91 days. Also, note that the commonly used mean monthly climate change signals are a subsample of the 31d MA. The same similarity exists between seasonal climate change signals and the 91d MA. Thus, the effects of sampling variability presented below are transferable to these common averaging intervals as well.

The results are shown in Fig. 2.2. Since both time series of each pair have been generated with the same rainfall generator settings, the asymptotic solution is one, indicating no change. The dots in Fig. 2.2 display one randomly chosen realisation of the synthetically generated climate change signal ∆$P_{\text{synth}}$ using different MA window widths. The 15d MA estimate shows large fluctuations in every experiment. The wider the MA window becomes, the smaller the fluctuations get. The 31d MA, corresponding to a monthly resolution, is a standard averaging window length in many impact studies. In the 31d MA estimates, the amplitudes of the ∆$P_{\text{synth}}$ fluctuations are typically in the order of 0.2, but spikes can be as large as 1.3 or 0.7 as in the case of CHS$\text{STATION}$. The grey bands depict the 10th–90th % quantile range of the 500 realisations.

Comparison of the upper and lower panels in Fig. 2.2 shows that the spatial averaging does not reduce the band width of the 10th–90th % quantile range substantially. In the CHN experiments, spatial averaging reduces the width of the 31d MA band averaged over the year from 0.87–1.15 to 0.89–1.12 whereas in the CHS cases, the width is reduced from 0.82–1.22 to 0.84–1.19.

The results indicate that on station scale as well as on regional scales relevant for hydrological impact studies, the fluctuations of ∆$P$ arising from sampling variability alone have to be considered in interpretations of climate change signals. The exact range of the sampling variability is dependent on the averaging window width, the spatial scale, the region of interest and the length of the climate records. For 31d MAs, our analysis shows for representative climate regions of Switzerland, that ∆$P$ values in the range of 0.8 to 1.2 could be solely caused by sampling variability and do not necessarily contain a climate change signal. Furthermore, the spikes within the annual cycles of ∆$P$ call for estimation methods that produce smoother climatological annual cycles than MAs.
2 Spectral representation of the annual cycle in the climate change signal

Figure 2.3: Mean over all station’s MSE_{CV} of harmonic models with increasing harmonic order (HO 1 to HO 12) for observed daily T (top) and Box-Cox transformed P (bottom) series at Swiss climate monitoring stations (see Fig. 2.1). Results of five 30 year periods are shown in different colours. The MSE_{CV} have been normalised by the mean MSE_{CV} for display reasons. In the case of T, the MSE_{CV} of HO 0 is much larger compared to higher order harmonic models and is therefore not shown.

2.5 Analysis of the climate change signal from regional climate models at Swiss station sites

The stochastic analysis in Sect. 2.4 revealed that for variables having similar characteristics as P, like e.g. a clustering of events and a heavily skewed intensity distribution, estimates of the climate change signal using MAs are prone to substantial artificial fluctuations caused by natural variability. In particular, such fluctuations lead to an impaired representation of the minima and maxima in the annual cycle of the climate change signal. Harmonic smoothing is able to filter these fluctuations. However, the maximal order of the harmonic smoothing model (see Eq. 2.6) needs to be chosen. In this section we first define the optimal order of the harmonic smoothing model for T and P. We then present a qualitative comparison between the harmonic and the 31d MA estimates of the climate change signal at station sites, since monthly averaging periods are often employed in climate impact studies. We also discuss the limitations of the spectral smoothing methodology in detail. Finally, we show the climate change signals of T and P estimated by harmonic smoothing for 10 GCM-RCM chains at station sites in Switzerland.
2.5 Analysis of the climate change signal at Swiss station sites

2.5.1 Estimation of the optimal harmonic model

We use long-term observational station records of the Swiss climate monitoring network to constrain the maximum order of the harmonic smoothing model which is then applied to GCM-RCM series. This approach implicitly assumes that signal components from GCM-RCM time series having a higher frequency than the optimal harmonic order are considered as being noise. We use a cross-validation technique to specify the harmonic order for the annual cycle that optimally represents the time series. The methodology is described in detail in Narapasetty et al. (2009). Here, we give only a brief introduction and present our specific setup.

We extracted 30 yr time slices from 25 temperature and 26 precipitation station records with daily resolution, and split them into ten blocks of three year lengths. Five different 30 yr time windows are analysed in order to test the robustness of the results with respect to decadal variability. At each station and for each order of the harmonic smoothing model, we carry out a 10-fold cross-validation by calibrating the harmonic model on 9 of the 10 blocks and validating it on the remaining block. The goal of the cross-validation is to estimate the harmonic model that has the lowest estimated prediction error (EPE). The EPE is a measure of the model error in an independent data set that was not used for calibration. It therefore penalises models that are overfitting the data (see, e.g. Chapter 7 in Hastie et al., 2009). We use the mean squared error (MSE) as a measure for the EPE and call it the cross-validated MSE (MSE\textsubscript{CV}). The MSE\textsubscript{CV} is optimal for normally distributed and independent residuals. P time series however show strongly non-normal residual distributions. This might cause the estimation of the optimal model to be biased. We therefore carried out the cross-validation on the original data, on root-transformed and on Box-Cox transformed data (Wilks, 2006, pp. 43) in order to test how sensitive the results are with respect to the distributional characteristics of the P data. Since the Box-Cox transformation only works on positive definite variables, we replaced zeros in the P data by 0.0001.

Figure 2.3 shows the results of the cross-validation. For T, the MSE\textsubscript{CV} drops to a low level at the harmonic order (HO) of 2 and remains on this low level up to HO 8. Within this plateau, the differences between the models in terms of the MSE\textsubscript{CV} are small. Depending on the analysis period, the order with the lowest MSE\textsubscript{CV} varies between HO 2 and HO 8. The MSE\textsubscript{CV} starts to consistently increase again for higher orders than HO 8. For P, the MSE\textsubscript{CV} has a minimum at HO 2, but the difference to HO 3 is very small. This result is robust for different analysis periods. We only show the results of the Box-Cox transformed data. The results based on the original and root-transformed data are very similar so are not shown here.

The above analysis yields different optimal harmonic orders for T and P. However, as the two atmospheric variables are linked through dynamical and thermodynamical processes,
2 Spectral representation of the annual cycle in the climate change signal

The optimal order for both variables should preferably be the same. We thus chose HO 3 as the optimal order for $T$ and $P$. With a higher joint HO, we would accept higher-frequency precipitation fluctuations that could stem from natural variability rather than climate change.

2.5.2 Comparison between the moving average and the spectral estimation of the climate change signal

Based on the results in Sect. 2.5.1, we use a third order harmonic model (HO 3) to estimate the annual cycle of $T$ and $P$ in the CTL and SCE periods and compare it to 31d MA estimates. We expect the HO 3 estimates to be characterised by smoother annual cycles and smaller peaks in the annual cycle than 31d MA estimates.

For illustration, Fig. 2.4 displays annual cycles of $T$ and $P$ at the two station sites BER and LUG as modelled by ETHZ-HadCM3Q0-CLM in the CTL period and the SCE period 2070–2099 (upper panels) as well as the climate change signal (lower panels). These two stations and the selected model chain represent typical results.
2.5 Analysis of the climate change signal at Swiss station sites

Figure 2.5: Examples of $\Delta P$ as estimated by a 31d MA (dashed lines) and the spectral smoothing (solid lines) of various model chains at different station sites as indicated in each panel’s title. Examples are shown for the SCE period 2070–2099.

In the case of $T$, the annual cycle in the CTL period is well captured at both stations, although biases of up to 2 K arise for individual months. The fluctuations in the 31d MA estimate of $\Delta T$ have a time scale of typically one month. The amplitudes of these fluctuations are in the order of 0.5–1 K. The HO3 estimate treats these fluctuations as noise and results in a smooth annual cycle of $\Delta T$.

In the CTL period, the depicted precipitation shows a large bias in winter on the northern side of the Alps (BER), whereas in southern Switzerland (LUG), the GCM-RCM is able to reproduce the two precipitation peaks in the annual cycle but has a biased amplitude. Such biases are not uncommon in regions of complex topography. For a detailed evaluation of the ENSEMBLES GCM-RCMs, we refer to Klein Tank et al. (2009) and references therein. The estimates of the climatological annual cycle using a 31d MA show high frequency fluctuations in the CTL and SCE periods, which are amplified in the annual cycle of $\Delta P$ due to the division of SCE by CTL values. A spurious amplification can be seen at the station LUG in mid October, when a decrease of $P$ in the CTL period and a rapid increase in the SCE period occur, leading to a spike in $\Delta P$. The HO3 estimate is not influenced by such high frequency fluctuations and results in a smooth annual cycle.

Figure 2.5 shows further examples of strong fluctuations in the 31d MA around the spectrally smoothed annual cycle $\Delta P$ at different station sites and for various model chains. The fluctuations of the 31d MA estimates relative to the spectrally smoothed annual cycles are in the same order of magnitude as the climate change signal.

In climate studies, an important figure is the ensemble mean of the climate change signal. Due to the averaging effect, it is expected that the fluctuations in the climate change signal of the ensemble mean are smaller than for an individual GCM-RCM. Figure 2.6 shows the ensemble mean of the $\Delta P$ signals of the individual GCM-RCMs as estimated by the spectral method and by a 31d MA. The fluctuations of the 31d MA estimate around the HO3 estimate are still substantial and are often in the range of the climate change signal itself. Thus, smoothing might also be necessary for the ensemble mean climate change signal. For applications to impact models such as, for e.g. hydrological models, we
2 Spectral representation of the annual cycle in the climate change signal

Figure 2.6: Examples of the ensemble mean $\Delta P$ as estimated by a 31d MA (dashed lines) and the HO3 spectral smoothing (solid lines) at the station sites BER and LUG. Examples are shown for the SCE period 2070–2099.

recommend not to use the ensemble mean but rather the individual GCM-RCMs climate change signal, and to derive the ensemble mean at the end of the entire impact modelling chain. Impact models are usually non-linear and thus do not yield the same results whether the averaging over the ensemble is done at the GCM-RCM stage or at the end of the impact modelling chain.

2.5.3 Limitations of the methodology

The employed harmonic smoothing model uses a sharp spectral low-pass filter as it removes all harmonic components above the order of 3 and retains the orders 0 to 3 without any damping. Such sharp spectral filters are prone to overshootings which can occur in situations when sudden changes in a time series take place within a time scale that cannot be resolved by the spectral model. This is also known as the Gibbs phenomenon. Overshootings are problematic as, e.g., negative precipitation values can occur, so results need to be scanned for such overshootings. For the application to the delta change methodology, there are three types of overshooting cases: overshootings occur in the smoothed mean annual cycle of (a) observed time series, (b) climate model time series in the CTL period (and possibly also in the SCE period) and (c) climate model time series in the SCE period only. Case (a) indicates that the spectral smoothing model is not appropriate for the climate in the region of interest. Case (b) indicates that the climate model is spectrally biased since its time series contain more spectral frequencies than estimated from the observed time series. In case (c), the climate model is not spectrally biased but changes its spectral characteristics substantially in the transition from the CTL to the SCE period. In case (a), one should not use the spectral model but resort to other smoothing methods such as, e.g., generalized linear models or smoothing filters that do not produce overshootings. In cases (b) and (c), one could use a less sharp spectral filter to estimate the mean annual cycle.

Figure 2.7 shows the application of the sharp (i.e. having a sharp cutoff) and a less sharp (i.e. having a gradual cutoff) HO3 filter as well as the 31d MA to the time series of the DMI-ARPEGE-HIRHAM GCM-RCM at the station LUG. The inset in Fig. 2.7 shows the response function of the 31d MA and the two spectral filters constructed following Duchon (1979). The response function of the 31d MA is a damped sinus oscillation (black line).
2.5 Analysis of the climate change signal at Swiss station sites

![Illustration of the oversooting problem](image)

**Figure 2.7:** Illustration of the oversooting problem for the example of the GCM-RCM DMI-ARPEGE-HIRHAM at the station LUG. Annual cycle of $P$ in the CTL (top) and SCE 2070–2099 (middle) period as well as its climate change signal (bottom) as estimated by use of different filters. The inset shows the response function of the filters.

The sharp filter has a cutoff at HO 3 (red line). The gradual cutoff filter (green line) has a linear transition from the response value of 1 to 0. We subjectively chose a transition range of ±2 harmonic orders around HO 3.

In the example shown, the GCM-RCM has a pronounced summer drying in the CTL period. In the SCE period, the summer drying still exists, but precipitation starts to increase a bit earlier in the year. As a result, the 31d MA shows a very high and unrealistic multiplicative climate change signal of up to 4.5 in August. The sharp cutoff HO 3 filter has problems resolving the pronounced low precipitation period in summer and produces overshootings that cause negative precipitation. The climate change signal is therefore not meaningful. The gradual cutoff HO 3 filter does not produce overshootings and results in a reasonable climate change signal.

It remains up to the users of the presented methodology to choose whether GCM-RCMs that show problems such as those in cases (b) and (c) should be included in the ensemble or not. We chose not to use GCM-RCMs that have severe overshootings. In principle though, the approach of a spectral filter with a gradual cutoff would be a suitable workaround if such GCM-RCMs should be included.
2.5.4 Climate change signal at Swiss station sites

Annual cycle of the climate change signal

For brevity, we show results of the climate change signal’s annual cycle only at the two exemplary stations BER and LUG (see Fig. 2.1). Figure 2.8 shows each GCM-RCM’s annual cycle of $\Delta T$ and $\Delta P$ respectively for both SCE periods relative to the CTL period 1980–2009. To compare the changes to the natural variability range, we resampled each station’s observed precipitation record of the CTL period. We constructed 100 realisations with a length of 30 yr by resampling with replacement the years of the observed record. From the 100 realisations, we randomly chose 500 pairs and estimated the climate change signal between the pairs. The range of $\pm 1$ standard deviation ($\sigma$) of the 500 resampled realisations is shown as a grey band in Fig. 2.8.

In the case of $\Delta T$, the ensemble mean shows peaks in winter and summer in both SCE periods. The model spread is largest in summer, which is mainly due to a strong summer warming of HadCM3Q0-driven experiments. Generally, the $\Delta T$ signal for both SCE periods is distinctively above the estimated natural variability range. The natural variability range of $\Delta P$ relative to the projected $\Delta P$ values is much larger than in the case of $\Delta T$. Also, the range of natural variability strongly differs from station to station. Only the decrease of $P$ in the summer of the later scenario period as projected by a majority of the GCM-RCMs substantially exceeds the range of the natural variability.
2.5 Analysis of the climate change signal at Swiss station sites

The results also indicate that for $\Delta T$ and to a lesser extent for $\Delta P$, GCM-RCMs belonging to the same GCM family show similar patterns in the annual cycle of the climate change signal.

### Spatial patterns of seasonal mean changes

Figure 2.9 shows the mean seasonal pattern of $\Delta T$ and $\Delta P$ for both scenario periods. Only the results for the seasons DJF and JJA are shown since the analysis of the annual cycles showed these seasons to have stronger climate change signals than the transition seasons.

For $\Delta T$, the spatial pattern is homogenous across Switzerland with the exception of the Alpine ridge region in JJA that generally shows higher $\Delta T$ than other regions in Switzerland. The strongest warming is projected for JJA. The median of all station’s ensemble mean $\Delta T$ for JJA is 1.35 K for 2021–2050 and 3.72 K for 2070–2099. At most stations, the ensemble mean $\Delta T$ is larger than 2 times the standard deviation of the natural variability for both scenario periods.

The strongest seasonal $\Delta P$ signal is projected for JJA in the SCE period 2070–2099 with ensemble mean $\Delta P$ values around 0.8 in large parts of Switzerland. Southern Switzerland
is projected to have the strongest decrease of summer precipitation, with $\Delta P$ values around 0.7. For DJF, an increase of $P$ can be expected but the strength of the signal is smaller than for JJA. In the period 2021–2050, the $\Delta P$ values generally do not exceed the range of estimated natural variability.

The ensemble mean’s projected seasonal changes for the SCE period 2070–2099 are consistent with the results from the PRUDENCE project (Christensen et al., 2007). In the PRUDENCE project, the ensemble mean $\Delta T$ in the Alpine region for the SCE period 2070–2100 relative to the CTL period 1961–1990 was $+2\,\text{K}$ for winter and $+4\,\text{K}$ for summer. The estimated ensemble mean $\Delta P$ was $+10\%$ and $-30\%$ for winter and summer respectively (Christensen and Christensen, 2007).

### 2.6 Summary and conclusions

The delta change method commonly used in climate impact modelling studies requires a representation of the climate change signal’s annual cycle. This implies the estimation of the annual cycle of $T$ and $P$ both in the CTL and the SCE period. Using a stochastic rainfall generator, we showed that climate change signals of mean precipitation derived by MAs are strongly affected by sampling artefacts. Spatial aggregation to a region corresponding to the area of a few RCM grid cells does not reduce the effect of sampling variability on the climate change signal substantially.

Climate change signals estimated using narrow averaging intervals, such as monthly means, should thus be regarded with caution, since associated artificial peaks in the annual cycle can lead to undesirable effects when used in combination with non-linear impact models.
We used a spectral smoothing to ameliorate the effects of natural variability on artificial fluctuations in the annual cycle. Compared to 31d MA estimates, the spectral smoothing successfully filters intraannual fluctuations. In a few cases when a strong amplitude of the annual precipitation cycle is paired with a large relative precipitation change, the spectral smoothing produces overshootings.

The derived climate change signal for the ENSEMBLES GCM-RCM chains is particularly clear for the later SCE period 2070–2099. The peak in the ensembles mean’s $\Delta T$ is around 4 K. In the case of $P$, a pronounced decrease of summer precipitation is projected for the whole of Switzerland. In the other seasons, precipitation is projected to increase.

In this study we focussed on changes in the annual cycle of mean $T$ and $P$ as used in the delta change method. This is a statistical model of low complexity in the whole variety of statistical post-processing methods. More complex models, such as, e.g. bias-correction methods involving a mean and variance scaling or quantile-mapping, are possibly even more sensitive to natural variability than the delta change method. The quantification of the sensitivity, however, requires further study. The study presented here indicates that the representation of the annual cycle in any statistical post-processing or downscaling method should be addressed with care.
2.7 Acknowledgements

This study was partly funded by swisselectric research, the Federal Office for the Environment (FOEN) and NCCR climate, a research instrument of the Swiss National Science Foundation. The ENSEMBLES data used in this work was funded by the EU FP6 Integrated Project ENSEMBLES (Contract number 505539) whose support is gratefully acknowledged. The observational station data were provided by MeteoSwiss, the Swiss Federal Office of Meteorology and Climatology. We thank the Center for Climate Systems Modeling (C2SM) for its technical support. Also, we would like to thank Lukas Rosinus from the statistical seminar of the ETH Zurich for his assistance.
3 Quantifying uncertainty sources in hydrological climate impact projections

Thomas Bosshard¹, Maria Carambia², Klaus Goergen³, Sven Kotlarski¹, Peter Krahe², Massimiliano Zappa⁴, Christoph Schär¹

¹Institute for Atmospheric and Climate Science, ETH Zurich, Universitätsstrasse 16, 8092 Zürich, Switzerland
²German Federal Institute of Hydrology, Am Mainzer Tor 1, 56068 Koblenz, Germany
³Centre de Recherche Public - Gabriel Lippmann, 41 rue du Brill, Belvaux, Luxembourg
⁴Swiss Federal Institute for Forest, Snow and Landscape Research WSL, Zürcherstrasse 111, 8903 Birmensdorf, Switzerland

(Submitted to Water Resources Research)

Abstract

The quantification of the uncertainties in projections of climate-impacts on river streamflow is highly important for water management and climate adaptation purposes. In this study, we present a methodology to separate uncertainties arising from the climate model, the statistical post-processing scheme, and the hydrological runoff model. The study is carried out for the Alpine Rhine in Eastern Switzerland. We used data from eight regional climate models of the ENSEMBLES project, two statistical post-processing methods and two hydrological models to analyze projections of hydrological changes for the near-term and far-term scenario periods 2021-2050 and 2070-2099 with respect to 1961-1990. For the later scenario period, the model ensemble projects a decrease of daily mean runoff in summer (−32.2 [−45.5 to −8.1] %) and an increase in winter (+41.8 [+4.8 to +81.8] %). We applied two statistical methods for variance decomposition in order to assess the importance of different uncertainty sources. The first method, the direct variance estimation, assumes that individual uncertainties are additive and independent, while the second method, the ANOVA approach, considers interactions between the uncertainty sources. Both methods identified the climate models to be generally the dominant uncertainty source in summer and autumn. In winter and spring of the far-term scenario period, the hydrological models are the dominant source of uncertainty. The ANOVA method further suggests that uncertainties due to interactions between the impact modeling chain elements contribute between 5 and 40 % to the total ensemble uncertainty. The time varying contribution of different uncertainty sources as well as their interactions indicate that it is the combination of the modeling chain elements that needs to be addressed in climate-impact studies.
3 Quantifying uncertainty sources in hydrological climate impact projections

3.1 Introduction

The projected impacts of climate change on river streamflow are associated with large uncertainties. For water management, they constitute an important source of uncertainty in addition to, e.g., natural variability or changes in water demand (Kundzewicz et al., 2008). While it has been debated how useful climate-impact projections are for water management planning purposes, most proposed planning strategies make use of information gained from climate-impact projections (Dessai and Hulme, 2007; Milly et al., 2008; Kundzewicz and Stakhiv, 2010). In particular, the quantification of uncertainties in climate-impact projections is of major interest for most decision-making strategies. So far, the overall uncertainty of climate-impacts is probably underestimated which is partly due to an imbalanced sampling of the uncertainty sources (Knutti, 2008; Wilby, 2010). Improving the knowledge about the importance of different uncertainty sources might help to design future climate-impact studies in order to aim for a more complete assessment of the uncertainties.

Impact modeling systems that include a cascade of different models are commonly used to assess climate-impacts and to provide information for water management. Elements of this cascade are an emission scenario, a global circulation model (GCM), a dynamical downscaling step by means of a regional climate model (RCM), a statistical post-processing (PP) and a hydrological model (HM). Alternatively, the dynamical downscaling and the PP steps can be replaced by a statistical downscaling. In the remainder of this article, we call this cascade of emission scenarios and models an impact modeling chain. Uncertainties in climate runoff projections arise due to different assumptions and model combinations in the whole impact modeling chain. For a complete analysis of uncertainty in runoff projections, it is therefore important to investigate the contributions of all various sources. Numerous previous studies have investigated climate runoff projections and their sensitivity to different uncertainty sources. Focussing on the Rhine basin, Shabalova et al. (2003) compared two PP methods and found that for the end of the 21st century, both methods agree on a decrease of summer runoff and an increase of winter runoff but the two methods lead to different increases of winter flood risk. Lenderink et al. (2007) also investigated uncertainties due to PP and confirmed the results of Shabalova et al. (2003) that the choice of the PP rather effects the changes in runoff extremes than in the mean runoff. Jasper et al. (2004) investigated the projected impact on runoff of an ensemble of 17 climate scenarios derived from 7 GCMs and 4 emission scenarios in two Swiss catchments. They found that changes in the seasonality of runoff are robust but the magnitude of the changes is strongly affected by the choice of the climate scenario. Combined uncertainties from emission scenarios, GCMs and RCMs were investigated by Graham et al. (2007) who found that the choice of the GCM has a larger impact on projected hydrological changes than the choice of the RCM or emission scenario.

More recently, studies that systematically investigate multiple uncertainty sources along
3.1 Introduction

the whole impact modeling chain have been published. Wilby and Harris (2006) assessed uncertainties from emission scenarios, GCMs, statistical downscaling, hydrological model structure and hydrological model parameters. Using a probabilistic framework, they showed that GCMs and the downscaling step were the most important sources of uncertainty in simulating changes of low flows in the Thames River (UK). Prudhomme and Davies (2009a) assessed uncertainties due to emission scenarios, GCMs, downscaling methods and hydrological models in four mesoscale British catchments and concluded that the driving GCM is the dominant source of uncertainty. They further stated that uncertainty due to PP and the choice of the emission scenario are of comparable magnitude while uncertainty due to the hydrological models is negligible in two out of four basins. Kay et al. (2009) investigated the same uncertainty sources as Wilby and Harris (2006) and also included the effect of internal variability in a case study which assessed changes of flood frequency in two British catchments. They found GCMs being the dominant source of uncertainty. However, after excluding one outlier from the GCM ensemble, other uncertainty sources such as RCMs and internal variability became more important than GCMs. For two catchments in Oregon (USA), Jung et al. (2011) found the natural variability and the driving GCM to be the major sources for uncertainty with respect to flood frequency changes.

In our study, we perform an ensemble of hydrological climate-impact projections for an Alpine river catchment and the two scenario periods (SCE) 2021-2050 and 2070-2099 with respect to the control period (CTL) 1961-1990. From this ensemble, we aim to infer the role of several uncertainty sources. The three uncertainty sources considered are (i) climate models (CMs) consisting of a GCM and a RCM, (ii) PP and (iii) HMs. The Alpine study area is a challenging region for all three impact modeling chain elements. Climate models, either GCMs or RCMs, for instance cannot fully resolve the complex topography, PP methods have to correct for potentially large biases and HMs are challenged by complicated and spatially highly variable hydrological processes such as accumulation and melt of snow.

We include fewer uncertainty sources than some of the mentioned previous studies but instead of performing single-propagation runs, we conduct a multi-propagation study, i.e. we vary the different CMs, PP methods and HMs in all possible combinations. This approach allows for an assessment of interactions between the uncertainty sources (Kay et al., 2009). The limited number of uncertainty sources and models included into the ensemble results in an underestimation of the overall uncertainty associated with the hydrological climate-impacts (i.e. the uncertainty if all possible uncertainty sources are fully sampled). Throughout the paper, we call the spread in our ensemble the total ensemble uncertainty in order to clearly distinguish it from the overall uncertainty.

Our two research questions are: (1) How large is the total ensemble uncertainty in the runoff projections and (2) how do different uncertainty sources contribute to the total ensemble uncertainty in the runoff projections for both SCEs. In particular, we are in-
interested in how the contributions vary throughout the annual cycle and how they affect
the uncertainty in changes of different runoff quantiles. We investigate two approaches to
split up the total ensemble uncertainty of the runoff projections into contributions from
individual sources. In the first approach we calculate the variances directly and neglect
any interactions between the different uncertainty sources (see e.g. Hawkins and Sutton,
2009). This method will here be referred to as the \textit{direct variance decomposition} (DVD)
approach. Second, we use the statistical theory of the analysis of variance (ANOVA) that
allows for the consideration of interactions between the uncertainty sources (Déqué et al.,
2007). These interactions represent uncertainty contributions that do not behave linearly.
For instance, a snow melt bias of an HM may depend upon the temperature bias of the
driving CM which could lead to a non-linear response in the impact on runoff. Both
approaches are complemented with a sub-sampling scheme in order to account for the
different sample sizes of the three uncertainty sources.

The paper is structured as follows: Section 3.2 briefly describes the study area and the
data. Section 3.3 introduces the employed HMs, explains the PP methods and presents
the sub-sampling procedure in combination with the two methods used for the variance
decomposition. In section 3.4, we present the results of the climate runoff projections and
the two variance decomposition methods. Section 3.5 summarizes our study and its main
findings.

\section*{3.2 Study region & Data}

\subsection*{3.2.1 Study region}

The study region consists of the Alpine Rhine catchment down to the gauge Diepoldsau
(see Fig. 3.1). It encompasses an area of 6119 km\(^2\) and has a mean elevation of about
1800 m a.s.l. The runoff regime is nival, i.e. snow-dominated, but altered to some extent
by hydropower production. The hydropower’s main effect is a seasonal redistribution of
water from summer to winter (Verbunt et al., 2005). The evolution of the hydropower
storage capacity in the period 1952-2009 is depicted in Fig. 3.3 (black line). At the end
of the period, the storage capacity amounts to about 10\% of the annual runoff volume.

\subsection*{3.2.2 Observational and climate model data}

Throughout the study, we use the control period 1961-1990 (CTL) and the two scenario
periods 2021-2050 (SCE1) and 2070-2099 (SCE2) as temporal subsets of the data series.
The two hydrological models HBV and PREVAH (see section 3.3.1) require different kinds
of input data with respect to the spatial resolution and the variables (see Tab. 3.1). HBV
3.2 Study region & Data

Figure 3.1: Map showing the studied catchment of the Alpine Rhine river down to Diepoldsau. The sub-basin structures used in HBV and PREVAH is shown in the left and right panels, respectively. The hydropower reservoirs are shown as red areas.

uses basin-averaged daily time series of precipitation, temperature and global radiation or sunshine duration. Basin-averaged observational data have been provided by the International Commission for the Hydrology of the Rhine basin (CHR; referred to as OBS\textsubscript{CHR} in the remainder of this article, Görgen et al. (2010)). PREVAH uses daily station data of the hydrometeorological variables precipitation (80 stations), temperature (36 stations), relative humidity (41 stations), sunshine duration (28 stations) and wind speed (43 stations) from the measurement network of MeteoSwiss (referred to as OBS\textsubscript{ST} in the remainder of this article). Both observational data sets cover the whole CTL period 1961-1990.

For the calibration of the hydrological models, we used daily runoff data of the gauges depicted in Fig. 3.1. The data was provided by the Swiss Federal Office for the Environment (see www.hydrodaten.admin.ch).

For the climate data, we used eight transient climate modeling chains of the ENSEMBLES project (van der Linden and Mitchell, 2009) as shown in Fig. 3.2. All of them assume the A1B emission scenario, have a horizontal resolution of about 25 km and cover the period 1951-2099. We use sub-basin averaged climate model time series (HBV) and climate model data interpolated to station locations (PREVAH; see section 3.3.2).
3.3 Methods

Figure 3.2 depicts the modeling chain combination scheme employed in this study. In the following, we describe the HMs and the PP methods and explain the approaches to attribute an uncertainty contribution to each of the modeling chain elements.

3.3.1 Hydrological Models (HM)

Both the HBV and the PREVAH model are semi-distributed conceptual rainfall-runoff models. Both models use the hydrological response type (HRU) approach to cluster the spatial units according to their hydrological characteristics. Table 3.1 summarizes the characteristics of the two hydrological models in the Alpine Rhine catchment and Fig. 3.1 shows the different sub-basin structure of the two HMs. For the HBV model, we use the HBV134 setup of the German Federal Institute of Hydrology. The HBV134 requires basin-averaged hydrometeorological time series as input data and applies a lapse rate correction to disaggregate basin-averaged temperature to zones of different elevation. For the PREVAH model, we use the setup by Verbunt et al. (2006) which has been recalibrated in the period 1985-1990 in order to use sunshine duration instead of global radiation as an input variable from which shortwave radiation is derived according to Schulla (1997). PREVAH requires hydrometeorological station data as an input. The station data are interpolated using inverse distance weighting. For temperature, a height dependent re-
gression is applied (detrended inverse distance weighting). Both HMs correct the input precipitation by a linear correction factor in order to close the water balance. The linear precipitation correction factor has been estimated by the calibration. For further details about the HMs, we refer to the references listed in Tab. 3.1.

**Table 3.1**: List of key characteristics of the two employed HMs.

<table>
<thead>
<tr>
<th></th>
<th>HBV</th>
<th>PREVAH</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Model type</strong></td>
<td>Conceptual, semi-distributed</td>
<td>Conceptual, semi-distributed</td>
</tr>
<tr>
<td><strong>Meteorological input data</strong></td>
<td>precipitation (P), temperature (T), global radiation (R) or sunshine duration (S)</td>
<td>precipitation (P), temperature (T), sunshine duration (S), cloud cover (C), relative humidity (H), wind speed (V)</td>
</tr>
<tr>
<td><strong>Number of sub-basins</strong></td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td><strong>Spatial resolution of the underlying digital elevation model</strong></td>
<td>1000 m x 1000 m</td>
<td>500 m x 500 m</td>
</tr>
<tr>
<td><strong>Land use classes</strong></td>
<td>4</td>
<td>29</td>
</tr>
<tr>
<td><strong>HRU definition</strong></td>
<td>elevation, land use</td>
<td>elevation, land use, aspect, soil type</td>
</tr>
<tr>
<td><strong>Snow/glacier melt modeling approach</strong></td>
<td>Degree-day factor</td>
<td>Degree-day factor with aspect and slope correction</td>
</tr>
<tr>
<td><strong>Evapotranspiration parameterization</strong></td>
<td>Penman-Wendling</td>
<td>Penman-Monteith</td>
</tr>
<tr>
<td><strong>Calibration period</strong></td>
<td>1970-1984</td>
<td>1985-1990</td>
</tr>
</tbody>
</table>

### 3.3.2 Statistical Post-Processing (PP)

We use a bias-correction (BC) and a delta change (DC) approach for the PP step in the impact modeling chain. These two PP methods differ distinctively regarding the treatment of changes in the variability. In the BC, the time series used to drive the HM are based on climate model data, which are corrected towards the climatological mean of the observations in the CTL. Thus, the variability in the climate model data determines the variability in the bias-corrected forcing data, e.g., the succession of wet and dry days. The DC approach on the other hand scales observed time series according to a climate change signal estimated from climate model data. The scaled observational time series are used to force the HMs in the SCE. It is therefore the variability in the observed time series that determines the variability in the HM’s forcing data. Also, the two PP methods
differ in the number of variables they include. The BC approach corrects all required variables from the output of the CMs while the DC method is applied to temperature and precipitation, only. Note that HBV and PREVAH use OBS\textsubscript{CHR} and OBS\textsubscript{ST} as the observational reference, respectively. Thus, also the PP methods use OBS\textsubscript{CHR} and OBS\textsubscript{ST} for modeling chains containing the HBV and PREVAH model, respectively (see above).

### Bias-correction (BC)

For the bias-correction of the climate model data, we use a linear scaling as employed by Lenderink et al. (2007) and Görgen et al. (2010). It corrects the climatological monthly mean of the CM output to match the observed climatological monthly mean in a control period.

Let $X$ be a meteorological input variable for the HM. As the observational reference, we use basin mean time series and call it $X_{\text{obs}}^{\text{avg}}$. For OBS\textsubscript{ST} used by PREVAH, we first derive basin averages from the station data which is done by elevation detrended inverse distance weighting for temperature and inverse distance weighting for the other variables. OBS\textsubscript{CHR} used by HBV is a basin-averaged data set, already. Next, for each CM $j$, we calculate a basin mean time series as area weighted grid cell averages and denote it with $X_{j}^{\text{avg}}$. We estimate the correction parameters $a^{X}(m)$ between the observed basin mean time series $X_{\text{obs}}^{\text{avg}}$ and $X_{j}^{\text{avg}}$ for each month $m$ in the annual cycle. We use the following linear correction models for the various meteorological variables (see Tab. 3.1 for an explanation of the abbreviations):

\begin{align}
    P_{j}^{\text{avg}} &= a^{P} P_{j}^{\text{avg}}, \\
    T_{j}^{\text{avg}} &= T_{j}^{\text{avg}} + a^{T}, \\
    S_{j}^{\text{avg}} &= \min\{S_{0}, a^{S} S_{j}^{\text{avg}}\}, \\
    R_{j}^{\text{avg}} &= a^{R} R_{j}^{\text{avg}}, \\
    H_{j}^{\text{avg}} &= \max[0, 100 - a^{H}(100 - H_{j}^{\text{avg}})], \\
    V_{j}^{\text{avg}} &= a^{V} V_{j}^{\text{avg}},
\end{align}

where the overbar denotes climatological monthly means in the CTL and the superscript $X$ is defined as

\begin{align}
    a^{P}(m) &= \frac{P_{\text{obs}}^{\text{avg}}(m)}{P_{j}^{\text{avg}}(m)}, \\
    a^{T}(m) &= \frac{T_{\text{obs}}^{\text{avg}}(m) - T_{j}^{\text{avg}}(m)}{T_{j}^{\text{avg}}(m)}, \\
    a^{S}(m) &= \frac{S_{\text{obs}}^{\text{avg}}(m)}{S_{j}^{\text{avg}}(m)}, \\
    a^{R}(m) &= \frac{R_{\text{obs}}^{\text{avg}}(m)}{R_{j}^{\text{avg}}(m)}, \\
    a^{H}(m) &= \frac{100 - H_{\text{obs}}^{\text{avg}}(m)}{100 - H_{j}^{\text{avg}}(m)}, \\
    a^{V}(m) &= \frac{V_{\text{obs}}^{\text{avg}}(m)}{V_{j}^{\text{avg}}(m)},
\end{align}
3.3 Methods

* stands for the bias-corrected basin-averaged daily time series. The equations in the left column show the linear bias-correction model whereas the equations in the right column explain how the correction parameter is derived. Note that, although not explicitly indicated in the left column, the correction factor $a^X$ varies according to the month. As indicated in Tab. 3.1, each HM uses only a selection of the meteorological variables listed above.

Some CMs do not provide sunshine duration. Thus, we use cloud fraction (C) as a proxy and derive $S$ according to

$$S_{avg}^j = S_0(1 - C_{avg}^j) \quad (3.7)$$

with $S_0$ representing the maximum possible sunshine duration.

After the bias-correction on the basin level, a further spatial disaggregation step is necessary for the PREVAH model in order to account for the finer spatial input structure. In the CTL, we derive for every month $m$ and every meteorological sub-area $i$ a disaggregation relation based on the observations according to

$$r_i(m) = \frac{X_{obs}^i(m) - X_{avg}^i(m)}{X_{obs}^i(m)} \quad \text{for temperature, and} \quad (3.8)$$

$$r_i(m) = \frac{X_{obs}^i(m)}{X_{avg}^i(m)} \quad \text{for the other variables,} \quad (3.9)$$

and scale the bias-corrected basin mean time series like

$$X_{i}^{*j} = X_{avg}^{*j} + r_i \quad \text{for temperature and} \quad (3.10)$$

$$X_{i}^{*j} = r_i X_{avg}^{*j} \quad \text{for the other variables.} \quad (3.11)$$

**Delta change (DC)**

We use the DC method as described in the work of Bosshard et al. (2011) and CH2011 (2011) which includes a spectral smoothing of the time series in order to damp the influence of natural variability on estimates of the climate change signal. Thus, in the DC approach, the annual cycle is resolved continuously as opposed to the BC method where monthly steps are used. The delta change scaling is applied to temperature and precipitation data only. For HBV, we derive the climate change signal from the basin-averaged time series $X_{avg}^j$ and scaled the OBS$^{CHR}$ accordingly. For PREVAH, we first interpolate the climate model data to the station locations using inverse distance weighting, derive the climate change signal at the station sites and scale OBS$^{ST}$ data accordingly (see Bosshard et al., 2011).
3 Quantifying uncertainty sources in hydrological climate impact projections

3.3.3 Variance decomposition

In the variance decomposition, we aim to decompose the total ensemble uncertainty into contributions from different elements of the impact modeling chain and interactions between them. In our study, the total ensemble uncertainty is the spread of the climate change signal in the mean annual cycle of runoff $Q_{AC}$ (estimated by a 31 day moving average; 31d MA), and in different runoff quantile levels ($Q_Q$). The full model ensemble consists of 32 impact modeling chain combinations (Fig. 3.2). For each of the 32 impact modeling chain combinations, we estimate $Q_{AC}$ and $Q_Q$ in the CTL and SCE (i.e. SCE1 or SCE2). Then, we calculate the climate change signal $Y$ as

$$Y_{AC} = Q_{SCE}^{AC} - Q_{CTL}^{AC}$$

$$Y_Q = Q_{SCE}^Q - Q_{CTL}^Q.$$  \hspace{1cm} (3.12)

In the following, we neglect the subscripts referring to the mean annual cycle or the runoff quantiles, and describe the methodology for the general variable $Y$. In order to relate the target variable $Y$ to the uncertainty sources, we use superscripts in $Y_{j,k,l}$ with $j$, $k$ and $l$ representing the different samples of CMs, PP methods and HMs, respectively. In total, we have eight different CMs, two different PP methods and two different HMs. We use a sub-sampling of the CMs in order to yield equal sample sizes for every uncertainty source.

Sub-sampling of the CMs

The different sample sizes for the three uncertainty sources might affect the uncertainty analysis (see also Déqué et al. (2007) for a short discussion of the problem). Therefore, we sub-sample the eight different CMs. In each sub-sampling iteration, we select two CMs out of the eight which results in a total of 28 possible CM pairs. For each of the 28 sub-sampling iterations $i$, we end up with two CMs, two PP methods and two HMs that define our model combination matrix for the variance decomposition. In order to differ between the full set of eight CMs and the sub-sampled CM pair, we replace the superscript $j$ with $g(h, i)$. The superscript $g$ is a $2 \times i$ matrix which contains the selected CMs for the particular sub-sampling iteration $i$:

$$g = \begin{pmatrix} 1 & 1 & \cdots & 1 & 2 & 2 & \cdots & 6 & \cdots & 7 \\ 2 & 3 & \cdots & 8 & 3 & 4 & \cdots & 8 & \cdots & 8 \end{pmatrix}. \hspace{1cm} (3.14)$$
3.3 Methods

Direct variance decomposition (DVD)

We apply two different methods for the decomposition of the total ensemble uncertainty into contributions of the three uncertainty sources, referred to as the direct variance decomposition (DVD) and the ANOVA method. In the DVD approach, we estimate the uncertainty contributions in terms of variances following Hawkins and Sutton (2009). For the estimation of the contribution of a particular modeling chain element, we vary only the samples of the particular element and calculate the variance based on the simulated changes in the output variable $Y$.

The variances of the individual elements with respect to $Y$ are estimated as,

$$
\sigma^2_{CM} = \frac{1}{I} \sum_{i=1}^{I} \frac{1}{H-1} \sum_{h=1}^{H} (Y^{g(h,i),\circ,\circ} - \bar{Y}^{g(\circ,\circ),\circ,\circ})^2
$$

(3.15)

$$
\sigma^2_{PP} = \frac{1}{I} \sum_{i=1}^{I} \frac{1}{K-1} \sum_{k=1}^{K} (Y^{g(\circ,i),k,\circ} - \bar{Y}^{g(\circ,i),\circ,\circ})^2
$$

(3.16)

$$
\sigma^2_{HM} = \frac{1}{I} \sum_{i=1}^{I} \frac{1}{L-1} \sum_{l=1}^{L} (Y^{g(\circ,i),\circ,l} - \bar{Y}^{g(\circ,i),\circ,\circ})^2
$$

(3.17)

where the superscript $\circ$ in place of $h$, $k$, or $l$ indicates averaging over all the samples of the superscript. $I$ is equal to the total number of iterations ($I = 28$). $H$ denotes the total number of sub-sampled CMs per iteration ($H = 2$). $K$ and $L$ are the number of levels for PP and HM, respectively ($K = 2$, $L = 2$). The DVD approach assumes independence of the individual elements. In this case, the fraction of variance $f$ of each element is given as

$$
f_{CM} = \frac{\sigma^2_{CM}}{\sigma^2_{CM} + \sigma^2_{PP} + \sigma^2_{HM}}
$$

(3.18)

$$
f_{PP} = \frac{\sigma^2_{PP}}{\sigma^2_{CM} + \sigma^2_{PP} + \sigma^2_{HM}}
$$

(3.19)

$$
f_{HM} = \frac{\sigma^2_{HM}}{\sigma^2_{CM} + \sigma^2_{PP} + \sigma^2_{HM}}
$$

(3.20)

The variance fraction $f$ is a measure for the contribution of an individual uncertainty source to the total ensemble uncertainty in $Y$. Values of 0 and 1 correspond to a contribution of 0 and 100% respectively.

ANOVA approach

We use the statistical theory of the analysis of variance (ANOVA) to provide an alternative estimate of the variance decomposition (see e.g., von Storch and Zwiers (1999) for an introduction and Déqué et al. (2007) for an application in climate modeling). In our
Quantifying uncertainty sources in hydrological climate impact projections

Experimental setup, there is one data point $Y$ for every possible combination of the three modeling chain elements CM, PP and HM. In the terminology of the ANOVA, the modeling chain elements are “effects”. An effect is a variable that one suspects could have an influence on the variability of the variable $Y$. In other words, we construct an ANOVA model based on the hypothesis that the CM, PP and HM elements of the modeling chain have an influence on the variability of the variable $Y$ and we want to quantify the influence. We define the model as

$$Y^{j,k,l} - \overline{Y^{o,o,o}} = a^j + b^k + c^l + ab^{j,k} + ac^{j,l} + bc^{k,l} + abc^{j,k,l},$$

(3.21)

where $a$, $b$ and $c$ are the effects corresponding to CM, PP and HM, the expressions $ab$, $ac$, $bc$ and $abc$ are the interaction terms and $j$, $k$ and $l$ indicate samples of the different effects. According to the ANOVA theory, the model allows us to split the total sum of the squares (SST) into sums of squares due to the individual effects (SSA, SSB, SSC) and their interactions (SSAB, SSAC, SSBC, SSABC) as

$$SST = SSA + SSB + SSC + SSI,$$

(3.22)

$$SSI = SSAB + SSAC + SSBC + SSABC.$$

(3.23)

In this model, we summarize all interaction terms into the term SSI. A substantial contribution of interactions to the SST indicates that it is important to test all possible impact modeling chain combinations in order to sample the total ensemble uncertainty. Neglecting the interactions would lead to a considerable underestimation of the total ensemble uncertainty.

We estimate the terms in Eq. 3.22 and 3.23 using the sub-sampling procedure introduced in section 3.3.3 as

$$SST_i = \sum_{h=1}^{H} \sum_{k=1}^{K} \sum_{l=1}^{L} (Y^{g(h,i),k,l}_i - \overline{Y^{g(o,i),o,o}})^2,$$

(3.24)

$$SSA_i = K \cdot L \cdot \sum_{h=1}^{H} (Y^{g(h,i),o,o}_i - \overline{Y^{g(o,i),o,o}})^2,$$

(3.25)

$$SSB_i = H \cdot L \cdot \sum_{k=1}^{K} (Y^{g(o,i),k,o}_i - \overline{Y^{g(o,i),o,o}})^2,$$

(3.26)

$$SSC_i = H \cdot K \cdot \sum_{l=1}^{L} (Y^{g(o,i),o,l}_i - \overline{Y^{g(o,i),o,o}})^2,$$

(3.27)

and

$$SSI_i = \sum_{h=1}^{H} \sum_{k=1}^{K} \sum_{l=1}^{L} (Y^{g(h,i),k,l}_i - \overline{Y^{g(h,i),o,o}} - \overline{Y^{g(o,i),k,o}} - \overline{Y^{g(o,i),o,l}} + 2\overline{Y^{g(o,i),o,o}})^2.$$

(3.28)
3.4 Results

Then, for each effect, the variance fraction $\eta^2$ is derived as

$$\eta_{CM}^2 = \frac{1}{I} \sum_{i=1}^{I} \frac{SSA_i}{SST_i},$$
(3.29)

$$\eta_{PP}^2 = \frac{1}{I} \sum_{i=1}^{I} \frac{SSB_i}{SST_i},$$
(3.30)

$$\eta_{HM}^2 = \frac{1}{I} \sum_{i=1}^{I} \frac{SSC_i}{SST_i}, \text{ and}$$
(3.31)

$$\eta_{Interactions}^2 = \frac{1}{I} \sum_{i=1}^{I} \frac{SSI_i}{SST_i}.$$  
(3.32)

Values of 0 and 1 for the variance fraction $\eta^2$ correspond to a contribution of an effect to the total ensemble variance (uncertainty) of 0 and 100%, respectively.

### Comparison of the two approaches

The DVD and the ANOVA approach differ in their characteristics. The DVD approach assumes independence (e.g., no interactions) of the different uncertainty sources and the total ensemble uncertainty is the sum of the contributions of the different uncertainty sources. The ANOVA approach accounts for interactions between the uncertainty sources in the SSI term and states that the total ensemble uncertainty, measured in terms of the SST, is the sum of the contributions of individual uncertainty sources and their interactions. In case of interactions, the SST in the ANOVA approach will be larger than the total variance in the DVD approach, which leads to smaller contributions of individual uncertainty sources to the total ensemble uncertainty. The interpretation of interactions is that the uncertainty due to, e.g., the eight CMs depends on the chosen combination of the PP and HM. However, since we only have one data point per modeling chain combination, we cannot disentangle the interaction terms from residual errors (von Storch and Zwiers, 1999). Thus, the interaction term SSI should be interpreted with caution as it does not formally prove but only indicate contributions of interactions to the total ensemble uncertainty.

### 3.4 Results

#### 3.4.1 Validation

The validation is done for the whole CTL which partly overlaps with the two different calibration periods of the HMs (see Fig. 3.3. In addition to the CTL simulations, we
Figure 3.3: Nash Sutcliffe Efficiency (NSE, solid lines), 99% (dash-dotted lines) and 5% (dashed lines) runoff quantile levels (left-hand scale) series of the two hydrological models HBV and PREVAH at the gauge Diepoldsau. The values are calculated for 3-year periods shifted in yearly intervals throughout the whole simulation period. The thick lines indicate the performance of the CTL run from 1961-1990 with the first 3 years being cut off. The thin lines show the performance of the same model configurations in a longer simulation covering the period 1954-2006 (HBV) and 1951-2009 (PREVAH). The shaded areas depict the calibration period for each hydrological model. The evolution of the total hydropower reservoir volume is shown as a solid thick black line (right-hand scale).

Conducted two long simulations covering the period 1954 to 2006 (HBV) and 1951 to 2009 (PREVAH) with the same parameter setup of the HMs. These long simulations are meant to give some insight into the model performance over a longer period, during which also the build-up of the hydropower storage volume took place.

Within the CTL, the NSE (Nash and Sutcliffe, 1970) values of 3-year-periods shifted in yearly intervals vary in the range of 0.81 and 0.91. With regard to the 99% runoff quantile values, the model HBV follows closely the observations. PREVAH tends to overestimate the high runoff values. The 5% runoff quantile values are simulated quite well by HBV whereas PREVAH has a systematic underestimation. This is probably due to the influence of the hydropower operation (Margot et al., 1992). Besides diurnal and weekly hydropower production cycles, runoff is typically used to fill up the reservoirs during the snow-melt season (Verbunt et al., 2005). In winter, the low runoff is artificially increased by hydropower production. This effect is also shown in Fig. 3.3 in the time series of the observed
5% runoff quantiles. Compared to the period after the major part of the storage volume has been installed, say around 1974, they are lower than in the preceding period. The CTL DC panel in Fig. 3.6 correspond to the CTL simulations used for the validation as in the DC approach, the observed time series of the input variables are used to drive the HMs. The analysis of the mean annual cycle shows that HBV simulates a slightly delayed onset of the snow-melt season while PREVAH simulates the timing correctly. Thus, the validation results indicate that the two HMs, despite their structural similarities, have distinctly different characteristics. The long simulation runs further show that there is a high degree of variation in the model performance. Generally speaking, such variations are due to model simplifications, which lead to an imperfect representation of the reality. For example, parameter non-stationarity has been mentioned to lead to drifts in model performance over a longer period of time (Merz et al., 2011). Also, the varying influence of hydropower operation might lead to variation in the model performance.

Overall, the results indicate that the hydrological models perform reasonably well in the CTL. As we cannot be sure if a good performance during the past implies a good performance in a future climate (e.g. Blöschl and Montanari, 2010), we here use both models equivalently and try to answer the question how much of the uncertainty is due to the choice of the hydrological model. The effect of the model imperfections on the climate change signal clearly needs further systematic study and we refer to the references given above and to further studies such as Finger et al. (2012); Krahe et al. (2011); Lustenberger and Schädler (2011); Schaeffli et al. (2007a), which investigate in particular the role of hydropower in the climate change context.

3.4.2 Hydrological climate-impact projections

For the analysis, we discard the first 3 years in the 30 year time slices in order to prevent spin-up effects. Thus, we use the period 1964-1990 as CTL and the periods 2024-2050 and 2073-2099 as SCE1 and SCE2, respectively. We first discuss the two main input variables temperature and precipitation, followed by the two target variables which are the mean annual cycle of runoff and the runoff quantiles.

Temperature

Fig. 3.4 shows the results for the basin average temperature projections. In the CTL, there is a clear annual cycle with a maximum in summer of around 11 °C and a minimum in winter of around -6 °C. The differences between the two input data sets OBS\textsubscript{CHR} and OBS\textsubscript{ST} are smaller than 1 °C (see CTL DC panel). The spread of the BC runs in the CTL is smaller than 1 °C throughout the whole annual cycle. In the two SCEs, the pattern of the annual cycle remains the same but is shifted to higher temperatures. The HadCM3Q0
driven CMs are at the higher end of the ensemble range in both SCEs whereas SMHI-BCM-RCA is at the lower end. The temperature increase is in the range of 0.3 to 3.4 °C for SCE1 and in the range of 1.7 to 5.9 °C for SCE2. In the SCE1, there is no clear pattern in the annual cycle of the temperature change signal except for the HadCM3Q0 driven CMs that show a peak increase in winter and summer. In the SCE2, all CMs except DMI-ECHAM-HIRHAM and SMHI-BCM-RCA show a summer peak of the temperature increase.

Precipitation

Fig. 3.5 shows the annual cycle of the precipitation projections. The annual cycle for the DC runs in the CTL do not fully agree because the two HMs use slightly different precipitation datasets as input (see section 3.2.2). This applies also to HBV-BC and PREVAH-BC in the CTL since the bias-correction is done using the different precipitation reference data shown in the top-left panel. The bias-correction is able to correct the CM precipitation to match the observed mean annual cycle (see BC panels for the CTL). Note that the BC is carried out on a monthly basis. Thus, natural variability causes the 31d MA precipitation of the individual CMs to diverge between the monthly correction steps. Also, the bias-correction was made for the full period 1961-1990 while we here only analyze 1964-1990 which causes some further deviations between the CMs as can be seen in, e.g., June.

In the SCEs, the ensembles of the BC runs show a larger spread than the DC ensembles. The bias-corrected DMI-ECHAM-HIRHAM runs project considerably higher precipitation peaks in summer than the rest of the ensemble. Also, the BC runs driven by HadCM3Q0 show a clear peak around May and a minimum around July. In the climate change signal as well, the mentioned CMs show some deviations from the rest of the ensemble. DMI-ECHAM-HIRHAM projects the strongest increase in summer precipitation while the HadCM3Q0 driven CMs show the strongest increase in late spring. The DC runs generally show smaller variations in the annual cycle which is partly due to the spectral smoothing employed in the DC post-processing as opposed to the monthly based correction used in the BC. The climate change signals of the two SCEs are similar except for a stronger precipitation decrease in summer and a slightly stronger increase in the winter for the SCE2.

Annual cycle of runoff

Fig. 3.6 shows the annual cycles of simulated runoff in the CTL and the SCEs as well as the climate change signal. Annual and seasonal values are listed in Tab. 3.2. The results for the BC runs in the CTL show similar characteristics as the validation results which correspond to the DC runs in the CTL. For example, HBV-BC runs tend to simulate a
Table 3.2: Annual and seasonal changes in mean runoff expressed in differences and percentages. The ensemble mean is shown followed in brackets by the range of the 32 hydrological climate runoff scenarios. The first 3 years of each SCE were cut off in order to exclude spin-up effects from the analysis.

<table>
<thead>
<tr>
<th></th>
<th>SCE1 2024-2050</th>
<th></th>
<th>SCE2 2073-2099</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SCE-CTL [mm/d]</td>
<td>%</td>
<td>SCE-CTL [mm/d]</td>
<td>%</td>
</tr>
<tr>
<td>Annual</td>
<td>−0.06 [−0.19 - 0.18]</td>
<td>−2.04 [−5.96 - 6.11]</td>
<td>−0.15 [−0.39 - 0.21]</td>
<td>−4.75 [−12.72 - 7.06]</td>
</tr>
<tr>
<td>DJF</td>
<td>0.24 [0.11 - 0.51]</td>
<td>15.73 [5.48 - 40.68]</td>
<td>0.61 [0.09 - 1.02]</td>
<td>41.76 [4.78 - 81.70]</td>
</tr>
<tr>
<td>MAM</td>
<td>0.24 [−0.01 - 0.59]</td>
<td>8.58 [−0.39 - 23.93]</td>
<td>0.59 [0.29 - 0.90]</td>
<td>21.04 [10.12 - 37.66]</td>
</tr>
<tr>
<td>JJA</td>
<td>−0.77 [−1.46 - 0.20]</td>
<td>−14.64 [−26.03 - 3.86]</td>
<td>−1.70 [−2.40 - 0.42]</td>
<td>−32.15 [−45.49 - −8.10]</td>
</tr>
<tr>
<td>SON</td>
<td>0.05 [−0.19 - 0.29]</td>
<td>1.72 [−7.22 - 11.22]</td>
<td>−0.08 [−0.55 - 0.31]</td>
<td>−2.98 [−20.06 - 11.37]</td>
</tr>
</tbody>
</table>

delayed onset of the snow-melt period in May instead of in April, whereas PREVAH-BC runs all simulate a too low mean runoff in winter. Furthermore, the relative ordering of the CMs is almost the same in the HBV-BC and PREVAH-BC simulations. For example, the HadCM3Q0 driven CMs both are at the upper end of the ensemble in the winter months, but at the lower end in summer. In the SCE1, the peak in the annual cycle occurs about half a month earlier compared to the CTL for all model combinations except for the runs using the CM DMI-ECHAM-HIRHAM which projects a later peak and higher runoff. In the SCE2, the runoff peak occurs about one month earlier and has a smaller amplitude in comparison to the CTL. In the climate change signal, all model combinations agree on an increase of winter runoff. All PP-HM combinations (except those driven by DMI-ECHAM-HIRHAM for SCE1) show a decrease of summer runoff. These signals are stronger for SCE2 than for SCE1. Most BC runs project a peak in the increase of runoff in the snowmelt season around May. The model spread is highest in summer. The climate change signals of the DC runs are smoother than the ones of the BC runs which is due to spectral smoothing in the DC as opposed to monthly correction intervals in the BC.

**Runoff quantiles**

The runoff quantiles in the CTL DC simulations show the performance of HMs. PREVAH underestimates low runoff quantiles but overestimates high runoff quantiles. HBV matches the observed low and high runoff quantiles better, but slightly underestimates the intermediate quantiles. In the BC results, larger deviations from the observations are apparent in the high quantile range. This is due to the bias-correction that corrects for biases in the mean but not for biases in the whole quantile distribution. In the climate change signal for SCE1, all the DC runs project slight decreases of runoff in the medium and moderate high quantile range and most DC model combinations agree on an increase of the 99.9 % runoff
quantile level. For SCE2, the decrease in the moderate to high quantile ranges is larger and some model chains simulate a decrease of the 99.9% runoff quantile level. The BC runs also show slight decreases of runoff in the medium to moderate high quantile range but for quantiles above 90%, the BC runs disagree on the sign of the change. Furthermore, the spread of projected changes in the high quantile range is similar in both SCEs.

### 3.4.3 Variance decomposition

As described in section 3.3.3, we analyze the contribution of the three sources CM, PP and HM to the total ensemble uncertainty of the hydrological climate projections. In principle, such a variance decomposition is feasible for any variable of interest. We conducted the analysis for differences in the annual cycle of mean runoff (see bottom row panels in Fig. 3.6) and for runoff quantiles (see bottom row panels in Fig. 3.7).

#### Changes in the mean annual cycle of runoff

Figure 3.8 shows the variance decomposition of the climate change signal in the mean annual cycle of runoff. Both the DVD and the ANOVA approach determine the CM to be the dominant source during summer and autumn. In May, the variance explained by the PP shows a peak which is clearer in SCE2 than in SCE1. The high contribution of PP is apparent neither in the variance decomposition of precipitation nor of temperature (analysis not shown). Thus, we suspect that the differences between the PP methods build up during the snow-accumulation period (e.g. through differences in the covariance between temperature and precipitation) and take effect only during the melt period. The HM is the least important individual source of uncertainty in SCE1 but it becomes the dominant source of uncertainty in winter and early spring of SCE2. This indicates that the HMs are indifferent to the projected temperature and precipitation changes for SCE1, but react differently to the projected changes for SCE2. PREVAH seems to respond more sensitively to the large temperature changes as projected precipitation changes are mostly below +0.5 mm/d during winter for SCE2 but PREVAH simulates a runoff increase of close to +1 mm/d whereas HBV projects runoff changes just slightly above +0.5 mm/d. Using the ANOVA approach, we can also quantify the interaction term. It varies between approximately 0.1 and 0.4. The high variance fraction around April might be due to strong non-linearities in the snow-melt process.

#### Changes in runoff quantiles

Figure 3.9 depicts the variance decomposition of the changes in different runoff quantiles using the DVD and the ANOVA approach. The results of the two methods show the same
3.4 Results

Figure 3.4: Mean annual cycle of basin-averaged temperature in the CTL period (top), the two SCE periods (middle) and climate change signal as the differences SCE-CTL (bottom panel). The coloring of the CMs is the same for all PP-HM combination panels. The PP-HM combination is indicated in the title of each panel. For the DC runs, there is only one CTL run per HM and thus, the HBV and PREVAH runs are combined in the top left panel. The annual cycle has been low-pass filtered using a 31d MA.
Figure 3.5: The same as in Fig. 3.4 but for water balance corrected precipitation.
Figure 3.6: The same as in Fig. 3.4 but for mean runoff. The observed mean runoff is added to the CTL period panels.
Figure 3.7: The same as in Fig. 3.6 but for different runoff quantiles.
Figure 3.8: Variance decomposition of the uncertainty in changes of the mean runoff for the three modeling chain elements CM, PP and HM in the course of the annual cycle. The fraction of variances are computed by using the DVD approach (left panel) and the ANOVA approach (right panel). The ANOVA panel also contains the interaction term.

pattern but the DVD approach attributes slightly higher variance fractions to the PP and HM than the ANOVA approach does. For the analysis, we divide the quantile bins in a low (5 to 35%), an intermediate (35 to 80%) and a high range (80 to 99.9%). The low range is closely connected to the runoff in the winter months. The high runoff quantile range is related to the runoff in the summer months. Runoff values in the intermediate quantile range occur most often in the snow-melt season as well as in late summer and autumn. This quantile range is though not as clearly linked to a particular season in the year as the other two quantile ranges.

In the low quantile range, the contributions of the different uncertainty sources to the total ensemble uncertainty change considerably from SCE1 to SCE2. While in SCE1 the CMs and the interactions are the dominant sources of uncertainty, it is the HMs that explain about 50% of the total ensemble uncertainty in SCE2. Please see the discussion of the winter period in section 3.4.3 for an explanation. In the high quantile range, the CMs and PP methods are the most important variance sources in both SCEs. Weather events leading to such high runoff values are linked to the precipitation frequency and intensity, both of which are determined by the CM and PP. The interpretation of the results for the intermediate quantile range is not straight-forward. As this quantile range is not clearly determined by a particular process or season in the year, the results reflect the time-varying contribution of the different uncertainty sources that we see in Fig. 3.8 as well.
3.5 Summary and conclusions

We investigated hydrological climate projections based on eight different climate models, two statistical post-processing methods and two hydrological models. An Alpine study region was chosen as a challenging test area for all the three modeling chain elements. For the two scenario periods 2021-2050 and 2070-2099 with respect to 1961-1990, we estimated the total ensemble uncertainty of the hydrological projections and quantified the contributions of different uncertainty sources.

For both scenario periods, all modeling chain combinations show an increase in winter runoff and most of the model combinations project a decrease of summer runoff. The resulting spread is highest in summer and for the high (99 and 99.9 %) runoff quantiles. For such high quantiles, the hydrological projections disagree on the sign of the changes. We used the direct variance decomposition (DVD) and the ANOVA approach to decompose the total ensemble uncertainty into contributions from single impact modeling chain elements. Both approaches identify the CMs to be the dominant source of uncertainty during summer and autumn for both SCEs. This fully agrees with the results of previous studies (Wilby and Harris, 2006; Jasper et al., 2004; Graham et al., 2007; Kay et al., 2009; Prudhomme and Davies, 2009a; Jung et al., 2011). For the far-term SCE period in winter and spring, however, the dominance of the CMs diminishes and the PP methods and HMs become more important than the CMs. The analysis of the changes in different runoff quantiles revealed that the HMs become the dominant uncertainty source for changes in the low runoff quantiles for the far-term SCE period. For high runoff quantiles, both the PP methods and the CMs contribute between 30 and 60 % to the total ensemble uncertainty. Furthermore, the ANOVA approach indicates that interactions between the uncertainty sources become more important.

Figure 3.9: Same as Fig. 3.8 but for variance decomposition of changes in different runoff quantiles.
sources are important as well. The presence of the interactions and the time varying contributions of the different uncertainty sources confirm the need for future impact modeling studies to include multiple uncertainty sources and to test all possible impact modeling chain combinations in order to aim for a more comprehensive representation of the uncertainty.

This study presents a methodological framework for the attribution of the total ensemble uncertainty in hydrological climate-impact projections to individual uncertainty sources. It was embedded in the AdaptAlp project (Adaptation to Climate Change in the Alpine Space; (Korck et al., 2011), see www.adaptalp.org) which aimed to develop adaptation practice strategies based on scientific knowledge and risk management assessments. The results are based on a modeling experiment using three different uncertainty sources and a limited number of CMs, PP methods and HMs. Thus, the resulting total ensemble uncertainty in the hydrological climate-impact projections is likely underestimated. In particular, we neglect the uncertainty due to the emission scenario (Prein et al., 2011) and the hydrological model parameters (Beven and Binley, 1992). Further research is required to include more uncertainty sources and to employ more diverse PP methods and HMs. The results are based on a modeling ensemble of opportunity. All elements of the modeling chain (i.e. CM, PP and HM) are combined in all possible combinations, but it should be noted that the two employed PP methods have been adapted to the specific requirements of the two HMs, and thus the PP methods are likely not completely independent from the HMs. Clearly, the results depend on the study region and setup and therefore do not allow for a generalization to other study areas or other impact modeling chains. However, knowledge about the contribution of different uncertainty sources to the total ensemble uncertainty may help to design future impact modeling studies. While the results as such are not transferable to catchments with different hydrological properties, the presented methods for quantifying different uncertainty sources are versatile and adaptable to other experiment designs and study regions.

3.6 Acknowledgement

The ENSEMBLES data used in this work was funded by the EU FP6 Integrated Project ENSEMBLES (Contract number 505539) whose support is gratefully acknowledged. We are also thankful for the RheinBlick2050, the KLIWAS, the AdaptAlp and the CCHydro projects that fostered collaboration. The observational data were provided by MeteoSwiss, the Swiss Federal Office of Meteorology and Climatology, the Swiss Federal Office for the Environment and the International Commission for the Hydrology of the Rhine Basin. We would also like to thank the Seminar for Statistics of the ETH Zurich for its assistance.
4 Hydrological climate-impact projections for the Rhine river: GCM-RCM uncertainty and separate temperature and precipitation effects

Thomas Bosshard\textsuperscript{1}, Sven Kotlarski\textsuperscript{1}, Massimiliano Zappa\textsuperscript{2}, Christoph Schär\textsuperscript{1}

\textsuperscript{1}Institute for Atmospheric and Climate Science, ETH Zurich, Switzerland
\textsuperscript{2}Swiss Federal Institute for Forest, Snow and Landscape Research WSL, Zürcherstrasse 111, 8903 Birmensdorf, Switzerland

\textit{(Submitted to Journal of Hydrometeorology)}

Abstract

Climate change is expected to affect the hydrological cycle with significant impacts on water resources. Climate induced changes in the hydrology of the Rhine river (Europe) are of potentially major importance for the riparian countries, as the Rhine river is the most important European waterway, serves as source of freshwater water supply, and is prone to floods and droughts. Here, we use regional climate model data from the ENSEMBLES project to drive the hydrological model PREVAH and to assess the impact of future climate change on the hydrology in the Rhine basin. Results suggest increases in monthly mean runoff during winter, and decreases in summer. At the gauge Cologne and for the period 2070-2099 under the SRES A1B scenario, projected decreases in mean summer runoff vary between -9 to -40\% depending on the climate model used, while increases in winter are in the range of +4 to +51\%. These projected changes in runoff are generally consistent with earlier studies, but the derived spread in the runoff projections appears to be larger. It is also shown that temperature and precipitation effects upon runoff changes are approximately additive for a large part of the Rhine basin. Analysis demonstrates that temperature (i.e. snow-melt) effects dominate in the Alpine tributaries and propagate downstream along the Rhine, while precipitation effects dominate in the lower Rhine tributaries. Analyses are also presented for selected runoff extreme indices.
4 Hydrological climate-impact projections for the Rhine river

4.1 Introduction

There is a broad consensus that future global climate change will intensify the global hydrological cycle, although the magnitude and regional details of the projected change often remain unclear (Allen and Ingram, 2002; Wild et al., 2008; Meehl et al., 2007). On the regional scale, the changes in the hydrological cycle can considerably deviate from the global pattern, due to changes in atmospheric circulation and moisture supply, as well as due to local factors related to the topography and hydrological characteristics of the catchment considered. In snowmelt-dominated basins, for instance, the expected temperature increase will impact on snow accumulation and melt dynamics and can cause important changes of the annual discharge regime (Barnett et al., 2005).

The Rhine catchment is located in central Europe. The elevation of the catchment ranges up to 4274 m a.s.l. The Rhine river is extensively used for water supply, as a waterway and for energy production. It is also prone to large-scale floods in winter and early spring (Disse and Engel, 2001). The Rhine basin is expected to experience climate change induced precipitation decreases in summer and increases in winter. Temperature is expected to increase faster than the global average (Meehl et al., 2007). Given the diversity of potential climate change impacts and the importance of the Rhine, regional climate-impact studies are of high relevance for water management, navigation and several other sectors.

Numerous studies have investigated climate induced changes of the hydrology in the Rhine basin. Kwadijk and Middelkoop (1994) conducted a climate sensitivity study and found that within the tested scenario range, temperature changes have a smaller effect on changes in peak discharge than precipitation changes. Middelkoop et al. (2001) used monthly changes of temperature and precipitation as simulated by two general circulation models (GCMs) and applied it to a suite of hydrological models in the Rhine catchment. The results for the whole Rhine basin were a decrease of runoff in summer and an increase of winter runoff. They concluded that the flood and drought risks both increase and that projected changes, despite their uncertainties, are so large that they should be considered in water management planning. Shabalova et al. (2003) were to our knowledge the first ones to use regional climate model (RCM) data, i.e. dynamically downscaled GCM scenarios, for a hydrological impact study in the Rhine catchment. They used two methods to statistically post-process the RCM data. The first one accounted for changes in the mean whereas the second also accounted for changes in the variability. They found the same change pattern of mean runoff as previous studies. Furthermore, they showed that both post-processing methods agree on an increase of runoff extremes but that the magnitude of the changes is sensitive to the choice of the post-processing method. Kleinn et al. (2005) evaluated the performance of a hydrological model driven by dynamically downscaled reanalysis data and found that runoff biases are particularly large in Alpine catchments. Menzel et al. (2006) used a non-linear statistical downscaling of two GCMs to study changes in runoff extremes. They found that the associated uncertainties are too
large to allow for drawing any conclusions. Graham et al. (2007) used climate scenarios of the PRUDENCE project and found for the end of the 21st century a decrease of monthly mean runoff in summer and autumn in the range of -20 to -40% and an increase in late winter and early spring in the range of +10 to +20% (with one outlier simulating +40% in February). Hurkmans et al. (2010) investigated hydrological changes in the Rhine basin induced by three dynamically downscaled high resolution climate scenarios assuming different emission scenarios. They found that peak flow increases over the whole 21st century but drought magnitudes only increase in the second half of the 21st century. Previous work also found runoff changes in snow-dominated basins to be stronger affected by changes in temperature than in precipitation (Barnett et al., 2005). Nijssen et al. (2001) conducted a climate-impact study in macroscale river basins around the world and found that transitional catchments (e.g. catchments with average winter temperatures just below the freezing temperature) will experience a decrease of the snowmelt peak and an increase of winter and a decrease of summer runoff. They also conducted a sensitivity study and applied seasonal changes of temperature and precipitation separately to the modelled catchments. Temperature increases generally lead to a runoff decrease but snow-dominated catchments experienced a shift in the seasonality of runoff. Precipitation increases lead to increases of runoff in all catchments but in snow-dominated catchments, additional precipitation during winter was first stored in the snow pack and became effective in spring and summer. For the Alpine part of the Rhine basin, Graham et al. (2007) analyzed the snow cover duration as a function of mean winter temperature and precipitation and found that projected changes according to the PRUDENCE data will lead to a decrease of snow cover duration by about 75-100 days. Also, the snowmelt peak in spring occurs one month earlier by the end of the 21st century. At the gauge Maxau, located along the Upper Rhine, Hurkmans et al. (2010) found the snowmelt season to start about 6-8 weeks earlier in the second half of the 21st century. Recently, the Rheinblick2050 project (Görgen et al., 2010) compiled a database of hydrological climate-runoff projections based on studies of the riparian countries of the Rhine basin and various research institutions. A common research framework facilitated the comparison among the projections. Most of the hydrological climate-runoff projections were driven by RCM data provided by the ENSEMBLES project (van der Linden and Mitchell, 2009).

In our study, we aim to assess climate induced changes in the hydrology of the Rhine river for the two scenario periods (SCE) 2021-2050 and 2070-2099 compared to the control period (CTL) 1980-2009. We use an ensemble of the latest available regional climate model projections from the ENSEMBLES project, an improved version of the delta change approach (Gleick, 1986; Bosshard et al., 2011) and the hydrological model PREVAH which has been extensively evaluated in the Alpine area (Gurtz et al., 1999; Zappa and Gurtz, 2003; Viviroli et al., 2009). We will also quantitatively assess the contributions of temperature and precipitation changes to the projected changes in runoff.
Figure 4.1: Map of the Rhine basin down to Cologne. The sub-basins used by the hydrological model PREVAH are indicated by thin black lines.

The paper is organized as follows: Section 4.2 presents the study area and the data, section 4.3 introduces the methods of the hydrological impact modeling study followed by section 4.4 that shows the results. In section 4.5, we summarise the results and draw conclusions.

## 4.2 Study area and data

### 4.2.1 Study area

The study area encompasses the whole Rhine catchment down to the gauge Cologne (see Fig. 4.1). The total catchment area amounts to 144,231 km². The catchment covers elevations ranging from 35 m a.s.l. at Cologne in the North to 4274 m a.s.l. at the Finsteraarhorn located in Alpine region in the South. Major tributaries are the Neckar, Main and Moselle rivers. Downstream from Basel (close to Rheinfelden), the Rhine is heavily used as a waterway. The population in the whole catchment amounts to about 50 millions (Disse and Engel, 2001).
4.2.2 Data

For the setup of the hydrological model in the area downstream of Rheinfelden, we use soil and land use data from the European Soil Database (ESDB v2.0, 1x1 km²) and a high resolution digital elevation model (DEM) with a resolution of 75x75 m² provided by the International Commission for the Hydrology of the Rhine Basin (CHR). For the Alpine and High Rhine area, we employ the hydrological model setup by Verbunt et al. (2006) (see section 4.3.1). The meteorological station data to force the hydrological model are provided by MeteoSwiss (Switzerland), the German Weather Service (Germany), Météo France (France), the Federal Ministry of Agriculture, Forestry, Environment and Water Management (Austria) and the European Climate Assessment & Dataset Project (Luxembourg). Figure 4.1 shows the locations of the precipitation stations. For Germany, a new extensive precipitation data set compiled by the German Weather Service is used (see Zolina et al. (2008)). Since temperature is spatially less variable than precipitation, we use the grid-ded temperature data set of the ENSEMBLES project (Haylock et al., 2008) instead of station data for the area downstream of Rheinfelden, while for the Alpine area upstream of Rheinfelden, we use station data. All the meteorological data sets cover the whole CTL period 1979-2008. Daily runoff data for calibration and validation are provided by the responsible national and federal state agencies of the riparian countries.

As climate model data, we use data of 10 climate modeling chains provided by the ENSEMBLES project (van der Linden and Mitchell, 2009). Each modeling chain consists of a GCM which is dynamically downscaled by a RCM (GCM-RCM) (see Tab. 4.1). The 10 GCM-RCMs assume the A1B emission scenario, cover at least the period 1961-2099 and have a horizontal resolution of about 25 km.

4.3 Methods

4.3.1 Hydrological model

In this study, we adopt the hydrological model PREVAH which has been widely used for hydrological studies in Switzerland (see e.g. Gurtz et al., 1999; Zappa et al., 2003; Verbunt et al., 2007; Jaun et al., 2008; Rotach et al., 2009; Zappa et al., 2011). PREVAH is a semi-distributed conceptual model based on hydrological response units. In our setup, it requires daily meteorological input data of precipitation, temperature, sunshine duration, relative humidity and wind speed. The runoff generation module contains an upper water storage that is non-linearly filled by infiltration and emptied by direct runoff, interflow and percolation to the second lower subsurface storage. The second lower subsurface storage produces baseflow. All the runoff generation processes are governed by storage time
**Table 4.1**: List of the employed climate model chains from the ENSEMBLES project. The naming follows the following pattern: Institution that provided the data - GCM - RCM.

<table>
<thead>
<tr>
<th>Institution</th>
<th>GCM</th>
<th>RCM</th>
</tr>
</thead>
<tbody>
<tr>
<td>DMI</td>
<td>ECHAM5</td>
<td>HIRHAM</td>
</tr>
<tr>
<td>ICTP</td>
<td>ECHAM5</td>
<td>REGCM</td>
</tr>
<tr>
<td>KNMI</td>
<td>ECHAM5</td>
<td>RACMO</td>
</tr>
<tr>
<td>MPI</td>
<td>ECHAM5</td>
<td>REMO</td>
</tr>
<tr>
<td>SMHI</td>
<td>ECHAM5</td>
<td>RCA</td>
</tr>
<tr>
<td>ETHZ</td>
<td>HadCM3Q0</td>
<td>CLM</td>
</tr>
<tr>
<td>HC</td>
<td>HadCM3Q0</td>
<td>HadRM3Q0</td>
</tr>
<tr>
<td>SMHI</td>
<td>HadCM3Q3</td>
<td>RCA</td>
</tr>
<tr>
<td>CNRM</td>
<td>ARPEGE</td>
<td>ALADIN</td>
</tr>
<tr>
<td>SMHI</td>
<td>BCM</td>
<td>RCA</td>
</tr>
</tbody>
</table>

constants (Zappa and Gurtz, 2003). PREVAH also includes snow and glacier modules. Melt processes are parameterized by an extended degree-day approach including an adjustment for incoming radiation (Zappa et al., 2003). Evapotranspiration is estimated by the Penman-Monteith equation (Penman, 1956) and radiation is parameterized via sunshine duration according to Schulla (1997). A precipitation adjustment factor is used in order to close the water balance. The routing of channel runoff is represented by a storage-translation model. This routing model does not include the effect of spill-over during flood events.

For the area upstream of the gauge Rheinfelden at the German-Swiss border, we use the setup of Verbunt et al. (2006) with a 500x500 m$^2$ spatial resolution. Downstream of Rheinfelden, we use a 75x75 m$^2$ DEM and 1x1 km$^2$ gridded data of soil type, soil depth and land use to set up PREVAH with a basic resolution of 400x400 m$^2$. We calibrated the model in the 6-year period 1985-1990 and validated it in the remaining 24 years of the CTL period 1979-2008. A total of 14 parameters were calibrated for each sub-basin separately using daily runoff data of 76 gauges. The calibration algorithm PEST (Doherty et al., 2005) was implemented to estimate an optimal parameter set. As objective function, we used the sum of the squares of the differences between modelled and observed daily runoff values with a weight of 0.2 plus the sum of the squares between the observed and modelled mean annual cycle estimated by a 31 day moving average (31d MA) with a weight of 1. Due to the 6-year long calibration, daily runoff values and the mean annual cycle have approximately the same integrated weight in the total sum of the squares.
4.3 Methods

4.3.2 Statistical post-processing of climate model data and coupling with PREVAH

We use the delta change approach to statistically post-process the climate model data and couple it with the hydrological model PREVAH. The delta change approach, although having the well-known caveat of not allowing for changes in variability, is widely used in hydrological climate-impact studies (Gleick, 1986; Arnell, 1992; Lenderink et al., 2007; Bosshard et al., 2011). In the delta change approach, observed meteorological data are scaled according to a climate change signal.

Here, we apply the delta change post-processing only to temperature \( T \) and precipitation \( P \) data. For the other meteorological variables required by PREVAH, we use the unscaled observed time series. Following Bosshard et al. (2011), we first interpolate the RCM data to the observational station sites using inverse distance weighting. At each site \( i \) and for every RCM \( j \), we estimate the mean annual cycle \( X(d) \) in the CTL and SCE period by use of a spectral filter and derived the climate change signal \( \Delta X \) according to

\[
\Delta T_{i,j}(d) = \frac{\bar{T}_{i,j}^{\text{SCE}}(d)}{\bar{T}_{i,j}^{\text{CTL}}(d)} - \bar{T}_{i,j}^{\text{CTL}}(d)
\]

\[
\Delta P_{i,j}(d) = \frac{\bar{P}_{i,j}^{\text{SCE}}(d)}{\bar{P}_{i,j}^{\text{CTL}}(d)} - \bar{P}_{i,j}^{\text{CTL}}(d)
\]

where \( d \) stands for a day in the annual cycle and the superscripts indicate the period (either CTL or SCE). The observed time series at the site are then scaled according to

\[
T_{i,j}^{\text{SCE}}(y,d) = T_{i,\text{obs}}^{\text{CTL}}(y,d) + \Delta T_{i,j}(d)
\]

\[
P_{i,j}^{\text{SCE}}(y,d) = P_{i,\text{obs}}^{\text{CTL}}(y,d) \cdot \Delta P_{i,j}(d).
\]

The variable \( y \) denotes the years in the CTL or SCE period and observed time series are indicated by the subscript \( \text{obs} \).

The spectral filter is applied in order to represent the annual cycle of the climate change signal with daily resolution and to reduce the effect of natural variability. Bosshard et al. (2011) found that the commonly used averaging on a monthly basis produces artificial peaks in the annual cycle of the climate change signal. The spectral filter removes such peaks effectively.

4.3.3 Separating the temperature and precipitation effect

A set of 4 simulations was conducted in which all 4 possible combinations of unscaled or scaled temperature and precipitation data were tested:
4.4 Results

4.4.1 Validation

For model validation, we only show the validation results for selected gauges along the Rhine river and at major tributaries (see Fig. 4.1). Figure 4.2 shows the validation of the modelled mean annual runoff cycle. The runoff regime gradually changes from nival...
at Diepoldsau with a peak of mean runoff in summer, to pluvial at Cologne with a peak in early spring. The model slightly underestimates mean runoff in spring and summer. In autumn, the model shows a systematic positive bias. A comparison with validation results of previous studies showed that the large relative biases of PREVAH during short periods in the year are within the range of other modeling studies in the Rhine basin (Shabalova et al., 2003; Kleinn et al., 2005; Lenderink et al., 2007). The volume error for the mean annual runoff at Cologne amounts to 0.8 % in the calibration and to 1.0 % in the validation period. The systematic bias in autumn is likely associated with inaccuracies in the seasonal evolution of terrestrial water storage. It could be explained by an overestimation of soil moisture storage or an underestimation of evapotranspiration. At the gauge Diepoldsau, winter runoff is strongly underestimated. This is at least partly due to the influence of hydropower production that artificially increases the low runoff in winter. PREVAH does not include a hydropower module and therefore fails to capture the anthropogenic influence on winter low flow. Hence, estimated relative changes of low flow indices at Diepoldsau have thus to be interpreted with large caution.

The validation using the Nash-Sutcliffe Efficiency (NSE; Nash and Sutcliffe, 1970) shows increasing performance with the upstream area (Fig. 4.3). The NSE value in the validation period is generally lower than in the calibration period. The difference decreases with the area upstream of the gauge. Also, the spread of yearly NSE values decreases the further downstream the gauge lies. The smaller the spread, the less sensitive is the model performance to the interannual variations in the CTL period, and the more likely the model also performs well in a different climate.
4 Hydrological climate-impact projections for the Rhine river

4.4.2 Spatial pattern of seasonal changes

Here, we only show the spatial patterns of seasonal changes for five out of ten impact modeling chains. These five modeling chains are driven by the five different GCMs used in the GCM-RCM ensemble. We show results for changes in temperature ($T$), precipitation ($P$), evapotranspiration ($E$) and runoff ($R$) for summer and winter of the later SCE period 2070-2099. A complete suite of results for the whole model ensemble, both SCE periods and all the seasons is compiled in the supplementary material.

Temperature

Figure 4.4 shows the seasonal changes of sub-basin-averaged $T$ for winter and summer. All GCM-RCM model chains agree on a temperature increase in both seasons. The magnitude of the increase though differs within the ensemble. In winter, the ensemble range is smaller and the HadCM3Q0 and ECHAM5-driven chains roughly agree on the magnitude. There are however some exceptions. SMHI-HadCM3Q3-RCA uses the low-sensitivity run of the HadCM3Qx GCM-runs. A low sensitivity signifies that on a global scale, the model projects lower temperature increases for a doubling of CO$_2$ than the other HadCM3Qx versions. Nevertheless, for the area of the Rhine catchment, SMHI-HadCM3Q3-RCA projects the strongest increase of winter temperature in the whole ensemble. CNRM-ARPEGE-ALADIN shows the weakest temperature increase in the ensemble.

In summer, the HadCM3Q0 driven model chains as well as CNRM-ARPEGE-ALADIN project a substantially stronger temperature increase than e.g. most of the ECHAM5 driven chains. With CNRM-ARPEGE-ALADIN projecting a low temperature increase in winter, this model chain thus has the largest amplitude in the annual cycle of temperature changes. SMHI-BCM-RCA, SMHI-HadCM3Q3-RCA and DMI-ECHAM5-HIRHAM (see supplementary material) are at the lower bound of the projected temperature increase. In summer, all model chains exhibit a North-South gradient in the temperature change. This gradient is probably connected to both the continental-scale pattern of warming (stronger summer warming in Southern Europe) and the amplified warming at high elevations (Kotlarski et al., 2011).

Precipitation

Figure 4.5 shows the relative seasonal changes of sub-basin-averaged $P$ for winter and summer. In winter, all model chains except for CNRM-ARPEGE-ALADIN agree on an increase in $P$ for almost the whole study area apart from some high alpine catchments that inherit the $P$ change signal from the Southern Alpine region. The HadCM3Q0-driven model chains have a similar $P$ change pattern. Among the ECHAM5 driven model chains, there are substantial differences in the spatial pattern of precipitation changes (see sup-
4.4 Results

Figure 4.4: Change of winter (top) and summer (bottom) temperature ($T$) in the sub-basins for the SCE-period 2070-2099 vs. CTL-period 1979-2008 [$^\circ$C] as projected by 5 representative GCM-RCMs modeling chains, each of them having a different driving GCM.

Figure 4.5: Change of winter (top) and summer (bottom) precipitation ($P$) in the sub-basins for the SCE-period 2070-2099 vs. CTL-period 1979-2008 [%] as projected by 5 representative GCM-RCMs modeling chains, each of them having a different driving GCM.
Figure 4.6: Change of winter (top) and summer (bottom) runoff ($R$) in the sub-basins for the SCE-period 2070-2099 vs. CTL-period 1979-2008 [%] as projected by 5 representative hydrological climate-impact modeling chains, each of them driven by a different driving GCM.

Figure 4.7: Change of winter (top) and summer (bottom) evapotranspiration ($E$) in the sub-basins for the SCE-period 2070-2099 vs. CTL-period 1979-2008 [%] as projected by 5 representative hydrological climate-impact modeling chains, each of them driven by a different driving GCM.
4.4 Results

DMI-ECHAM5-HIRHAM projects smaller increases of $P$ than the rest of the ECHAM5-driven model chains while KNMI-ECHAM5-RACMO and SMHI-ECHAM5-RCA both are at the upper bound of the ECHAM5-driven GCM-RCMs. The three GCM-RCMs provided by SMHI using the RCM RCA show distinctly different patterns of changes in $P$. This provides further evidence that the driving GCM can have a strong effect on the dynamically downscaled climate change signal.

In summer, most model chains agree on a precipitation decrease. However, the spatial pattern is more heterogeneous than for temperature. DMI-ECHAM5-HIRHAM (see supplementary material) as well as SMHI-BCM-RCA show slight increases of summer precipitation in parts of the catchment. These two model chains also project temperature changes at the lower end of the ensemble.

Runoff

Figure 4.6 shows the relative seasonal changes of sub-basin-averaged $R$ for winter and summer. For winter, most impact-modeling chains agree on an increase of runoff. The Alpine sub-basins generally show larger increases than the other sub-basins, despite lower precipitation increases. This is related to the temperature increase which causes more of the winter precipitation to become runoff effective in winter instead of being stored in the snow-pack until the melt season. CNRM-ARPEGE-ALADIN is the only modeling chain that projects runoff decreases in winter in a large fraction of the Rhine basin. This is consistent with its precipitation change pattern.

For summer, the spatial patterns of changes in $R$ deviate substantially between the different GCMs. The HadCM3Q0-driven chains and CNRM-ARPEGE-ALADIN project large runoff decreases, consistent with the pronounced decrease of summer precipitation. The ECHAM5-driven chains also project runoff decreases in the Alpine region, along the Upper and Middle Rhine and parts of the Moselle catchment in the North-Western part of the basin. In the Main and Neckar catchments, however, the ECHAM5-driven chains project runoff increases, despite decreasing precipitation in this region. This pattern can probably be explained by groundwater processes which are of major importance particularly in the Main catchment (CHR/KHR, 1978): The ECHAM5-driven chains project larger precipitation increases in winter and spring than the HadCM3Q0-driven chains and CNRM-ARPEGE-ALADIN. This additional surplus of winter and spring precipitation refills the soil water storage. The elevated soil water storage levels last until summer and cause increasing summer runoff by increased baseflow. The same also applies to SMHI-BCM-RCA and SMHI-HadCM3Q0-RCA, accompanied by less pronounced decreases of summer precipitation than projected by the ECHAM5-driven chains.
Evapotranspiration

Figure 4.7 shows the seasonal changes of basin-averaged $E$ for winter and summer. In winter, all model chains simulate an increase of evapotranspiration, though relative to a low reference value. The spatial patterns are very similar among the model ensemble. In summer and in the Southern part of the basin, the majority of the ensemble agrees on a decrease in $E$. Thus, although temperature increases suggest increasing potential evapotranspiration, the real $E$ as estimated by PREVAH is moisture-limited and decreases over large areas of the Rhine basin. Decreasing $E$ over parts of the Rhine basin was also found by Hurkmans et al. (2010) who employed a physically-based hydrological model.

4.4.3 Changes in the annual cycle of water balance components

At selected gauges in the Rhine basin, we show the annual cycles of the water balance quantities $P$, $E$ and $R$ in the CTL period (Fig. 4.8) and their changes for the SCE periods 2021-2050 and 2070-2099 with respect to CTL (Fig. 4.9). For results at additional gauges,
4.4 Results

please see the supplementary material. Figure 4.8 shows how the relative fraction of the water balance quantities changes within the basin and throughout the annual cycle. At gauges close to the Alps, $P$ and $R$ are always larger than $E$. Further downstream and in the tributaries, $E$ almost balances $P$ in summer which causes the summer low flow season. In Fig. 4.9, results are shown for the 10 different driving GCM-RCM chains, and the coloring of the hydrological impact scenarios is grouped according to the driving GCM. Red to yellow colors stand for impact modeling chains driven by ECHAM5, blueish colors represent the HadCM3Qx family whereas the greenish colors show two modeling chains that use either ARPEGE or BCM as driving GCM. The figures show the spatial integral over the whole area upstream of the respective gauge. Thus, for the gauges along the Rhine river (left column in Fig. 4.9), the change signals resemble each other since their upstream areas partly overlap.

For precipitation and the SCE period 2021-2050, none of the gauges shows a clear tendency. The change signal of $P$ is slightly clearer for the later SCE period. Most models agree on a decrease of $P$ in summer at all gauges. The change signal varies considerably between the GCM families. At all the gauges in the tributaries (right column in Fig. 4.9), the HadCM3Q0-driven modeling chains and CNRM-ARPEGE-ALADIN show a more pronounced and prolonged decrease of $P$ in summer than the rest of the ensemble. This also applies to the furthest downstream gauge at Cologne where the maximal decrease varies between -5 and -40%. Also, there is a broad agreement among the ensemble for an increase of $P$ in the rest of the year with the largest increase generally occurring in late autumn and early winter.

For evapotranspiration, the clear tendencies for the SCE period 2070-2099 are already visible in the first SCE period. All modeling chains agree on an increase of $E$ in winter at all gauges. In summer of the second SCE period, all gauges except for Frankfurt a.M. show a decrease of $E$. In the Alpine area, the decrease is smaller than in the other areas. The impact-modeling chains showing the largest $P$ changes (HadCM3Q0-driven chains and CNRM-ARPEGE-ALADIN) also show the largest decreases of $E$.

For runoff, there are some tendencies visible in the first SCE period at the gauges along the Rhine river, but not at the gauges of the tributaries. At the Rhine gauges, $R$ is projected to increase in winter and decrease in summer. This is the typical pattern of change found in snow-dominated catchments (Barnett et al., 2005). It indicates that the temperature effects are already visible in the first SCE period. For the second SCE period, these tendencies become clearer at the gauges along the Rhine. In the tributaries, however, there is still no clear tendency but the ensemble is split up into two groups. The first group consists of the ECHAM5-driven modeling chains, the SMHI-HadCM3Q3-RCA and the SMHI-BCM-RCA. They generally project increases of $R$ during the whole year. The second group, consisting of the HadCM3Q0-driven modeling chains plus the CNRM-ARPEGE-ALADIN, projects decreases of $R$ during most of the annual cycle except for late autumn and early winter. These modeling chains project large decreases of $P$ in the
Figure 4.10 shows the climate induced changes of the runoff regime along the Rhine river in absolute terms. While the relative changes of Fig. 4.9, they act on different regimes. At gauges in the Alpine Rhine, the summer decrease takes place in a period of high mean flows whereas the winter increase of $R$ coincides with the period of low flows. Thus, the amplitude of the mean annual cycle of $R$ decreases. The opposite is true for gauges further downstream where the summer decrease and winter increase coincide with the low flow and high flow season, respectively.

At Cologne, the two-group patterns of the tributary areas is superimposed upon the Alpine pattern of decreasing and increasing $R$ in summer and winter, respectively. The projected maximal decreases of $R$ in summer range from -9 to -40% and the maximal increases in winter are in the range of +4 to +51%.
4.4 Results

Figure 4.10: Mean annual cycles of runoff at gauges along the Rhine river for the CTL period (dashed lines) and SCE period 2070-2099 (solid lines). The shading indicates the range of the 10 hydrological scenarios whereas the solid lines show the ensemble mean.

As a consequence, the amplitude of the mean annual cycle at downstream gauges is larger in the future SCE period which might enlarge the risk for floods and droughts.

4.4.4 Changes of runoff indices

Note that due to the limitations of the impact modeling chain, projected changes in low and high flows are less reliable than changes in mean flow. For instance, the employed delta change post-processing method does not account for changes in the spatio-temporal variability of $T$ and $P$. Also, the hydrological model PREVAH employs a simple runoff routing scheme and includes a simplified representation of soil moisture and groundwater storage. Nevertheless, we include this analysis in order to compare it to previous studies using other post-processing methods and hydrological models. The analysis shows how the extremes would change assuming no changes in variability.

Following Görgen et al. (2010), changes in three standard runoff indices are presented in Fig. 4.11. For low flows, we use the 90 % quantile of the flow duration curve (FDC90) and the mean of the lowest 7-day mean runoff per year in the simulation period (NM7Q). For high flows, we use the mean highest 1-day runoff per year in the simulation period (MHQ).

Regarding low flow indices, both the FDC90 and the NM7Q increase at gauges close to the Alps (Diepoldsau, Rheinfelden). Please note that due to the underestimated low flows in winter, the relative increase is probably overestimated. For gauges further downstream, there is no clear tendency in the SCE period 2021-2050. While the HadCM3Q0-driven scenarios and CNRM-ARPEGE-ALADIN project slight decreases, the rest of the ensemble shows small increases. For 2070-2099, the signal of the HadCM3Q0-driven scenarios and CNRM-ARPEGE-ALADIN amplifies towards a more pronounced decrease.

The reason for the gradient of the changes in the low flow indices along the river course is
the changing runoff regime. At gauges close to the Alps, the low flow season is in winter when increasing $T$ combined with increasing $P$ will lead to an increase of $R$. Further downstream, the low flows typically occur in summer when decreasing $P$ leads to decreasing mean $R$. However, it seems that the substantial decrease of mean $R$ in summer as shown in Fig. 4.9 is not transferable to changes in the low flow indices as e.g. the ECHAM5-driven modeling chains show an increase of FDC90 and NM7Q. A possible explanation might be that the extreme low flows are constrained by the groundwater level. The groundwater levels in the lower reaches of the Rhine basin depend on the amount of $P$ in winter and spring. In the ECHAM5-driven modeling chains, the increase of $P$ in winter and spring is larger than in e.g. the HadCM3Q0-driven chains, lifting the groundwater level all year round and causing increasing low flows.

The MHQ is projected to increase at all gauges in both SCE periods. The increase is clearer for the later SCE period. The HadCM3Q0-driven scenarios as well as CNRM-ARPEGE-ALADIN are generally at the lower end of the ensemble and even show slight decreases in a few cases. At Cologne, the changes range from -7 to +37 %. This is considerably larger than the range estimated in e.g., Görgen et al. (2010) who found changes of the MHQ at Cologne in the range of -5 to +25 % for 2070-2099. They used a bias-correction for
the statistical post-processing which allows for changes in variability. Thus, the differing results suggest that changes in variability diminish the increase of MHQ at the gauge Cologne. This hypothesis is supported by Lenderink et al. (2007) who found annual runoff maxima to increase less if bias-corrected GCM data is used compared to the scenarios derived from applying the delta change method. However, also differences of the employed hydrological models and the different CTL period (here: 1979-2008, Rheinblick2050: 1961-1990) could cause the differing results.

4.4.5 Separating $T$ and $P$ effects

Here, we present an assessment of the importance of changes in temperature and precipitation upon derived runoff changes. For simplicity, we restrict this analysis to the GCM-RCM chain ETHZ-HadCM3Q0-CLM, but we expect similar results for the other model chains. Figure 4.12 shows the changes in the mean annual cycle of runoff as projected by the $T$ and $P$ effect experiments as, well as by the fully nonlinear simulation at selected gauges along the Rhine river. Qualitatively speaking, the closer the additional experiments (using either the $T$ or $P$ effect) follow the nonlinear simulation, the higher the contribution of temperature or precipitation to the runoff changes. At Diepoldsau, the changes in the annual cycle of runoff are largely determined by the $T$ effect. At Rheinfelden, the $T$ effect determines the runoff change signal in winter while in late summer and early autumn, the $P$ effect is more relevant. At Kaub and Cologne, both the $T$ effect and $P$ effect contribute to the increase in mean runoff during winter, while in summer, the changes in runoff are almost exclusively determined by the $P$ effect.

In Fig. 4.12, also the sum of the $T$ and $P$ effect is shown by the dashed line. The closer this sum follows the nonlinear simulation, the more linear (e.g. additive) the $T$ and $P$ effects are. The results show that at all gauges except for Cologne, the $T$ and $P$ effect are almost perfectly additive. This is surprising since the hydrological model includes nonlinear processes that are expected to be sensitive with regard to changes in $T$ or $P$. For instance, the hydrological model includes many threshold effects (among others e.g. related to snowmelt or interception storage), which make the runoff response at least in principle nonlinear. The results thus indicate that these nonlinear effects do not imply a significant nonlinear dependence of the mean runoff response, although for extremes, a nonlinear effect still is in principle possible.

Figure 4.13 shows the spatial pattern of the contribution of the single effects ($T$ or $P$) to the changes in the mean annual cycle of the routed runoff using the ANOVA model (see section 4.3.3) for the quantification. Since the routed runoff is analyzed, the result for e.g. the most downstream catchment at Cologne is not the contribution of the single effect on the runoff changes in this individual catchment, but rather the integrated contribution of the single effect for the whole upstream area.
Figure 4.12: Difference in mean annual cycle of runoff between the SCE period 2070-2099 and the CTL period at 4 gauges along the Rhine, resulting from temperature ($T$) and precipitation ($P$) change experiments. The full nonlinear effect of applied temperature and precipitation change ($TP$) is shown as a black solid line. The dashed line represents the sum of the $T$ and $P$ results. The close agreement between the dashed and the solid black lines indicates that $T$ and $P$ effects are approximately additive.

The results show that the $T$ effect has a high contribution in the Alpine region. This influence is routed along the Upper and Middle Rhine almost down to Cologne. In the Neckar, Main and Moselle catchments, however, the $P$ effect is more important than the $T$ effect. Thus, the propagated importance of the $T$ effect from the Alpine to the Middle Rhine is diluted at the confluences of the tributaries with the Rhine. At Cologne, the $P$ effect is slightly more important than the $T$ effect for the changes in the mean annual cycle of runoff. The pattern in Fig. 4.13 suggests that the $T$ effect might largely be explained by snowmelt processes, as the respective signal is strongest in the Alpine region and the downstream gauges, while the low-altitude tributaries are dominated by the $P$ effect. This result is of significance in terms of uncertainty assessment. It is suggesting that projected temperature changes, which are considered more reliable than projected precipitation changes, explain a substantial fraction of the response in the seasonality of runoff along the Rhine river.

4.5 Summary and conclusions

This study assesses the impact of climate change on the hydrology of the Rhine down to Cologne using GCM-RCM model chains of the ENSEMBLES project. The methodology
4.5 Summary and conclusions

Figure 4.13: Contributions of the $T$ effect (left) and $P$ effect (right) in the changes of the annual cycle of the mean routed runoff. The coefficient $\eta^2$ measures how much of the projected changes are explained by a single effect ($T$ or $P$). A value of 1 corresponds to 100% of explained variance. See Fig. 4.1 for geographical details of the catchment. The black squares denote the runoff gauging stations as in Fig. 4.1.

Involves the delta change post-processing method and the hydrological model PREVAH. We found a decrease of mean runoff in summer and an increase in winter. This result is consistent with previous studies. The magnitudes of the changes in mean runoff are, however, dependent on the employed GCM-RCM. For Cologne and winter in the SCE period 2070-2099, maximal increases of mean runoff as estimated by a 31d MA range between +4 to +51%. Also, the period in the annual cycle during which increases in runoff are projected, varies considerably from December-January to September-June, depending upon the model considered.

Using PRUDENCE GCM-RCM data, Graham et al. (2007) found increases in mean monthly runoff only between January and May and maximal increases in the range of +5 to 20% with one outlier projecting +40%. The smaller spread could be related to the fact that most RCMs used by Graham et al. (2007) are driven by one single GCM. Our results clearly demonstrate that the driving GCM has a strong influence on the runoff scenarios. The HadCM3Q0-driven chains, for instance, show smaller runoff increases during a shorter period than the ECHAM5-driven scenarios. Also regarding extreme runoff indices, the driving GCM strongly characterises the estimated changes. This result highlights the importance of a multi-GCM ensemble for climate-runoff impact studies. To some extent this result is related to the fact that temperature changes provide the dominant influence upon runoff changes, at least for the river Rhine (see section 4.4.5). For temperature changes, it has been established that the dominant uncertainty is the large-scale atmospheric circulation provided by the GCM (Déqué et al., 2005). Nevertheless,
differences between different RCMs (driven by the same GCM) are significant, and may strongly affect the total uncertainty. Regarding the high flow index MHQ, our results show stronger increases than the scenarios compiled within the Rheinblick2050 project (Görgen et al., 2010). For low flow indices, our results show a clearer tendency towards a decrease than the more extensive scenario ensemble of the Rheinblick2050 project. Both results could be caused by the different statistical post-processing method (here: delta change, Rheinblick2050: bias-correction). If this hypothesis was true, projected variability changes damp the changes of extreme runoff indices both for low and high flows compared to the delta change scenarios that do not account for changes in spatio-temporal variability. This hypothesis is also supported by the studies of Shabalova et al. (2003) and Lenderink et al. (2007) who found high flow extremes to increase less strongly for impact modeling chains using post-processing methods that account for variability changes than for such that ignore variability changes.

We also quantified the contribution of the temperature and precipitation changes to the projected changes in the mean annual cycle of runoff. Consistent with previous studies (e.g. Nijssen et al., 2001; Graham et al., 2007; Hurkmans et al., 2010), we found that for the Alpine basins the effect of temperature changes ($T$ effect) determines to a large extent the changes in the annual cycle of mean runoff. In the lower tributaries, however, the effect of precipitation changes ($P$ effect) is more important. Thus, the dominance of the $T$ effect propagates downstream along the Upper and Middle Rhine but is diluted at the confluences of the tributaries with the Rhine. At Cologne and in the annual mean, the $P$ effect is more important than the $T$ effect. In winter, however, the projected increase in runoff is caused by $T$ and $P$ with contributions of about 50% each.

This study uses climate model data from the ENSEMBLES project, one post-processing method and one hydrological model. It extends the database of previous studies which are partly coordinated within the Rheinblick2050 project. The general pattern of significant changes agrees among all studies, but the here documented uncertainties are larger than previously found. Given the overall clear pattern of changes and considerable uncertainties, the principles of a “no-regret” and “flexibility” adaptation strategy as stated by Middelkoop et al. (2001) appears valid for water management planning in the Rhine catchment.

4.6 Acknowledgments

We would like to acknowledge all the persons and institutions that provided data for the study. Regarding meteorological data, these were: Olga Zolina (University of Bonn, Germany) and Hermann Mächel (German Weather Service), the German Weather Service, MeteoSwiss, MétéoFrance, the Federal Ministry of Agriculture, Forestry, Environment and Water Management (Austria) and the European Climate Assessment & Dataset
Regarding runoff data, these were: For Germany: The Federal Institute of Hydrology; Wasser- und Schifffahrtsverwaltung des Bundes; Landesanstalt für Umwelt, Messungen und Naturschutz Baden-Württemberg; Bayerisches Landesamt für Umwelt; Ministerium für Umwelt, Landwirtschaft, Ernährung, Weinbau und Forsten Rheinland-Pfalz; Hessisches Landesamt für Umwelt und Geologie; Landesamt für Umwelt- und Arbeitsschutz Saarland. For France: DREAL Lorraine, MEDD/DE; DREAL Alsace, MEDD/DE. For Switzerland: Federal Office for the Environment. The climate model data is provided by the ENSEMBLES project which was funded by the EU FP6 Integrated Project ENSEMBLES (Contract number 505539) and whose support is gratefully acknowledged. Regarding spatial data, we would like to thank the Commission for the Hydrology of the Rhine Basin (CHR) and the European Soil Database (ESDB). We are very grateful for the advises and help of Dr. Peter Krahe, Maria Carambia and Eric Sprokkereef, as well as of the team of the statistical seminar at ETH Zurich. The Center for Climate Systems Modeling (C2SM) at ETH Zurich is acknowledged for providing technical and scientific support. This study has partly been funded by the Swiss National Science Foundation through the NCCR Climate programme.
5 Synthesis and outlook

The studies compiled in this thesis investigate two aspects of a hydrological climate-impact modelling system in detail and apply the modelling system to the impacts of climate change on the hydrology of the river Rhine. In the following, the results are synthesised and their significance for climate-impact modelling is evaluated. Remaining open questions will be discussed, and a subjective view on some lessons learned in the history of climate-impact research is presented.

This work has tried to shed light on

a) how natural variability may affect estimates of the annual cycle of the climate change signal which is an important figure in climate-impact modelling studies,

b) how different uncertainty sources interact and affect the total uncertainty in climate-runoff projections and

c) how the hydrology of the Rhine river is projected to change according to the newest available multi-model GCM-RCM ensemble provided by the ENSEMBLES project, which has been used to drive the hydrological model PREVAH.

Concerning a) it has been found that the estimation of the annual cycle of the climate change signal is not trivial at all, even when restricting attention to surface temperature and precipitation. If moving averages are used, as done in many studies, substantial artificial fluctuations may arise in the annual cycle of the climate change signal. In an analysis of synthetically generated precipitation change signals, such fluctuations caused the change signal to deviate from the true prescribed change by more than 20%. Thus, smoothing techniques need to be applied that are able to filter out the effects of natural variability. Here, a spectral smoothing approach has been proposed. The estimated annual cycle of the climate change signal is a key element in the delta change approach which is the simplest statistical post-processing method for use in climate-impact studies. In principle, however, natural variability effects every statistical post-processing method. In the future, the presented methodological framework might therefore be extended to other post-processing methods as well.

Regarding b) it has been found that for projections of changes in runoff regime and runoff quantiles, climate models are the most important individual source of uncertainty in the impact modelling chain for changes in the summer runoff. During winter, the hydrological models are the dominant source of uncertainty for projections for the end of the twenty-first century. The uncertainty in the changes of the high runoff quantiles is determined to a large extent by the climate models and the statistical post-processing. The results further suggest that interactions between the different uncertainty sources contribute a
considerable fraction to the total uncertainty. This means that projected uncertainties due to, e.g., the choice of the statistical post-processing method, depend on the combination with other modelling chain elements. The verification of these results requires further research that includes more uncertainty sources and more diverse models for each impact modelling chain element.

The results for c) have revealed a clear tendency towards less runoff in summer and more runoff in winter in the Rhine basins down to Cologne. This result is consistent with previous studies (Middelkoop et al., 2001; Shabalova et al., 2003; Lenderink et al., 2007; Graham et al., 2007; Hurkmans et al., 2010). However, the estimated uncertainty range of the projected regime changes is larger than in previous studies. The results indicate that the driving GCM in the modelling chain determines the pattern of the regime changes to a large extent. The same is true for changes in both extreme low and high runoff indices. Furthermore, in regions close to the Alps, temperature changes have been found to have a larger effect on changes in the runoff regime than changes in precipitation. Further downstream, the precipitation changes become more important as tributaries with a pluvial runoff regime join the Rhine river. At Cologne, precipitation is more important than temperature. However, in winter both variables contribute equally to the projected runoff changes.

The presented suite of climate-impact studies complements the long tradition of hydrological climate-impact research. Today, about 35 years after the first hydrological climate-impact studies have been conducted, we have much more powerful tools available than in the early phase. The GCMs today include much more physical processes and are even coupled to non-atmospheric systems such as the ocean, the land surfaces and the carbon cycle. Dynamical downscaling is becoming a standard application to translate the GCM information to regional levels, and hydrological models typically work on daily or sub-daily time steps and describe the hydrological processes in a distributed manner. Still, in many aspects we are facing the same problems as 35 years ago. Interestingly, these problems match the criteria list of Gleick (see chapter 1) remarkably well. The spatio-temporal gap between climate models and hydrological models is still not fully closed, the climate models and to a lesser extent also the hydrological models still show large biases, and we are still not sure if a skillful hydrological model in present climate is also skillful in a future climate. Also, the uncertainties in hydrological climate scenarios could neither be substantially narrowed nor fully quantified.

Thus, from a scientific point of view, there are still numerous open questions that need to be addressed. The horizontal resolution of GCM-RCMs will soon reach scales that allow for a more realistic simulation of the climate in the Alpine region. Here, realistic means that the representation of the physical processes is more complete and adequate in the model. It does not mean that the models perfectly reproduce the reality. Biases will still remain. In particular, there is a need for more advanced statistical post-processing meth-
ods to enable a sound analysis of climate impacts on extreme events. These methods need to correct for biases not only in the mean but also in the spatio-temporal variability of the RCM data. Also, bias-correction methods need to allow for an accurate representation of the annual cycle and need to be robust with regard to natural variability (see chapter 2). Next, the assumption that the calibrated hydrological models also perform well in a future climate has to be addressed more thoroughly. Merz et al. (2011) showed that using a calibrated parameter set based on a 5-year period, a conceptual hydrological model produced runoff trends in a 30-year past modelling period that have not been observed. Transient parameter calibration from 1976 to 2006 revealed the parameters for the snow and soil moisture processes to be particularly non-stationary. They concluded that a fixed parameter set might induce artificial runoff changes. A solution could be to derive relations between hydrological model parameters and climate variables such as mean temperature or precipitation. Another solution could be to use probabilistic parameter sets (see chapter B).

Also, we are far away from a complete quantification of the uncertainty in hydrological climate-impact projections (see chapter 3). There is a quickly growing suite of available regional climate model ensembles, probabilistic post-processing tools etc., that specifically allow for a quantification of uncertainties (Tebaldi et al., 2005; Buser et al., 2009; van der Linden and Mitchell, 2009). It is therefore foreseeable that climate-impact studies will try to include more uncertainty sources to progress towards a complete assessment of uncertainty. One option could be to link the modelling chain elements in a probabilistic manner to fully quantify the uncertainty associated with climate-impact scenarios (Wilby and Harris, 2006). Consequently, the uncertainties in climate-impact projections might in principle enlarge in future studies, despite reducing the biases of the participating climate and hydrological models. This is not because the models have become worse but because until now, the estimated uncertainty ranges were overconfident due to an incomplete representation of all uncertainties involved in the impact modelling chain (Knutti et al., 2010).

Climate-impact modelling involves transdisciplinary research. After having discussed some open research questions, let’s ask what have we learned from 35 years of transdisciplinary climate-impact research. Of course, every climate-impact scientist will answer this question differently. Thus, the answer given here is a very subjective view based on roughly 3 years of experience in the field, and considerable ignorance in many aspects. I personally think we have learned that climate-impact modelling does not work as a one-way road. A lot of knowledge transfer between the climate modelling community and the hydrological modelling community is necessary in order to establish sound hydrological climate-impact modelling chains. The wish list of impact modellers is becoming more and more extensive, of course driven by socio-economic considerations. Today, it is not the changes in the mean annual runoff in large basins with simple topography that is interesting to climate-impact researchers. They are rather interested in the climate-impact on
5 Synthesis and outlook

runoff extremes or on small-scale catchments with complex topography. However, not all of the necessary input variables for such studies can be provided by climate models with an acceptable accuracy. Thus, climate and impact modellers need to find the optimal way to balance the wish list, taking into account the capabilities and weaknesses of state-of-the-art climate models and hydrological models.

Further, the literature shows that there is a vast amount of knowledge in climate science about how to statistically post-process climate model data in order to ameliorate their biases and transfer their information to the local scale. This knowledge though has not been used to a full extent by impact modellers. It seems that complex statistical methods do not provide the data needed by hydrological modellers. E.g., many statistical downscaling studies developed advanced statistical models for precipitation. However, hydrological models require other input variables, too. Thus, in the development of statistical post-processing methods, the requirements of climate-impact models need to be considered more thoroughly.

Also, climate scientists have become aware of the importance of natural variability in the climate system. In this framework of thinking, the climate system that we observe is just one realisation from all possible climate evolutions and climate models are tools to sample the range of all climate evolutions. If only one climate model run is analysed, it might deviate from the observed time series as it represents another realisation of the climate that does not necessarily coincide with the observed one (indeed, it will most likely not, as the climate system has a chaotic nature). This is a non-deterministic view of the climate system. One needs to be aware of this effect when comparing climate models to observations or when interpreting the climate change signal. Hydrologists are used to observed data and therefore have a deterministic view on the climate system. Consequently, hydrologists do not trust a climate model that deviates strongly from observations. In climate-impact modelling, it is necessary to bring together these markedly different views of reality.

Recently, some promising initiatives have been launched to foster the knowledge transfer between the climate modelling and the impact modelling community. In the UK, the UK Climate Impacts Programme (UKCIP) was established in 1997. Since then, it has developed an extensive knowledge how to design the interfaces of scientific research, policy makers and stakeholders. In the Netherlands, the fourth generation of national climate scenarios is currently available. The Royal Netherlands Meteorological Institute (KNMI) also provides a climate service that advises the sound usage and interpretation of climate scenarios. In Switzerland, the OcCC - an advisory board on climate research to the Swiss federal government - was established in 1996. Its purpose is to monitor climate research and provide recommendations to the stakeholders. The OcCC has published first national climate scenarios in 2007. Shortly after, in 2008, the Center for Climate Systems Modelling (C2SM) was founded. Regarding climate-impact research, its goal is to develop tools and methods to bridge the spatio-temporal gaps between the modelling systems used in
climate-impact research. The OcCC’s 2007 assessment will be updated by a new report in 2011 (CH2011). In the update, the spectral method for estimating the annual cycle of the climate change signal as presented in chapter 2 will be applied. Also, the new EU COST-Action VALUE (Validating and Integrating Downscaling Methods for Climate Change Research) that is going to be launched in autumn 2011 will explicitly address the collaboration between the climate-downscaling community and climate-impact researchers.

Within this PhD thesis, a set of climate scenarios for Switzerland using the delta change method has been compiled (see chapter 2). Originally, the work was initiated by the research projects “Klimawandel und Wasserkraftnutzung” (funded by swisselectric, the Swiss Federal Office of Energy and the canton Valais) and “CCHydro” (funded by the Federal Office for the Environment). These climate scenarios will be maintained and disseminated by C2SM and will be used in a suite of further climate impact studies. The experience gained in the collaboration with impact modellers will hopefully further improve the climate services provided by C2SM and lead to new climate scenarios using more advanced post-processing methods that meet the requirements of the impact modellers.

All the above mentioned climate service initiatives are confounded to national perspectives. Environmental systems, however, are not bound to any national boundaries. As the study about the climate-impact on the Rhine river shows, also international cooperation is needed (Görgen et al., 2010). It is therefore essential that the national initiatives also extend their scope and reserve a portion of their resources for international climate-impact projects.

Above all, results of climate-impact studies remain projections of a possible climate future, based on assumptions about future greenhouse gas emissions and physical process representation in the models. Such projections are best estimates based on our current knowledge and currently available observations. As the modelling tools improve and as the emission scenario assumptions change, one needs to reassess the climate impacts. Considerable effort has been invested to improve the impact-modelling tools, and current institutional and national initiatives are promising in order to foster an internationally coordinated transdisciplinary climate-impact research. We can look forward to the coming decades of climate-impact research. Let’s see if Gleick’s list can be proved dispensable.
A Supplementary material
Hydrological climate-impact projections for the Rhine river: GCM-RCM uncertainty and separate temperature and precipitation effects.

A.1 Introduction

This online supplementary material contains additional figures for the article “Hydrological climate-impact projection for the Rhine river: GCM-RCM uncertainty and separate temperature and precipitation effects”.
A.2 Spatial pattern of changes of atmospheric variables

A.2.1 SCE period 2021-2050

Temperature

Figure A.1: Change of winter temperature ($T$) in the sub-basins for the SCE-period 2021-2050 vs. CTL-period 1979-2008 [°C]. Each panel shows the values for a specific GCM-RCM chain (indicated in the panel’s title).

Figure A.2: Same as in Fig. A.1 but for spring.
A.2 Spatial pattern of changes of atmospheric variables

Figure A.3: Same as in Fig. A.1 but for summer.

Figure A.4: Same as in Fig. A.1 but for autumn.
A Supplementary material: Hydrological climate-impact projections for the Rhine river

**Precipitation**

**Figure A.5:** Change of winter precipitation ($P$) in the sub-basins for the SCE-period 2021-2050 vs. CTL-period 1979-2008 [%]. Each panel shows the values for a specific GCM-RCM chain (indicated in the panel’s title).

**Figure A.6:** Same as in Fig. A.5 but for spring.
A.2 Spatial pattern of changes of atmospheric variables

Figure A.7: Same as in Fig. A.5 but for summer.

Figure A.8: Same as in Fig. A.5 but for autumn.
Runoff

Figure A.9: Change of winter runoff ($R$) in the sub-basins for the SCE-period 2021-2050 vs. CTL-period 1979-2008 [%]. Each panel shows the values for a specific GCM-RCM chain (indicated in the panel’s title).

Figure A.10: Same as in Fig. A.9 but for spring.
A.2 Spatial pattern of changes of atmospheric variables

Figure A.11: Same as in Fig. A.9 but for summer.

Figure A.12: Same as in Fig. A.9 but for autumn.
Evapotranspiration

Figure A.13: Change of winter evapotranspiration ($E$) in the sub-basins for the SCE-period 2021-2050 vs. CTL-period 1979-2008 [%]. Each panel shows the values for a specific GCM-RCM chain (indicated in the panel’s title).

Figure A.14: Same as in Fig. A.13 but for spring.
A.2 Spatial pattern of changes of atmospheric variables

Figure A.15: Same as in Fig. A.13 but for summer.

Figure A.16: Same as in Fig. A.13 but for autumn.
A.2.2 SCE period 2070-2099

Temperature

Figure A.17: Change of winter temperature ($T$) in the sub-basins for the SCE-period 2070-2099 vs. CTL-period 1979-2008 [°C]. Each panel shows the values for a specific GCM-RCM chain (indicated in the panel’s title).

Figure A.18: Same as in Fig. A.17 but for spring.
A.2 Spatial pattern of changes of atmospheric variables

Figure A.19: Same as in Fig. A.17 but for summer.

Figure A.20: Same as in Fig. A.17 but for autumn.
Precipitation

Figure A.21: Change of winter precipitation ($P$) in the sub-basins for the SCE-period 2070-2099 vs. CTL-period 1979-2008 [%]. Each panel shows the values for a specific GCM-RCM chain (indicated in the panel’s title).

Figure A.22: Same as in Fig. A.21 but for spring.
A.2 Spatial pattern of changes of atmospheric variables

Figure A.23: Same as in Fig. A.21 but for summer.

Figure A.24: Same as in Fig. A.21 but for autumn.
Runoff

Figure A.25: Change of winter runoff ($R$) in the sub-basins for the SCE-period 2070-2099 vs. CTL-period 1979-2008 [%]. Each panel shows the values for a specific GCM-RCM chain (indicated in the panel’s title).

Figure A.26: Same as in Fig. A.25 but for spring.
Figure A.27: Same as in Fig. A.25 but for summer.

Figure A.28: Same as in Fig. A.25 but for autumn.
Evapotranspiration

Figure A.29: Change of winter evapotranspiration ($E$) in the sub-basins for the SCE-period 2070-2099 vs. CTL-period 1979-2008 [%]. Each panel shows the values for a specific GCM-RCM chain (indicated in the panel’s title).

Figure A.30: Same as in Fig. A.29 but for spring.
A.2 Spatial pattern of changes of atmospheric variables

Figure A.31: Same as in Fig. A.29 but for summer.

Figure A.32: Same as in Fig. A.29 but for autumn.
A.3 Changes in the annual cycle

Here, changes in the mean annual cycle of water balance quantities are shown at the additional gauges Brugg, Worms and Kaub.

Figure A.33: Mean annual cycle of the water balance quantities precipitation, evapotranspiration and runoff in the CTL period 1979-2008 for additional gauges along the Rhine river. The depicted precipitation includes a water balance correction as derived by the hydrological model PREVAH. Evapotranspiration and runoff are estimated by PREVAH. Scales are identical.
Figure A.34: Relative changes in the mean annual cycle of precipitation ($P$), evapotranspiration ($E$) and runoff ($R$) for additional gauges along the Rhine river. Each panel pair shows the changes relative to the CTL period 1979-2008 for the SCE period 2021-2050 on the left side and for the SCE period 2070-2099 on the right side. The gauge name is indicated in the top left corner of each panel pair. The scales are identical.
Regional parameter allocation and predictive uncertainty estimation of a rainfall-runoff model in the poorly-gauged Three Gorges Area (PR China)

Thomas Bosshard¹, Massimiliano Zappa²

¹Institute for Atmospheric and Climate Science, ETH Zurich, Switzerland
²Swiss Federal Institute for Forest, Snow and Landscape Research WSL, Zürcherstrasse 111, 8903 Birmensdorf, Switzerland

(Published in Physics and Chemistry of the Earth)

Abstract

In the framework of the IAHS initiative on Predictions in Ungauged Basins, the predictive uncertainty in hydrological simulations constitutes a key issue. The Three Gorges Area located in central China is a poorly gauged macro-scale catchment with an area of about 57’000 km² which is frequently hit by floods. The semi-distributed hydrological model PREVAH was implemented in this catchment as a part of the Changjiang Flood Assistance Project. The precipitation correction, being the most sensitive tuneable parameter in the basin, was chosen to be regionally allocated in order to cope with regionally varying precipitation measurement errors due to differences in the measurement network setup. The model was calibrated on one single discharge time series at the basin’s outlet by means of the Adaptive Metropolis algorithm. The estimated posteriori parameter probability distribution revealed that the regional allocation of the precipitation correction induced a strong parameter interdependence. The calculated predictive uncertainty is large but nevertheless suggests that additional uncertainty sources should be included to get a sound probabilistic simulation. Despite the large uncertainty and parameter interdependence, the model performance in the flood season of the year 2007 shows that the Adaptive Metropolis algorithm successfully inferred a well behaving bet-bin parameter set.
B.1 Introduction

The International Association of Hydrological Sciences (IAHS) launched the Decade on Predictions in Ungauged Basins (PUB) 2003-2012 (Sivapalan et al., 2003). The overall goal of the PUB initiative is to reduce the predictive uncertainty in modelling ungauged and poorly gauged basins regarding hydrological variables of interest such as runoff, groundwater level, sediments etc. Predictive uncertainty arises mainly from uncertainty about the model structure, data and parameters.

The Three Gorges Area (TGA) which constitutes the study area, can be classified as a partially poorly gauged basin due to the rather low and very unevenly distributed rain gauge density. It is a macro-scale catchment compromising an area of about 57,000 km². Within the Changjiang Flood Forecasting Assistance Project (Zappa et al., 2005), one objective was the implementation of an operational modelling system (Zappa et al., 2005) based on the semi-distributed hydrological model PREVAH (Gurtz et al., 1999; Zappa and Gurtz, 2003) in the TGA. The project started in 2003 as a reaction to the devastating floods in 1998 and 2002 that affected the Changjiang river basin in central China. The mountainous upper and middle reaches of the Changjiang that the TGA belongs to have been the sources of several flood waves in 1998 (Wu et al., 2006). Given the importance of the Changjiang basin to China as a settlement area for about 400 million inhabitants, a reliable flood forecasting system (FFS) is necessary. The already existing FFS is run by the Changjiang Water Resource Commission (CWRC).

In this article we present the methodology and results of the calibration of the PREVAH model in the TGA. Such a calibration of a conceptual rainfall-runoff model ideally has to deal with the three above-mentioned sources of uncertainty. However, most work has been conducted on including parameter uncertainty into calibration algorithms whilst input data (Kavetski et al., 2002) and model (Butts et al., 2004) error prove harder to be estimated. In our study, we introduced a crude input data error model consisting of a regionally allocated precipitation correction. The other parameters subject to the calibration are rainfall-runoff model parameters.

The estimation of the parameter uncertainty requires a knowledge about the parameter probability distribution (PPD) within the whole parameter space. Most parameter spaces of hydrological models are non-linear and thus algorithms able to cope with multiple local optima are needed. Given a particular model structure, multiple local optima or ridges within the parameter space signify that different parameter sets perform equally well. This phenomenon is called equifinality (Beven, 2001).

In this study, the Adaptive Metropolis (AM) algorithm has been used to infer the PPD. The AM is a Monte Carlo Markov Chain scheme which have frequently been used for hydrological model calibration throughout the last decade (see e.g. Kuczera and Parent, 1998; Bates and Campbell, 2001; Marshall et al., 2004; Schaefli et al., 2007b; Feyen et al., 2007). MCMC schemes are theoretically not restricted to a maximal number of parameters. Nash
and Sutcliffe (1970) stated though that the number of parameters representing the model complexity should be chosen parsimoniously (principle of parsimony). The complexity of a model should only be enhanced if the quality of a model increased substantially. In this study, we increased the spatial complexity by regional allocation of the most sensitive model parameter to eleven sub-regions (SR) in the TGA. The objective was to address the question whether or not regionally adjusted parameters and their uncertainties can successfully be inferred from a single discharge time series. In case of a positive result, more reliable flood predictions for the sub-regions could be provided to the local authorities. The precipitation correction as being the most sensitive parameter (Bosshard, 2007) has been chosen for the regional allocation. Its a priori identification is particularly challenging in poorly gauged mountainous areas (Sevruk and Nevenic, 1998) with a very limited number of representative gauges. An estimation of regional allocated precipitation correction factors by means of a calibration might be a possible way out of such a dilemma and represents the focus of this paper. Section 2 introduces the TGA area, the data sources, the adopted modelling framework and the setup of the MCMC inference. The results are presented in section 3 and discussed in section 4. The conclusions and an outlook of this study close this paper in section 5.

Figure B.1: Map of the Three Gorges Area. The eleven sub-regions are indicated with different grey tones and numbered from SR 1 to SR 11.
B.2 Methods

B.2.1 Study area and data sources

The Three Gorges Area (TGA) is situated in Central China and belongs to the upper reaches of the Changjiang laying between the two cities of Chongqing and Yichang and covering an area of about 57'000 km$^2$ (Fig. B.1). It is strongly heterogeneous in terms of hydrological relevant characteristics as e.g. slope and aspect of the topography. The elevation within the TGA ranges from 66 to 3193 m.a.s.l. Average annual rainfall accounts to 1000 to 1600 mm (Yihui, 1994). The region is dominated by a monsoon climate. The summer monsoon period with a prevailing southwesterly wind typically lasts from June to September (Yihui, 1994). It is this period of the year that receives most of the mean annual and extreme precipitation (Zhang et al., 2005; Su et al., 2005). The Meiyu rainfall season (plum rainfall season) with heavy rainfall events lasts from June to July. It temporally corresponds with the stationary rainfall belts over the Changjiang river valley (Yihui, 1994). The generated runoff drains to the Changjiang river that enters the TGA at Chongqing and leaves it at Yichang. Extreme flood discharges at Yichang can be as large as 71'000 m$^3$s$^{-1}$ (Wu et al., 2006).

Within the framework of the Changjiang Flood Forecasting Assistance Project (CFFAP) we could access a database of spatial information with a grid resolution of 90x90 m$^2$ (Zappa et al., 2005). The dataset consists of a digital elevation model, a land use map, a map of the flow accumulation and flow directions, and a map of the soil classes (FAO, 1988) for the entire TGA. Being aware that the identification of the most suitable grid size for the rainfall-runoff simulation is still an unresolved problem for hydrologists (Vazquez et al., 2002; Booij, 2003) we carefully determined the resolution for our simulation experiment. In a pre-study focussed on the well gauges sub-basin Daning (Sonderegger, 2004) an optimal grid size of 630 meters resolution could be identified. Such resolution has been adopted for all thirty-four mesoscale sub-basins defined within TGA (see section B.2.2). The Changjiang Water resources Commission (CWRC) provided data on rainfall and runoff for the period 2003 to 2007. The gauging network is depicted in Fig. B.1. The density of the rainfall gauging stations is about one per 1’000 km$^2$. There are nine discharge gauges in the area. The gauges in Chongqing and Wulong measure the flow into the TGA, the gauge in Yichang records the outflow. The discharge gauges within the TGA have a poor data quality except for the gauge in the Daning sub-basin. Rainfall and discharge data are available on a time resolution of 6 hours at 2:00, 8:00, 14:00 and 20:00 each day. During floods discharge observations are sampled more frequently. Concerning the other meteorological variables needed to force the hydrological model (Gurtz et al., 1999) no data are available within the investigated area. The CWRC provided data on temperature, wind speed, relative humidity, sunshine duration and global radiation from a network of few stations in a larger area for the period 1990-2002. Based on these records a
climatology has been derived and adopted for our investigations. Such climatology consists of an interannual mean value for every calendar day and each sub-basin within the TGA.

### B.2.2 PREVAH model structure

**Table B.1:** List of the parameters. SP stands for the starting parameter sets of the three Markov chains.

<table>
<thead>
<tr>
<th>Parameter name</th>
<th>Parameter abbr.</th>
<th>Min.</th>
<th>Max.</th>
<th>Regional allocation</th>
<th>SP 1</th>
<th>SP 2</th>
<th>SP 3</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precipitation correction</td>
<td>cprec</td>
<td>-20</td>
<td>100</td>
<td>Yes</td>
<td>22</td>
<td>10.5</td>
<td>70</td>
<td>%</td>
</tr>
<tr>
<td>Non-linearity parameter for the infiltration module</td>
<td>cbeta</td>
<td>1</td>
<td>6</td>
<td>No</td>
<td>3.0</td>
<td>3.2</td>
<td>2.25</td>
<td>-</td>
</tr>
<tr>
<td>Treshold for surface runoff generation</td>
<td>sgrluz</td>
<td>10</td>
<td>50</td>
<td>No</td>
<td>20.5</td>
<td>25</td>
<td>20</td>
<td>mm</td>
</tr>
<tr>
<td>Storage time parameter for surface runoff</td>
<td>k0h</td>
<td>10</td>
<td>50</td>
<td>No</td>
<td>20.4</td>
<td>18</td>
<td>20</td>
<td>h</td>
</tr>
<tr>
<td>Storage time parameter for interflow</td>
<td>k1h</td>
<td>50</td>
<td>150</td>
<td>No</td>
<td>60.0</td>
<td>74</td>
<td>75</td>
<td>h</td>
</tr>
<tr>
<td>Maximal percolation rate</td>
<td>perc</td>
<td>0.04</td>
<td>0.35</td>
<td>No</td>
<td>0.12</td>
<td>0.045</td>
<td>0.115</td>
<td>mm/h</td>
</tr>
<tr>
<td>Error lag parameter</td>
<td>$\rho$</td>
<td>0</td>
<td>1</td>
<td>-</td>
<td>0.8</td>
<td>0.8</td>
<td>0.8</td>
<td>-</td>
</tr>
</tbody>
</table>

The PREVAH model belongs to the type of deterministic, distributed, conceptual rainfall-runoff models (Refsgaard, 1997). Since its development it has been applied to various alpine catchments (Gurtz et al., 2003; Verbunt et al., 2006; Zappa and Kan, 2007). For a detailed presentation of PREVAH please refer to Viviroli et al. (2007). PREVAH is based on the Hydrological Response Unit (HRU) approach instead of a regular raster-grid solution. The chosen basic grid resolution of 630 m for the spatial data is clustered into HRUs with the aim to summarise areas of the basin where similar hydrological behaviour is expected. The recommended procedure to generate HRU’s in mountainous environments consists of assigning all the grid elements located in the same meteorological sub-unit (e.g. in the same sub-basin and range of elevation) that show similar aspect, the same land-use classification and similar soil properties to the same HRU (Gurtz et al., 1999). The tuneable model parameters are assigned to all the HRUs within a modelled basin. Thus, the tuneable parameters are not HRU but sub-basin based.

PREVAH contains up to 17 tuneable parameters. Because the TGA experience hardly any snow during the flood season, the parameters governing snow accumulation and snowmelt (Zappa et al., 2003) and icemelt (Gurtz et al., 2003) are not sensitive in the here invest-
igated TGA. Furthermore, the baseflow parameters were excluded from the calibration as well because only measured discharge values above 324 m$^3$s$^{-1}$ were used for the calibration (see section B.2.4). Thus, six model parameters are considered within this optimisation study (Table B.1). Besides the precipitation correction ($c_{prec}$), the other five tuneable parameters are found in the modules describing the soil’s unsaturated zone (Fig. B.2, Zappa and Gurtz, 2003). It is important to notice that the parameters differ in their nature regarding their affiliation to error sources. Introducing $c_{prec}$ into the calibration procedure implies the integration of a crude input error model. $C_{prec}$ therefore is not a pure model parameter but accounts for input uncertainties. The correct choice of an input error model is not trivial. More elaborate input error models are discussed in (Kavetski et al., 2002). The other five parameters are rainfall-runoff model parameters. In this article, we call both the input error model and the rainfall-runoff model parameters simply parameters, keeping in mind that they are associated with different error sources. The following explanations describe the function of the parameters in the model.

The precipitation correction ($c_{prec}$) is the percentage which the measured and spatially interpolated precipitation is increased or decreased with. $C_{prec}$ corrects for errors coming from rainfall point measurements as well as for errors in the evapotranspiration. It is therefore dependent on the rainfall measurement network. The corrected precipitation is the amount of rainfall above the canopy. The actual inflow to the soil and runoff generation modules $P_b$ is supplied by the precipitation reaching the soil and by snowmelt (Zappa et al, 2003). Fig. B.2 presents the conceptual structure of the soil model, including the module of runoff generation in the upper zone SUZ (Bergström, 1976). The link between the loss of water by evapotranspiration ET and runoff is given by the plant-available water storage in the aeration zone of the soil SSM (Zappa and Gurtz, 2003). The parameterisation of the maximum plant available water storage SFC depends on soil depth, effective root-depth, and the plant-available field capacity of the soil (Viviroli et al., 2007). The non-linearity parameter for infiltration $c_{beta}$ and the actual ratio between SSM and SFC controls the redistribution of $P_b$ between SSM and SUZ. $P_b$ is split into a component flowing into the runoff generation module DSUZ and a component contributing the soil moisture recharge SMR. SMR increases with increasing $c_{beta}$ (Zappa and Gurtz, 2003). SUZ is emptied by deep percolation PERC, as governed by the tuneable parameter $maxperc$ and by generation of surface runoff RS and interflow RI. Surface runoff occurs when the storage of water within SUZ exceeds the tuneable threshold parameter $sgrluz$. The two separate tuneable storage coefficients $k0h$ and $k1h$ are introduced for estimating surface runoff and interflow (Viviroli et al., 2007).

The TGA is subdivided into thirty-four sub-basins. The six tuneable parameters described above could principally be individually assigned to each of these thirty-four sub-basins. If all the parameters were allocated to the sub-basins, a total of 204 parameters would have to be estimated. To prevent possible overparameterisation, we decided to reclassify
the sub-basins into eleven sub-regions (SR) and to regionally allocate only one tuneable parameter to these SR’s. We chose $c_{prec}$ for the regional allocation because of two reasons: First, the uneven gauging station density and the topography within the TGA (see fig. B.1) is likely to cause varying $c_{prec}$ between the SR’s. Second, $c_{prec}$ proved to be the most sensitive parameter in the PREVAH model structure (Sonderegger, 2004; Bosshard, 2007). The sub-basins were reclassified to the eleven SR’s according to their geographical proximity and similar valley directions. Rain gauges are often situated in valleys and thus, the precipitation measurement error due to wind induced effects is dependent on the valley direction of the basin (Sevruk and Nevenic, 1998). Eleven regionally allocated $c_{prec}$ and five homogeneously parameters were eventually subjected to the calibration in the described experimental setting. Tab. B.1 gives further details on the units and ranges adopted for the six tuneable parameters and Fig. B.1 depicts the eleven sub-regions.

### B.2.3 Parameter inference and predictive uncertainty

Every conceptual rainfall-runoff model can be written as (e.g. Schaeffli et al. (2007b); Yaffee and McGee (2000))

\[
q_t = h(x_t, \Theta) + e_t \quad \text{(B.1)}
\]

\[
v_t = e_t - \sum_{l=1}^{p} \rho_l e_{t-l} \quad \text{for } p > 0 \quad \text{(B.2)}
\]

where $q_t$ is the observed runoff at time step $t$ and $h(x_t, \Theta)$ is the simulated runoff for the given input data $x_t$ and a given parameter set $\Theta$. The term $e_t$ is the error at time step $t$. This error term can be modified by an autoregressive error model of the order $p$ and the lag-coefficients $\rho_l$. A preliminary examination of the error time series revealed a
significant first order autocorrelation (Bosshard, 2007). A first order lag-coefficient was therefore added to the sixteen model parameters subjected to the parameter inference. The Adaptive Metropolis algorithm (AM) belonging to the family of Monte Carlo Markov Chain algorithms was chosen to numerically infer the parameter space (Haario et al., 2001). Marshall et al. (2004) reported a superior convergence performance of the AM compared to three variations of the Metropolis-Hastings algorithm with block updating or single site updating (Hastings, 1970; Bates and Campbell, 2001). The AM is usually used in a bayesian framework. Bayes’ theorem combines the prior knowledge about the parameter probability distribution \( p(\Theta) \) with a likelihood deduced from the observational data \( p(y|\Theta) \) to calculate the posteriori parameter probability distribution \( p(\Theta|y) \).

\[
p(\Theta|y) = \frac{p(y|\Theta) \cdot p(\Theta)}{P(y)} \tag{B.3}
\]

The observational data are represented by \( y \) while \( \Theta \) stands for the parameter vector. \( P(y) \) is a normalising constant representing the total probability of \( y \). If the model errors \( v_t \) are normally distributed, have a stationary variance and do not show autocorrelation the likelihood function \( p(y|\Theta) \) takes the form (Bates and Campbell, 2001):

\[
p(y|\Theta) = \prod_{t=1}^{n} p(v_t) \propto \sigma^{-n} e^{-\frac{\sum_{t=1}^{n} v_t^2}{\sigma^2}} \tag{B.4}
\]

The prior distribution in eq. B.3 was chosen to be uniform.

**Figure B.3:** Traceplot and autocorrelation function for the precipitation correction (cprec) in the sub-region SR 9 of the Markov chain 1.
In our study we set up the AM for three parallel Markov chains. The chain lengths lay between 731 and 860 iterations depending on the computational power of the computer the AM was running on. The first fifty accepted parameter sets were discarded from the analysis. The starting parameter sets (SP 1, SP 2, SP 3) are listed in Table B.1. SP 1 values are the calibrated values of the Daning setting (Sonderegger, 2004). SP 2 are the best parameter values derived in a pre-study (Bosshard, 2007) and SP 3 is chosen in a manner that the model will surely simulate too much peak flow.

The inferred PPD is equal to the true posteriori PPD if the AM has reached a convergent state. Since the Markov chains are considerably short, the convergence state has to be carefully investigated. The convergence behaviour of the AM was monitored by the traceplots of single parameter chains, their autocorrelation function, interdependence inspection of the posteriori parameter probability distribution (PPD) and the scale reduction factor suggested by Gelman and Rubin (1992).

The posteriori PPD summarises the information about the parameter uncertainties. The predictive uncertainty arising from the parameter uncertainty of both input error and rainfall-runoff model parameters was estimated by generating one hundred parameter sets according to their posteriori PPD, employing them for model simulations and eventually calculating the desired discharge quantiles of the simulation results. The predictive uncertainty was analysed with regard to its width relative to the observed discharge and its accuracy.

![Figure B.4: Two-dimensional section through the posteriori probability distribution for the parameter pairs SR 5-SR 1, SR 5-SR 8 and SR 5-SR 9. The gray scale from black to white stands for increasing pixel sample density representing the posteriori probability.](image)

**B.2.4 Calibration and verification**

The split sample calibration-verification method has been applied to observe a possible overparameterisation (Refsgaard, 1996). PREVAH was calibrated in the period May to September of the years 2003 and 2004 and subsequently verified in the same season of the years 2005-2007. The calibration and verification periods are very short. Wagener and
Table B.2: Goodness-of-fit measures for the calibration (2003-2004) and verification (2005-2007) period. The year 2007 is analysed separately due to different water storage level in the Three Gorges dam. BBP is the best-bin-parameter set and MP the median parameter set.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>BBP</td>
<td>MP</td>
<td>BBP</td>
</tr>
<tr>
<td>TGAI(NSE\textsubscript{lin})</td>
<td>0.624</td>
<td>0.594</td>
<td>0.589</td>
</tr>
<tr>
<td>TGAI(E\textsubscript{vol})</td>
<td>-4.26</td>
<td>-1.93</td>
<td>-12.4</td>
</tr>
<tr>
<td>YI(NSE\textsubscript{lin})</td>
<td>0.976</td>
<td>0.974</td>
<td>0.983</td>
</tr>
<tr>
<td>YI(NSE\textsubscript{log})</td>
<td>0.977</td>
<td>0.976</td>
<td>0.978</td>
</tr>
<tr>
<td>TOT</td>
<td>0.788</td>
<td>0.778</td>
<td>0.750</td>
</tr>
</tbody>
</table>

Weather (2006) report necessary calibration time periods of 3 to 10 years depending on the model complexity. However, the split sample method requires a measurement data set without changes in the basins hydrological characteristics (Refsgaard and Storm, 1996). This requirement is only fulfilled for the years 2003-2006. Before 2003 the Three Gorges dam did not store any water, in 2003 the water level was elevated up to 139 m and at the end of 2006 it was further raised up to 154 m (Ponseti and López-Pujol, 2006). These water level changes have a major impact on the hydrological and hydraulic properties of the basin. The performance of the model in the year 2007 was therefore separately analysed. Two deterministic parameter sets were inferred in the calibration procedure. The best-bin parameter set (BBP) consists of the mean values of the parameter bins with the highest posteriori probability. The Median parameter set (MP) consists of the medians derived from the PPD.

PREVAH was calibrated only by means of the likelihood function described in eq. B.4. However, different goodness-of-fit measures have varying meanings and it is therefore not recommendable to rely on just one goodness-of-fit measure (Madsen, 2000; Yapo et al., 1998). The linear and logarithmic Nash-Sutcliffe efficiency (NSE\textsubscript{lin}, NSE\textsubscript{log}) (Nash and Sutcliffe, 1970) and the volume error (E\textsubscript{vol}) of the two parameter sets were estimated for the calibration and verification period. The NSE and the volume error were also aggregated to a multi-objective function (see eq. B.8).

The calculation of the four goodness-of-fit measures was based on two discharge time series. One is the total discharge at Yichang (YI) which was used for the estimation of the NSE and volume error. The other one is the differential discharge time series of the total discharge at Yichang minus the routed discharge of Chongqing and Wulong. It represents an inferred estimation of the runoff generated within the TGA (TGAI). The TGAI series was used for the likelihood calculation and also received a heavy weight in the multi-objective function (see eq. B.5-B.8).
B.2 Methods

\[
\begin{align*}
LIN &= \frac{1 \cdot YI(NSE_{lin}) + 5 \cdot TGAI(NSE_{lin})}{6} \quad (B.5) \\
LOG &= YI(NSE_{log}) \quad (B.6) \\
VOL &= \frac{100 - |TGAI(E_{vol})|}{50} \quad (B.7) \\
TOT &= \frac{1}{8} \cdot (5 \cdot LIN + LOG + VOL) \quad (B.8)
\end{align*}
\]

The discharge in Yichang is measured downstream of the Three Gorges dam. Therefore, dam management and routing errors in the order of one timestep are likely to cause physically impossible negative inferred discharge rates for the TGAI. An obvious case of dam management can be seen in Fig. B.6 a) in the first half of September. At first, water has been held back in the Three Gorges dam producing a negative inferred discharge and afterwards, the previously stored water has been released which caused a clearly too high discharge. To prevent a calibration to obviously wrong data, data points with a discharge below the lower quartile of about 324 m$^3$ s$^{-1}$ were discarded from the calibration.

Figure B.5: Histograms of the sampled parameter values representing the posteriori parameter probability distributions. The abscissa stands for the parameter value and the ordinate shows the bin counts. Only three sub-regions are shown but they are representative for all the sub-regions.
B Regional parameter allocation and predictive uncertainty estimation of a RR model

B.3 Results

B.3.1 Inference monitoring

The inference behaviour of the AM is dominated by low sensitivity of the seventeen parameters. A low parameter sensitivity explains all the inference inspection results that are summarised in the following.

The acceptance rate was very high ranging from 83 % to 88% for the three Markov chains. This points either to a small jump width or to a low parameter sensitivity. The traceplot in Fig. B.3 is representative for all the sixteen model parameter chains. The traceplot proves that the jump width is large relative to the parameter range. Thus, only the low sensitivity can explain the high acceptance rate. Furthermore, the traceplot does not show any substantial narrowing jump width which would be a sign of a sound convergence. Only the traceplot of the lag parameter showed a narrow jumping band width which is in accordance to the substantially lower parameter uncertainty of this parameter (see the last row in Tab. B.3). The autocorrelation plot in Fig. B.3 shows that the memory of the Markov chain is short. This points to an efficient mixing which is beneficial for the convergence behaviour. The scale reduction factor of Gelman and Rubin (Gelman and Rubin, 1992) dropped to values below 1.15 at the end of each parameter chain. This indicates a Markov chain state close to convergence. However, the standard deviations of the posteriori PPD account to 25 % of the parameter range for most of the parameters. Thus, the scale reduction factor as the ratio of the between chain variance to the within chain variance is likely to be close to 1 even in a non-convergent case. In either case, convergent or not, the scale reduction factor of 1.15 indicates that a substantial decrease of the variances of the posteriori PPD cannot be expected even if the Markov chain would be continued.

One reason for the low parameter sensitivity could be parameter interdependence. Indeed, the parameters showed interdependence. Fig. B.4 depicts two-dimensional sections of the posteriori PPD. The SR 5-SR 9 pair shows a diagonal ridge in the PPD. This provides evidence that these two parameters are likely to compensate for each other. The SR5-SR 8 pair is an example of a section with a slightly higher parameter sensitivity while the SR 5-SR 1 pair shows a case with low interdependence. The three two-dimensional sections stand for the three kinds of sections encountered in the analysis of the parameter interdependences.

B.3.2 Calibration and verification

The model performance results of the two inferred parameter sets BBP and MP for the calibration and verification period are summarised in Table B.2.
Table B.3: Inferred parameter sets BBP and MP used for the calibration and verification comparison and list of the standard deviation $\sigma$ of the parameter distributions. SR1-SR11 denote the $c_{prec}$ values in the eleven sub-regions.

<table>
<thead>
<tr>
<th>Units</th>
<th>SR 1</th>
<th>SR 2</th>
<th>SR 3</th>
<th>SR 4</th>
<th>SR 5</th>
<th>SR 6</th>
<th>SR 7</th>
<th>SR 8</th>
<th>SR 9</th>
<th>SR 10</th>
<th>SR 11</th>
<th>cbeta</th>
<th>sgrluz</th>
<th>k0h</th>
<th>k1h</th>
<th>maxperc</th>
<th>$\rho$</th>
</tr>
</thead>
<tbody>
<tr>
<td>BBP</td>
<td>12.5</td>
<td>34.0</td>
<td>55.6</td>
<td>31.4</td>
<td>38.7</td>
<td>31.4</td>
<td>43.7</td>
<td>24.5</td>
<td>26.8</td>
<td>57.9</td>
<td>41.1</td>
<td>2.95</td>
<td>35.2</td>
<td>40.8</td>
<td>101</td>
<td>0.198</td>
<td>0.96</td>
</tr>
<tr>
<td>MP</td>
<td>39.9</td>
<td>35.4</td>
<td>38.8</td>
<td>37.3</td>
<td>36.1</td>
<td>32.5</td>
<td>32.2</td>
<td>25.9</td>
<td>36.2</td>
<td>44.4</td>
<td>38.4</td>
<td>3.52</td>
<td>33.8</td>
<td>31.5</td>
<td>101</td>
<td>0.185</td>
<td>0.95</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>29.3</td>
<td>31.2</td>
<td>30.7</td>
<td>30.1</td>
<td>30.6</td>
<td>31.1</td>
<td>29.2</td>
<td>27.5</td>
<td>30.9</td>
<td>32.0</td>
<td>30.0</td>
<td>1.21</td>
<td>9.65</td>
<td>10.1</td>
<td>24.3</td>
<td>0.079</td>
<td>0.031</td>
</tr>
<tr>
<td>$\sigma$/range</td>
<td>0.25</td>
<td>0.26</td>
<td>0.26</td>
<td>0.25</td>
<td>0.26</td>
<td>0.26</td>
<td>0.24</td>
<td>0.23</td>
<td>0.26</td>
<td>0.27</td>
<td>0.25</td>
<td>0.24</td>
<td>0.24</td>
<td>0.25</td>
<td>0.25</td>
<td>0.031</td>
<td>0.031</td>
</tr>
</tbody>
</table>

The NSE and $E_{val}$ values for the TGAI time series as well as the multi-objective function TOT show a decline from the calibration to the verification period of 2005-2006 thereby indicating a tendency for overparameterisation. The performance measures for the YI time series remain on the same high level in the two periods. This reflects the relative small contribution of the TGA catchment to the total discharge measured in Yichang. The BBP proved to perform better than the MP in most of the goodness-of-fit measures in all the analysed periods. The performance measures for the year 2007 are listed separately because the water storage level in the Three Gorges dam was raised from 139 m to 154 m (see section B.2.4). The model’s accuracy in the flood season of 2007 is substantially higher than in the other two periods.

### B.3.3 Parameter uncertainty

Fig. B.5 depicts some examples of parameter histograms produced by the AM. The uncertainty in $c_{prec}$ are associated with input data uncertainty while the other parameter uncertainties stand for rainfall-runoff model parameter uncertainty. Not all the seventeen histograms are shown in order to keep the graphic concise. The distributions of the regionally allocated $c_{prec}$ do not differ in a significant manner. The median $c_{prec}$-values vary from 25.9 to 44.4 % but they cannot be classified as distinctly differing due the large standard deviations. On the chosen parameter range none of the model parameters shows a high sensitivity as the $\sigma$-values of the parameter distributions relative to their range proves (see last row in Table B.3). Only the lag parameter for the correction of the autoregressive errors is highly sensitive. The parameter uncertainty therefore is very large with regard to the chosen parameter range.
Figure B.6: Rainfall and runoff of the years 2004 and 2005 for the TGA in a) and b) and the Daning sub-basin in c) and d). The grey bands are the 90% and 50% predictive uncertainty credibility interval (PUCI) band.
B.3 Results

Figure B.7: 90 % predictive uncertainty credibility interval width in dependence of the measured discharge of the TGA the Daning sub-basin for the summer season of the year 2004.

B.3.4 Predictive uncertainty

The simulation results for the predictive uncertainty arising from the inferred parameter uncertainty of the input error model and the soil moisture module are shown in Fig. B.6a) and b) for the whole TGA and in Fig. B.6c) and d) for the Daning sub-basin. The 50 % as well as the 90 % predictive uncertainty credibility interval (PUCI) are indicated. Only the years 2004 and 2005 are plotted. The year 2003 is strongly influenced by dam management and the year 2006 experienced no flood (Bosshard, 2007). First, we address the question of the width of the PUCI. In fact, the uncertainty is quite large. Fig. B.7 indicates that the credibility interval width increases with increasing discharge. For the TGA in the year 2004, the mean ratio of the width of the 90 % PUCI to the measured discharge for cases where the measured discharge exceeded \(87 \text{ l s}^{-1} \text{ km}^{-2}\) (corresponding to \(5000 \text{ m}^3 \text{s}^{-1}\)) was estimated to account for 62 % on average. The PUCI for the Daning sub-region is much larger as the steeper slope in Fig. B.7 proves. For most of the time the width of the 90 % band is larger than the actual measured discharge.

Second, we turn to the accurateness of the PUCI. The very large uncertainty in the Daning sub-region results in a full coating of the measured discharge by the 90 % PUCI band. This indicates that the estimated PUCI is too large. If its size was correct, only 90 % of the observed discharge data points would be coated. The picture is quite different for the whole TGA. Despite the well coated flood peaks, an analysis of the whole time series reveals that the estimated PUCI mantles the discharge points with a frequency being too low regarding the predictive uncertainty quantiles (see Table B.4). This indicates that apart from the estimated parameter uncertainties, other uncertainty sources like model uncertainty or other data error models have to be included to correctly simulate the discharge of the TGA in a probabilistic manner. The discharge time series for the TGA is likely to be
B Regional parameter allocation and predictive uncertainty estimation of a RR model

erroneous in low flow periods and high flow periods are inferred more accurately. Thus, an increase of the accurateness is observed if only high flow data points above 5000 m$^3$s$^{-1}$ are analysed.

Table B.4: Percentage of the observed discharge data points for the whole TGA that lay within the 50 % and 90 % predictive uncertainty credibility intervals (PUCI). The discharge data have been selected according to the limit used for the calibration (324 m$^3$s$^{-1}$) and a high flow limit (5000 m$^3$s$^{-1}$). Periods of obvious dam management have been excluded of the analysis.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;324 m$^3$s$^{-1}$</td>
<td>50 % PUCI</td>
<td>34.4 %</td>
<td>33.8 %</td>
</tr>
<tr>
<td></td>
<td>90 % PUCI</td>
<td>53.7 %</td>
<td>51.8 %</td>
</tr>
<tr>
<td>&gt;5000 m$^3$s$^{-1}$</td>
<td>50 % PUCI</td>
<td>46.4 %</td>
<td>52.8 %</td>
</tr>
<tr>
<td></td>
<td>90 % PUCI</td>
<td>70.8 %</td>
<td>76.7 %</td>
</tr>
</tbody>
</table>

Figure B.8: Rainfall and runoff in the heavy flood season of 2007. The inferred discharge time series of the TGA is shown in a) whereas the total discharge in Yichang is depicted in b).
B.3.5 The flood season 2007

The flood season in 2007 experienced four major flood events at Yichang (see Fig. B.8b). The contribution of the TGA to the total discharge varies between the four flood events from a negligible amount in the flood event around July 10, 2007 up to more than half of the total discharge in the flood peak on June 19, 2007. The rainfall patterns of the three events with heavy rainfall occurring in the TGA are shown in Fig. B.9. The central TGA receives most of the rainfall in all the three events. However, the overall intensity and the distribution varies between the three events. In the first two flood periods, the northern central TGA receives highest precipitation whereas in the last event, the southern central TGA received the highest precipitation and the rainfall was more distributed than in the first two flood events. Despite these spatial rainfall difference, the simulation based on the BBP parameter set was able to reproduce the observed flood discharge quite well. Dam management before June 5, 2007 and on August 1, 2007 caused some deviations between the simulation and the inferred observations. Also during the flood event round June 19, 2007 water was hold back and caused a short decline.

![Rainfall distributions for flood season 2007](image)

**Figure B.9:** Rainfall patterns of the three flood events in the flood season 2007 with rainfall occurring in the TGA. The scale is individual for each of the three patterns and ranges from 0 to the double mean rainfall amount of each event.

B.4 Discussion

The regional allocation of the \textit{cprec} to eleven SR’s within the TGA challenged the principle of parsimony stated by Nash and Sutcliffe (1970). The challenge meant to investigate the inference behaviour of regionally allocated parameters based on a single discharge time series for the whole TGA. It was the idea to test a way to attribute parameters to different SR’s in cases where regional differing parameters can a priori be assumed as in the case of \textit{cprec} but the poor knowledge about the catchment is hindering a sound regionalisation of the parameters. Our results show that the parameter distribution of the regionally allocated \textit{cprec} do not distinctly vary between the SR’s although the topography of the different SR’s and the varying gauge network density support the assumption of spatially
varying \(c_{\text{prec}}\) values. The most probable reason for this outcome is the detected parameter interdependence. The hypothesis of a strong parameter interdependence is also supported by the very low parameter sensitivity. Even though the low parameter sensitivity could principally lay in the function of the parameter, \(c_{\text{prec}}\) in particular should show high sensitivity due to the linear input data correction. Parameter interdependence has the effect, however, that an unrealistic high \(c_{\text{prec}}\) in one SR could be compensated by a low \(c_{\text{prec}}\) in other SR’s. The parameter uncertainty distribution therefore becomes artificially broadened. The overall effect is the lack of distinctly varying parameter distributions and low parameter sensitivity. This indicates that in the case of the TGA with eleven SR’s, the regional allocation of parameters does not necessarily bear more accurate information than a model with a lower spatial complexity. Or in other words, a setup with less parameters is likely to perform equally well as this chosen setup. With regard to the principle of parsimony, our degree of spatial model complexity proved to be not justified.

The enlarged parameter uncertainties due to interdependence caused large predictive uncertainty credibility intervals (PUCI) in the individual SR’s as the example of the Daning SR proves. Due to the artificial enlargement of the predictive uncertainty in the individual SR’s, the estimated PUCI are not representative for the SR’s. It is important to notice that the interdependence effect cancels out for the TGA as a whole. The inferred PUCI is therefore still representative for the whole TGA. The width of the estimated 90 % PUCI in the TGA in the average accounts for 62 % of the measured discharge. The accuracy of the 90 % PUCI is around 50 % instead of 90 %. This fact provides evidence that additional error sources like other input data uncertainty, model uncertainty, routing errors or dam management should be included in order to get a correct probabilistic simulation. The climatological data used for the forcing of PREVAH in particular is a source of uncertainty. It would be interesting to conduct a similar study about inference of regionally allocated parameters in a smaller catchment with a better data quality and a lower number of SR’s. The lower number of SR’s would reduce the effect of parameter interdependence.

The estimated PUCI cannot easily be compared to PUCI derived in other studies although such a comparison would be of high relevance in the PUB framework. The included uncertainty sources, data quality, model structure, study area characteristics and discharge affect the accurateness and width of the PUCI. However, no matter what inference setup has been used to estimate the PUCI, its width and accuracy can be analysed. Feyen et al. (2007) investigated the predictive uncertainty of the LISFLOOD model in a catchment with an area of 21'000 km\(^2\). In terms of size this catchment is comparable with the TGA. However, the discharge is by far lower as in the TGA and the catchment is not as mountainous. The peak discharges in the study period in Feyen et al. (2007) do not exceed 3000 m\(^3\)s\(^{-1}\). A graphical comparison of the time series including PUCI indicates that the predictive uncertainty due to parameter uncertainty is much smaller relative to the observed discharge in the study of Feyen et al. (2007) compared to our study. No indication of the accuracy is given in Feyen et al. (2007). The study of Schaefli et al. (2007b) investigated
a small alpine catchment with glacial properties. The PUCI derived for that catchment is almost as large as the one in our study but more accurate. It is difficult to find reasons for the differing PUCI accuracy and width. There is a lack of systematic comparative studies. The PUCI for the TGA proved to be large while lacking an optimal accuracy. Apart from the PUCI, also deterministic parameter sets have been derived from the posteriori PPD. The model simulation with the BBP parameter set showed remarkably good results especially in the peak flow determination during the flood season of the year 2007. The model was able to perform well in varying rainfall patterns in terms of spatial distribution and rainfall intensity. This indicates that the model setup with the parameter regionalisation by means of the HRU approach is able to cope with the poor data quality in the TGA. Furthermore, the parameter inference by means of an AM produced a well performing parameter set even though the number of iterations was very small compared to other studies that used Monte Carlo Markov Chain algorithms (Kuczera and Parent, 1998; Bates and Campbell, 2001; Marshall et al., 2004).

B.5 Conclusion

The regional allocation of the precipitation correction to the eleven sub-regions did not produce distinctly varying posteriori parameter probability distributions which have low parameter sensitivities. Parameter interdependence is likely to be the cause of these two effects. The results support the principle of parsimony which was challenged in this study. Despite the parameter interdependencies, the gained knowledge about the parameter uncertainties proved valuable to calculate predictive uncertainty credibility intervals (PUCI) in the TGA catchment as a whole. The PUCI for the TGA showed an underestimation of the uncertainty. Further uncertainty sources like choosing another data error model and model uncertainty should be included in order to get a sound probabilistic discharge simulation. Unfortunately, we have to acknowledge that model based warnings for the eleven sub-regions are not reliable because the regional discharge simulations are prone to very large predictive uncertainty as the example of the Daning sub-basins proves. In contrast, the obtained information about the predictive uncertainty of PREVAH for the whole TGA is of great importance for communicating the model’s reliability in the operational discharge nowcasting and forecasting for the TGA basin. The modelling setup presented in this paper is operationally implemented at the forecasting division of the CWRC. PREVAH is coupled in real time to the hydrometeorological database of the CWRC and to the numerical weather prediction model MM5 (Kunstmann and Stadler, 2005). MM5 provides operational 72h hour forecasts of the meteorological variables needed to force PREVAH two times per day. MM5 forecasts are operationally performed using 3 nested domains that dynamically downscale global forecasts to a 7 km horizontal resolu-
tion. Such forecasts are available since the 2006 flood season. Future efforts will focus on analysing the performance of the coupled forecasting systems. It is also aimed to test if MM5 requires a regionally differentiated bias correction. The experience gathered in the present study will provide a valuable basis for the planned activities.

### B.6 Acknowledgments

The authors would like to thank the Swiss Agency for Development and Cooperation (SDC) for initiating and financing the Changjiang Flood Forecasting Assistance Project. We are also thankful for the cooperation with the Bureau of Hydrology of the Changjiang Water Resource Commission in Wuhan (PR China). They provided us the data for this study and were kind hosts during our visits in Wuhan. Prof. Christoph Schär advised us in the numerical aspects of the work whereas the colleagues of the Statistical Seminar of the ETH Zürich, namely Prof. Künsch and Christoph Buser were of great help for the statistical part of the study.


van der Linden, P. and Mitchell, J., eds.: ENSEMBLES: Climate change and its impacts: Summary of research and results from the ENSEMBLES project, Met Office Hadley Centre, FitzRoy Road, Exeter EX1 3PB, UK, 2009.


Special thanks

Numerous people supported me during my PhD study. For me, it was an exceptional time. I therefore want to express my gratitude to all the people involved in my PhD study.
First of all, I want to thank Christoph Schär for accepting me as a PhD student. I especially appreciated his patience, the freedom he gave me to develop own ideas and his subtle guidance. Looking back, I am smilingly urged to believe that he already knew right from the beginning what odyssey this PhD study will be.
I also want to thank Sven Kotlarski who supervised me during the second half of my PhD. His skill to explain complex details in a simple and understandable way always impressed and helped me. Apart from work, I enjoyed a lot the various activities we undertook - see also below.
Further, my thanks go to Tracy Ewen who supervised me during the first part of my PhD. I could profit a lot from her expertise in mathematics and statistics. Her cheerful nature was catching and melted assumed problems away.
Next, I want to thank Rolf Weingartner for accepting to be in my PhD committee. I know him from national climate-impact projects. I was impressed by his ability to integrate different views from different research fields - a skill which is clearly needed in climate-impact research.
I know Massimiliano Zappa since summer 2006 when I looked for a scientific project that would allow me to visit China. When I contacted Massimiliano, he immediately agreed to offer me a traineeship within his cooperation project with the Changjiang Water Resource Commission. It turned out that this internship as well as the subsequent diploma thesis on the topic lead me to the PhD thesis. I’m extremely thankful for Massimiliano’s support and guidance through all the years and of course also for being a member of my PhD committee.
The icing on the cake of the PhD time was the whole bunch of friends and colleagues I got to know. I can hardly name all the people but rather want to name a few activities I will always remember. Sometimes, our group seemed to mainly exist of top-athletes. I enjoyed a lot the biking and hiking trips we made. Also, some memorable events involve helping colleagues to survive the first hours on skis. I also got to know almost forgotten hiking trails in Ticino, I remember a legendary party at the lake of Zurich with the most spectacular thunderstorm ever and I cannot count all the concerts I have been to with colleagues from the IAC. Thank you for all these (and the non-mentioned) special moments.
Very special thanks go to my family. They always supported me during all the years and let me the freedom of choice to do whatever I want.
And of course, I want to thank Nina who was my closest supporter, coach, personal chef and much more during the last few months of my PhD. Without her, there probably would not have been such a happy end.