REPORT SERIES IN AEROSOL SCIENCE
N:o 211 (2018)

Application of future climate projections in hydrological impact modeling

Olle Räty

Institute for Atmospheric and Earth System Research
Faculty of Science
University of Helsinki
Helsinki, Finland

Academic dissertation

To be presented, with the permission of the Faculty of Science of the University of Helsinki, for public criticism in auditorium E204, Gustaf Hällströmin katu 2, on June 8th, 2018, at 12 o'clock noon.

Helsinki 2018

Author's Address: INAR – Institute for Atmospheric and Earth System Research
P.O. Box 64
FI-00014 University of Helsinki
olle.raty@helsinki.fi

Supervisors: Professor Heikki Järvinen, Ph.D.
INAR – Institute for Atmospheric and Earth System Research
University of Helsinki

Docent Jouni Räisänen, Ph.D.
INAR – Institute for Atmospheric and Earth System Research
University of Helsinki

Reviewers: Associate Professor Abdel Hannachi, Ph.D.
Department of Meteorology
University of Stockholm

Senior Researcher Stefan Hagemann, Ph.D.
Institut für Küstenforschung
Helmholtz-Zentrum Geesthacht

Opponent: Post-Doctoral Researcher Claudia Teutschbein, Ph.D.
Department of Earth Sciences
Uppsala University

ISBN (978-952-7276-03-7) (printed version)
ISSN 0784-3496
Helsinki 2018
Unigrafia Oy

ISBN (978-952-7276-04-4) (pdf version)
http://ethesis.helsinki.fi
Helsinki 2018
Helsingin yliopiston verkkojulkaisut

Acknowledgements

This research was carried out in the Department of Physics of the University of Helsinki. I would like to thank the former heads of the department, Prof. Hannu Koskinen and Prof. Juhani Keinonen for providing me the necessary working facilities. I acknowledge Academy of Finland and Finnish Academy for Science and Letters for financially supporting this thesis. I also express my gratitude for the Director of the Institute for Atmospheric and Earth System Research (formerly Division of Atmospheric Sciences), Academician, Prof. Markku Kulmala for providing a stimulating working environment in one of the world's leading research groups in atmospheric research. My thanks goes also for the dynamic meteorology group for providing an interesting working environment and all the discussions with more or less work-related topics.

I acknowledge the mentoring provided by my supervisors Doc. Jouni Räisänen and Prof. Heikki Järvinen. A special thanks goes to Jouni for helping me when I was taking my first steps with my PhD studies and also for sharing his vast knowledge on pretty much any climate-related matter. I will also acknowledge my co-authors, especially Thomas whose help was invaluable to me. I would like to thank my coworkers and friends who have supported me throughout these years. In particular, I would like to thank Olli and Anton for their patience when sitting in the same room with me (I am not the most quiet person when working).

I would like to express my deepest gratitude to my sister for her support and to my parents, whose example sparked my interest for nature and its different phenomena already at a young age and who have also been supporting me throughout my studies. Finally, I would like to dedicate this thesis to my children, Ellen and Aaron.

Olle Einari Räty
University of Helsinki, 2018

Abstract

Increasing temperature is expected to intensify the hydrological cycle. The accompanied changes in (e.g.) snow cover, water availability and river flows impose challenges for several socio-economical sectors, whose operations are affected by these changes. The main question in these sectors is: how to cost-effectively adapt and mitigate to changes in the hydrological cycle?

To answer this question, hydrological modeling run with regional climate model (GCM-RCM) simulations as input is typically required. To reduce biases in the GCM-RCM simulations before using them as input in hydrological modeling, adjustments with statistical tools, such as model output statistics (MOS), is required. While MOS methods have been tested comprehensively in the present-day climate, not much has been known about how they perform in transient climatic conditions.

This gap is bridged in this thesis by evaluating a set of MOS methods designed for daily mean temperature and precipitation in a changing climate in Europe, using climate model simulations as proxies, i.e., pseudo-realities for the future. The main message from these exercises is that the identification of a single universally well performing method is practically impossible, as the best performing method varies in time, space and between different distributional aspects. This conclusion is further strengthened, when the selected methods are evaluated from the hydrological modeling perspective. Overall, several methods should ideally be used in impact studies.

In addition to the pseudo-reality tests, the relative importance of MOS-method differences was compared against GCM-RCM differences as uncertainty sources both in real-world climate projections in Europe and hydrological simulations in Scandinavia. As the second message of this thesis, although climate GCM-RCM differences explain a larger part of the spread in future projections, MOS-method differences are non-negligible, when the high and low extremes are considered. In line with this result, MOS-method uncertainty has the largest contribution to the spread in projected changes of low and high flows.

Keywords: climate modeling, hydrological modeling, bias correction, model output statistics, climate projection, uncertainty

Contents

List of publications

This thesis consists of an introductory review of the candidate's PhD study topic, followed by 4 research articles. **Papers I** and **II** are reprinted by permission from Springer Nature (license numbers 4335750781926 and 4335751091156). **Paper III** is reproduced with permission from the copyright holder, IWA Publishing, while **Paper IV** is reprinted under the Creative Commons Attribution License. In the introductory part, these papers are cited according to their Roman numerals.

I Räisänen, J. and O. Räty (2013). Projections of daily mean temperature variability in the future: cross-validation tests with ENSEMBLES regional climate simulations. *Climate Dynamics*, **41**(5), 1553–1568, doi:10.1007/s00382-012-1515-9

II Räty, O., J. Räisänen, and J. S. Ylhäisi (2014). Evaluation of delta change and bias correction methods for future daily precipitation: intermodel cross-validation using ENSEMBLES simulations. *Climate Dynamics*, **42**(9), 2287–2303, doi:10.1007/s00382-014-2130-8

III Räty, O., H. Virta, T. Bosshard, and C. Donnelly (2017). Regional climate model and model output statistics method uncertainties and the effect of temperature and precipitation on future river discharges in Scandinavia. *Hydrology Research*, **48**(5), 1363–1377, doi:10.2166/nh.2017.127

IV Räty, O., J. Räisänen, T. Bosshard, and C. Donnelly (2018). Intercomparison of Univariate and Joint Bias Correction Methods in Changing Climate From a Hydrological Perspective. *Climate*, **6**(2), 33, doi:10.3390/cli6020033

1 Introduction

Atmospheric concentrations of carbon dioxide and other greenhouse gases are increasing due to mankind's greenhouse gas emissions (IPCC, 2013). As a result, the global climate is warming, and changes are also expected in daily-scale temperature variability. As a major consequence of the warming, horizontal and vertical transport of water (i.e., hydrological cycle) is expected to intensify, inducing regionally varying changes in precipitation (e.g. Held and Soden, 2006; Huntington, 2006). Over the mid- and high latitude continents, the intensification is also manifested as changes in evapotranspiration, water flows above and within soil as well as in the amount of water stored in snow pack and soil (Barnett T. P. et al., 2005; Nohara et al., 2006; Adam et al., 2009; Vicente-Serrano et al., 2010; Kay et al., 2013; Liu et al., 2013; Räisänen, 2016).

It is apparent that the regionally varying changes in hydrological cycle will affect several socio-economic sectors such as hydropower production, flood protection, water management and agriculture (e.g. Oki and Kanae, 2006; Schewe et al., 2014; van Vliet et al., 2015; Arnell and Gosling, 2016). It is essential to provide reliable estimates of future changes as well as properly communicate the related uncertainties to policy makers and other stakeholders, in order to establish plausible long-term mitigation and adaptation strategies in the relevant sectors (Wilby and Dessai, 2010; Hewitson et al., 2014). Otherwise the selected strategies might fail and lead to substantial economical losses and even harmful societal effects.

Although the resolution of regional climate model (RCM) simulations driven with global climate models (GCMs) has improved substantially in recent years ($\approx$ 10 km), off-line hydrological modeling is typically required to properly evaluate regional scale changes in hydrological processes. It is well known that the direct use of climate model data in impact models is substantially hampered by biases in relation to the observed climate caused by insufficient description of physical processes on different spatial scales, smoothed representation of orography and the land-water distribution as well as errors in the driving GCMs, which propagate to the regional-scale simulations through the lateral boundary conditions (e.g. Räisänen, 2007; Knutti, 2008; Rummukainen, 2016). Figure 1 illustrates some of the typical issues related to the modeling of temperature and precipitation. In this particular example, the GCM-RCM simulation has a cold bias in comparison to the observations, while the daily variability is overestimated. On the other hand, the GCM-RCM overestimates the mean precipitation, has a smaller spread in daily values and underestimates the highest precipitation intensities. Also the fraction of wet days is overestimated due to excessive weak precipitation ("drizzle"), a common issue in many GCM-RCM simulations. In addition to biases in the marginal aspects, model simulations also exhibit biases in spatial, temporal and inter-variable correlation structures (Vrac and Friederichs, 2014; Maraun, 2016). A particular interest for this study are biases in inter-variable correlations between temperature and precipitation fields, which have been linked to biased simulations of different hydrological processes such as evapotranspiration, snowmelt and runoff generation (Kay et al., 2013). In the illustrated case, the GCM-RCM simulates too strong correlation between temperature and precipitation.

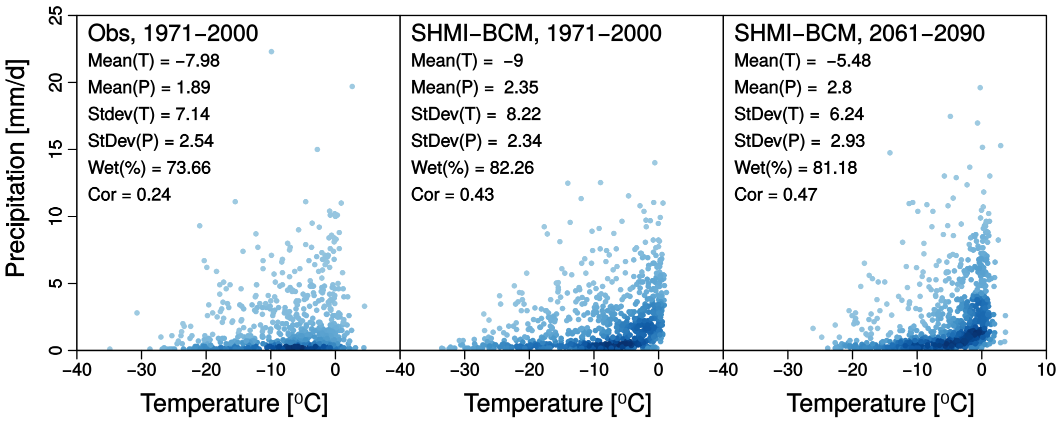

Figure 1: Scatter plot of the observed January daily mean temperature and daily precipitation in Jyväskylä (62.4°N, 25.7°E) in years 1971-2000 shown together with plots for the SMHI-BCM simulation (see Table 2) in the same period and in the late 21st century conditions (2061-2090). Dark blue indicates highest point densities in the plots. The basic summary statistics are also shown in each panel.

Due to the aforementioned issues, there is a practical need to improve the usability of both GCM and RCM simulations as input in hydrological modeling studies (Sharma et al., 2007). To this end, a substantial amount of different types of methods have been documented in the literature for temperature (Engen-Skaugen, 2007; Déqué, 2007; Yang et al., 2010; Amengual et al., 2012; Piani and Haerter, 2012) and particularly for precipitation (Engen-Skaugen, 2007; Li et al., 2010; Maraun et al., 2010; Piani et al., 2010; Yang et al., 2010). Lately, methods designed for adjusting the joint distribution of temperature and precipitation have also been developed (Piani and Haerter, 2012;

Li et al., 2014; Vrac and Friederichs, 2014; Cannon, 2016, 2018). While the ability of these methods to reduce biases in comparison to observational data has been extensively tested both from climate modeling and hydrological modeling perspectives (Wood et al., 2004; Themeßl et al., 2011; Gudmundsson et al., 2012; Watanabe et al., 2012; Chen et al., 2013; Lafon et al., 2013), an inherent limitation of these assessments is that they do not give information on the relative performance of bias correction methods in non-stationary conditions imposed by anthropogenic climate change. As a partial remedy, recent studies have taken an alternative approach, whereby statistical methods have been evaluated using climate model simulations as proxies for future climate (Vrac et al., 2007; Bracegirdle and Stephenson, 2012; Maraun, 2012). This approach provides additional pathways to evaluate the applicability of these methods in a changing climate and to address the possible shortcomings of these methods such as their "blindness" to underlying physical processes and the related assumption of bias stationarity, lately questioned by several studies (Christensen et al., 2008; Ehret et al., 2012; Bellprat et al., 2013; Maraun, 2016; Van Schaeybroeck and Vannitsem, 2016; Maraun et al., 2017).

This thesis aims to bring further insights to the use of MOS methods when constructing daily mean temperature and precipitation projections for impact study purposes, the main weight being on hydrological impact studies. In short, the following questions have been addressed:

- How to best combine information obtained from observations and GCM-RCM simulations when constructing future climate projections of daily mean temperature and daily precipitation with univarite and bi-variate methods?
- What is the relative importance of uncertainties related to the choice of MOS method in comparison to uncertainties related to differences between GCM-RCM simulations and how are these reflected in hydrological future simulations?
- How do the effects of temperature and precipitation changes on river discharges vary between different GCM-RCM and MOS-method combinations?

To answer these questions, a set of MOS methods developed in the literature and applied in this study are used to construct temperature and precipitation projections at daily scale in the European region using GCM-RCM simulations available from the latest international joint modeling efforts. Previous assessment studies, which are typically based on validations in the present-day climate, are complemented with tests in non-stationary climatic conditions both from climate model and hydrological modeling point of view. Furthermore, additional information on inherent uncertainties related to the variations between the future projections obtained using different MOS methods is given and compared against other uncertainty sources. A more detailed view of these uncertainties and their spatio-temporal variations is given for the Scandinavian region, particularly from hydrological modeling perspective.

2 Model output statistics

This thesis concentrates on evaluations of statistical methods belonging to so-called model output statistics (MOS). This group of methods has lately become somewhat of a $defacto$ standard post-processing tool for many impact modeling applications. MOS has its origin in numerical weather prediction (Glahn and Lowry, 1972), where the rationale has been to derive a statistical relationship (typically using multiple linear regression) between observations and NWP model output at specific forecast lead times. This allows to reduce systematic biases and when combined with geographical information can also be used to produce downscaled site-specific forecasts. In climate study context, MOS is used to derive either an event- or a distribution-wise relationship between the observed and the model simulated climate, which is then used to adjust model projections for the future typically in conjunction with a downscaling step due to the coarse resolution of GCM and RCM simulations. The studies conducted in this thesis are limited to distribution-wise methods due to their simplicity and widespread application in impact studies. Furthermore, direct evaluation of issues related to downscaling aspects (e.g. Maraun, 2013) has been omitted.

In this section, we illustrate the basic formulations for several types of distribution-wise MOS methods, starting from simple methods designed for the adjustment of time-mean climate, then continuing to more sophisticated methods, which adjust the full marginal distribution and finally introducing two bi-variate algorithms applied in **Paper IV**. In total, 10 methods for daily mean temperature, 9 for daily precipitation and 2 methods adjusting the joint distribution of temperature and precipitation are covered in this thesis (Table 1).

2.1 Univariate methods

The formulation of the problem is simple when biases only in the time mean climate are of interest. Let X_o denote the observed time series of either daily mean temperature or daily precipitation and let X_c and X_s be the simulated time series in the baseline and scenario period, respectively. In case of temperature, bias adjusted values in the scenario period are simply obtained with

Table 1: Short descriptions of the MOS methods used in this thesis. The first column describes briefly the addressed distributional aspects, while the second and third column show the shorthand used hereafter in the text for both delta change (DC) and bias correction (BC) versions. The papers in which a particular method is applied are given in the fourth column.

Method	DC	BC	Paper
Daily mean temperature			
Mean	T1	T6	**I, III**
Mean and spread	T2	T7	**I**
Mean, spread and skewness	T3	T8	**I**
Non-parametric quantile mapping with smoothing	T4	T9	**I, III, IV**
Parametric quantile mapping, linear fit	T5	T10	**I**
Daily precipitation			
Mean	P1		**II, III**
Mean and spread, non-linear scaling	P2	P6	**II**
Mean and spread, power transformation	P3	P7	**II**
Non-parametric quantile mapping with smoothing	P4	P8	**II, III, IV**
Parametric quantile mapping, double gamma fit	P5	P9	**II**
Bi-variate distribution			
Copula-based bias correction		B1	**IV**
N-pdft algorithm with T9 and P8		B2	**IV**

$$X_p = X_o + (\overline{X}_s - \overline{X}_c) = X_s + (\overline{X}_o - \overline{X}_c), \tag{1}$$

where $\overline{()}$ denotes the temporal average. For precipitation, relative changes are usually considered to avoid negative values:

$$X_p = X_o \frac{\overline{X}_s}{\overline{X}_c} = X_s \frac{\overline{X}_o}{\overline{X}_c}. \tag{2}$$

The two forms on the right-hand-side in Eq. 1 and 2 are commonly known as delta change (T1 and P1 in Table 1) and bias correction (T6) approaches, respectively. Both forms are mathematically equivalent; they give exactly the same change in the mean as simulated by the model. However, previous studies as well as the results in **Paper I** indicate that the preservation of the model-simulated mean change is not necessarily optimal even when only time-mean projections are of interest, as biases in daily variability might affect time mean projections (Boberg and Christensen, 2012). In the following, the equations for the remaining methods are given in the bias correction form. Delta change forms can be simply obtained by switching the subscripts of observations (o) and the scenario period simulation (s).

A natural extension of the simple time mean adjustment is to take the spread in the daily values into account. For temperature (T2/T7), this adjustment can be written in the bias correction form as

$$X_p = \overline{X_s} + (\overline{X_o} - \overline{X_c}) + (X_s - \overline{X_s}) \frac{s_o}{s_c}. \tag{3}$$

Here, s_o and s_c denote the observed and modeled standard deviation in the baseline period, respectively. A similar type of method (P2/P6) as described in Eq. 3 but tailored for precipitation was developed by Engen-Skaugen (2007). The main drawback of the bias correction form of this method (P6) is that, when the observed coefficient of variation (CV) is larger than the modeled one in the baseline period, some of the precipitation values become negative. These values need to be set to zero and the adjustment repeated iteratively until satisfactory results are obtained.

An alternative method (P3/P7) to correct biases in the precipitation distribution spread developed by Leander and Buishand (2007) was also tested. It is based on the power transformation of precipitation time series according to

$$X_p = a{X_s}^b, \tag{4}$$

where $a = \frac{\overline{X_o}}{\overline{X_c}}$ and b is a constant, which is found (in method P7) iteratively by modifying the CV of the scenario period simulation to match the observed one. In addition to the distribution spread, also the mean and higher distribution moments are changed. Multiplying the simulated precipitation values with a finally corrects the distribution mean.

Equation 3 can be further extended to take biases in the distribution skewness (γ) into account. In this thesis this was only made for temperature (methods T3/T8), and in the bias correction form was accomplished by finding a solution to the following equation

$$\gamma_p = \gamma_s + (\gamma_o - \gamma_c). \tag{5}$$

While Eq. 5 could be solved in several different ways, the algorithm described by Ballester et al. (2010) was applied in this thesis (**Paper I**).

A generalization of the aforementioned approaches is the so-called quantile mapping, which also originates from weather prediction (Panofsky and Brier, 1968). This method and its several variants are currently perhaps the most widely used MOS methods when

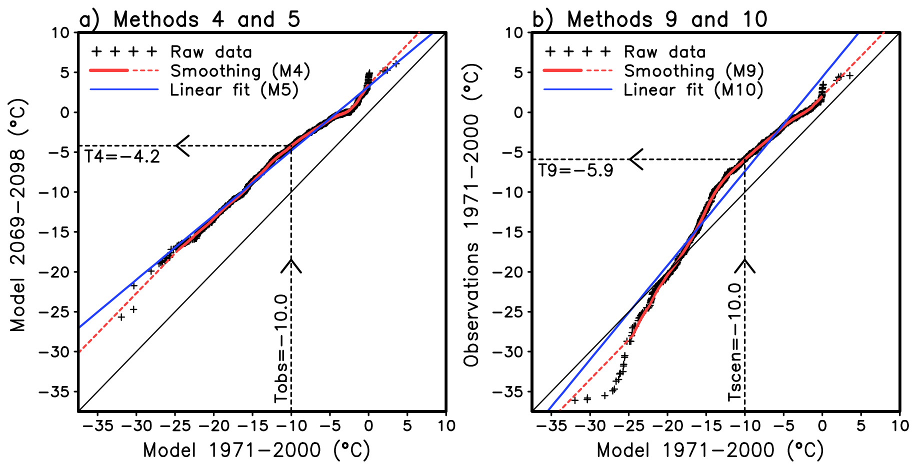

Figure 2: An illustration of quantile mapping when applied in (left) the delta change and (right) bias correction mode to daily mean temperature distributions. Red (blue) curves show the quantile-quantile relationship for the non-parametric (linear-fit) version of quantile mapping. Adapted from **Paper I**.

bias adjusted (and potentially downscaled) climate simulations are needed in impact studies. Two variants of a non-parametric quantile mapping algorithm were developed and tested in **Papers I and II** separately for daily mean temperature (methods T4 and T9) and daily precipitation (methods P4 and P8). In the traditional bias correction form, quantile mapping can be formally expressed as

$$X_p = F_o^{-1}\left[F_c\left(X_s\right)\right], \tag{6}$$

where F_c is the cumulative distribution of the baseline period model simulation and F_o^{-1} the inverse of the observed cumulative distribution.

The basic idea of quantile mapping is illustrated for temperature in Fig. 2, when applied both in the delta change (left) and bias correction (right) mode. These figures show that the raw quantile-quantile relationship is rather noisy, particularly in the tails of the distribution. To reduce the effect of random variations to the estimation of model biases or simulated changes, values for individual quantiles are smoothed by taking a running average over a specific range of quantiles (a) of $F_o^{-1}(x)$ according to

$$\tilde{F}_o^{-1}(x) = \int_{\max(x-a,0)}^{\min(x+a,1)} F_o^{-1}(x)dx / \int_{\max(x-a,0)}^{\min(x+a,1)} dx. \tag{7}$$

Similar smoothing is applied to $F_c^{-1}(x)$. Tests with different values for a showed that daily mean temperature benefits from slightly stronger smoothing ($a = 0.05$) than daily precipitation ($a = 0.02$). For daily mean temperature, this means that $\tilde{F}_o^{-1}(100)$ is obtained as the average of the top 5% of the distribution. As smoothing contracts the distribution, the smaller smoothing parameter value for precipitation is likely explained by the deterioration of the estimated quantile-quantile relationship in the upper tail of the precipitation distribution with larger values of a. Another issue with non-parametric quantile mapping, which is partially exacerbated by the contracting effect of the smoothing, is the need to extrapolate the quantile-quantile relationship if the future simulation is outside the control period climate. For temperature, it is assumed that the difference $\tilde{F}_o^{-1} - \tilde{F}_c^{-1}$ remains constant for $x < 0$ and $x > 1$ (red dashed lines in Fig. 2). When the quantile-quantile relationship for daily precipitation needs to be extrapolated, the ratio between the observed and modeled quantiles is used.

An additional issue that requires a particular attention are equal values, which hamper the estimation of the quantile-quantile relationship particularly for daily precipitation due to the singularity at 0. To avoid equal values and to handle dry days in the precipitation distribution, small uniformly distributed random values are added to both observed and simulated time series. This approach, when applied specifically to zero precipitation values, is called singularity stochastic removal by Vrac et al. (2016) and has also been applied in other studies (Zhang et al., 2009; Cannon et al., 2015). This approach is particularly useful for precipitation, as it allows to modify the fraction of wet days both when GCM-RCMs under- or overestimate it and is also more flexible than the commonly taken approach to match the fraction of simulated wet days to the observed one before applying quantile mapping.

As an alternative strategy to reduce the effect of sampling noise in the quantile-quantile relationship, a parametric fit can also be used. This approach also helps to avoid extrapolation that would otherwise occur when the simulated future values fall outside the observed and modeled region in the control period. For daily mean temperature, this was implemented in **Paper I** by fitting a simple linear regression line to the quantile-quantile relationship (blue lines in Fig. 2, methods T5 and T10). The quantile-quantile relationship can also be modeled based on assumptions on the underlying distribution. Yang et al. (2010) described a parametric version of quantile mapping both for temperature and precipitation (distribution based scaling), in which the quantile-quantile relationship between the observed and simulated time series in the baseline period is modeled by fitting suitable distributions to them. The algorithm for precipitation was tested in **Paper II** (methods P5 and P9). First, the fraction of wet days (threshold 0.1 mmd^{-1}) was corrected in the modeled time series to match the observed one. Next, separate gamma distributions (Wilks, 2006) were fitted below and above the 95^{th} percentile of the wet-day distributions, and precipitation values from the future simulation were adjusted based on the parameters for these distributions. Yang et al. (2010) argue that this allows to better capture the high daily precipitation intensities, which are not necessarily captured by a single fit. On the other hand, this is achieved with the expense of introducing potential discontinuity to the upper tail of the precipitation distribution.

Although both bias correction and delta change approaches are mathematically equivalent (see **Papers I and II** for a detailed discussion), there is a fundamental difference between them. On one hand, projections constructed using the delta change approach have the same temporal and spatial structure with the observations. Thus, it could be expected that they perform well in constructing near-term projections (**Papers I** and

II). However, this also sets a strong limitation as delta change methods are incapable to take simulated changes in different correlation structures into account. On the other hand, projections constructed using bias correction methods inherit different types of correlation structures directly from the underlying model simulation, thus incorporating simulated changes in these aspects (but also their biases, if not explicitly corrected for) in the future projections. Whether delta change or bias correction methods are preferable ultimately depends on which is more worrisome: the biases in those statistical aspects of simulated climate that are not accounted for by the bias correction, or the changes in those aspects that are assumed to remain unchanged in delta change methods?

2.2 Bi-variate methods

A common limitation in all the previously introduced methods is that they only adjust marginal distributions without explicitly taking temporal, spatial or inter-variable correlations into account, although some of these correlation structures might be altered indirectly by univariate bias correction methods too (Rajczak et al., 2016). As discussed above, the realistic description of inter-variable correlations might be crucial for certain end user applications. For example, increase/decrease in snow pack around the melting point depends on co-variations in temperature and precipitation, i.e., whether precipitation falls as rain or snow.

Bi- and multivariate extensions of MOS aim to address these issues by incorporating inter-variable correlation structures in their adjustments. Here, we focus on two methods (**Paper IV**), which were both applied to adjust the joint distribution of temperature and precipitation. The first algorithm is based on the decomposition of the joint distribution of daily mean temperature and precipitation into separate marginal distributions and the so-called copula, which describes the dependence structure between the marginal distributions (Fig. 3). The (2-dimensional) copula-based approach is based on Sklar's theorem (Sklar, 1959), which states that, given marginal distributions F and G, there exists a copula C such that

$$\begin{aligned} H(x,y) &= C\left[F(x),G(y)\right] \\ &= C\left[u,v\right],\ u,v \in [0,1], \end{aligned} \tag{8}$$

where $H(x,y)$ denotes the joint distribution of random variables X and Y. In other words, copula is the joint distribution of $F(x)$ and $G(y)$, with uniform marginals and

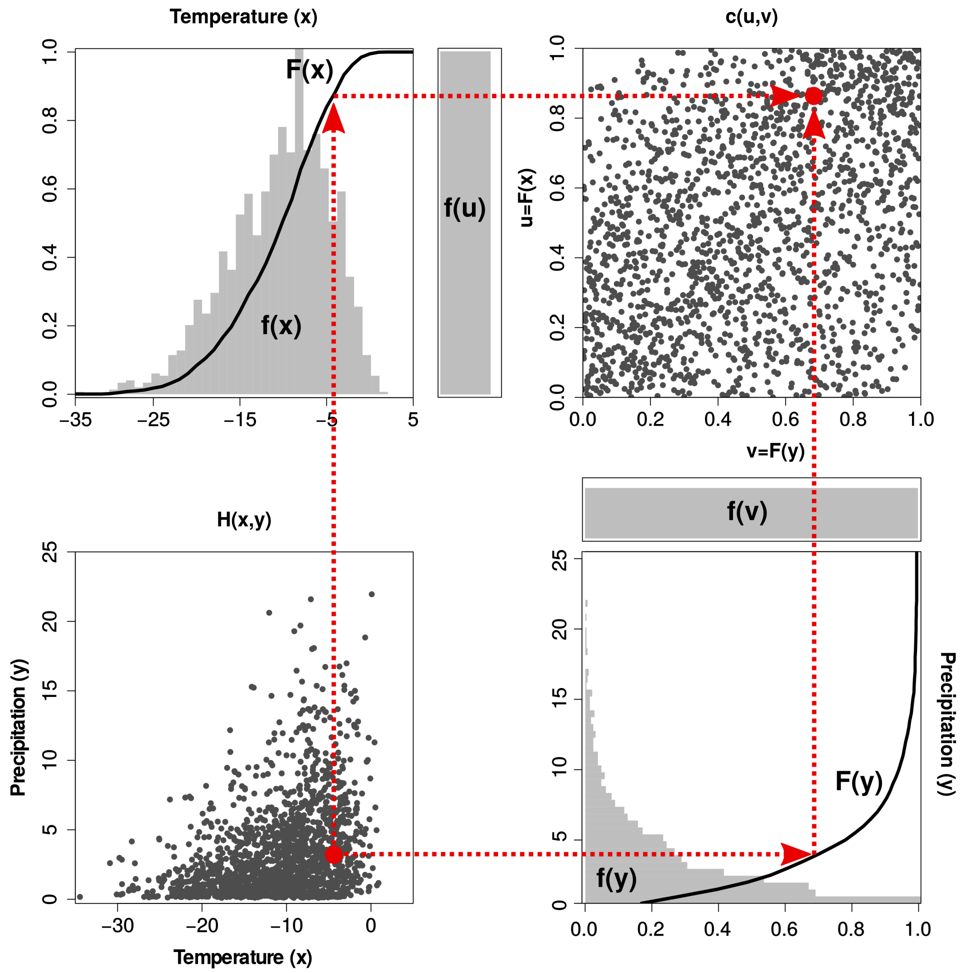

Figure 3: A schematic figure showing the decomposition of the joint distribution of daily mean temperature and precipitation (bottom left) to temperature (top left) and precipitation (bottom right) marginals and the empirical copula *density* (top right) structure. Red arrows illustrate how a point taken from the joint distribution is located on the cumulative marginals ($F(x)$ and $F(y)$) together with the corresponding location in the copula. In this particular example, the positive correlation between temperature and precipitation is indicated by the higher point density on the region extending from the bottom left corner to the top right corner in the copula panel. The *cumulative* copula probability is (empirically) obtained for a particular point by counting the fraction of points having smaller or equal u and v values in the copula plot.

values distributed in a unit rectangle. The copula-based correction exploits the fact that the required conditional probabilities can be expressed as partial derivatives of the cumulative copula. Copula models are popular in economics and have recently gained more interest also in the climate research community (see Schölzel and Friederichs (2008) for an overview).

The implementation of the copula-based algorithm (method B1) follows Li et al. (2014) with some modifications taken from Gennaretti et al. (2015). Although Li et al. (2014) originally applied their method to monthly temperature and precipitation values, here the method is applied to daily time series as in Gennaretti et al. (2015). First, the full joint distribution of daily mean temperature (X) and precipitation (Y) is decomposed into wet- and dry-day components according to

$$H(x, y) = (1 - p_w)F_d(x) + p_w C\left(F_w(x), G(y)\right), \tag{9}$$

where p_w is the wet-day probability ($\mathrm{P}(Y > 0.1 \text{ mmd}^{-1})$), $F_d(x)$ and $F_w(x)$ denote the cumulative distribution of daily mean temperature on dry and wet days, $G(y)$ is the cumulative distribution of daily precipitation and $C()$ stands for the copula between X and Y. Temperature is assumed to follow normal distribution, while gamma distribution is fitted to precipitation. In the original algorithm, temperature was assumed to be identically distributed on wet and dry days. Here, wet- and dry-day distributions were estimated separately as in Gennaretti et al. (2015). Furthermore, wet-day frequencies were adjusted to correspond with the observed one before fitting the gamma distribution. Although several parametric copula models could in principle be used, Gaussian copula was adopted due to its relatively simple formulation and as it is capable of describing both positive and negative correlations. As corrections on wet days need to be conditioned on either temperature or precipitation, two options to apply this method are possible. Based on initial tests with the algorithm, a decision was made to first adjust temperature and then correct precipitation conditionally on temperature, as this approach resulted in slightly better results.

The second bi-variate method (B2) tested in this thesis has been recently developed by Cannon (2018) based on a multidimensional probability distribution function transformation (N-pdft) algorithm originally developed for image processing purposes (Pitié et al., 2007). The algorithm is a genuine multivariate method in the sense that it can be applied to any N-dimensional probability distribution and is therefore not restricted to temperature and precipitation. The motivation behind the algorithm is to reduce the multidimensional problem into a sequence of one-dimensional corrections using an iterative algorithm. Let us define $N \times 2$ matrices containing the observed ($\mathbf{X}_o$) and

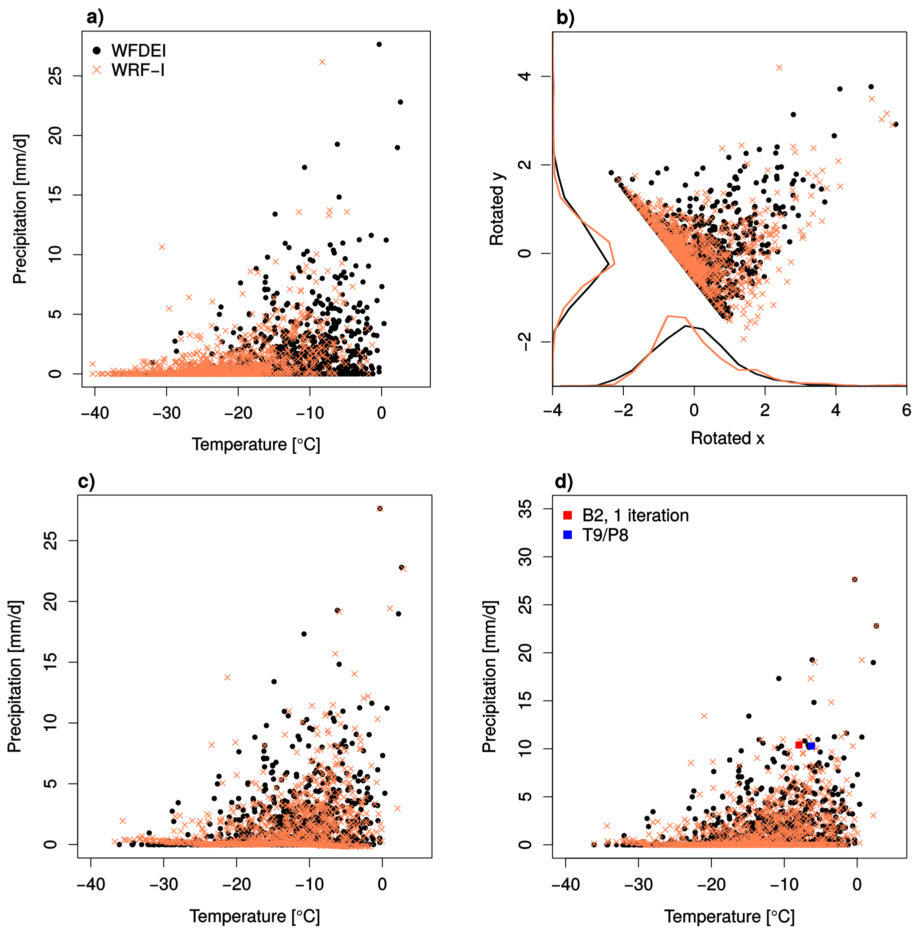

Figure 4: An illustration of how method B2 adjusts a modeled (WRF-I) joint distribution towards a target (WFDEI) distribution (panel a) within one iteration step. Panel b): apply standard quantile mapping to the modeled marginal distributions, normalize both joint distributions and transform them to a new orthogonal system (e.g., 45° rotation). Match the projected marginals (red and black curves) using quantile mapping (T9). Panel c): rotate both joint distributions back to the original coordinate system. This returns the actual temperature and precipitation marginals as a linear combination of the projected marginal distributions. Panel d): replace the quantiles of the adjusted GCM-RCM distribution, shown in panel c), with those obtained using standard quantile mapping (T9/P8) to get the marginals correct (e.g., remove the negative precipitation values in panel c)) but simultaneously keep the corrected dependence structure, illustrated as the different locations of red (B2) and blue (T9/P8) points in panel d).

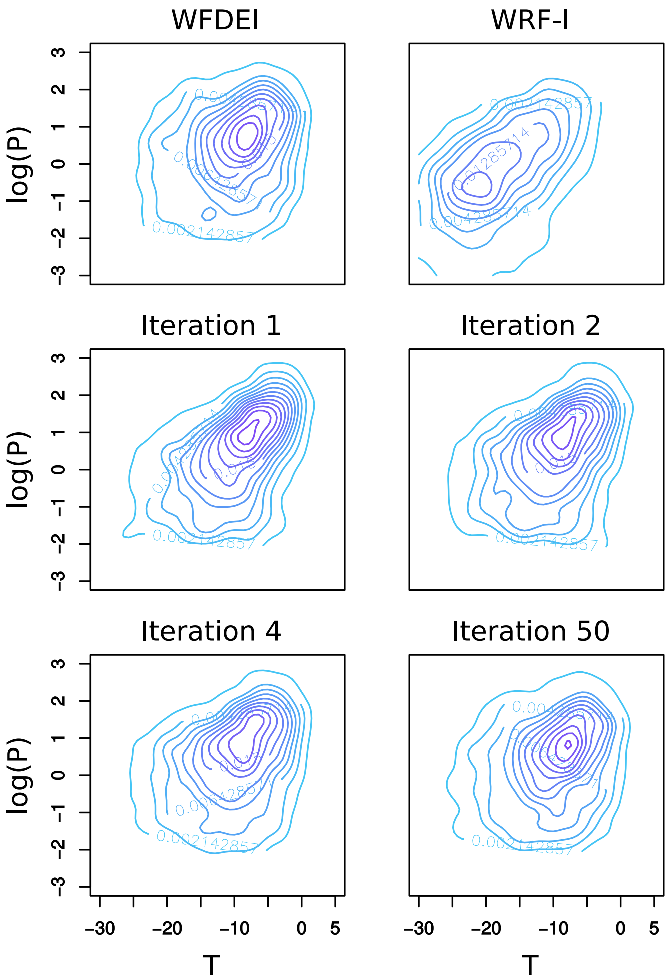

Figure 5: An example illustrating the convergence of B2, when correcting the joint distribution of daily mean temperature and daily precipitation on wet days (P$>$ 0.1 mmd^{-1}) simulated by a GCM-RCM (WRF-I) against a reference (WFDEI).

simulated temperature and precipitation in the baseline ($\mathbf{X}_c$) and scenario period ($\mathbf{X}_s$) with variables arranged as columns. As temperature and precipitation have different physical dimensions, both variables are normalized to have zero mean and unit standard deviation before applying the corrections to them. The algorithm for the jth iteration cycle is then as follows (see Fig. 4 for illustration). First, the observed and modeled joint distributions of temperature and precipitation are transformed to a new orthogonal coordinate system by rotating the marginal distributions with a random orthogonal matrix $\mathbf{R}$ following

$$\begin{aligned}\tilde{\mathbf{X}}_o^j &= \mathbf{X}_o^j\mathbf{R}^j\\ \tilde{\mathbf{X}}_c^j &= \mathbf{X}_c^j\mathbf{R}^j\\ \tilde{\mathbf{X}}_s^j &= \mathbf{X}_s^j\mathbf{R}^j.\end{aligned} \tag{10}$$

The rotation step essentially provides a projection of the original marginal distributions from a particular direction. Secondly, the rotated matrices containing data for the baseline and scenario period simulations are adjusted against $\tilde{\mathbf{X}}_o^j$ using traditional quantile mapping (method T9) to obtain $\hat{\mathbf{X}}_c^j$ and $\hat{\mathbf{X}}_s^j$, respectively. The adjusted marginal distributions are then transformed back to the original coordinates and used as input in the next iteration cycle:

$$\begin{aligned}\mathbf{X}_o^{(j+1)} &= \mathbf{X}_o^j\\ \mathbf{X}_c^{(j+1)} &= \hat{\mathbf{X}}_c^j\mathbf{R}^{j\,(-1)}\\ \mathbf{X}_s^{(j+1)} &= \hat{\mathbf{X}}_s^j\mathbf{R}^{j\,(-1)}.\end{aligned} \tag{11}$$

This step provides the adjusted joint distribution as a linear combination of the original marginals, which allows to modify the 2-dimensional distribution. The iteration is repeated until the full joint distribution $\mathbf{X}_c^{(j+1)}$ has converged sufficiently close to the target distribution (i.e., $\mathbf{X}_o$). In practice, both marginal distributions are first corrected separately using quantile mapping. Based on convergence properties inspected by Cannon (2018), the algorithm is terminated after 50 iterations (Fig. 5). A negative side effect of the rotation step is that the ratio property of precipitation is lost (i.e., not bounded by zero) after the iterative transformation. As the final step, quantiles of the adjusted marginal distributions are replaced with the quantiles obtained with traditional quantile mapping, which restores this property, simultaneously preserving the corrected ranks of both variables. In the original paper Cannon (2018) used quantile delta mapping, which retains the modeled change separately for each quantile. However, as practically any form of quantile mapping can be used, the same form of quantile mapping (methods T9/P8) used in **Papers I and II** was applied for consistency reasons in the final step.

2.3 Uncertainties related to the use of MOS methods

Projected hydrological (or any climatic) impacts are enveloped by a multi-layered cascade of uncertainty, which expands at each step in the impact modeling chain. The

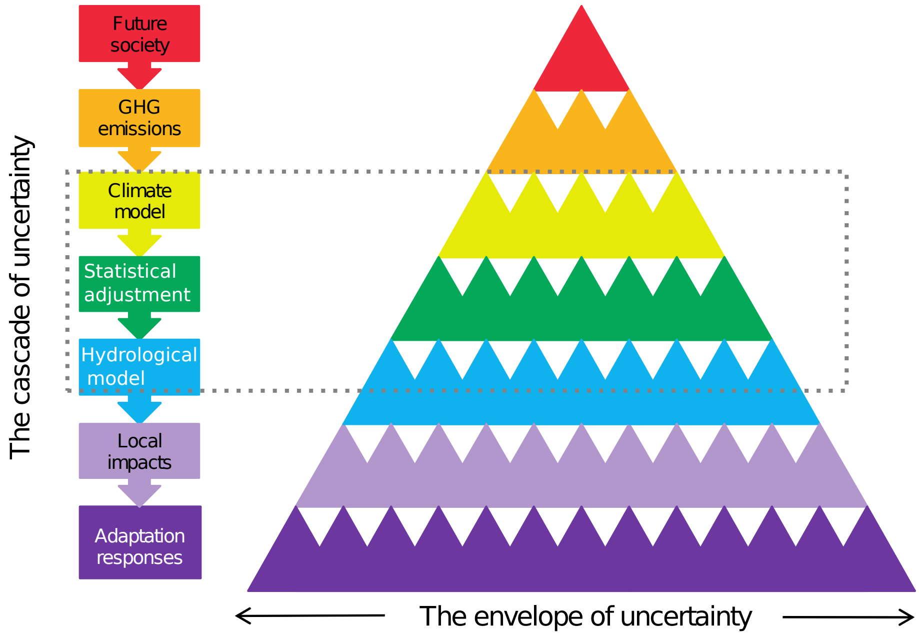

Figure 6: A top-down view of the cascade of uncertainty in hydrological climate change impact assessments. Layers relevant for this thesis are indicated with the gray box. Modified from Wilby and Dessai (2010).

proper communication of these uncertainties is an essential part of delivering climate change information for stakeholders as well as other users of this information (Hewitson et al., 2014). A simplified top-down view of identified sources of uncertainty and how they cascade downwards in hydrological climate change impact modeling chain is provided in Fig. 6. At the top levels of the pyramid, uncertainties are related to the uncertain evolution of anthropogenic greenhouse gas emissions, didacted by how the future human society will develop. In the next step, additional uncertainties arise from differences in GCM and RCM structures and parameters, differences in the response between the climate models to external forcing as well as uncertainties caused by the internal dynamics of the climate system (e.g., Räisänen, 2007; Hawkins and Sutton, 2011; Lafaysse et al., 2014). Furthermore, different approaches to statistically post-process climate model simulations for hydrological modeling purposes introduce an additional layer of uncertainty, which is reflected in the hydrological simulations, estimated impacts and adaptation responses as well.

The most relevant layers in the pyramid for this thesis (gray box) are those describing climate modeling and statistical post-processing uncertainties. These have been documented to certain extent in the literature both in global (Chen et al., 2011; Hagemann et al., 2011) and regional (Dobler et al., 2012; Teutschbein and Seibert, 2012; Bosshard et al., 2013; Lafaysse et al., 2014) scales. Additionally, it has been shown that although the climate model component is in many cases the dominant part, uncertainties related to MOS methods should also be covered, not the least due to the difficulty to select a single universally well performing method (**Paper I** and **II**). However, regional scale spatial variations in the relative importance of these two components in Europe, and how they are reflected in hydrological climate change simulations in the Scandinavian region have received less attention.

Part of this thesis concentrates on evaluating the relative importances of GCM-RCM and MOS-method uncertainties to future projections, with the main weight being on the aforementioned aspects. In **Papers I** and **II** the potential to reduce MOS-method uncertainty by excluding less well performing methods when constructing future projections is also tested. To be able to quantitatively estimate the contributions of GCM-RCM and MOS-method differences to the overall spread in the future projections, a simple analysis of variance (ANOVA) framework was applied in **Papers I** and **II** similarly to Yip et al. (2011) and Déqué et al. (2012). In this approach, the overall variance (V_{TOT}) in the future projections is decomposed according to the equation

$$V_{TOT} = V_{MOD} + V_{MET} + V_{INT}, \tag{12}$$

where V_{MOD} denotes the fraction of variance due to GCM-RCM differences, V_{MET} the fraction due to the MOS-method differences and V_{INT} denotes the interaction term caused by non-linear dependences between the GCM-RCMs and MOS methods. The main shortcoming of this approach is that internal climate variability is not explicitly separated from the actual GCM-RCM and MOS-method uncertainties and is confined in the three terms on the right-hand side. It should be stressed that the results of the variance decomposition should be associated only with the limited ensemble of model simulations method used in this thesis, with potential dependences between the GCM-RCM simulations and particularly between individual methods.

A more detailed evaluation of the relative contributions of GCM-RCM and MOS-method uncertainties to hydrological climate changes was made in the Scandinavian region in **Paper III** by inspecting spatial variations of the partition described in Eq. 12 on a seasonal level. Instead of inspecting the absolute values in the future climate, variance decomposition of daily mean temperature, precipitation and a set of hydrological variables was done on projected changes following Bosshard et al. (2013). In addition to decomposing changes in the annual cycle of river discharges to the aforementioned contributions, also total runoff and evapotranspiration were studied to cover the most important parts of the water balance.

2.4 Separation of temperature and precipitation effects on river discharges

A similar ANOVA framework was applied to decompose the annual cycle of the climate change signal in river discharges into separate parts caused by temperature and precipitation changes, respectively. This separation was done in order to assess, which one of the driving variables contributes more to the simulated changes in river discharges and how this differs between different methods and GCM-RCMs. To this end, three sets of simulations were needed in addition to the hydrological simulation for the present-day conditions: two sets of simulations, in which either the temperature or precipitation changes were included in the forcing data and a set of simulations including projected changes for both temperature and precipitation. The decomposition was only applied to projections made with two delta change methods (T1/P1 and T4/P4). One should note that due to non-linear processes in the hydrological model, the simulated changes in (e.g.) river discharges cannot be exactly divided into components arising solely from temperature and precipitation changes. However, the results showed that the non-linear part in simulated river discharge changes was in most cases relatively small, so only the linear part of this decomposition was assessed in **Paper III**. The details of the separation of the effects of temperature and precipitation can be found from Bosshard et al. (2014).

3 Data sets

3.1 Reference data

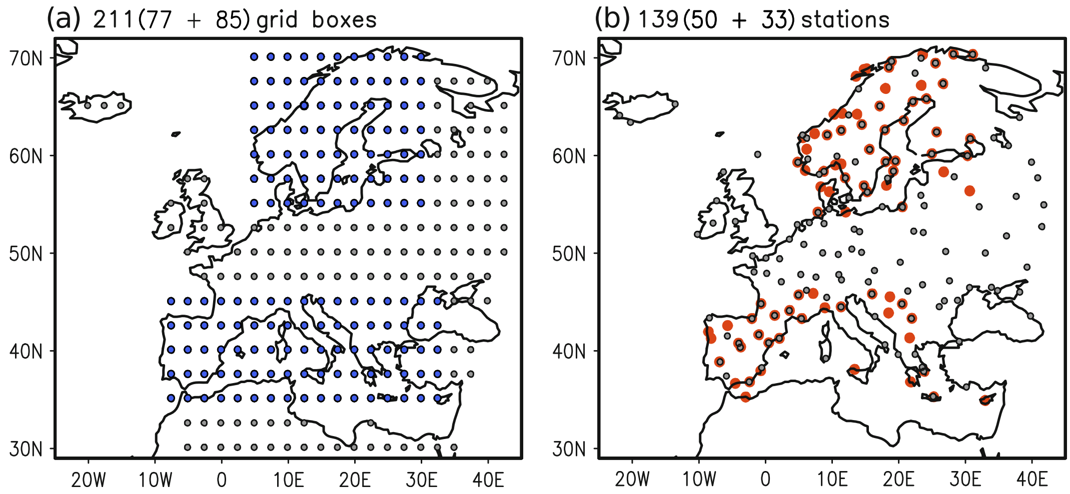

Figure 7: Grid boxes sampled for the (left) cross-validation exercises and stations selected for constructing real-world projections in **Papers I and II**. The gray (colored) dots denote the grid boxes used for daily mean temperature (daily precipitation). Adapted from **Papers I and II**.

Tests with real-world projections conducted in **Papers I** and **II** were based on blended station time series downloaded from the European Climate Assessment & Data archive available through the Royal Netherlands Meteorological Institute Climate Explorer (`climexp.knmi.nl/`). The locations of the used stations are shown in Fig. 7. In **Paper III** the reference data has been taken from the ERA-interim re-analysis (Dee et al., 2011), with additional adjustments to the monthly precipitation made using Global Precipitation Climatology Centre gridded precipitation (Schneider et al., 2011). The reference data used in **Paper IV** was taken from the ERA-interim based WATCH forcing data set (WFDEI). Monthly mean values have been adjusted to biases in comparison to gridded observations with additional elevation correction applied for temperature. The details of the WATCH methodology are documented in Weedon et al. (2014).

3.2 GCM-RCM simulations

In **Papers I** and **II**, six GCM-RCM simulations obtained from the ENSEMBLES project data base (`http://ensemblesrt3.dmi.dk/`) were used when constructing future projections (van der Linden and Mitchell, 2009). Each model simulation had a different GCM and RCM component and had been run using Special Report on Emissions Scenarios (Nakicenovic et al., 2000) A1B emission scenario forcing (Table 2). To reduce the computational needs, grid boxes were sampled with 2.5° interval in both longitudinal and latitudinal direction. The same model simulations were used as forcing in hydrological simulations conducted for the Scandinavian region in **Paper III**.

In **Paper IV**, latest high-resolution ($0.11° \times 0.11°$) simulations produced by the European branch of the Coordinated Regional Climate Downscaling Experiment (Jacob et al., 2014; Kotlarski et al., 2014) available in the Earth System Grid Federation (ESGF) nodes (`https://esgf.llnl.gov/`) were used as future forcing in hydrological simulations. In total, five GCM-RCM simulations, each again having a different GCM and RCM part and run with Representative Concentration Pathways (Moss et al., 2010) RCP4.5 scenario forcing, were selected from the data base. Although the number of "independent" models was initially larger, the most poorly performing models were dropped after initial tests due to their highly biased simulations. It should be noted that the GCM-RCM simulations used in **Paper III** and **Paper IV** might not have optimal performance from hydrological modeling perspective due to the constraints posed by other selection criteria (i.e., having separate GCM and RCM components in each model).

3.3 Hydrological simulations

Hydrological simulations analyzed in **Papers III** and **IV** were made with the Hydrological Predictions for the Environment (HYPE) model developed in the hydrology research group in Swedish Meteorological and Hydrological Institute (SMHI). The model is a process-based, semi-distributed model, designed for hydrological simulations at varying spatial scales. A detailed description of the model can be found from Lindström et al. (2010). The model settings used in these studies were extracted from the European scale application of HYPE (E-HYPE). This application has a median sub-basin size of 215 km^2 and has been built to cover also ungauged regions (Don-

Table 2: A List of the GCM-RCM simulations used in this thesis.

Institution	GCM	RCM	Abbreviation
ENSEMBLES ($0.25° \times 0.25°$ resolution), **Papers I**, **II** and **III**			
CNRM	ARPEGE	ALADIN	CNRM-A
ETHZ	HadCM-Q0	CLM	ETHZ-H0
Met Office	HadCM-Q3	HadRM-Q3	METO-H3
Met Office	HadCM-Q16	HadRM-Q16	METO-H16
MPI	ECHAM5	REMO	MPI-E5
SMHI	BCM	RCA3	SMHI-BCM
EURO-CORDEX, ($0.11° \times 0.11°$ resolution), **Paper IV**			
CNRM	CNRM-CM5	ALADIN	CNRM-A
CLM Community	MPI-ESM-LR	CCLM4.8.17	CCLM-MPI
KNMI	EC-EARTH	RACMO22E	RACMO-EC
SMHI	HadGEM	RCA4	RCA4-H
IPSL-INERIS	IPSL-CMA5-MR	WRF	WRF-I

nelly et al., 2016). All the necessary topographical, land-use and soil data have been extracted from freely available data sets, as described in Donnelly et al. (2016). The source code for the model, along with the relevant model documentation, is available at `http://hypeweb.smhi.se/`.

In **Paper III**, a sub-model covering Scandinavia and the adjacent regions was extracted from the full E-HYPE domain with settings available at that time (version 2.1). While the original model setup was mostly used, an attempt was made to partially re-calibrate the model due to the negative bias in river discharges, apparently caused by the overestimation of evapotranspiration in this particular model version. For conducting pseudo-reality tests in the hydrological modeling phase, four sub-models from different hydroclimatic conditions were selected from E-HYPE (version 2.5) in **Paper IV** (Fig. 8). These sub-models were selected from basins with natural flow conditions in order to ensure that the model is capable to capture the observed flow conditions without a need to re-calibrate the model. This also avoided the hampering effects of flow regulation on the assessment of the relative MOS-method performance from a

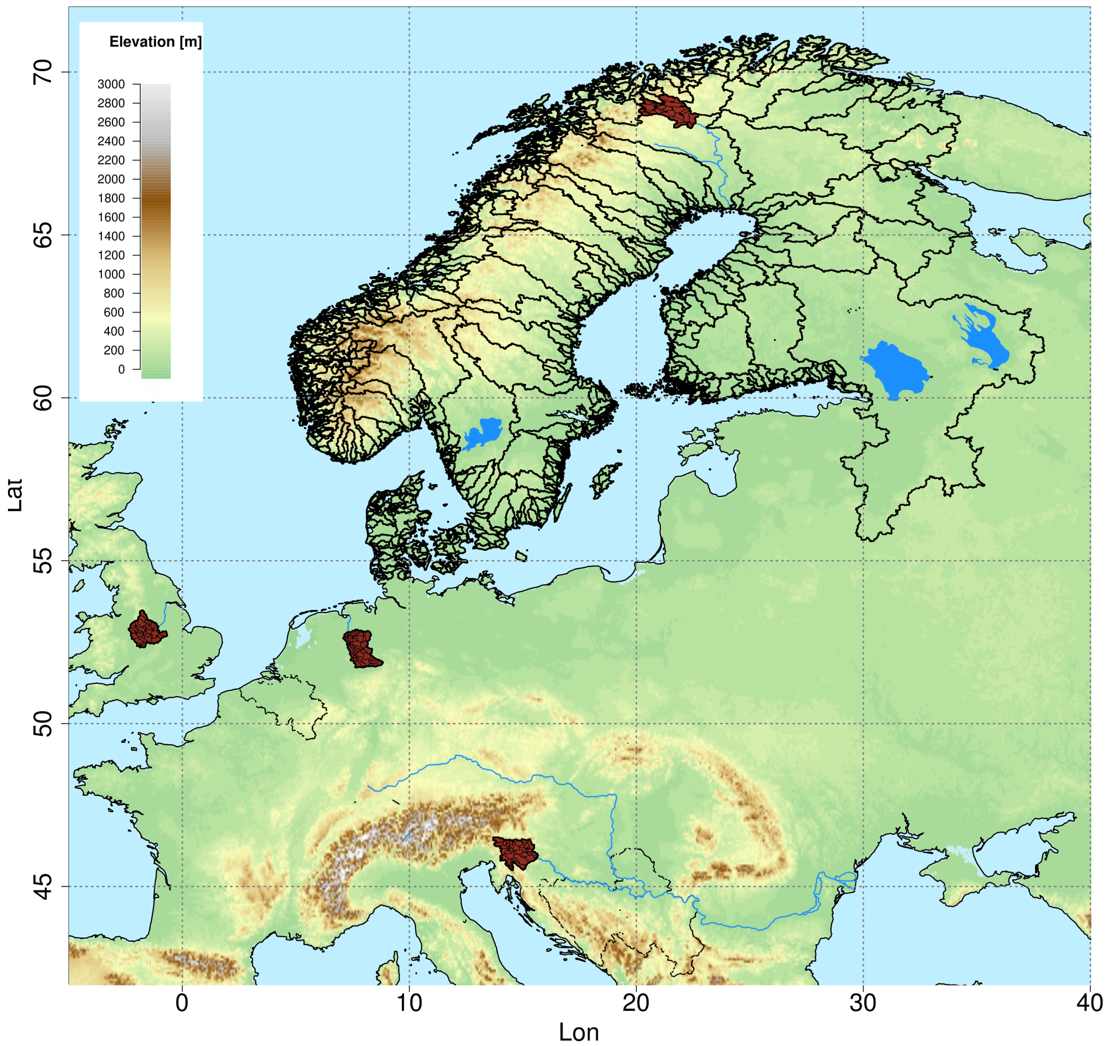

Figure 8: An illustration of the hydrological modeling domains in **Papers III** and **IV**. The black lines denote the individual river catchments selected from the full-scale E-HYPE in **Paper III**, while the four sub-domains used in the pseudo-reality tests (**Paper IV**) are marked with red color in the figure.

hydrological modeling point of view.

In both papers the model was run at daily time step using only daily mean temperature and daily precipitation as input. In **Paper III** the study period was constrained to years 2041-2070, while **Paper IV** covers two periods from early (2011-2040) and late (2061-2090) 21^{st} century conditions.

4 Cross-validation and pseudo-reality approach

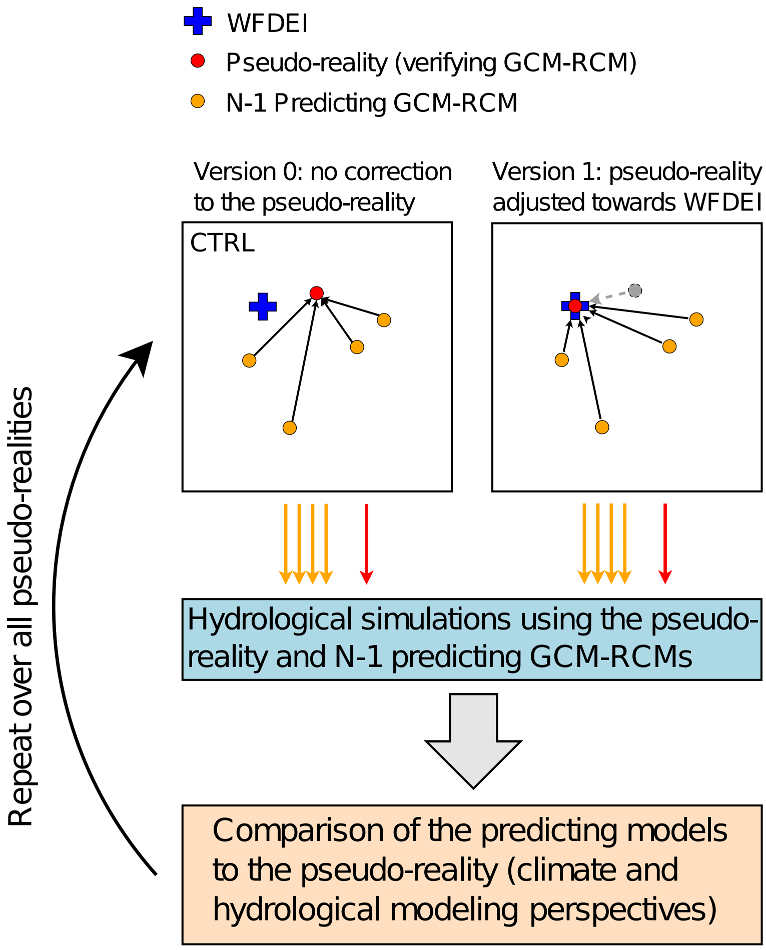

Figure 9: A schematic description of the pseudo-reality framework, covering both climate modeling and hydrological modeling perspectives. Adapted from **Paper IV**.

In most inter-comparison studies, the evaluation of MOS-method performance has been based on split-sample tests or, equivalently, cross-validation using observational data sets to validate the methods. If the aim is to assess the relative method performance in future climatic conditions, observations cannot obviously be used as the reference. However, some conclusions can be drawn using climate model simulations as proxies for future climate. This approach was used to cross-validate the studied methods in **Papers I**, **II** and **IV**. The main assumption behind this approach is that the model simulations are independent from each other and are plausible realizations of the actual future climate (i.e., the statistically indistinquishable ensemble paradigm (Annan and

Hargreaves, 2010)). From an ensemble of GCM-RCM simulations, each model at its time is selected to represent observations for the future, against which the rest of the GCM-RCM simulations are then validated. When this procedure is cycled over all GCM-RCM permutations, cross-validation statistics can be used to compare the relative ability of the studied methods in capturing future climatic conditions.

As the bias corrected climate simulations are eventually used as input in hydrological simulations, the evaluation should ideally cover also the impact modeling step. To this end, the pseudo-reality framework applied in **Papers I** and **II** was extended in **Paper IV** to cover the evaluation of the selected MOS methods in a cross-validation manner from hydrological modeling perspective. The full procedure is illustrated in Fig. 9. An attempt to further improve the applicability of the pseudo-reality approach in the impact modeling step was made based on the fact that systematic biases in uncorrected GCM-RCM simulations might lead to unrealistic shifts in hydrological regimes (e.g. Wood et al., 2004; Rojas et al., 2011). In addition to the standard pseudo-reality approach (version 0 in Fig. 9), each day of the annual cycle of both input variables in pseudo-reality was first adjusted towards WFDEI using a 30-day sliding window when assessing the bias at each day (version 1 in Fig. 9). This simple adjustment removed systematic biases both from temperature and precipitation time series in the pseudo-reality, although with the expense of slightly modifying the spread of the pseudo-reality precipitation distributions. The only study known by the author in which the pseudo-reality approach covered hydrological modeling results when assessing bias correction method performance was made by Velázquez et al. (2015), although with a smaller set of MOS methods and GCM-RCMs and with the slightly different aim to directly estimate the implications of bias non-stationarity to the use of bias corrected time series as input in future hydrological simulations.

To comprehensively cover different aspects of the projected temperature and precipitation distributions, several verification statistics were used in **Papers I**, **II** and **IV**, when assessing the relative MOS-method performance. For clarity, the results are here illustrated in terms of both the mean squared error (MSE) and mean absolute error (MAE) between the verifying model (i.e., pseudo-reality) and the average projection of the predicting models.

5 Results

5.1 Intercomparison of univariate MOS methods for daily mean temperature and precipitation

The relative performance of MOS methods designed for daily mean temperature in changing climatic conditions was investigated in **Paper I**. When looking at the results for years 2069-2098 (Fig. 10) it is evident that the simplest time mean delta change and bias correction (T1 and T6) perform poorly in comparison to other methods and have the largest overall MSE in their projections. When the spread of the distribution is taken into account, the results are markedly improved (T2 and T7). On the other hand, the inclusion of distribution skewness improves the results rather modestly (T3 and T8). Overall, the MSE is smallest for the four quantile mapping methods (T4-T5 and T9-T10), non-parametric quantile mapping applied in the bias correction mode (T9) having the best performance out of all methods.

To obtain further insights to reasons for the differences in the relative method performance, dark bars in Fig. 10 show the MSE which would have been obtained if predicting the 30-year mean temperature perfectly in the scenario period. They readily highlight the fact that most of the MSE in daily mean temperature projections is related to biases in time-mean temperature. Furthermore, the slightly better performance of quantile mapping methods and particularly T9 is apparently due to their ability to modify the projected mean temperature, whereas other methods provide identical results for changes in the temporal mean. This is a generic difference to other methods as discussed in **Paper I** and might be a desirable property, if biases in daily variability affect the time mean change (Boberg and Christensen, 2012).

To assess whether an increased sample size would reduce the effect of random noise on future projections, one-, two- and three-month time windows were tested when estimating monthly biases and simulated changes. Figure 10 shows that using two-month time window typically results with better cross-validation statistics than when one-month time window is used. However, differences between the two- and three-month windows are unsystematic, which indicates that a relatively small window size is sufficient for daily mean temperature when constructing monthly projections.

Although quantile mapping applied in the bias correction mode (T9) has a superior

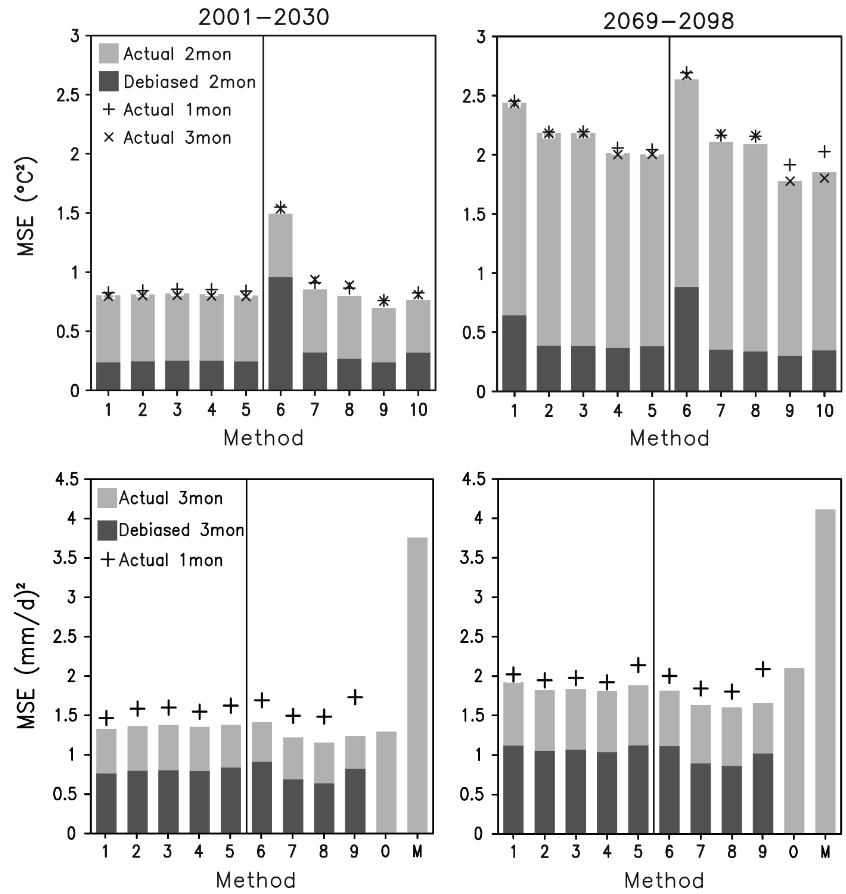

Figure 10: Cross-validated MSE for (top) daily mean temperature and (bottom) daily precipitation in years (left) 2001-2030 and (right) 2069-2098 when averaged over the whole distribution, 12 months and the whole region. For temperature, bars denote the values for two-month time window. Statistics obtained with (plusses) one-month and (crosses) three-month time window are also shown. In the case of precipitation, only (plusses) one-month and (bars) three-month time window has been used. In addition to the overall MSE (light bars), MSE that would have been obtained if the mean value had been perfectly captured by the methods are also shown (dark gray bars). For precipitation, cross-validation statistics obtained using uncorrected model simulations (M) and the pseudo-reality observations with no adjustment for climate change (O) when predicting the future climate are shown for comparison. Adapted from **Papers I and II**.

performance in years 2069-2098, the differences are substantially smaller in the earlier scenario period (2001-2030). Furthermore, the less complicated methods, particularly the simple time mean delta change (T1), show a much better performance in comparison to quantile mapping methods, as changes in the shape and the spread of the distribution emerge more slowly from the background noise. Yet, simple mean bias correction (T6) has a substantially poorer performance than the other methods due to the unrealistic assumption of same bias in all parts of the temperature distribution.

While distribution-averaged statistics illustrate the overall competitive ability of T9, a more detailed inspection of the relative method performance in different parts of the distribution illustrates the inherent difficulty of selecting a universally well performing method. In the extreme upper tail, the relative performance of methods T9 and T10 deteriorates and they are outperformed by the two quantile mapping delta change methods (T4 and T5), as shown in Fig. 8 of **Paper I**. This is likely related to extrapolation issues, as the late 21st century temperatures exceed the present-day range by a substantial margin. In addition to variations between the different parts of the distribution, the details of the results depend on region and season considered.

The relative performance of MOS methods designed for adjusting the frequency distribution of daily precipitation (**Paper II**) is shown in the bottom row of Fig. 10. To better infer the added value of the MOS step in comparison to using raw model data, bars showing the MSE for the uncorrected model simulations (M) and for the case where the verifying model (i.e., pseudo-reality) simulation in the control period is used as an estimate of future climate (O) were added. The MSE values are roughly twice as large for M as they are even for the poorest performing methods in both scenario periods, showing the substantial reduction in distribution errors with MOS. On the other hand, although using "observations" directly as the estimate of the future climate is not reasonable in the later scenario periods, the difference is much smaller in the near-future projections. This is not surprising, as the climate change signal in daily precipitation has been shown to emerge rather slowly from the background noise in the European region (Maraun, 2013), which suggests that the added value of MOS might be relatively small when estimating short-term climate changes, particularly in precipitation.

When compared against the MSE in daily mean temperature distribution, it is evident that precipitation benefits more strongly from the use of a wider time window when estimating simulated changes or model biases; all methods have a substantially smaller MSE when three-month time window is used. In particular, the parametric quantile mapping (P5/P9) has a smaller MSE, as the estimation of distribution parameters is more robust. Similarly to daily mean temperature, non-parametric quantile mapping applied in the bias correction mode (P8) tends to outperform other methods. However, differences between individual methods are typically smaller than for daily mean temperature, and methods P7 and P9 have MSE values relatively close to P8. Another difference to daily mean temperature is the relatively larger contribution of errors in the distribution around the mean (the dark gray bars in Fig. 10), which explains over half of the overall MSE in the projections. Thus, the relatively good performance of P8 is likely related to its ability to flexibly adjust the full distribution shape.

The results in the earlier scenario period (2001-2030) are broadly similar, P8 again having the best performance. Although the relative performance of delta change methods improves in comparison to bias correction methods, the improvement is smaller than what was seen in the case of daily mean temperature. The largest exception is the simple time mean delta change (P1), which actually has the best performance within its group (P1-P5) at that time. This is probably related to the aforementioned low signal-to-noise ratio, which more generally reduces the added value of near-term climate change information.

Partially in contrast to daily mean temperature, variations in the relative performance of MOS methods between different parts of the frequency distribution are noisier for daily precipitation. Although P8 has the best performance in most parts of the distribution of daily precipitation, methods P7 and P9 outperform P8 above the 96th percentile (Fig. 9 in **Paper II**). The spatial and temporal variation in the relative method performance were also studied in more detail **Paper II**. The main message from these tests is that it is difficult to define a single universally well performing method, which suggests that at least part of the uncertainty related to the statistical post-processing of GCM-RCM simulations is unavoidable. Thus, (ideally) several methods should be used when constructing future projections for end user needs.

An intuitive approach to further reduce random errors in the projected temperature and precipitation distributions and also to improve projections in probabilistic terms would be to combine projections obtained with several well-performing methods. Tests with different method combinations in **Papers I** and **II** showed that the errors are indeed reduced when combining several methods, the combination of the four quantile mapping methods (T4/T5 and T9/T10) having the best statistics in the case of daily mean temperature. Broadly similar conclusions were obtained in **Paper II** for daily precipitation, although improvements in comparison to individual methods, especially P8, were less clear. Although differences between the tested method combinations were small, the results showed that using a larger set of methods would likely provide better results even for daily precipitation, with the expense of potentially introducing larger errors in the extreme upper tail of the distribution.

5.2 Intercomparison of uni- and bi-variate methods

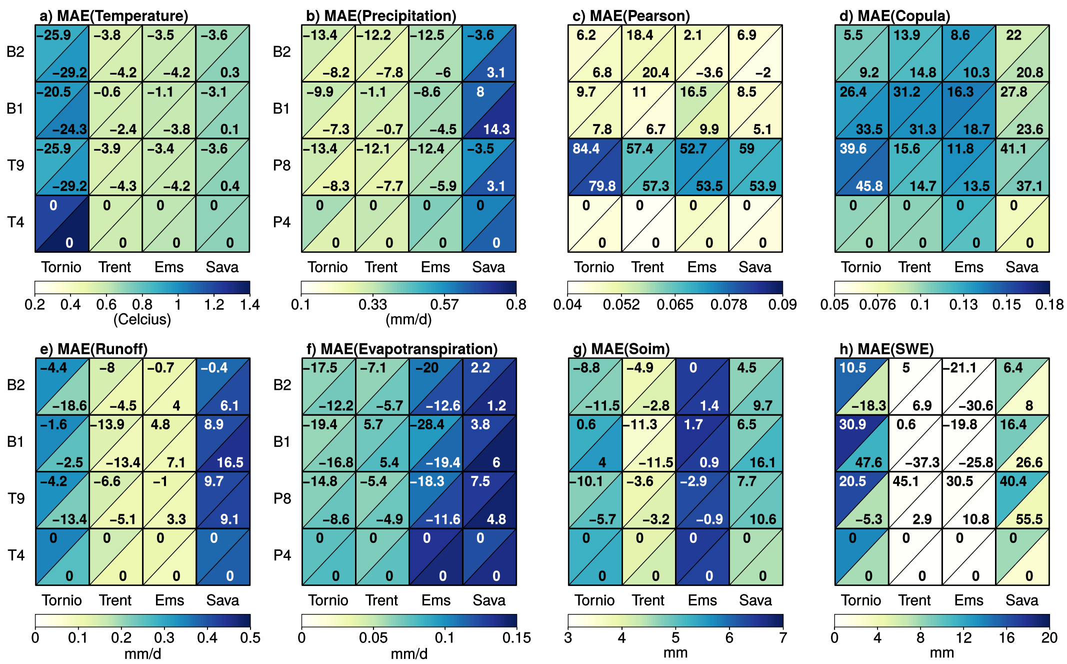

Figure 11: Cross-validated MAE for (a) daily mean temperature and (b) daily precipitation, (c) Pearson correlation between temperature and precipitation, and (d) full copula structure in years 2061-2090 displayed separately for each HYPE sub-model. Two-month time window has been used to estimate model biases and simulated changes. Similar panels are also shown for the MAE in monthly mean (e) total runoff, (f) evapotranspiration, (g) soil moisture and (h) snow water equivalent simulated by HYPE. The values in each triangle show the per cent difference in MAE when compared against quantile mapping applied in the delta change mode (T4 and P4). Results for the pseudo-reality approach without (with) corrections to WFDEI are shown in the top (bottom) triangles. Based on data from **Paper IV**.

Comparisons between uni-variate quantile mapping (T4/P4 and T9/P8) and the two bi-variate methods (B1 and B2) (Fig. 2 in **Paper IV**) showed that, by design, bi-variate bias correction methods are capable of reducing errors in those correlation aspects they have been designed to address, when compared against WFDEI in the present-day climate. However, adjustment of the full copula structure using copula-based bias correction (B1) fails in some cases. For example, Fig. 2 in **Paper IV** shows for that B1 tends to introduce additional errors to the dependence structure in winter months in the HYPE Tornio sub-model. This, together with larger biases in the marginal distributions (Fig. 3 in **Paper IV**) is also reflected as poorer performance in the hydrological modeling step, which indicates a limited applicability of this method in cold climates. In this particular example, method B2 improves some aspects of the hydrological simulations such as total runoff and SWE in comparison to univariate quantile mapping (T9/P8).

When looking at the cross-validation results for temperature and precipitation distributions in years 2061-2090, the relative method performance shows substantial spatial (Fig. 11) and temporal variations. On a general level, both univariate and bi-variate bias correction methods tend to perform better in the three northernmost domains, while quantile mapping delta change (T4/P4) has a relatively better performance in the southernmost HYPE sub-model Sava. The most striking feature in Fig. 11 is the better performance of T4/P4 in capturing the Pearson correlation coefficient and particularly the copula structure of temperature and precipitation in comparison to the other methods. This suggests that, in certain conditions, inter-variable dependences between daily mean temperature and precipitation might not change substantially with climate change.

One of the main hypotheses in **Paper IV** was that the inclusion of inter-variable correlations of daily mean temperature and precipitation in the adjustment step would improve the corresponding hydrological simulations. The results shown in the lower row of Fig. 11 indicate that this hypothesis is only partially true. The improvements in cross-validation statistics calculated from the hydrological modeling outputs are in many cases relatively modest for both bi-variate methods and show substantial temporal and spatial variations. When first looking at the total runoff, some improvements in comparison to quantile mapping (T4/P4 and T9/P8) are obtained using method B2, especially in Tornio. This is most likely related to its ability to better capture future SWE, which is further reflected in other aspects of simulated surface hydrology. Copula-based bias correction (method B1) is not without merits either. For example, it generally has the best performance in Trent, where it outperforms other methods with the exception of evapotranspiration. One of the most interesting results obtained in **Paper IV** is the competitive performance of quantile mapping applied in the delta change mode (T4/P4), even when compared against B2. While not the best performing method when adjusting the marginal distributions, cross-validation statistics for the hydrological simulations show that quantile mapping applied in the delta change mode (T4/P4) performs equally well or even better than the other methods in Ems and Sava.

The effect of bi-variate adjustments on future errors in simulated river discharges was further evaluated by calculating additional cross-validation statistics directly for the monthly river discharge distributions in the outlet model block of each HYPE submodel (Fig. 7 in **Paper IV**). Backing up the previous results, the differences were found out to be small in most distributional aspects and particularly when the distribution-averages were considered. A small improvement was obtained with B2 in the simulation of high flows, which is likely related to the improved simulation of SWE in Tornio and Sava. This further backs up the aforementioned results and underlines the difficulty to show the added value of bi-variate bias correction from hydrological modeling point of view.

5.3 Relative importances of GCM-RCM and MOS-method uncertainties

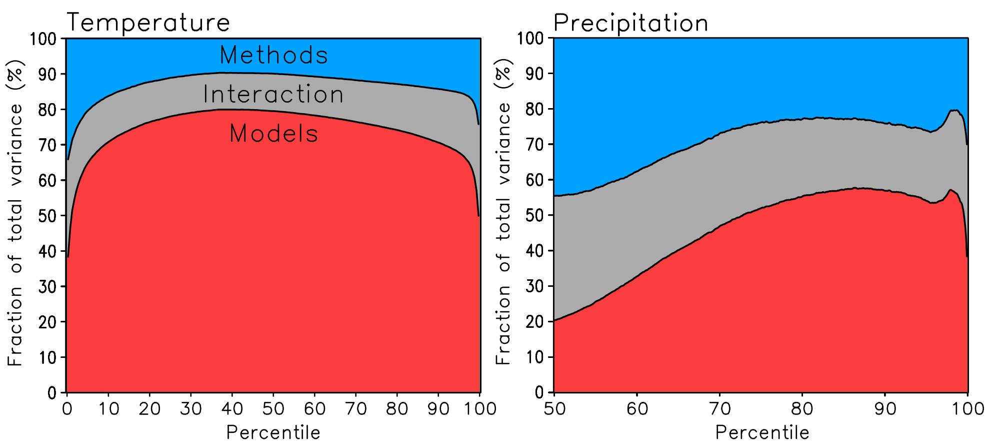

Figure 12: ANOVA partition of the total variance in the projected daily mean temperature and precipitation distributions into contributions from (red) GCM-RCM, (blue) MOS-method differences and (gray) the interaction term, when averaged over the 12 months and all stations shown in Fig. 7. Modified from **Papers I** and **II**.

Contributions of MOS-method and GCM-RCM uncertainties to the overall spread in different parts of projected frequency distributions were inspected from the climate modeling perspective in **Papers I** and **II** using ANOVA as described in Section 2.3. When averaged over all stations in Fig. 7, GCM-RCM uncertainties explain to large extent the spread between the different projections (Fig. 12), particularly in the middle parts of the distribution, where method differences are smallest. When averaged over the whole temperature distribution, V_{MOD} and V_{MET} explain 73 and 13 % of the total variance in the future projections. When the tails of the distribution are considered, however, the fraction of total variance explained by method differences increases noticeably and exceeds 30 % in the lower tail of the distribution. The results are qualitatively similar for precipitation, although with a generally larger contribution from MOS-method differences to the overall spread. On average, V_{MOD} and V_{MET} explain 46 and 29% of the overall variance. The interpretation of the results is somewhat complicated by the larger interaction term especially in the lower parts of the distribution, where V_{INT} tends to become the largest component.

As discussed in Section 5.1 and illustrated in **Papers I** and **II**, it is probably sensible to exclude methods, which perform poorly for well-identified reasons. When only the four best performing methods for daily mean temperature (T4, T5, T9 and T10) and the six best methods for daily precipitation (P2-P4 and P7-P9) are included, the fraction of variance explained by MOS-method differences decreases for both variables throughout the distribution. However, V_{MET} is still non-negligible for precipitation and in the tails of the temperature distribution. The general conclusion from these exercises is that while GCM-RCM differences are clearly of larger importance for applications where the distributional details are unimportant, MOS-method differences become increasingly important when the tails of the temperature and precipitation distributions are of particular interest for an end user.

To test to what extent the GCM-RCM and MOS-method uncertainties are reflected in hydrological simulations in the Scandinavian region, a similar decomposition for the simulated changes in river discharges was done using a subset of three method pairs for daily mean temperature and precipitation (T1/P1, T4/P4 and T9/P8) and the same set of GCM-RCM simulations as in **Papers I** and **II**. The results showed (Fig. 13) that the contributions of GCM-RCM and MOS-method differences are qualitatively similar to those obtained in **Paper I** and **II**, V_{MOD} explaining the major part of the overall variance in future changes. Figure 13 also indicates that the relative importance of

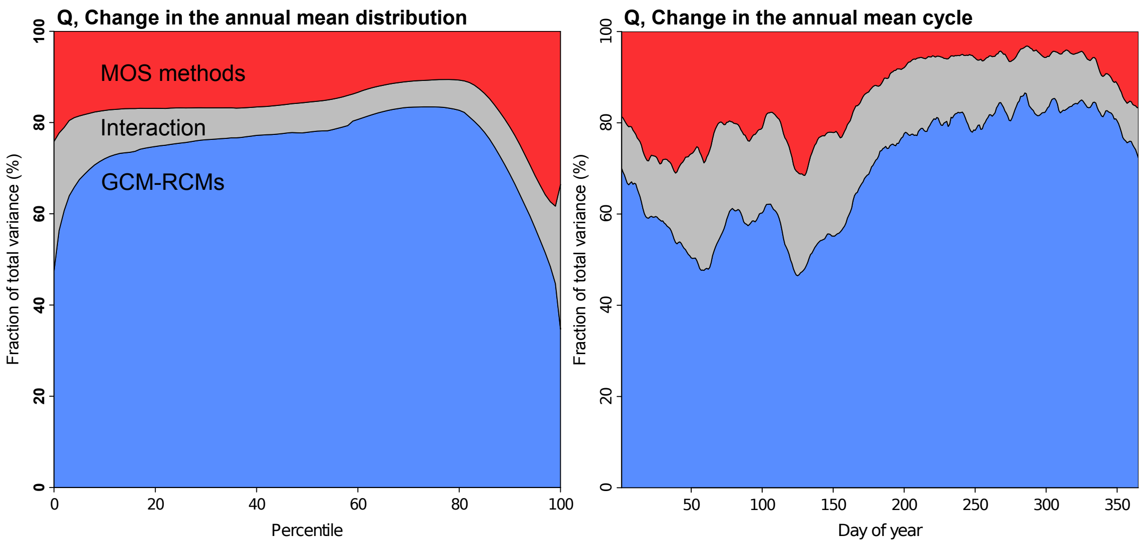

Figure 13: Similar to Fig. 12 but shown for the changes in river discharge distribution and the annual cycle of simulated changes in river discharges in years 2041-2070, when averaged over the whole model domain. Taken from **Paper III**.

MOS-method differences is particularly important for high and low flows. In addition to the distributional aspects, the contributions of the two sources of uncertainty on changes in the annual cycle were also inspected (right panel in Fig. 13). The relative importance of MOS-method uncertainty has a noticeable maximum in winter and early spring, which is likely explained by the sensitivity of the modeling of snow processes to the selected MOS method. For example, there are substantial differences between the simple mean bias correction and quantile mapping approaches when evaluated in terms of the snow cover duration, suggesting that the exclusion of daily variability from the adjustment step leads to exaggerated changes in snow water equivalent.

Figure 14 shows an example of spatial variations in the ANOVA components, when applying the partitioning to total runoff, evapotranspiration and the two input variables. MOS-method component has a maximum in both total runoff and evapotranspiration in the northern parts of the domain, which at least in the case of total runoff is likely linked to snow processes. This is further backed up by the absence of the maximum in summer. Also the relatively good inter-model agreement on temperature changes around this region in comparison to other parts of the domain might have a role in explaining this. For temperature and precipitation, the GCM-RCM component is visibly larger than the MOS-method component, which is in line with the results of **Papers I and II**.

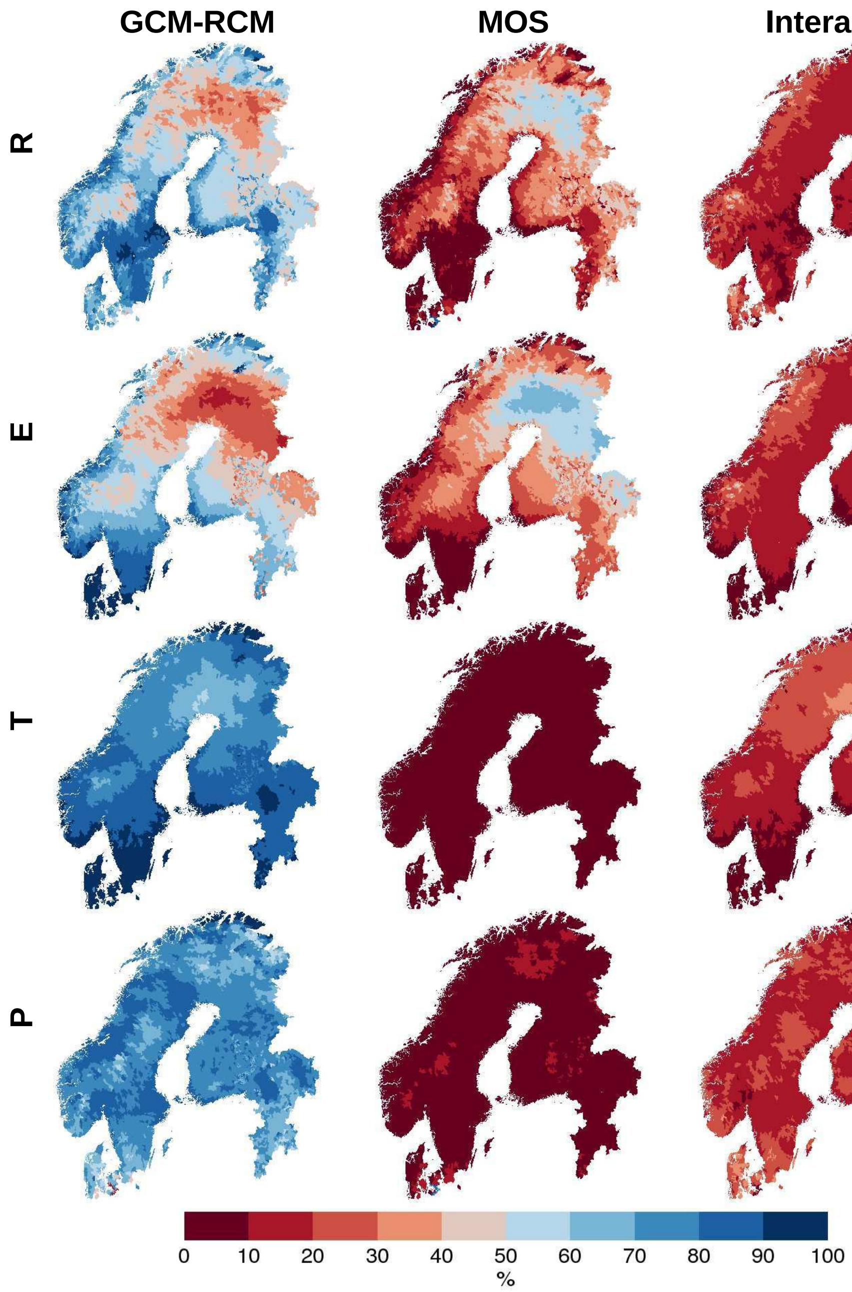

Figure 14: Spatial distributions the ANOVA components when calculated for the HYPE-simulated changes in total runoff (R), evapotranspiration (E) as well as daily mean temperature (T) and daily precipitation (P), shown for the winter months (Dec-Jan-Feb) in years 2041-2070. Taken from **Paper III**.

5.4 Temperature and precipitation effects to river discharge changes in Scandinavia

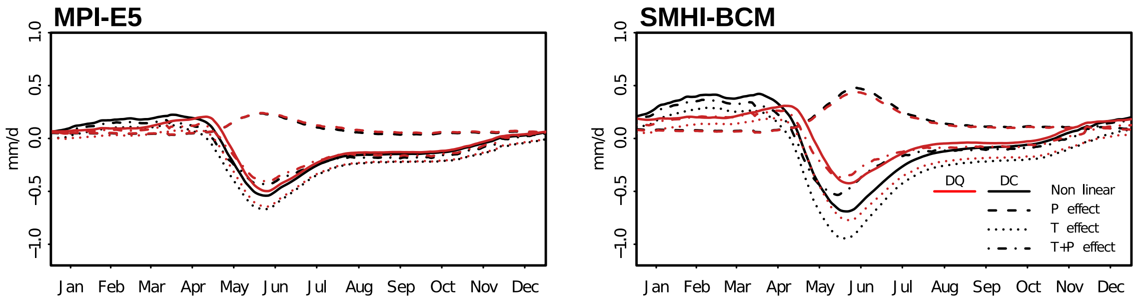

Figure 15: An example of the contributions of separate (short-dashed line) temperature and (long-dashed line) precipitation effects on changes in river discharges in the Scandinavian region from years 1980-2009 to 2041-2070, shown separately for (black) the simple delta change method and (red) quantile mapping applied in the delta change mode. Curves for (short-long dashed line) the sum of the separate temperature and precipitation effects and (solid lines) the overall non-linear change are also plotted in the figure. Adapted from **Paper III**.

Tests with the effect of separate adjustment of daily mean temperature and daily precipitation on river discharge changes showed that, in the Scandinavian region, temperature changes have a dominating effect with imposed changes both in timing and the magnitude (Fig. 15). The effect is positive in winter and early spring due to changes in snow accumulation and melting. On the other hand, increasing temperature reduces flow rates in summer and autumn, which is associated with enhanced evapotranspiration in a warmer climate. Daily precipitation changes have a positive effect on river discharges and act as to increase river discharges throughout the year, with the largest changes occurring in spring and early summer. The maximum in the response at that time is explained by the larger amount of water stored in snow pack, which is then released during the snow melt season. Differences between the two panels in Fig. 15 also illustrate that the magnitude and relative importance of both effects naturally depends on the GCM-RCM.

The results in Fig. 15 are shown separately for quantile mapping delta change T4/P4 (red curves) and the simple time mean delta change methods T1/P1 (black curves). The difference in the magnitude of the temperature effect between these two methods is mostly small, except in spring when the simple delta change method projects a larger decrease in river discharges. This indicates that changes in river discharge seasonality are projected slightly differently by these methods. Precipitation effect is almost identical for both methods with only small differences being visible in spring. The relative importance of temperature and precipitation effects on river discharges also shows substantial spatial variations, with the largest contributions from temperature changes occurring in large rivers in the northern parts of the model domain in winter and over most of the area in summer (Fig. 7 in **Paper III**). Conversely, the relative importance of precipitation changes is most important in the coastal areas and southern parts of the domain in winter. Overall, the temperature effect depends more strongly on whether the daily variability is included in the MOS adjustment step or not.

6 Conclusions and future directions

This thesis provides some of the first attempts to evaluate the relative performance of MOS methods, which aim to combine information obtained from observations and GCM-RCM simulations, in changing climatic conditions, using climate model simulations as pseudo-realities for the future when cross-validating the relative method performance. The main result obtained from the pseudo-reality studies (**Papers I**, **II** and **IV**) is that the identification of the best performing method is practically impossible; the relative method performance varies between different locations and seasons and also in different parts of the frequency distribution. Additional tests with more complex methods, which aim to cover inter-variable correlations in their adjustments, improve the correlation statistics only to a limited extent, as the improvement in cross-validation statistics is modest at the best. When compared against quantile mapping applied in the delta change mode, which retains the present-day correlation between temperature and precipitation, this simpler univariate method has comparable or even better performance than the tested bi-variate methods.

This study also represents one of the first attempts to incorporate the impact modeling perspective in the future assessments using hydrological modeling as an example. The results of these idealized tests indicate that care should be taken, when selecting suitable methods for specific impact study purposes due to similar reasons mentioned above. Again, bi-variate bias correction of temperature and precipitation provides only modest improvements, which are mostly limited to snow-dominated, cold climatic regions. In some cases, the simple assumption of unchanged correlations, as is associated with quantile mapping applied in the delta change mode, might be sufficient. It should be noted, however, that for a larger set of variables with stronger dependences between them, the added value of multivariate corrections is potentially larger.

It cannot be stressed enough that the previous evaluations should be accompanied with a proper assessment of MOS-method uncertainties in comparison to other uncertainty sources. This also provides some guidelines to what extent MOS-method differences should be covered, when assessing potential climate change impacts. When the relative importance of GCM-RCM and MOS-method differences to the overall spread in the future projections is compared, the GCM-RCM component generally has a larger contribution to the spread of both temperature and precipitation projections. However, the relative importance of MOS-method differences tends to increase when the extreme high and low temperatures and high precipitation intensities are considered. This has implications for those applications where the extremes are particularly important; ideally, several MOS methods should be included when constructing projections for the future climate. In line with the results for temperature and precipitation, the relative importance of MOS-method differences is largest for high and low flows, when future changes in river discharges are considered (**Paper III**).

A common limitation to all the papers in this thesis is the relatively small number of model simulations available for these studies. Thus, the results should not be generalized to cover other climate models and MOS methods. Another shortcoming in **Paper IV** is that although the adjustment of inter-variable correlations tends to modify the temporal sequencing, the pseudo-reality approach does not unfortunately allow the direct estimation of how this affects the method performance when assessed from a hydrological modeling perspective. Previous studies suggest that the preservation of temporal structure could, at least partially, explain the surprisingly good ability of T4/P4 to capture the future hydrological conditions (Chen et al., 2013).

Because it is unlikely that raw climate model simulations will be directly usable for most end user purposes in the near future, there is continuing need to use and further develop MOS and other types of statistical methods to bridge the gap between the end user needs and the information provided by climate model simulations. In particular, there is an on-going discussion to what extent bias correction should be guided by a better understanding of physical mechanisms behind the prevailing model biases; as a purely statistical tool bias correction is "blind" to the underlying physics. This topic has been discussed extensively by Maraun et al. (2017), who illustrate several issues related to process-uninformed bias corrections.

Overall, this thesis is hoped to bring additional insights on how different methods could be evaluated using a more demanding setup and to better cover end user perspectives. In particular, **Paper IV** serves as a proof-of-concept for future studies, which aim to evaluate the relative method performance from hydrological (and potential other impact) modeling perspective in non-stationary conditions. For example, the same framework can potentially be applied to study more specific aspects of method performance, such as those schemed in the VALUE experimental framework (Maraun, 2016).

7 Summary of papers and the author's contribution

This thesis consists of four peer-reviewed papers (Fig. 16) that study the relative performance of different types of MOS methods using a novel evaluation framework (**Papers I**, **II** and **IV**) and attempt to assess the relative importance of MOS-method uncertainties both from climate modeling and hydrological modeling perspectives in the European region (**Papers I**, **II** and **III**).

Paper I evaluates the relative performance of a set of MOS methods, ranging from the adjustment of time-mean climate to more sophisticated methods which take the daily variability into account from distributional perspective. The evaluation is done in a cross-validation manner using the pseudo-reality approach. In addition to the inter-comparison of the selected MOS methods, uncertainties are also assessed by applying them to real-world observations in the European region. The author contributed to the development of the analysis code and made a small contribution to the manuscript.

Paper II presents a similar study to **Paper I** for daily precipitation. In contrast to **Paper I**, more emphasis is put on temporal and spatial variations in the relative MOS-method performance. Similarly to **Paper I**, uncertainties in the real-world projections in late 21st century conditions were also assessed. The author performed most of the analysis and was also the main responsible for writing the paper.

Paper III investigates the relative importance of GCM-RCM simulations and MOS methods as uncertainty sources for hydrological simulations in the Scandinavian region. A reduced set of MOS methods for daily mean temperature and precipitation was selected from **Papers I** and **II** for this study. Hydrological simulations were performed with the HYPE model using ENSEMBLES simulations for mid 21st century conditions as input. The author performed the MOS adjustments of GCM-RCM simulations, ran the hydrological simulations, performed most of the analysis of the results and was the main responsible for writing the paper.

Paper IV extends the analysis made in **Papers I** and **II** by including two bi-variate MOS methods, which take the inter-variable relationships between daily mean temperature and precipitation into account in their adjustments. The pseudo-reality approach used in **Paper I** and **Paper II** was extended to cover also the hydrological modeling step to assess the added value of bi-variate MOS from hydrological modeling perspective. The author selected and pre-processed the required data, and applied/wrote the required bi-variate methods. MOS-adjustments and hydrological simulations were also conducted by the author. The author was the main responsible for analyzing the results and writing the manuscript.

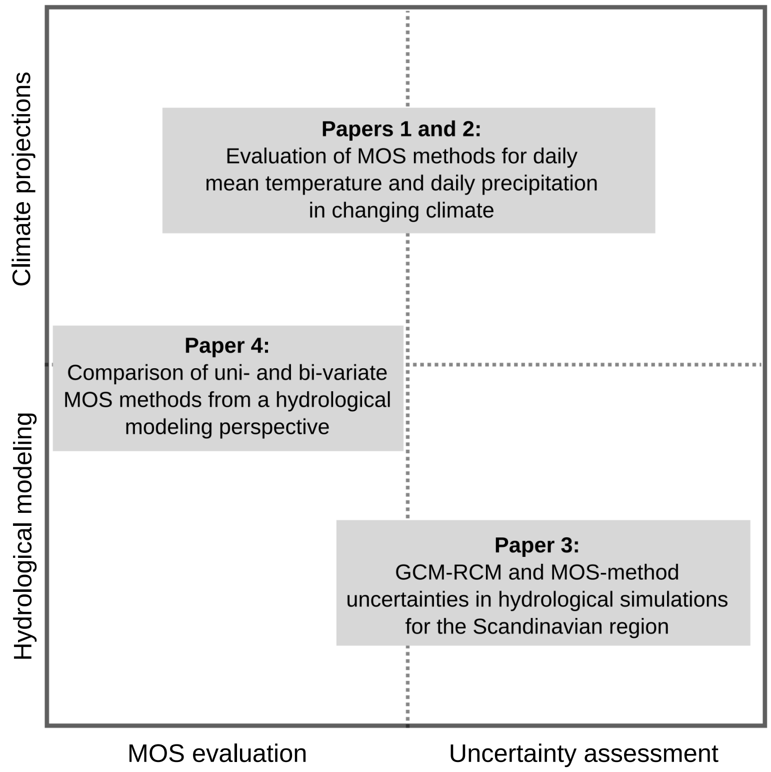

Figure 16: An overview of the articles and their relation to the studied aspects, shown on the axes.

References

Adam, J. C., A. F. Hamlet, and D. P. Lettenmaier (2009). Implications of global climate change for snowmelt hydrology in the twenty-first century. *Hydrological Processes 23*(7), 962–972.

Amengual, A., V. Homar, R. Romero, S. Alonso, and C. Ramis (2012). A Statistical Adjustment of Regional Climate Model Outputs to Local Scales: Application to Platja de Palma, Spain. *Journal of Climate 25*(3), 939–957.

Annan, J. D. and J. C. Hargreaves (2010). Reliability of the CMIP3 ensemble. *Geophysical Research Letters 37*(2), L02703.

Arnell, N. W. and S. N. Gosling (2016). The impacts of climate change on river flood risk at the global scale. *Climatic Change 134*(3), 387–401.

Ballester, J., F. Giorgi, and X. Rodó (2010). Changes in European temperature extremes can be predicted from changes in PDF central statistics. *Climatic Change 98*(1), 277–284.

Barnett T. P., Adam J. C., and Lettenmaier D. P. (2005). Potential impacts of a warming climate on water availability in snow-dominated regions. *Nature 438*, 303–309.

Bellprat, O., S. Kotlarski, D. Lüthi, and C. Schär (2013). Physical constraints for temperature biases in climate models. *Geophysical Research Letters 40*(15), 4042–4047.

Boberg, F. and J. H. Christensen (2012). Overestimation of mediterranean summer temperature projections due to model deficiencies. *Nature Climate Change 2*(6), 433–436.

Bosshard, T., M. Carambia, K. Goergen, S. Kotlarski, P. Krahe, M. Zappa, and C. Schär (2013). Quantifying uncertainty sources in an ensemble of hydrological climate-impact projections. *Water Resources Research 49*(3), 1523–1536.

Bosshard, T., S. Kotlarski, M. Zappa, and C. Schär (2014). Hydrological Climate-Impact Projections for the Rhine River: GCM–RCM Uncertainty and Separate Temperature and Precipitation Effects. *Journal of Hydrometeorology 15*(2), 697–713.

Bracegirdle, T. J. and D. B. Stephenson (2012). Higher precision estimates of regional polar warming by ensemble regression of climate model projections. *Climate Dynamics 39*(12), 2805–2821.

Cannon, A. J. (2016). Multivariate Bias Correction of Climate Model Output: Matching Marginal Distributions and Intervariable Dependence Structure. *Journal of Climate 29*(19), 7045–7064.

Cannon, A. J. (2018). Multivariate quantile mapping bias correction: an N-dimensional probability density function transform for climate model simulations of multiple variables. *Climate Dynamics 50*(1), 31–49.

Cannon, A. J., S. R. Sobie, and T. Q. Murdock (2015). Bias Correction of GCM Precipitation by Quantile Mapping: How Well Do Methods Preserve Changes in Quantiles and Extremes? *Journal of Climate 28*(17), 6938–6959.

Chen, C., J. O. Haerter, S. Hagemann, and C. Piani (2011). On the contribution of statistical bias correction to the uncertainty in the projected hydrological cycle. *Geophysical Research Letters 38*(20), L20403.

Chen, J., F. P. Brissette, D. Chaumont, and M. Braun (2013). Finding appropriate bias correction methods in downscaling precipitation for hydrologic impact studies over north america. *Water Resources Research 49*(7), 4187–4205.

Christensen, J. H., F. Boberg, O. B. Christensen, and P. Lucas-Picher (2008). On the need for bias correction of regional climate change projections of temperature and precipitation. *Geophysical Research Letters 35*(20), L20709.

Dee, D. P., S. M. Uppala, A. J. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M. A. Balmaseda, G. Balsamo, P. Bauer, P. Bechtold, A. C. M. Beljaars, L. van de Berg, J. Bidlot, N. Bormann, C. Delsol, R. Dragani, M. Fuentes, A. J. Geer, L. Haimberger, S. B. Healy, H. Hersbach, E. V. Hólm, L. Isaksen, P. Kållberg, M. Köhler, M. Matricardi, A. P. McNally, B. M. Monge-Sanz, J.-J. Morcrette, B.-K. Park, C. Peubey, P. de Rosnay, C. Tavolato, J.-N. Thépaut, and F. Vitart (2011). The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society 137*(656), 553–597.

Déqué, M. (2007). Frequency of precipitation and temperature extremes over France in an anthropogenic scenario: Model results and statistical correction according to observed values. *Global and Planetary Change 57*(1), 16–26.

Déqué, M., S. Somot, E. Sanchez-Gomez, C. M. Goodess, D. Jacob, G. Lenderink, and O. B. Christensen (2012). The spread amongst ENSEMBLES regional scenarios: regional climate models, driving general circulation models and interannual variability. *Climate Dynamics 38*(5), 951–964.

Dobler, C., S. Hagemann, R. L. Wilby, and J. Stötter (2012). Quantifying different sources of uncertainty in hydrological projections in an Alpine watershed. *Hydrology and Earth System Sciences 16*(11), 4343–4360.

Donnelly, C., J. C. Andersson, and B. Arheimer (2016). Using flow signatures and catchment similarities to evaluate the E-HYPE multi-basin model across Europe. *Hydrological Sciences Journal 61*(2), 255–273.

Ehret, U., E. Zehe, V. Wulfmeyer, K. Warrach-Sagi, and J. Liebert (2012). Hess opinions "should we apply bias correction to global and regional climate model data?". *Hydrology and Earth System Sciences 16*(9), 3391–3404.

Engen-Skaugen, T. (2007). Refinement of dynamically downscaled precipitation and temperature scenarios. *Climatic Change 84*(3), 365–382.

Gennaretti, F., L. Sangelantoni, and P. Grenier (2015). Toward daily climate scenarios for Canadian Arctic coastal zones with more realistic temperature-precipitation interdependence. *Journal of Geophysical Research: Atmospheres 120*(23), 11,862–11,877.

Glahn, H. R. and D. A. Lowry (1972). The Use of Model Output Statistics (MOS) in Objective Weather Forecasting. *Journal of Applied Meteorology 11*(8), 1203–1211.

Gudmundsson, L., J. B. Bremnes, J. E. Haugen, and T. Engen-Skaugen (2012). Technical Note: Downscaling RCM precipitation to the station scale using statistical transformations; a comparison of methods. *Hydrology and Earth System Sciences 16*(9), 3383–3390.

Hagemann, S., C. Chen, J. O. Haerter, J. Heinke, D. Gerten, and C. Piani (2011). Impact of a Statistical Bias Correction on the Projected Hydrological Changes Obtained from Three GCMs and Two Hydrology Models. *Journal of Hydrometeorology 12*(4), 556–578.

Hawkins, E. and R. Sutton (2011). The potential to narrow uncertainty in projections of regional precipitation change. *Climate Dynamics 37*(1), 407–418.

Held, I. M. and B. J. Soden (2006). Robust Responses of the Hydrological Cycle to Global Warming. *Journal of Climate 19*(21), 5686–5699.

Hewitson, B. C., J. Daron, R. G. Crane, M. F. Zermoglio, and C. Jack (2014). Interrogating empirical-statistical downscaling. *Climatic Change 122*(4), 539–554.

Huntington, T. G. (2006). Evidence for intensification of the global water cycle: Review and synthesis. *Journal of Hydrology 319*(1), 83–95.

IPCC (2013). *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.* Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press. 1535pp.

Jacob, D., J. Petersen, B. Eggert, A. Alias, O. B. Christensen, L. M. Bouwer, A. Braun, A. Colette, M. Déqué, G. Georgievski, E. Georgopoulou, A. Gobiet, L. Menut, G. Nikulin, A. Haensler, N. Hempelmann, C. Jones, K. Keuler, S. Kovats, N. Kröner, S. Kotlarski, A. Kriegsmann, E. Martin, E. van Meijgaard, C. Moseley, S. Pfeifer, S. Preuschmann, C. Radermacher, K. Radtke, D. Rechid, M. Rounsevell, P. Samuelsson, S. Somot, J.-F. Soussana, C. Teichmann, R. Valentini, R. Vautard, B. Weber, and P. Yiou (2014). EURO-CORDEX: new high-resolution climate change projections for european impact research. *Regional Environmental Change 14*(2), 563–578.

Kay, A. L., V. A. Bell, E. M. Blyth, S. M. Crooks, H. N. Davies, and N. S. Reynard (2013). A hydrological perspective on evaporation: historical trends and future projections in Britain. *Journal of Water and Climate Change 4*(3), 193–208.

Knutti, R. (2008). Should we believe model predictions of future climate change? *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences 366*(1885), 4647–4664.

Kotlarski, S., K. Keuler, O. B. Christensen, A. Colette, M. Déqué, A. Gobiet, K. Goergen, D. Jacob, D. Lüthi, E. van Meijgaard, G. Nikulin, C. Schär, C. Teichmann, R. Vautard, K. Warrach-Sagi, and V. Wulfmeyer (2014). Regional climate modeling on European scales: a joint standard evaluation of the EURO-CORDEX RCM ensemble. *Geoscientific Model Development 7*(4), 1297–1333.

Lafaysse, M., B. Hingray, A. Mezghani, J. Gailhard, and L. Terray (2014). Internal variability and model uncertainty components in future hydrometeorological projections: The Alpine Durance basin. *Water Resources Research 50*(4), 3317–3341.

Lafon, T., S. Dadson, G. Buys, and C. Prudhomme (2013). Bias correction of daily precipitation simulated by a regional climate model: a comparison of methods. *International Journal of Climatology 33*(6), 1367–1381.

Leander, R. and T. A. Buishand (2007). Resampling of regional climate model output for the simulation of extreme river flows. *Journal of Hydrology 332*(3), 487–496.

Li, C., E. Sinha, D. E. Horton, N. S. Diffenbaugh, and A. M. Michalak (2014). Joint bias correction of temperature and precipitation in climate model simulations. *Journal of Geophysical Research: Atmospheres 119*(23), 13,153–13,162.

Li, H., J. Sheffield, and E. F. Wood (2010). Bias correction of monthly precipitation and temperature fields from Intergovernmental Panel on Climate Change AR4 models using equidistant quantile matching. *Journal of Geophysical Research: Atmospheres 115*(D10), D10101.

Lindström, G., C. Pers, J. Rosberg, J. Strömqvist, and B. Arheimer (2010). Development and testing of the HYPE (Hydrological Predictions for the Environment) water quality model for different spatial scales. *Hydrology Research 41*(3-4), 295–319.

Liu, M., J. C. Adam, and A. F. Hamlet (2013). Spatial-temporal variations of evapotranspiration and runoff/precipitation ratios responding to the changing climate in the Pacific Northwest during 1921-2006. *Journal of Geophysical Research: Atmospheres 118*(2), 380–394.

Maraun, D. (2012). Nonstationarities of regional climate model biases in european seasonal mean temperature and precipitation sums. *Geophysical Research Letters 39*(6), L06706.

Maraun, D. (2013). When will trends in European mean and heavy daily precipitation emerge? *Environmental Research Letters 8*(1), 014004.

Maraun, D. (2016). Bias Correcting Climate Change Simulations - a Critical Review. *Current Climate Change Reports 2*(4), 211–220.

Maraun, D., T. G. Shepherd, M. Widmann, G. Zappa, D. Walton, J. M. Gutiérrez, S. Hagemann, I. Richter, P. M. Soares, A. Hall, et al. (2017). Towards process-informed bias correction of climate change simulations. *Nature Climate Change 7*(11), 764–773.

Maraun, D., F. Wetterhall, A. M. Ireson, R. E. Chandler, E. J. Kendon, M. Widmann, S. Brienen, H. W. Rust, T. Sauter, M. Themeßl, V. K. C. Venema, K. P. Chun, C. M. Goodess, R. G. Jones, C. Onof, M. Vrac, and I. Thiele-Eich (2010). Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user. *Reviews of Geophysics 48*(3), RG3003.

Moss, R. H., J. A. Edmonds, K. A. Hibbard, M. R. Manning, S. K. Rose, D. P. Van Vuuren, T. R. Carter, S. Emori, M. Kainuma, T. Kram, et al. (2010). The next generation of scenarios for climate change research and assessment. *Nature 463*(7282), 747–756.

Nakicenovic, N., J. Alcamo, A. Grubler, K. Riahi, R. Roehrl, H.-H. Rogner, and N. Victor (2000). *Special report on emissions scenarios (SRES), a special report of Working Group III of the intergovernmental panel on climate change*. Cambridge University Press. 599 pp.

Nohara, D., A. Kitoh, M. Hosaka, and T. Oki (2006). Impact of Climate Change on River Discharge Projected by Multimodel Ensemble. *Journal of Hydrometeorology 7*(5), 1076–1089.

Oki, T. and S. Kanae (2006). Global Hydrological Cycles and World Water Resources. *Science 313*(5790), 1068–1072.

Panofsky, H. and G. Brier (1968). *Some Applications of Statistics to Meteorology*. Earth and Mineral Sciences Continuing Education, College of Earth and Mineral Sciences. 224 pp.

Piani, C. and J. O. Haerter (2012). Two dimensional bias correction of temperature and precipitation copulas in climate models. *Geophysical Research Letters 39*(20), L20401.

Piani, C., G. Weedon, M. Best, S. Gomes, P. Viterbo, S. Hagemann, and J. Haerter (2010). Statistical bias correction of global simulated daily precipitation and temperature for the application of hydrological models. *Journal of Hydrology 395*(3), 199–215.

Pitié, F., A. C. Kokaram, and R. Dahyot (2007). Automated colour grading using colour distribution transfer. *Computer Vision and Image Understanding 107*(1), 123–137.

Räisänen, J. (2007). How reliable are climate models? *Tellus A 59*(1), 2–29.

Räisänen, J. (2016). Twenty-first century changes in snowfall climate in Northern Europe in ENSEMBLES regional climate models. *Climate Dynamics 46*(1), 339–353.

Rajczak, J., S. Kotlarski, and C. Schär (2016). Does Quantile Mapping of Simulated Precipitation Correct for Biases in Transition Probabilities and Spell Lengths? *Journal of Climate 29*(5), 1605–1615.

Rojas, R., L. Feyen, A. Dosio, and D. Bavera (2011). Improving pan-European hydrological simulation of extreme events through statistical bias correction of RCM-driven climate simulations. *Hydrology and Earth System Sciences 15*(8), 2599–2620.

Rummukainen, M. (2016). Added value in regional climate modeling. *Wiley Interdisciplinary Reviews: Climate Change 7*(1), 145–159.

Schewe, J., J. Heinke, D. Gerten, I. Haddeland, N. W. Arnell, D. B. Clark, R. Dankers, S. Eisner, B. M. Fekete, F. J. Colón-González, S. N. Gosling, H. Kim, X. Liu, Y. Masaki, F. T. Portmann, Y. Satoh, T. Stacke, Q. Tang, Y. Wada, D. Wisser, T. Albrecht, K. Frieler, F. Piontek, L. Warszawski, and P. Kabat (2014). Multimodel assessment of water scarcity under climate change. *Proceedings of the National Academy of Sciences 111*(9), 3245–3250.

Schneider, U., A. Becker, P. Finger, A. Meyer-Christoffer, B. Rudolf, and M. Ziese (2011). GPCC full data reanalysis version 6.0 at 0.5°: monthly land-surface precipitation from rain-gauges built on GTS-based and historic data. doi:10.5676/DWD_GPCC/FD_M_V6_050.

Schölzel, C. and P. Friederichs (2008). Multivariate non-normally distributed random variables in climate research – introduction to the copula approach. *Nonlinear Processes in Geophysics 15*(5), 761–772.

Sharma, D., A. Das Gupta, and M. S. Babel (2007). Spatial disaggregation of bias-corrected GCM precipitation for improved hydrologic simulation: Ping River Basin, Thailand. *Hydrology and Earth System Sciences 11*(4), 1373–1390.

Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris 8*, 229–231.

Teutschbein, C. and J. Seibert (2012). Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods. *Journal of Hydrology 456*, 12–29.

Themeßl, M. J., A. Gobiet, and A. Leuprecht (2011). Empirical-statistical downscaling and error correction of daily precipitation from regional climate models. *International Journal of Climatology 31*(10), 1530–1544.

van der Linden, P. and J. Mitchell (Eds.) (2009). *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. 160 pp.

Van Schaeybroeck, B. and S. Vannitsem (2016). Assessment of calibration assumptions under strong climate changes. *Geophysical Research Letters 43*(3), 1314–1322.

van Vliet, M. T., C. Donnelly, L. Strömbäck, R. Capell, and F. Ludwig (2015). European scale climate information services for water use sectors. *Journal of Hydrology 528*, 503–513.

Velázquez, J. A., M. Troin, D. Caya, and F. Brissette (2015). Evaluating the Time-Invariance Hypothesis of Climate Model Bias Correction: Implications for Hydrological Impact Studies. *Journal of Hydrometeorology 16*(5), 2013–2026.

Vicente-Serrano, S. M., S. Beguería, and J. I. López-Moreno (2010). A Multiscalar Drought Index Sensitive to Global Warming: The Standardized Precipitation Evapotranspiration Index. *Journal of Climate 23*(7), 1696–1718.

Vrac, M. and P. Friederichs (2014). Multivariate—Intervariable, Spatial, and Temporal—Bias Correction. *Journal of Climate 28*(1), 218–237.

Vrac, M., T. Noël, and R. Vautard (2016). Bias correction of precipitation through singularity stochastic removal: Because occurrences matter. *Journal of Geophysical Research: Atmospheres 121*(10), 5237–5258.

Vrac, M., M. L. Stein, K. Hayhoe, and X.-Z. Liang (2007). A general method for validating statistical downscaling methods under future climate change. *Geophysical Research Letters 34*(18), L18701.

Watanabe, S., S. Kanae, S. Seto, P. J.-F. Yeh, Y. Hirabayashi, and T. Oki (2012). Intercomparison of bias-correction methods for monthly temperature and precipitation simulated by multiple climate models. *Journal of Geophysical Research: Atmospheres 117*, D23114.

Weedon, G. P., G. Balsamo, N. Bellouin, S. Gomes, M. J. Best, and P. Viterbo (2014). The WFDEI meteorological forcing data set: WATCH Forcing Data methodology applied to ERA-Interim reanalysis data. *Water Resources Research 50*(9), 7505–7514.

Wilby, R. L. and S. Dessai (2010). Robust adaptation to climate change. *Weather 65*(7), 180–185.

Wilks, D. S. (2006). *Statistical Methods in the Atmospheric Sciences* (2nd ed.). Burlington: Elsevier. 627 pp.

Wood, A. W., L. R. Leung, V. Sridhar, and D. P. Lettenmaier (2004). Hydrologic Implications of Dynamical and Statistical Approaches to Downscaling Climate Model Outputs. *Climatic Change 62*(1), 189–216.

Yang, W., J. Andréasson, L. Phil Graham, J. Olsson, J. Rosberg, and F. Wetterhall (2010). Distribution-based scaling to improve usability of regional climate model projections for hydrological climate change impacts studies. *Hydrology Research 41*(3-4), 211–229.

Yip, S., C. A. T. Ferro, D. B. Stephenson, and E. Hawkins (2011). A Simple, Coherent Framework for Partitioning Uncertainty in Climate Predictions. *Journal of Climate 24*(17), 4634–4643.

Zhang, X., F. W. Zwiers, and G. Hegerl (2009). The influences of data precision on the calculation of temperature percentile indices. *International Journal of Climatology 29*(3), 321–327.