# Empirical-statistical downscaling and error correction of temperature, precipitation, and derived extremes in Europe

Matthias Jakob Themeßl

December 2011

The **Wegener Center for Climate and Global Change** combines as an interdisciplinary, internationally oriented research center the competences of the University of Graz in the research area „Climate, Environmental and Global Change". It brings together, in a dedicated building close to the University central campus, research teams and scientists from fields such as geo- and climate physics, meteorology, economics, geography, and regional sciences. At the same time close links exist and are further developed with many cooperation partners, both nationally and internationally. The research interests extend from monitoring, analysis, modeling and prediction of climate and environmental change via climate impact research to the analysis of the human dimensions of these changes, i.e, the role of humans in causing and being effected by climate and environmental change as well as in adaptation and mitigation. (more informationen at www.wegcenter.at)

The present report is the result of Phd thesis work completed in June 2011.
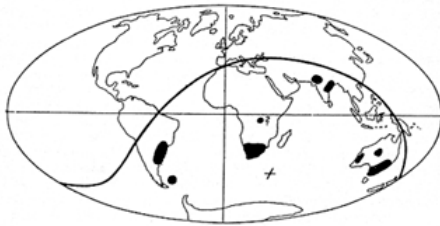


**Alfred Wegener** (1880-1930), after whom the Wegener Center is named, was founding holder of the University of Graz Geophysics Chair (1924-1930) and was in his work in the fields of geophysics, meteorology, and climatology a brilliant, interdisciplinary thinking and acting scientist and scholar, far ahead of his time with this style. The way of his ground-breaking research on continental drift is a shining role model — his sketch on the relationship of the continents based on traces of an ice age about 300 million years ago (left) as basis for the Wegener Center Logo is thus a continuous encouragement to explore equally innovative scientific ways: *paths emerge in that we walk them* (Motto of the Wegener Center).

*Contact: Matthias Themeßl*
*matthias.themessl@uni-graz.at*

Wegener Center for Climate and Global Change
University of Graz
Leechgasse 25
A-8010 Graz, Austria
www.wegcenter.at

**Mag. Matthias Jakob Themeßl**

# Empirical-statistical downscaling and error correction of temperature, precipitation, and derived extremes in Europe

Dissertation

zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

an der Naturwissenschaftlichen Fakultät der

Karl-Franzens-Universität Graz

Betreuer:

Univ.-Prof. Mag. Dr. Gottfried Kirchengast

Ass.-Prof. Mag. Dr. Andreas Gobiet

Graz, Juni 2011



Wegener Zentrum für Klima und Globalen Wandel
und
Institutsbereich Geophysik, Astrophysik und
Meteorologie/Institut für Physik
Karl-Franzens-Universität Graz

# Acknowledgements

I would like to thank Univ.-Prof. Mag. Dr. Gottfried Kirchengast and Ass.-Prof. Mag. Dr. Andreas Gobiet for their scientific support in conceiving this PhD topic as well as for their helpful advices during the PhD work. I am also grateful to the Regional and Local Climate Analysis and Modeling Group (ReLoClim) team as well as to all roommates of office 41-EG-0018 and 41-01-0014 and the administration and the technical support group at the Wegener Center for Climate and Global Change for their discussions and risen ideas. They made it as exciting and interdisciplinary as it can be in science.

In addition, I would like to acknowledge the Munich Re and the University of Graz for their scholarships as well as the EU FP6 Climate Change and Variability: Impact on Central and Eastern Europe (CLAVIER) project for supporting this PhD project. Further specific acknowledgements, including data providers, are contained at the end of Paper I and Paper II in chapters 4 and 5, respectively.

Besides my colleagues at work, I would like to mention my family and my friends and thank them for their great support throughout the last years.

Finally, I would like to thank myself. Although this may sound selfish, this PhD work obviously shaped me and anybody who ever spent nights preparing a manuscript for submission, knowing that in fact he or she won't find any new bug in there but wants to submit "the perfect one" knows what I am talking about. Within this PhD work I grew up professionally as well as personally and I am very grateful for all the experiences I made.

# Abstract

Although RCMs have already proven their capability to simulate regional climate and its variability, they still feature systematic errors compared to observations. Besides their steady enhancement, empirical-statistical post-processing, based on the concept of model output statistics (MOS), provides a ready opportunity to mitigate RCM error characteristics and to further downscale climate model data to the point-scale.

In the course of this PhD work, seven empirical-statistical downscaling and error correction methods (DECMs) are inter-compared for their applicability to and error correction potential for daily precipitation, temperature, and derived extreme indices from RCMs in Europe. Furthermore, error corrected climate scenarios for the respective parameters are generated for Europe and the impact of DECMs on the climate change signal (CCS) is investigated.

Overall, the findings of this PhD work strongly emphasize the combination of RCMs and DECMs to provide suitable climate data for climate impact assessments and decision making. DECMs drastically reduced the error characteristics of RCMs regarding mean, variability, and extremes. Particularly, Quantile Mapping (QM) resulted in outstanding error correction potential and can be considered as highly recommendable due to its simplicity and flexibility.

In application to future climate scenarios QM only moderately modified the CCSs of mean, minimum, maximum temperature, and precipitation amount. In contrast, QM strongly changed the CCSs of non-linearly derived indices of extremes such as threshold indices in some cases. However, these modifications were considered as reasonable because the respective uncorrected parameters featured magnitude-dependent error characteristics and trends in the future scenarios.

Besides, this PhD work also defined useful climate data from the point of view of the climate impact community and decision makers in order to promote collaborations with the climate modeling community.

# Kurzbeschreibung

Obwohl regionale Klimamodelle (RCMs) regionale Klimabedingungen und deren Variabilität recht zufriedenstellend abbilden können, weisen sie nach wie vor systematische Fehler im Vergleich zu Beobachtungen auf. Neben ihrer stetigen Verbesserung bietet die empirisch-statistische Nachbearbeitung, basierend auf dem Konzept von Model Output Statistics (MOS), eine sofort einsetzbare Möglichkeit, Modellfehler zu verringern und Klimasimulationen auf der Punktskala anzubieten.

Im Zuge dieser Doktorarbeit wurden sieben empirisch-statistische Downscaling and Error Correction Methods (DECMs) auf ihr Fehlerkorrekturpotenzial für Tagesniederschläge, Tagestemperaturen sowie für abgeleitete Extremindizes von RCMs in Europa untersucht. Weiters wurden fehlerkorrigierte Klimaszenarien für Europa generiert und die Auswirkungen von DECMs auf das Klimaänderungssignal analysiert.

Zusammenfassend stellte sich heraus, dass DECMs großes Potenzial zeigten, um die Güte von RCM Simulationen, bezogen auf ihre Fehler im Mittelwert, in der Variabilität sowie bei Extremen stark zu verbessern. Insbesondere Quantile Mapping (QM) überzeugte dank herausragendem Fehlerkorrekturpotenzial, Einfachheit und Flexibilität.

Angewendet auf zukünftige Klimasimulationsergebnisse veränderte QM die Klimaänderungssignale von Tagestemperaturen sowie Tagesniederschlag nur geringfügig. Allerdings veränderte QM die Klimaänderungssignale abgeleiteter Extremindizes zum Teil beträchtlich. Diese Änderungen waren jedoch durch größenabhängige Fehler in den unkorrigierten Daten sowie zugrundeliegende Trends in den unkorrigierten Zukunftsszenarien erklärbar.

Zusätzlich beschäftigte sich diese Doktorarbeit mit der Definition von sinnvollen Klimadaten für die Klimafolgenforschung und für Entscheidungsträger, um die Zusammenarbeit zwischen Klimadatenproduzenten und -nutzern zu verbessern und dadurch mehr integrative Klimafolgenstudien zu ermöglichen.

# Contents

# 1 Motivation and introduction

The provision of fine-scale spatial and temporal climate data for the climate impact research community and decision makers is one of the highly topical research and working fields in climate science. For this purpose, regional climate model (RCM) simulations are usually performed as they represent the state-of-the-art methodologies to provide physically consistent regional and daily climate data. RCM simulations have become broadly available for Europe via projects such as PRUDENCE (Christensen and Christensen, 2007) and ENSEMBLES (van der Linden and Mitchell, 2009) and will be further complemented in the CORDEX project (http://wcrp.ipsl.jussieu.fr/SF_RCD_CORDEX.html). However, RCMs are still affected by partly significant error characteristics. Applied directly to, e.g., climate change impact modeling, such model errors would lead to subsequent errors in climate change impact assessments. Empirical-statistical downscaling and error correction methods (DECMs) provide a ready opportunity to mitigate such RCM errors.

Build on the experiences from my Master Thesis "*Downscaling of Temperature and Precipitation in the Alpine Region Hohe Tauern*", this PhD thesis is intended to

a) extend the knowledge of available DECMs for RCM post-processing
b) investigate the applicability of DECMs for downscaling and error correcting daily temperature and daily precipitation amount from RCM simulations
c) investigate the applicability of DECMs to derived extremes
d) provide temporal and spatial fine-scale and error corrected climate scenarios until 2050 for the climate impact community.

Besides establishing methodological knowledge in subitems a to c, subitem d also envisages the provision of useful data for further climate change impact assessments and thus deals with the interface between data producers and users.

The PhD thesis is structured as follows. Chapter 2 briefly introduces the climate system and climate models. Chapter 3 describes modeling techniques that enable to simulate regional and local climate (downscaling techniques). Chapter 4 then proposes a combination of

downscaling techniques to provide as accurate as possible climate data for the climate change impact community and inter-compares various DECMs for daily precipitation and derived moderate extremes in the Alpine region. Chapter 5 focuses on the applicability of the best performing DECM to further meteorological parameters (mean, minimum, and maximum temperature) in Europe, investigates the issue of error correction of extremes in more detail and examines the impact of error correction on the climate change signal. Chapter 6 defines general data needs of the climate change impact community before Chapter 7 summarizes and discusses the key findings of this PhD thesis.

Chapter 4 and Chapter 5 are based on either already published or accepted journal articles in the International Journal of Climatology and Climatic Change and are included in this PhD thesis with the acceptance of the publishers. Their correct citation is given in the respective chapters. The articles styles' are aligned to the synopsis of this PhD thesis.

# 2 The climate system and its representation via models

"*Weather is the fluctuating state of the atmosphere at a certain location experienced between seconds and a few days and characterized by meteorological variables such as temperature, precipitation, or wind. Almost all weather phenomena are limited to the lowest ~12 km of the atmosphere, the so-called troposphere. Typical weather states, e.g., are mid-latitude low pressure systems with associated frontal zones and showers.*

*Climate is the average weather in terms of statistics such as mean, variability, and extremes over a time-span of, e.g., 30 years*" (Baede et al., 2001; Häckel, 2008).
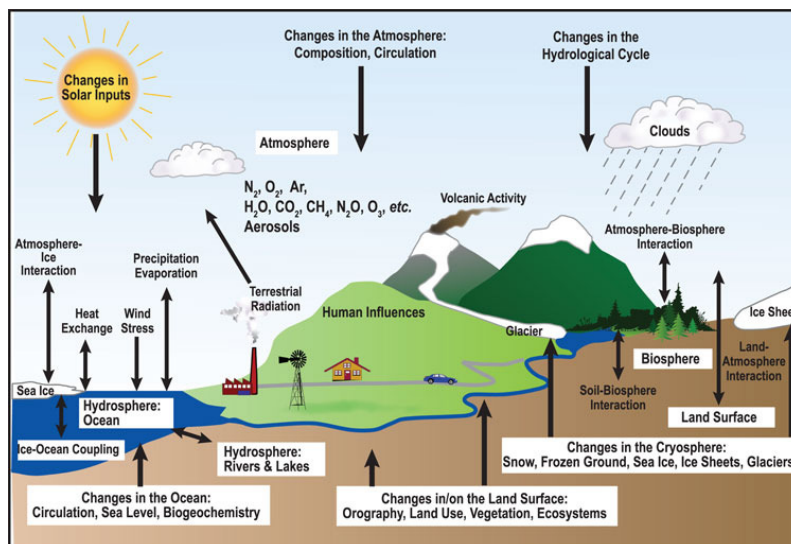


Figure 2.1 The climate system and its sub-systems (from IPCC, 2007).

Weather and climate determine to a certain degree our daily lives as well as the success of various economic sectors (Steininger et al., 2005). In particular, extreme conditions (compare Section 3.2.8 for detailed definition) can cause enormous economic damages (compare Munich Re, 2011) or may even endanger our lives (compare the heat wave in summer 2003; Poumadere et al., 2005).

Thus, aiming at suitable mitigation and adaptation strategies, we are highly interested in possible future climate conditions. For this purpose, historic trends can either be extrapolated, assuming that the processes leading to those trends remain unchanged also in future, or we try to understand the mechanism that determine what we experience as climate. For the latter purpose, climate can also be defined as the result or the attribute of the climate system (Bauer et al., 2001) created by its respective forcings, components, processes, and interactions.
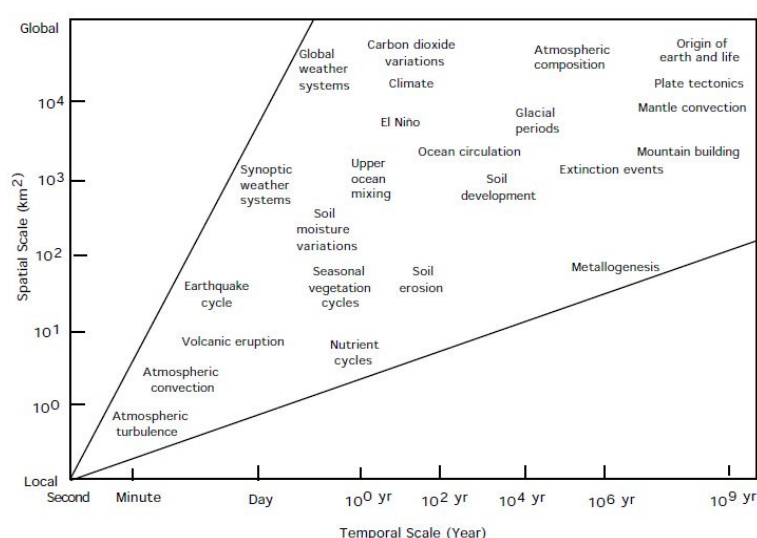


Figure 2.2 Characteristic spatial and temporal scales in the climate system (from Wu, 1999).

The climate system comprises five major sub-systems (components; compare Figure 2.1): the atmosphere, hydrosphere, cryoshere, biosphere, and geosphere. The atmosphere is the layer of gases surrounding the Earth. It constitutes of a certain compositions of gases including $N_2$, $O_2$, Ar, $H_2O$, and $CO_2$ and represents the most unstable and rapidly changing part of the system. The hydrosphere contains all liquid surface and subterranean water; the cryosphere comprises the ice sheets of Greenland and Antarctic as well as all continental glaciers and snow fields, sea ice and permafrost; the geosphere describes the stones and soils and finally the biosphere contains all flora and fauna.

4

The system is externally forced, e.g., by the Earth's inclination, but most dominant by the solar radiation. Possible influences of human beings can also be regarded as an external forcing.

Many of these sub-systems interact with each other via various physical, biological or chemical processes and dynamics (indicated by the black arrows in Figure 2.1) on various spatial and temporal scales (compare Figure 2.2). These interactions are partly non-linear including feedbacks (Baede et al., 2001; Stocker et al., 2001; Seneviratne et al., 2010). For example, if precipitation is reduced, the evapotranspiration of the soil and vegetation is reduced, which can lead to further reduced cloud formation and thus less precipitation. Such a feedback-interaction between geo-, bio- and atmosphere, e.g., took place thousands of years ago in the Sahara desert, then a landscape similar to the Serengeti today, and formed the landscape we know today (Brovkin, 2006).
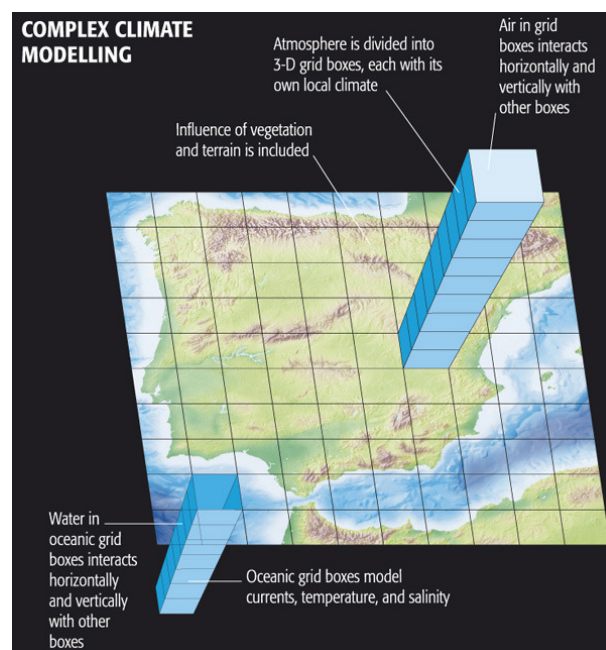


Figure 2.3 Schematic of an Atmosphere-Ocean coupled GCM
(from https://www.e-education.psu.edu/meteo469/node/140).

In climate sciences the representation of the climate system and its dynamics is accomplished by climate models. Climate models can either be one dimensional, conceptual models such as radiation balance models or three-dimensional highly complex models that aim to copy single sub-systems of the climate system or even the entire climate system (compare von Storch et al., 1999). Referring to the latter models, global climate models, also called general circulation models (GCMs), are the primary tools for simulating the global climate in a physically consistent way.

GCMs numerically simulate the climate system, its known properties and dynamics based on physical laws such as the conservation of momentum, mass and energy, the ideal gas law or boundary layer physics (Giorgi and Mearns, 1991; Baede et al., 2001). The model equations are solved on a three dimensional grid (e.g., Warner, 2011) with state-of-the-art grid point distances of ~200 km and several vertical layers as depicted in Figure 2.3 for an Atmosphere-Ocean coupled GCM.
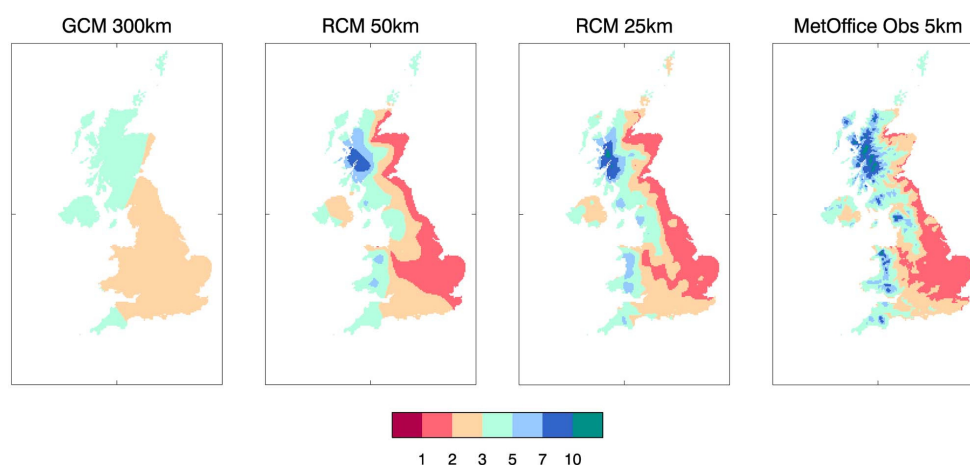


Figure 2.4 Comparison of climate models with different horizontal resolution to observation for winter precipitation (in mm/day) in the UK (from Maraun et al., 2010).

GCMs have proven to reliably simulate global distributions of, e.g., temperature, precipitation, radiation, wind, oceanic temperature, currents, and ice covers as well as important patterns of climate variability as the Hadley cell, monsoon systems, or storm tracks (Randall et al., 2007). GCMs, however, fail at providing realistic climate information at the

sub-grid or even regional scale (compare Figure 2.4) as well as at the daily/sub-daily scale (Baede et al., 2001; Giorgi et al., 2001; Prudhomme et al., 2002). This can be related to systematic model errors and the absence or simplifications (parameterization) of sub-grid-scale processes and forcings such as complex topography, inland water bodies, or land use characteristics (Randell et al., 2007; Benestad et al., 2008). Small scale climate data is nevertheless essential for climate change impact and adaptation assessments (see Chapter 6). The associated problem to get to this data is obviously related to a spatial incompatibility of GCMs and is often called regional climate problem.

# 3 Downscaling: providing fine scale climate model data

Following the definition of Giorgi et al. (2001), the regional scale is defined as an area between $10^4$ km² and $10^7$ km². Areas smaller than 100 km x 100 km are assigned to the local scale, areas greater than the regional scale are assigned to the sub-continental and afterwards to the planetary scale. For an appropriate description of the regional climate, processes at all scales have to be taken into account.

Two pathways are common in climate modeling in order to bridge the scale gap between globally available and regionally or locally demanded climate data. Both pathways cascade "down" information from a larger spatial domain ("the large scale") to a smaller spatial domain ("the small scale"; von Storch et al., 2000). Thus, these procedures are also referred to as downscaling procedures. The two downscaling approaches are

- dynamical downscaling (DD) and
- empirical-statistical downscaling (ESD).

## 3.1 Dynamical downscaling

Although this work focuses on ESD, DD is also briefly described. DD simulates regional or local climate conditions by nesting a finer-meshed dynamical climate model (typically between 10 km and 50 km grid spacing) over a limited area in coarser climate information, e.g., provided by GCMs (compare Figure 3.1). The GCM provides the boundary conditions (winds, temperature, humidity, pressure, and sea surface temperature) for the fine-mesh RCM, which profits from its higher spatial resolution via a more accurate representation of small scale forcings as listed in Chapter 2. Furthermore, processes which are parameterized in GCMs can be resolved hereby (Giorgi and Mearns, 1991, 1999; Laprise, 2008; Maraun et al., 2010 and the references therein). As a consequence, RCMs have proven to add regional information compared to GCM outputs (e.g., Wang et al., 2004; Feser et al., 2006; Laprise, 2008) including

enhanced mesoscale circulation patterns (Buonomo et al., 2007) and regional scale scenarios for extremes (e.g., Schmidli et al., 2006; Christensen and Christensen, 2007).
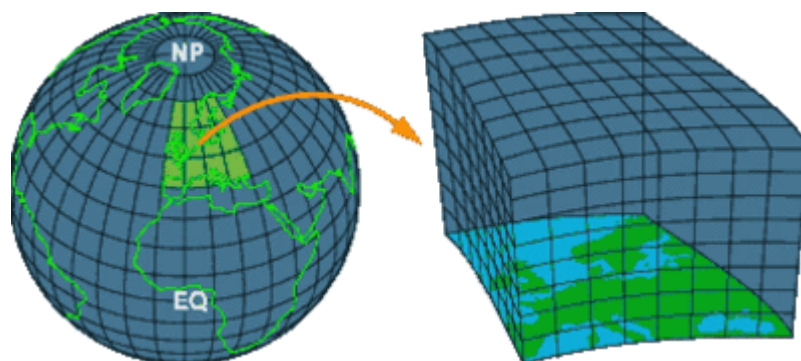


Figure 3.1 The RCM nesting scheme. In this example the fine-mesh RCM on the right hand side is nested in coarser resolved raster information over Europe on the left (from Deutsches Museum homepage: http://www.deutsches-museum.de/ausstellungen/energie/umwelt/klima/klimawandel/klimamodelle/gitternetz/).

Further information concerning RCMs and the skill of state-of-the-art RCMs can be found in the former listed articles as well as, e.g., in Frei et al. (2003), Truhetz et al. (2003), van der Linden and Mitchell (2009), Maraun et al. (2010), and Suklitsch et al. (2011).

## 3.2 Empirical-statistical downscaling

### 3.2.1 The history of ESD

Although ESD studies primarily emerged since the 1970s onwards associated with the development of GCMs, empirical-statistical techniques to translate across spatial scales are already known since the 1940s in the field of synoptic climatology (Hewitson and Crane, 1996). In Barry and Perry (1973) synoptic climatology is defined as one pathway to "*obtain insight into local or regional climates by examining the relationship of weather elements, individually or collectively, to atmospheric circulation processes*". The underlying assumption is that similar synoptic patterns result in similar regional weather conditions (compare classification of Grosswetterlagen; e.g., Baur et al., 1944). However, synoptic climatology differs from a recent definition of downscaling in its basis on a finite number of discrete

weather classes because this generalization of the atmospheric circulation does not accurately describe the evolution of weather or climate as a continuous function in time (Hewitson and Crane, 1996).

### 3.2.2 Definition of ESD

Based on the given definition of the regional climate, ESD is henceforth defined as the process of making the link between the state of some variable at the large scale and the state of some variable describing the small scale by using an empirically based statistical formulation (Benestad et al., 2008). This formulation, sometime denoted as transfer function, can be either deterministic or stochastic, whether it disregards any residual noise term in their model description or not. Formally, (in a deterministic way) this leads to

$$local\ climate\ response = f\left(large\ scale\ forcing\right)$$
$$predictand = f\left(predictor\right)$$
$$dependent\ varibale = f\left(independent\ variable\right)$$

where all three notations can be found in literature. Here, we will use the second formulation for all further notations.

According to these formulations, the small scale cannot feed back into the large scale. Thus, ESD represents a one way coupling approach. However, it should be considered, that although the small scale may follow the large-scale conditions to some degree, it can also exhibit a behavior independent of the large-scale situation, which then is attributed as noise and cannot be captured by the predictor (Benestad et al., 2008).

### *3.2.3 Assumptions of ESD*

In order to generate reliable regional climate scenarios based on ESD the following four general assumptions have to be fulfilled (compare Benestad et al., 2008).

1)  The relationship is strong and physically based.
2)  The predictors are of satisfying accuracy.
3)  The statistical model is stationary over time.
4)  The incorporated predictors fully represent the climate change signal.

Ad assumption 1) The found relationship should be physically meaningful and not result from any statistical coincidence (e.g., von Storch and Navarra, 1999). This means that the incorporated predictors are of relevance for the predictands (e.g., large scale humidity will condition local scale precipitation). In most cases, except distribution based methods (compare Chapter 4), the link is strong if the disparate scales show a matching time behavior, thus if they co-vary.

Ad assumption 2) If defective predictors are used for model calibration, ESD will result in invalid estimations. The choice, which predictors should be encountered, depends on spatial, temporal, as well as data availability aspects. As already mentioned GCMs should be considered with care on the daily and regional scale whereas RCMs, if nested in reanalysis data, are assumed to satisfactorily represent the regional scale daily weather evolution and even to represent extremes to some degree. At the sub-daily scale, however, also RCMs show only limited skill in, e.g., reproducing characteristics as the diurnal cycle (e.g., Prein et al., 2011)

Ad assumption 3) The stationarity assumption implies that the statistical link does not change over time. Thus, also small scale forcings such as changes in land-use characteristics or global characteristics such as the oceanic circulation are assumed to stay constant.

Ad assumption 4) If the used predictors do not fully account for the climate change signal, ESD won't produce reliable future scenarios. For example Wilby and Wigley (2000) demonstrated the importance and significant difference of precipitation scenarios with and without the inclusion of humidity as predictor variable. Therefore, in general it can be

advisable to include both thermodynamic as well as fluid dynamic predictors (e.g., Salathé, 2003).

Concerning assumption 3 and 4, split sample or cross validation tests are suitable tools for verification purposes (e.g., von Storch et al., 2000).

### 3.2.4 Model Output Statistics vs. Perfect Prognosis

Relating to the ESD classification by Rummukainen (1997), which is based on the origin of the used predictors, ESD can be divided in model output statistics (MOS) and perfect prognosis (PP; Wilks, 2006) applications (compare Figure 4.1)

PP downscaling represents the traditional ESD application. By this means, the statistical link is established between observed predictors and observed predictands. Instead of observed large scale data, reanalysis data are often found in application. Reanalyses combine various data sources as observational and numerical weather prediction data (e.g., Nakiekovic et al., 2000). The found linkage can then be applied to predictors from climate models under assumption 2 in Section 3.2.3, thus assuming that the model perfectly simulates (or prognosticates in the case of numerical weather prediction) the predictor (Wilks, 2006; Maraun et al., 2010).

MOS, besides already known from numerical weather prediction, has also become popular in climate sciences for downscaling purposes with the increasing accuracy and availability of GCM and RCM data. Contrary to PP, MOS establishes a link between simulated predictors and observed predictands. As a consequence, MOS calibrated relationships enable to downscale the model data to the observational resolution and additionally account for the simulated model errors. However, MOS calibrated relationships are only valid for the model they are calibrated on, contrary to PP relationships, which are in general valid for any model.

### *3.2.5 ESD techniques*

Methodologically, three major groups of statistical techniques can be identified in ESD reviews referring, e.g., to Wilby and Wigley (1997), von Storch et al. (2000), or Benestad et al. (2008). These groups are regression models, weather typing approaches and weather generators. Considering the often found hybrid approaches in literature and integrating more recently described MOS techniques, here a slightly different classification based on Bardossy (2000) with four classes is presented. These four classes are:

1) Regression-like methods (Hewitson and Crane, 1996; Murphy, 1999; Schoof and Pryor, 2001; Haylock and Goodess, 2004; Benestad et al., 2008)
2) Resampling or analogue methods (Cubasch et al., 1996; Brandsma and Buishand, 1998; Zorita and von Storch, 1999; Wetterhall et al., 2005; Matulla et al., 2008; )
3) Conditional probability methods/weather generators (Conway et al., 1996; Goodess and Palutikof, 1998; von Storch and Navarra, 1999; Wilks and Wilby, 1999; Wilby et al., 2003)
4) Scaling approaches (Wood et al., 2002; Schmidli et al., 2006; Déqué, 2007; Graham et al., 2007; Dobler and Ahrens, 2008).

In the course of this PhD, regression models, analogue methods, as well as scaling methods are investigated in more detail. Their methodological description can be found in Chapter 4 and Chapter 5. These techniques are chosen to cover linear as well as non-linear, parametric as well as non-parametric, and techniques based on point-wise as well as spatially distributed predictors.

Concerning conditional probability methods/weather generators, and, e.g., not applied non-linear regression techniques such as artificial neural networks (ANNs), interested readers are directed to the respectively listed journal articles for further information.

In general, the choice upon the applied ESD technique highly depends on the characteristics of the predictand. If, for example, the focus is on monthly time series at single observational stations, linear regression models in most cases will do a good job, as the predictand is expected to be normally distributed, also for precipitation amount, and the model can be applied station by station, without taking into account spatial correlations. If however,

non-normally distributed, daily precipitation fields are demanded, more sophisticated methods as analogue procedures, stochastic weather generators or scaling approaches might be advisable (von Storch et al., 2000; Themeßl 2006; Themeßl et al., 2010).

### *3.2.6 Predictors*

Predictor variables can represent either grid cell values or spatially coherent patterns (e.g., Huth, 1999). While the application of single grid cell time series aims at taking into account small scale characteristics such as convection, spatially coherent predictor patterns enables to put more emphasis on the large scale flow. Using single grid cells, here also referred to as point-wise application, the question of the skillful spatial scale of a climate model occurs. According to the literature, 3 x 3 grid cells represent the minimum number of grid cells, which should be interpreted. Less grid cells are likely to be inaccurate due to numerical noise or parameterization schemes applied (Benestad et al., 2008; Kapper et al., 2010). Applying spatially coherent fields, the question of the location and extension of the domain becomes crucial. The "optimal" domain should be decided upon the inclusion of the relevant processes affecting the region (Benestad et al., 2008; Wilby et al., 2004). This may also result in predictor domains remote of the predictand area as demonstrated in Wilby and Wigley (2000). Spatially coherent predictors can be either obtained subjectively or objectively by classifying reoccurring spatial atmospheric distributions with distinct meteorological attributes. Subjective approaches comprise well known synoptic classification schemes as the European Grosswetterlagen (Hess and Brezowsky, 1977), or the British Island Lamb Weather Types (Lamb, 1972). Objective weather classification is based on statistical methods such as principle component analysis (PCA), canonical correlation analysis (CCA), clustering (mostly in conjunction with PCA analysis) or on recent developments like ANN (Self organizing maps) and fuzzy rule based methods. Additionally, airflow indices and circulation indices as the North Atlantic Oscillation (NAO) or the El Niño- Southern Oscillation (ENSO) can be found in this context (Yarnal et al., 2001).

As the predictor–predictand link may vary temporally, it is often found in literature that predictor variables are selected, e.g., on seasonal basis (e.g., Boé et al., 2007). Besides empirically knowledge, statistical methods can then be applied in this context in order to

objectively select predictors. These so called screening approaches are described in Appendix A in more detail.

### 3.2.7 Predictands

Predictands can represent single station data, areal means, spatial coherent fields or probability density function (*pdf*) characteristics (Maraun et al., 2010). Usually, predictand and predictor(s) are different variables but not necessarily (compare Chapter 4). Using *pdf* characteristics in this context enables to implicitly capture the effects of climate change on the shape of the predictand's distribution (Benestad et al., 2005; Benestad, 2007; Wang and Zhang, 2008). However, transient, e.g., daily time series cannot be generated hereby, which restricts its application in climate impact studies.

Most ESD applications deal with standard mean meteorological parameters as rainfall (e.g., Salathé, 2003), temperature (e.g., Huth, 2003), wind (e.g., Kaas et al., 1996), solar radiation (e.g., Wilks and Wilby, 1999) or snow (e.g., Bednorz, 2004). However, ESD is also capable of inferring information about extreme events (e.g., Billet et al., 1997; Brandsma and Buishand, 1998; Pryor et al., 2005; Benestad 2007; Goodess et al., 2007) as well as non-meteorological parameters such as agricultural yields (e.g., Sun et al., 2007). This flexibility in application represents one major strengths of ESD compared to DD. A general comparison of strengths and weaknesses between ESD and DD is given in the Appendix B.

### 3.2.8 Downscaling of extremes

Extreme events and their impacts represent large threats to our society (OcCC, 2003). According to the IPCC (2001) an extreme weather event is an event that is rare within its statistical reference distribution at a particular place (compare Figure 3.2). Due to differing global climate zones (e.g., Köppen and Geiger, 1961) meteorological situations that are common in one part of the world can be regarded as devastating extremes in another (compare maximum precipitation amounts in middle Europe and the same parameter at Mount Waialeale, Hawaii). Thus, instead of using absolute thresholds to define rare events, distribution-based definitions as the $10^{th}$ or $90^{th}$ percentile may be applied. Many climatological studies have investigated these sometimes called soft extremes (e.g., Klein Tank and Können, 2003; Goodess et al., 2007). Additional definitions of extreme indices

are listed in Table 6.2. An extreme climate event, for completeness, is an average of a number of weather events over a certain period of time, an average which is itself extreme (e.g., rainfall over a season) (IPCC, 2001).
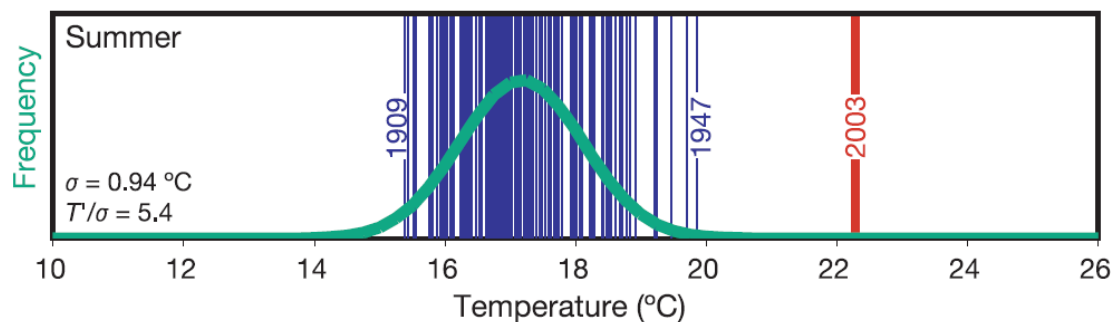


Figure 3.2 Distribution of Swiss summer temperatures for 1864–2003. The fitted Gaussian distribution is indicated in green. The values in the lower left corner lists the standard deviation and the 2003 anomaly normalized by the 1864–2000 standard deviation (T′/δ) (from Schär et al., 2004).

Especially since the Second Assessment Report of the International Panel on Climate Change (IPCC; IPCC, 2001) the question if the climate has become more extreme and how future conditions could look like has been issue to various studies. These studies comprise analyses of historical observational data (e.g., Alexander et al., 2006; Klein Tank and Können, 2003) as well as of future climate conditions (e.g., Booij, 2002; Räisänen et al., 2004; Frei et al., 2006). Although the latter studies mainly focus on dynamical climate modeling also ESD techniques can be used to estimate extreme conditions as already mentioned in the previous section. For this purpose, ESD is classified into two categories:

- direct approaches and
- indirect approaches.

Indirect approaches model daily time series of meteorological variables in a first step, and then derive extreme indices of this time series. Direct approaches calibrate their ESD models on extreme indicators (compare Schmidli et al., 2006) or attributes of extreme value distributions (Wang and Zhang, 2008). If direct approaches are build on extreme value

distributions, they can also account for problematic attributes of extremes for modeling, namely their rarity and their heavy tail distributions (e.g., Smith, 2001; Katz et al., 2002).

Goodess et al. (2007) compare 22 ESD methods for extreme conditions in a comprehensive inter-comparison study along with the Statistical and Regional dynamical Downscaling of Extremes for European regions (STARDEX) project. They conclude that the performance of the respective ESD methods varied across regions, seasons and indices. No systematic best performing method could be found. These results are confirmed, e.g., by Haylock et al. (2006) or Schmidli et al. (2006). The latter studies also compares dynamical downscaling and ESD for precipitation and results in overall comparable skill for both strategies. Goodess et al. (2007), thus, recommends taking into account an ensemble of techniques rather than one selected method, plus an ensemble of RCMs and GCM in order to sample the entire range of uncertainty. This recommendation of Goodess et al. (2007) is also, e.g., underlined by Murphy (2000).

In the presented papers in Chapter 4 and Chapter 5, indices of extremes will be obtained indirectly, derived from daily corrected time series. Appendix C additionally complements the results of Chapter 5 focusing on the generation of suitable error correction function of one selected DECM for extreme conditions.

# 4 Paper I: Empirical-statistical downscaling and error correction of daily precipitation from regional climate models

## 4.1 Abstract

Although regional climate models are powerful tools for describing regional and even smaller scale climate conditions, they still feature severe systematic errors. In order to provide optimized climate scenarios for climate change impact research, this study merges linear and non-linear empirical-statistical downscaling techniques with bias correction methods and investigates their ability for reducing regional climate model error characteristics. An ensemble of seven empirical-statistical downscaling and error correction methods (DECMs) is applied to post-process daily precipitation sums of a high resolution regional climate hindcast simulation over the Alpine region, their error characteristics are analyzed and compared to the raw RCM results.

Drastic reductions in error characteristics due to application of DECMs are demonstrated. Direct point-wise methods like quantile mapping and local intensity scaling as well as indirect spatial methods as non-linear analogue methods yield systematic improvements in median, variance, frequency, intensity, and extremes of daily precipitation. Multiple linear regression methods, even if optimized by predictor selection, transformation, and randomization, exhibit significant shortcomings for modeling daily precipitation due to their linear framework. Comparing the well-performing methods to each other, quantile mapping shows the best performance, particularly at high quantiles, which is advantageous for applications related to extreme precipitation events. The improvements are obtained regardless of season and region, which indicates the potential transferability of these methods to other regions.

## 4.2 Introduction

General circulation models (GCMs) are established tools for estimating the large-scale evolution of the Earth's climate, but due to their relative coarse horizontal resolution, they are not suited to properly represent regional scale climate characteristics. Therefore, dynamical downscaling techniques are often applied to derive regional-scale information from GCMs. Limited area regional climate models (RCMs) are forced by lateral boundary conditions of GCMs or reanalysis products and simulate the regional climate over a certain area on a finer grid (typical horizontal resolution 10 km to 50 km; Giorgi and Mearns, 1991, 1999; Wang et al., 2004). RCMs have considerably advanced in reproducing regional climate, but are nevertheless known to feature systematic errors (e.g., Frei et al., 2003; Hagemann et al., 2004; Suklitsch et al., 2008, 2010). Particularly, small-scale patterns of daily precipitation are highly dependent on model resolution and parameterization and can often not be used directly in climate change impact assessment studies (Fowler et al., 2007). Statistical post-processing of RCMs, according to the concept of model output statistics (MOS; Wilks, 1995), may help to overcome these problems, leading to qualitatively enhanced climate information. Such statistical post-processing of RCMs is mostly neglected in climatological studies as traditional empirical-statistical downscaling methods (ESDMs) are preferably applied according to the concept of perfect prognosis (perfect prog; Wilks, 1995). Perfect prog downscaling determines a statistical model (transfer function) between suitable large scale observation/reanalysis data and local observations (e.g., Wilby and Wigley, 1997; Murphy, 1999; Schmidli et al., 2006) which is applied directly to GCMs for generating regionalized climate scenarios (e.g., Schmidli et al., 2007) but without the intention of model error correction (compare Figure 4.1). RCMs are often favored to traditional empirical-statistical downscaling because they are capable to simulate regional-scale climate feedback effects and were already shown to create added value compared to GCMs on the meso- and regional scale for surface variables (e.g., Wang et al., 2004; Feser, 2006).
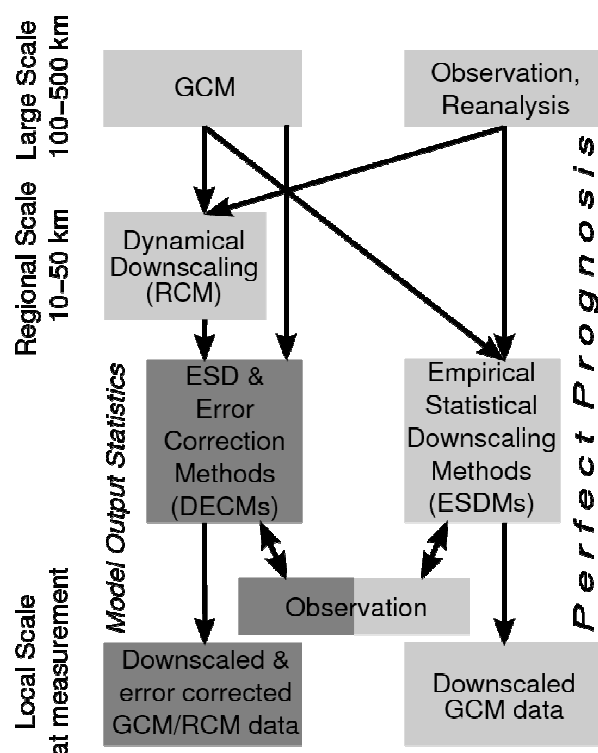
Figure 4.1 Scheme of different downscaling approaches. Darker grey color indicates the applied approach for this study. Traditional empirical-statistical downscaling (right pathway) calibrates the statistical transfer function between large-scale observation/reanalysis data and local scale observations. These empirical-statistical relationships can be used for downscaling of any GCM. DECMs (left pathway) are calibrated on RCM (or GCM) data and local observations, account for downscaling as well as model errors, but can only be applied to the model they are calibrated for.

Recently, the availability of regional RCM-based climate scenarios for Europe tremendously increased due to projects like ENSEMBLES (http://ensembles-eu.metoffice.com/) or PRUDENCE (http://prudence.dmi.dk/). However, due to the error characteristics of RCMs and when climate information at the point scale is needed, statistical transfer functions are inevitable to provide suitable climate scenario data for climate change impact research.

Aiming at a reduction of RCM error characteristics as well as of resolution at the same time, this study follows the principles of MOS and compares direct ESDMs that solely rely on modeled precipitation (e.g., Wood et al., 2004; Graham et al., 2007; Dobler and Ahrens, 2008)

to indirect ESDMs that derive fine-scale information by relating various model outputs (predictors − frequently upper air atmospheric data and not necessarily model precipitation) to observed surface variables (predictands − such as precipitation; e.g., Wilby and Wigley, 1997; Goodess et al., 2007; Benestad et al., 2008). All methods are applied to RCM results instead of their usual application to GCMs (compare Figure 4.1). In order to distinguish this application from perfect prog downscaling (which does not regard model errors), these methods are referred to as "empirical-statistical downscaling and error correction methods" (DECMs) henceforth. Their skill is assessed by analyzing their success in modeling daily precipitation on the station scale in the orographical complex Alpine region in Austria.

The study is organized as follows: Section 4.3 introduces the applied RCM and observational data as well as the study region. Section 4.4 describes the implemented DECMs, which are evaluated in Section 4.5. Finally, Section 4.6 summarizes the key findings of the study.

## 4.3 Data and study region

For this study the mesoscale limited area model MM5 (Dudhia et al., 2005) from the Penn State University (PSU) and the National Center of Atmospheric Research (NCAR) is used to provide the predictor data. MM5 is a non-hydrostatic, terrain-following sigma-coordinate model designed to simulate mesoscale atmospheric circulation. The simulations used in this study originate from the Austrian project "reclip:more─Research for Climate Protection: Model Run Evaluation" (Loibl et al., 2007) in which parts of the ERA-40 reanalysis (Uppala et al., 2005) were dynamically downscaled (hindcast simulation) in a two step nesting approach (Gobiet et al., 2006). The data covers the domain shown in Figure 4.2(a) with a horizontal grid spacing of 10 km.

Temporally, the data is given in 6 hour time steps for the time-span from 1981 to 1990 and the single year 1999. The year 1999 is treated as any other year in the 11 year period. For all parameters, except precipitation, daily mean values are calculated from 00:00 UTC to 23:59 UTC. For comparison to the observational data, daily precipitation is summed up between 06:00 UTC and 05:59 UTC the following day. All parameters used in the analyses are listed in Table 4.1.

22
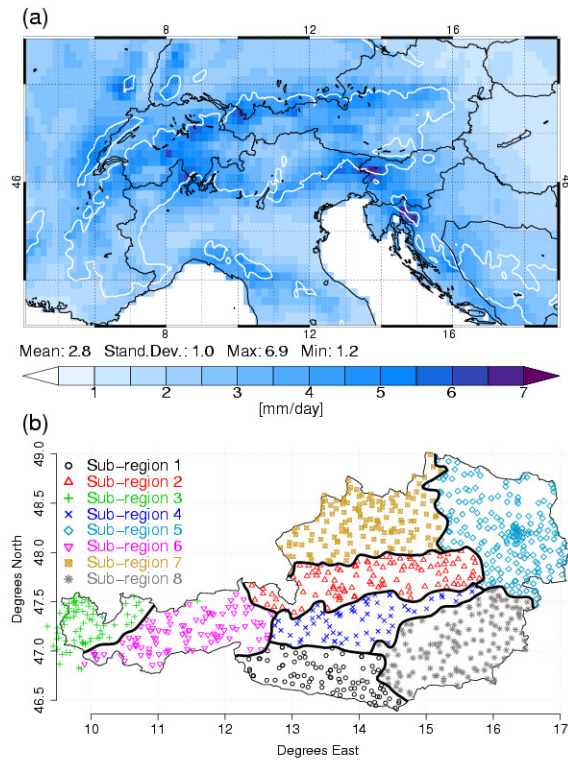
Figure 4.2 Upper panel (a): Spatial domain of the RCM data. Additionally, climatological precipitation conditions for the Alpine region including Austria from 1981 to 1990 according to HISTALP observations. The bold, white line indicates the 800 m isoline for better orientation. Lower panel (b): Location of the 919 observational stations within Austria and the clustered eight homogeneous precipitation regions.

The observed daily precipitation sums, provided by the Austrian Meteorological Service (ZAMG) and the Austrian Hydrological Service (HZB), are used as predictand for empirical-statistical error correction at 919 observation sites, which are evenly distributed across the entire area of Austria (Figure 4.2(b)). The data is quality checked, but not homogenized. The 919 observational stations fulfill the conditions of at least 80 % of data availability and insignificantly changed data distributions after station replacements. The latter condition was tested by a non-parametric Wilcoxon rank-sum test (Wilks, 1995).

| Variable | Abbreviation | Level |
|---|---|---|
| Geopotential height | zg | [700,500] hPa |
| Mixing ratio | q | [850,700] hPa |
| Eastward wind | u | [850,700] hPa |
| Northward wind | v | [850,700] hPa |
| Upward air velocity | w | [700] hPa |
| Vapour pressure | e | [850,700] hPa |
| Saturation vapour pressure | es | [850,700] hPa |
| Total surface precipitation | pre | surface (sfc) |
| Convective rain | accrcon | sfc |
| Advective rain | accrnon | sfc |
| Precipitable water | pwat | sfc |
| Atmosphere cloud condensed water content | iclc | sfc |
| Surface air pressure | psfc | sfc |
| Temperature | t2 | 2 m |
| Mixing ratio | q2 | 2 m |
| Vapour pressure | e2 | 2 m |
| Saturation vapour pressure | es2 | 2 m |
| Relative humidity | f2 | 2 m |
| Eastward wind | u10 | 10 m |
| Northward wind | v10 | 10 m |
| Sea level pressure | pslv | sea level (slv) |

Table 4.1 The MM5 predictor variables used in this study.

Although Austria is a rather small country, it features several climate provinces. These provinces originate from three main airflow directions (Atlantic, Mediterranean, and continental Eastern Europe) and their interaction with the Eastern Alps which cover large parts of the country and vertically range from basins and low altitude regions of a few hundred meters in the eastern foothills of the Alps to the mountainous western parts of the Eastern Alps up to nearly 4000 m. The central, northern, as well as the southern parts of the Eastern Alps block the Atlantic and Mediterranean moist airflows, force them to raise and rain out, resulting in annual precipitation maxima in the northerly and southerly upslope regions of the Alpine crest. Due to strong advection of wet Adriatic air masses in summer months, southern and south-eastern parts of Austria frequently feature severe thunderstorms as well as hail.

Decreasing precipitation amounts towards the flatter eastern parts result from the predominating continental air masses there. Additionally, the inner alpine valleys feature reduced precipitation amounts due to the rain shadowing of the mountain ranges (compare Figure 4.2(a); Cebon et al., 1998; Matulla et al., 2003). The observed climatology given in Figure 4.2(a) covers the period 1981 to 1990 according to the focus period of this study, but the spatial patterns as well as the magnitudes are comparable to the climatic conditions between 1971 and 1990 shown in Frei and Schär (1998). Accounting for the regional climatological differences, the observational stations are clustered into eight sub-regions (Figure 4.2(b)), based on correlated daily precipitation. The clustering method is described in Suklitsch et al. (2008). The clusters represent spatially homogenous regions, with only two displaced stations (both in sub-region 6). Despite their displacement, these two stations are allocated to their original cluster.

## 4.4 Methods

Seven statistical approaches are applied in this study. They are applied for each observational station separately. The approaches are selected to span indirect DECMs (Section 4.4.1) as well as direct DECMs (Section 4.4.2). The former comprise linear and non-linear techniques. Furthermore, the selected methods cover point-wise approaches as well as spatial approaches that use distributed predictors. Point-wise approaches relate 3 x 3 adjacent RCM grid cells to each station. Spatial approaches are based on meteorological fields and use principle components (PCs) from the entire RCM domain shown in Figure 4.2(a) to build transfer functions to the station scale. The PCs originate from principle component analysis (PCA), which is also referred to as empirical orthogonal function (EOF) analysis in geophysics (e.g., von Storch and Zwiers, 1999; Zorita and von Storch, 1999).

### 4.4.1 Indirect DECMs

Multiple Linear Regression (MLR)

In addition to its application in MOS, MLR is frequently found in statistical downscaling of GCM data (e.g., Kilsby et al., 1998; Huth, 1999; Murphy, 1999; Schoof and Pryor, 2001; Hay and Clark, 2003) as well as in climate change impact analyses (e.g., Alexandrov and Hoogenboom, 2000). In general, linear regression models establish a linear transfer function between one or more predictors and the predictand such that

$$Y^{\mathrm{MLR}} = \alpha + \sum_{p=1}^{l} \beta_p X_p + \varepsilon \qquad (4.1)$$

with $\alpha$ being the intercept, $\beta$ the regression coefficients, $\varepsilon$ the error term, $X_p$ the $p$ predictor variables, and $Y^{\mathrm{MLR}}$ the estimated predictand.

For this study, a MLR based on ordinary least squares (OLS) is applied on daily basis, for each season separately and point-wise since preliminary tests favored point-wise to spatial application. In these tests, point-wise and spatial MLR (EOF-based) yielded similar results for seasonal means. However, the variability of daily errors was mostly smaller and the shapes of the modeled distributions were closer to the observed distributions in point wise MLR than in EOF-based MLR. Furthermore, several data transformations were analyzed in preliminary tests to account for daily precipitation's non-normal distribution and non-linear predictors-predictand relationships (Wilks, 1995; Kidson and Thompson, 1998). Results are shown only for best performing cube root transformation of the predictand (denoted as MLRT; Helsel and Hirsch, 2002).

Predictors are chosen by a semi-objective procedure for each station. Physical meaningful variables are pre-selected empirically before a two-step objective predictor selection method is performed: Firstly, a stepwise regression, based on the Akaike Information Criteria (Wilks, 1995), is used to reduce the number of potential predictors; secondly, an all subset regression (von Storch and Zwiers, 1999) selects the most influencing combinations of predictors, limited

to a maximum number of four predictors per combination, according to the adjusted coefficient of determination (R²; Helsel and Hirsch, 2002).

Multiple Linear Regression with Randomization (MLRR)

Climate impact studies often need precipitation time series with realistic day-to-day variability. Since MLR models reduce variability because the regression line is fitted to pass through the centroid of the data (Helsel and Hirsch, 2002) and only a part of local climate variability is related to larger-scale variability in predictors (Fowler et al., 2007), von Storch (1999) proposed randomization of time series to recover their original variability. MLRR extends the MLR estimation $Y_{t,i}^{\text{MLRR,val}}$ in a given validation period (val) at station $i$ and day $t$ by adding noise $R_i^{\text{MLR,cal}}$ which represents the unexplained part of the regression model (compare Equation (4.1); Dehn and Buma, 1999) according to

$$Y_{t,i}^{\text{MLRR, val}} = Y_{t,i}^{\text{MLR, val}} + R_i^{\text{MLR,cal}} \quad . \tag{4.2}$$

$R^{\text{MLR,cal}}$ is obtained from classified MLR residuals of the calibration period (cal) and grouped in four classes corresponding to the respective quantity of $Y^{\text{MLR,cal}}$ between zero precipitation, the model (mod) wet-day threshold $WT^{\text{mod}}$, the 50th percentile of the time series, and the maximum estimated precipitation. $WT^{\text{mod}}$ is defined after Schmidli et al. (2006) that the number of days greater or equal $WT^{\text{mod}}$ in the calibration period equals the respective observed (obs) number of days greater or equal $WT^{\text{obs}}$ in the calibration period. $WT^{\text{obs}}$ is defined as 1 mm/day.

The Analogue Method (AM)

Resampling approaches can be classified as stochastic precipitation models (von Storch and Navarra, 1999). Their primary application area is hydrology where they are used to generate a large number of synthetic observations as input to hydrological models in order to assess their uncertainties (Mehrotra and Sharma, 2006). The analogue method (AM) represents a special case of resampling. In the context of AM, the resampling is conditioned on atmospheric states (predictors; e.g., Zorita and von Storch, 1999) which enables its application for downscaling purposes. Conceptually, AM compares the atmospheric state on the day under consideration ($t$) to an archive of historic atmospheric states and determines the most similar historic atmospheric state – the analogue – according to some measure of similarity. The local weather on this analogue date u is then resampled as an estimate of the predictand on day $t$ (Cubasch et al., 1996). Thus, for an adequate description of the local climate, particularly regarding extreme conditions, sufficiently long historic archives are necessary. The resampling limits AM to historical extremes, which can be considered as the method's main drawback. Further problems with consistency in the order of consecutive days may occur if atmospheric regimes are not well defined by the spatial predictors. In return, AM does not assume any particular probability distribution in the modeling process and enables to capture non-linear predictor-predictand relationships (Fernández and Sáenz, 2003; Benestad et al., 2008).

The crucial point of AM is the definition of similarity of atmospheric states (e.g., Wetterhall et al., 2005; Matulla et al., 2008). A method based on PCs (e.g., Zorita and von Storch, 1999) is implemented in this study. PCs are derived from fields of all RCM variables given in Table 4.1. Analogues are found for each season by minimizing a weighted Euclidean distance:

$$\sum_{k=1}^{n}\left\{d_i^k\left[R_{t,i}^{\mathrm{cal},k} - F_{t,i}^{\mathrm{val},k}\right]^2\right\} \to \min . \tag{4.3}$$

*F* represents the so-called feature vector at validation day *t* which takes into account *k* predictors. $R$ corresponds to the respective historical archive in the same PC phase space as the predictors. The weighting vector $d$ (Fernández and Sáenz, 2003; Imbert and Benestad, 2005) consists of normalized eigenvalues and reflects the importance of each considered predictor. If the same Euclidean distance is found several times in the historical sample, the temporally first condition is considered. Additionally, an unweighted distance was tested with slightly worse results (not shown).

Contrary to the standard AM application with multisite prediction, this study uses a site-specific AM. Such a site-specific approach weakens the merit of maintaining the spatial covariance structure of the predictand, but enables the selection of the most important predictors for each station separately. The first three PCs of each RCM variable are proposed as independent predictors to the automated predictor selection scheme (see MLR description). The PCs are standardized to unit variance. Limitations in AM's non-linearity due to linearly selected predictors are accepted. The predictor selection results in the seasonally prevailing predictor combination of four predictors (*k* = 4).

The Nearest Neighbor Analogue Method (NNAM)

Extending the search for the analogue towards a probabilistic approach, the nearest neighbor analogue method (e.g., Brandsma and Buishand, 1998; Mehrotra and Sharma, 2006; Moron et al., 2008) randomly chooses the analogue situation from the *nn* most similar historical conditions. Consequently, increased modeled variability can be expected with the drawback that equal predictor conditions on time *t* and *t+1* may result in different local predictands. The same predictors as applied to AM are considered. Equal to AM, a weighted Euclidean distance in PC phase space provides the measure of similarity. Instead of picking the most similar historic condition to prediction time *t* (*nn* = 1), the nn days with smallest Euclidean distances (nearest neighbors) are retained. The analogue is selected using a discrete probability distribution that weights the *nn*-days according to

$$p_j = \frac{\frac{1}{j}}{\sum\limits_{i=1}^{nn} \frac{1}{i}}, \ j = 1,...,nn \ , \tag{4.4}$$

with $p_j$ being the probability of the $j$ closest neighbors (Lall and Sharma, 1996). By this means, higher weights are given to closer neighbors. Several tests with $nn = 5$, $nn = 10$, $nn = \sqrt{n}$, with $n$ being the calibration sample size were performed (compare Beersma and Buishand, 2003). In this study $nn = 5$ is used. Unweighted random selection of the analogue was tested as well, but is outperformed by the presented approach (not shown).

### *4.4.2 Direct DECMs*

Local Intensity Scaling (LOCI)

LOCI is a direct DECM and represents one traditional bias correction method (also compare Graham et al., 2007; Leander and Buishand, 2007), which is based on the work of Widmann et al. (2003), suggested by Schmidli et al. (2006) and successfully applied by, e.g., Salathé (2003), Dobler and Ahrens (2008), or Moron et al. (2008). The basic idea of direct DECMs is that climate model precipitation integrates all relevant predictors. Deviations between climate model precipitation and regional- or local-scale precipitation observations are in first order due to systematic climate model errors and an incomplete or inaccurate representation of the orography (Schmidli et al., 2006). Thus, instead of using various predictors to create local weather, LOCI applies a spatially varying scaling to climate model precipitation accounting for its long term bias at the location of the observation.

Following the approach of Schmidli et al. (2006), a separate correction of wet-day frequency and wet-day intensity is applied point-wise and for each day of the year separately. With the definitions of $WT^{\text{mod}}$ and $WT^{\text{obs}}$ (see MLRR description) climate model precipitation $X_t^{\text{val}}$ is corrected by Equation (4.5) using scaling factor $S$ from Equation (4.6). $\overline{X}_t^{\text{wet,cal}}$ and $\overline{Y}_t^{\text{wet,cal}}$ represent climatological means on wet days (i.e., days with precipitation greater or equal $WT$) of the modeled and the respective observed precipitation data over the

calibration period at day $t$. Only pair-wise recorded modeled and observed data is used for calibration.

$$Y_{t,i}^{\text{val}} = \max\left(S_{t,i}\left(X_{t,i}^{\text{val}} - WT_{t,i}^{\text{mod,cal}}\right) + WT_{t,i}^{\text{obs,cal}}, 0\right) \tag{4.5}$$

$$S_{t,i} = \frac{\overline{Y}_{t,i}^{\text{wet,cal}} - WT_{t,i}^{\text{obs,cal}}}{\overline{X}_{t,i}^{\text{wet,cal}} - WT_{t,i}^{\text{mod,cal}}} \tag{4.6}$$

Contrary to previous applications, this study uses a moving window approach centered over focus day $t$ to calculate $\overline{X}_t^{\text{wet,cal}}$ and $\overline{Y}_t^{\text{wet,cal}}$. By this means, inhomogeneities at the end of fixed calibration periods (e.g., months, seasons) are avoided, and the dependence of model errors on the time of the year is included. Moving time windows between 15 days and 61 days were investigated. A window size of 61 days is chosen to enable an annual-cycle sensitive correction as well as a sufficient large sample size. Further, setting $WT^{\text{obs}}$ to 0.3 mm/day and 0.5 mm/day instead of the 1 mm/day standard value was investigated, but resulted in no significant differences (not shown).

Quantile Mapping (QM)

Extending the correction from means (LOCI) to the entire distribution, QM corrects for errors in the shape of the distribution and is therefore capable to correct errors in variability as well. This quantile based approach originates from the empirical transformation of Panofsky and Brier (1968) and was successfully implemented in hydrological applications (Dettinger et al., 2004; Wood et al., 2004; Boé et al., 2007) but recently also for error correction of RCMs (Dobler and Ahrens, 2008; Piani et al., 2009).

For this study, QM is based on point-wise and daily constructed empirical cumulative distribution functions (*ecdfs*; Wilks, 1995) of modeled and observed datasets in the calibration period. This is in contrast to other bias correction studies where theoretical *cdfs* are estimated only from wet days (e.g., Ines and Hansen, 2006; Dobler and Ahrens, 2008; Piani et al., 2009). Using *ecdfs*, QM is generally applicable to all possible meteorological parameters, whereas applications based on *cdfs* may become problematic for parameters that do not fit to theoretical

functions such as global radiation, where the *ecdf's* shape is changing with the season (compare Camuffo, 1978). Equal to LOCI, a 61 days moving window, centered over the focus day, is used for *ecdf* construction. The correction function transfers raw RCM output $X^{\text{val}}$ to the corrected estimate $Y^{\text{val}}$ such that

$$Y_{t,i}^{\text{val}} = ecdf_{t,i}^{\text{obs,cal}\,-1}\left(ecdf_{t,i}^{\text{mod,cal}}\left(X_{t,i}^{\text{val}}\right)\right), \qquad\qquad (4.7)$$

with $ecdf^{-1}$ indicating the inverse *ecdf*, thus a data quantile. As this purely empirical QM only maps modeled values to observed values, no new extremes (outside the observed range) can be obtained. This is a suitable approach for our study, since we apply the correction to a historical hindcast simulation. For applications to future climate simulations, however, some kind of extrapolation beyond the range of observations has to be added to allow for "new extremes" (e.g., Boé et al., 2007).

## 4.5 Results and discussion

### 4.5.1 Validation framework

Statistical approaches implicitly assume stationarity in their transfer functions in the case of indirect DECMs or in model error characteristics in the case of direct DECMs (Wilby, 1997; Benestad et al., 2008). If this assumption is violated, statistical models cannot account for changes described by predictor forcings. As this assumption cannot be approved in advance, a temporal cross-validation framework is applied which repeatedly divides the data period into a calibration (10 years) and independent validation period (1 year). By this means, each year is estimated and evaluated independently with the remaining 10 years used for model calibration (sometimes denoted as "leave one out" cross-validation; see Figure 4.3).
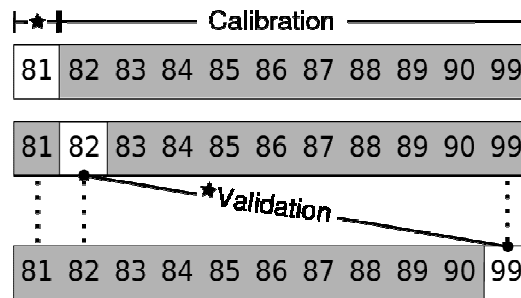
Figure 4.3 The "leave one out" cross validation scheme. Each of the 11 simulated years is post-processed once independently from the remaining 10 years used for calibration.

For evaluation purpose, model skill scores as well as model error characteristics as used. The models' performances are analyzed via mean skill scores and mean model error characteristics, averaged over all validation periods, as well as using the entire, not averaged, 11 year validation time series. Mean skill scores and model error characteristics are presented in Figures 4.5–4.9. The models' performances represent the station scale as each statistical model is calibrated and evaluated separately station by station. For graphical representation, the station-wise evaluation results are spatially averaged in sub-region 6, sub-region 8 (compare Figure 4.2(b)), and for entire Austria. The two sub-regions are selected because of their different climate characteristics. Sub-region 6 is mainly dominated by westerly flows from the Atlantic with high precipitation amounts, whereas sub-region 8 features more continental dry characteristics with additional influence from the Mediterranean Sea.

The results are divided in 3 parts. The first describes the general characteristics of the uncorrected RCM within all regions. The second focuses on the characteristics of each DECM and the third part analyses the effectiveness of DECMs compared to uncorrected RCM results.

### 4.5.2 Regional climate model evaluation

Gobiet et al. (2006) already compared the MM5 precipitation data from 1981 to 1990 on monthly scale to HISTALP observations (Auer et al., 2007). The same comparison is shown in Figure 4.4 on seasonal basis.
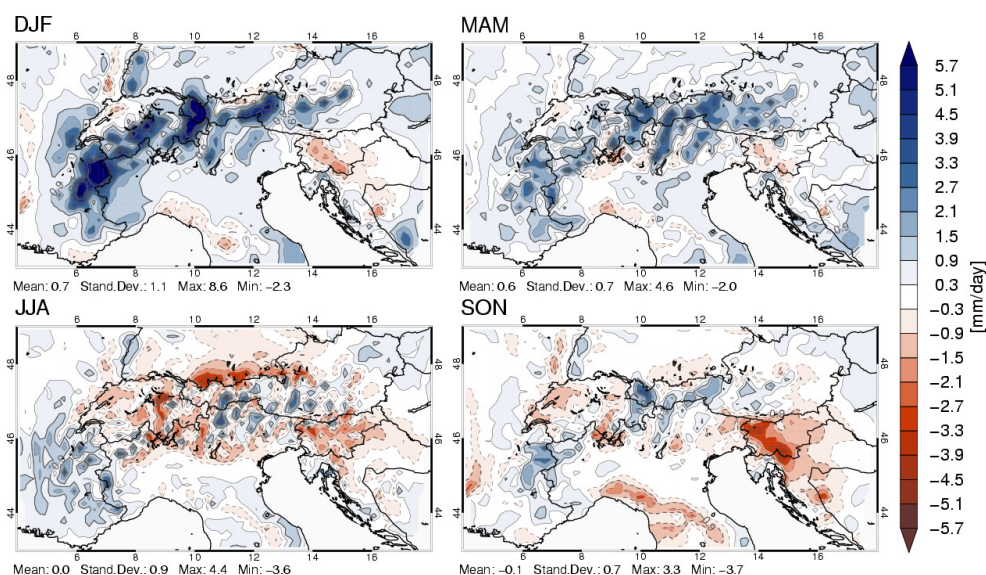
Figure 4.4 Mean seasonal differences of daily RCM precipitation sums compared to HISTALP observations. Continuous contours are positive; dashed contours are negative.

Regarding Austria, the RCM features seasonally and regionally varying error characteristics with strong precipitation overestimation along the Alpine crest in winter (DJF), an overall good performance in summer (JJA), and underestimation at the southern Austrian border in autumn (SON). Gobiet et al. (2006) argue that, besides possible model deficiencies, the well-known problematic precipitation measurement at high altitudes, especially in DJF, may partly cause the pronounced overestimation. Secondly, the reduced SON precipitation in south-eastern parts of Austria is probably related to an under-representation of northern Mediterranean cyclones and a consequent lack of humidity. These findings further motivate the selection of sub-regions 6 and 8 for evaluation as these regions cover the problematic areas.

### 4.5.3 Characteristics of the applied DECMs

Referring to the MLR predictor selection, Table 4.2 shows the most important seasonal predictors for the considered study regions. All three regions indicate precipitation (accrnon, accrcon, pre), humidity-related parameters at surface (q2, pwat), as well as eastward (u) and

northward (v) wind at 10 m and 750 hPa, and surface vapor pressure (e2) to be the dominant predictors for local precipitation. The composition of the predictor set varies seasonally with increased importance of the convective precipitation (accrcon) and northward wind in summer months, which reasonably corresponds to the regional climate characteristics (see Section 4.3). The dominance of RCM precipitation as predictor supports the assumption that RCM precipitation integrates large parts of the relevant information for local precipitation. Further, the frequently claimed integration of humidity as predictor (e.g., Giorgi and Mearns, 1991; Wilby and Wigley, 2000; Fowler et al., 2007) is supported.

Similar to point-wise predictor selection for MLR, Table 4.3 indicates that PCs of precipitation fields are by far most important for local precipitation for conditional resampling approaches. Further relevant predictors are pressure-related parameters at surface (es, psfc, pslv), geopotential height at 500 hPa (zg), and vertical velocity at 700 hPa (w).

| DJF | | | | | | MAM | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Sub-region 6** | | **Sub-region 8** | | **Entire Austria** | | **Sub-region 6** | | **Sub-region 8** | | **Entire Austria** | |
| Predictor | Prob [%] | Predictor | Prob [%] | Predictor | Prob [%] | Predictor | Prob [%] | Predictors | Prob [%] | Predictor | Prob [%] |
| pre_sfc | 19.5 | q2_2m | 16.2 | pre_sfc | 17.6 | q2_2m | 18.0 | accrnon_sfc | 19.6 | accrnon_sfc | 15.0 |
| accrnon_sfc | 17.0 | accrnon_sfc | 13.1 | accrnon_sfc | 12.1 | accrnon_sfc | 12.7 | accrcon_sfc | 14.4 | q2_2m | 13.3 |
| u_700hPa | 10.0 | e2_2m | 13.13 | q2_2m | 10.8 | pre_sfc | 12.7 | q2_2m | 8.6 | pre_sfc | 10.0 |
| iclc_sfc | 8.2 | pre_sfc | 12.9 | v_700hPa | 9.8 | e2_2m | 11.7 | e2_2m | 7.5 | e2_2m | 9.4 |
| e2_2m | 6.0 | v_700 hPa | 12.8 | e2_2m | 8.9 | u10_10m | 10.0 | zg_700hPa | 6.1 | accrcon_sfc | 6.6 |
| **JJA** | | | | | | **SON** | | | | | |
| **Sub-region 6** | | **Sub-region 8** | | **Entire Austria** | | **Sub-region 6** | | **Sub-region 8** | | **Entire Austria** | |
| Predictor | Prob [%] | Predictor | Prob [%] | Predictor | Prob [%] | Predictor | Prob [%] | Predictor | Prob [%] | Predictor | Prob [%] |
| pre_sfc | 14.5 | v_700hPa | 19.2 | v_700hPa | 17.4 | pre_sfc | 20.7 | q2_2m | 23.0 | accrnon_sfc | 16.6 |
| v_700hPa | 14.2 | pre_sfc | 17.6 | pre_sfc | 15.4 | accrnon_sfc | 16.0 | e2_2m | 22.8 | pre_sfc | 14.5 |
| u_700hPa | 10.7 | accrcon_sfc | 12.0 | accrcon_sfc | 7.7 | q2_2m | 11 | accrnon_sfc | 16.7 | q2_2m | 13.0 |
| accrnon_sfc | 10.2 | iclc_sfc | 8.1 | accrnon_sfc | 7.4 | e2_2m | 10.2 | accrcon_sfc | 9.7 | e2_2m | 12.0 |
| pwat_sfc | 8.5 | u_700hPa | 7.4 | v_850hPa | 6.3 | v10_10m | 7.2 | pre_sfc | 9.3 | accrcon_sfc | 8.6 |

Table 4.2 Seasonal predictor variables for MLR approaches in sub-region 6, sub-region 8, and for entire Austria according to their occurrence probability (Prob) after objective predictor selection. Only predictors with occurrence probability greater or equal 5 % are taken into account.

| DJF | | | | | | MAM | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Sub-region 6** | | **Sub-region 8** | | **Entire Austria** | | **Sub-region 6** | | **Sub-region 8** | | **Entire Austria** | |
| Predictor | Prob [%] | Predictor | Prob [%] | Predictor | Prob [%] | Predictor | Prob [%] | Predictor | Prob [%] | Predictor | Prob [%] |
| accrnon_sfc PC2 | 17.2 | accrnon_sfc PC3 | 15.6 | accrnon_sfc PC1 | 11.3 | accrnon_sfc PC2 | 16.5 | pre_sfc PC2 | 23 | pre_sfc_D2 PC2 | 15.6 |
| accrnon_sfc PC1 | 16.7 | es_2m PC2 | 14.7 | pre_sfc PC1 | 10.6 | accrcon_sfc PC1 | 16.2 | accrnon_sfc PC2 | 17.1 | accrnon_sfc PC2 | 14.1 |
| w_700hPa PC3 | 10.2 | accrcon_sfc PC1 | 13.8 | accrnon_sfc PC2 | 9.8 | pre_sfc PC2 | 11.5 | pre_sfc PC1 | 8.8 | accrcon_sfc PC1 | 10.1 |
| pre_sfc PC1 | 7.7 | pre_sfc PC1 | 8.3 | accrnon_sfc PC3 | 7.7 | accrnon_sfc PC1 | 6.5 | pre_sfc PC3 | 7.4 | pre_sfc PC1 | 5.8 |
| psfc_sfc PC2 | 6.7 | zg_500hPa PC2 | 6.1 | pre_sfc PC2 | 6.0 | | | iclc_sfc PC1 | 6.3 | | |
| JJA | | | | | | SON | | | | | |
| **Sub-region 6** | | **Sub-region 8** | | **Entire Austria** | | **Sub-region 6** | | **Sub-region 8** | | **Entire Austria** | |
| Predictor | Prob [%] | Predictor | Prob [%] | Predictor | Prob [%] | Predictor | Prob [%] | Predictor | Prob [%] | Predictor | Prob [%] |
| accrnon_sfc PC1 | 18.0 | pre_sfc PC2 | 18.9 | pre_sfc PC1 | 13.0 | accrnon_sfc PC3 | 22.5 | accrcon_sfc PC3 | 20.9 | accrnon_sfc PC3 | 22.1 |
| zg_500hPa PC2 | 10.7 | w_700hPa PC3 | 12.9 | pre_sfc PC2 | 12.1 | accrcon_sfc PC1 | 22.0 | accrnon_sfc PC3 | 18.5 | accrcon_sfc PC1 | 14.6 |
| w_700hPa PC3 | 10.5 | accrcon_sfc PC1 | 11.9 | accrnon_sfc PC1 | 10.8 | pre_sfc PC2 | 15.5 | accrcon_sfc PC1 | 15.5 | accrcon_sfc PC3 | 9.4 |
| pre_sfc PC1 | 8.0 | pre_sfc PC1 | 11.5 | accrcon_sfc PC1 | 6.0 | | | pre_sfc PC1 | 7.5 | pre_sfc PC2 | 7.6 |
| pslv_slv PC3 | 5.2 | accrnon_sfc PC1 | 7.9 | w_700hPa PC3 | 5.9 | | | pre_sfc PC2 | 6.6 | pre_sfc PC1 | 7.0 |

Table 4.3 As in Table 4.2 but with atmospheric predictor fields for the analogue methods. PC indicates the used principle component.

Figure 4.5 illustrates the annual evolution of the wet-day thresholds $WT^{mod}$ and the scaling factors $S$ used in LOCI. Both parameters feature distinct annual cycles, which indicate frequency overestimation in winter and intensity underestimation in summer in the RCM. $WT^{mod}$ ranges from 1.5 mm/day to 5 mm/day, which differs from the results of Schmidli et al. (2006), who found wet-day thresholds around 1 mm/day for the same region with LOCI, but applied to coarser ERA-40 reanalysis and calibrated on the entire year. $S$ varies around one with a reversed pattern compared to $WT^{mod}$ and shows comparable quantities to Schmidli et

al. (2006). The ranges of the magnitudes of $S$ and $WT^{mod}$ indicate that, besides for the summer season, the RCM precipitation error is overall dominated by a frequency overestimation error.
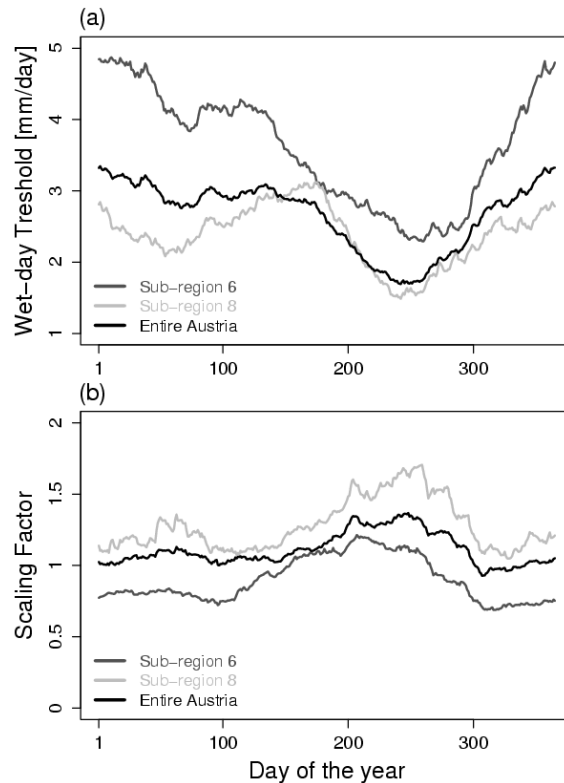


Figure 4.5 Annual cycles of wet-day threshold $WT^{mod}$ (a) and scaling factor $S$ (b) used for LOCI. Both parameters result from station-wise calibration and are spatially averaged.

The seasonal correction functions of QM in Figure 4.6 show differences of all percentiles between observed and modeled calibration *ecdfs* for all study regions. The respective precipitation quantities are indicated on the x-axes. Generally, and particularly in winter in sub-region 6, the RCM overestimates wet-day precipitation intensities which leads to partly significant negative correction values, especially at the highest precipitation intensities (i.e., at the highest percentiles). By contrast, particularly in summer in sub-region 8 significant positive correction values at the highest precipitation intensities indicate a lack of extreme precipitation events in the RCM data. The highest corrections are applied to the highest

percentiles and range from $-12$ mm/day (in winter in sub-region 6) to $+15$ mm/day (in summer and autumn in sub-region 8).
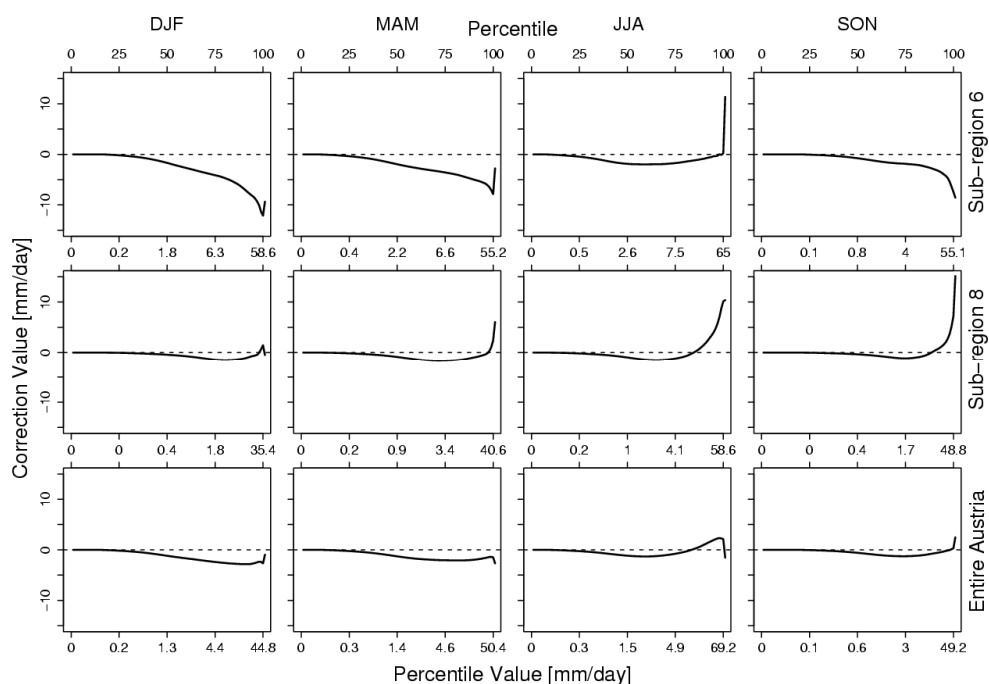


Figure 4.6 Seasonal correction functions derived from differences of all percentiles between observed and modeled *ecdfs* in sub-region 6 (upper panels), sub-region 8 (middle panels), and entire Austria (lower panels). The $0^{th}$ percentile represents the difference between *ecdfs'* minima, the $100^{th}$ percentile represents the difference between *ecdfs'* maxima. The precipitation quantities corresponding to these percentiles are indicated on the x-axes. The correction functions are obtained equally as described in Figure 4.5.

For entire Austria the correction function is strongly damped, which illustrates the importance of point-wise application where local error characteristics are taken into account instead of a broad spatial average. Abrupt changes of the correction function at highest modeled precipitation amounts, as illustrated in winter, spring, and summer in sub-region 6, are more probably related to statistical noise at these percentiles than to RCM error characteristics.

### *4.5.4 DECM evaluation*

For assessing the skill of the considered DECMs, their performances are evaluated regarding the median, variability, and indicators for extremes. Boxplots in Figure 4.7 display the median seasonal and annual differences between models and observations as lines in the middle of 25th and 75th quantile boxes derived from daily differences. Standardized Taylor diagrams (Figure 4.8; Taylor, 2001) show the normalized centered root-mean-square (RMS) difference of the different DECMs compared to observations as the distance to point 1 on the abscissa, the variance ratio between models and observations as the radial distance to the zero point, and the correlation between models and observations as the angle between the abscissa and the position vector (i.e. a perfect model would be displayed on point 1 of the abscissa). Error diagrams in Figure 4.9 illustrate the performances of the methods regarding precipitation intensity (SDII), wet-day frequency (Freq), the 95th percentile of all modeled days (Q95), and the 75th percentile on wet days (RQ75), where the latter two represent moderately extreme conditions. The results in Figure 4.9 are color-coded; lighter colors indicate smaller errors. Finally, a quantile-quantile plot in Figure 4.10 compares the 11 year seasonally and annually modeled to the observed distributions using all station time series within the respective region. This enables the analysis of the DECMs' performances for absolute extreme conditions. In the case of linear regression models, also negative precipitation values are produced. Though unphysical, we did not replace these negative values by zeros in order to avoid the introduction of biases or the reduction of variability in the evaluation statistics.

In Figure 4.7 the leftmost bars display the regional average RCM error characteristics. They indicate the largest error ranges in sub-region 6, as expected. The error range shows a high seasonality, which is related to overestimated temporal variability, shown in Figure 4.8. The results from Figure 4.5 and Figure 4.6, showing that higher modeled precipitation sums are positively biased, can be identified by the positive skewness of the difference bars.
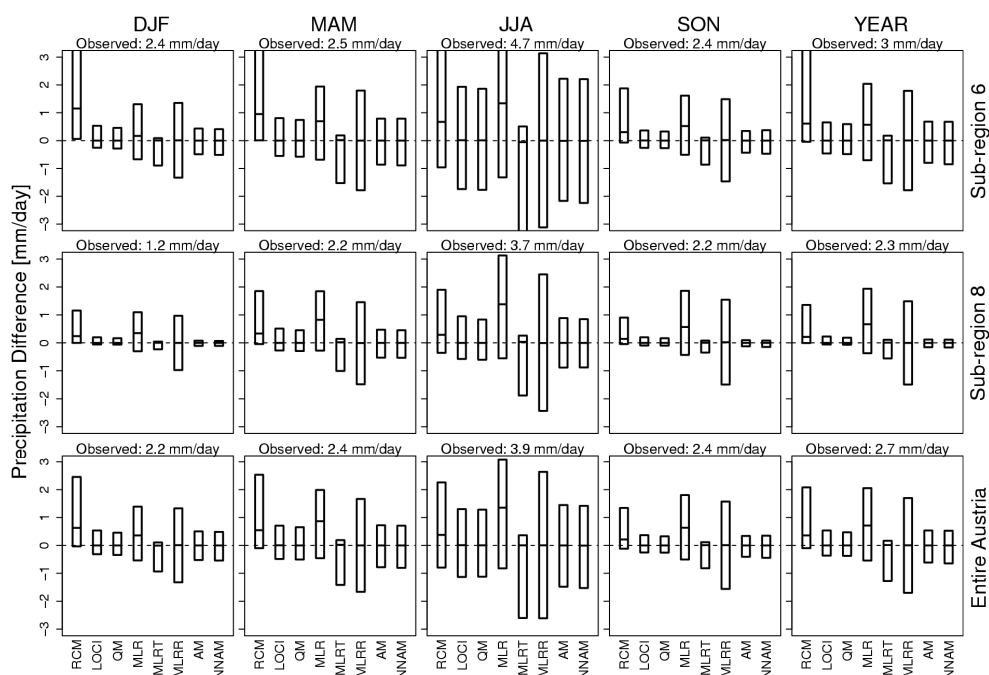
Figure 4.7 Seasonal and annual errors of the uncorrected RCM and the considered DECMs in sub-region 6 (upper panels), sub-region 8 (middle panels), and for entire Austria (lower panels). The boxes show the 75th percentile (upper limit), the median (line within the box), and the 25th (lower limit). The respective mean observed precipitation amount is given in the header of each panel. The statistics result from station-wise evaluation of daily precipitation data and are spatially averaged.

In comparison, all DECMs except MLR virtually correct the median error of daily precipitation to zero independent of season and region. QM systematically yields the best results followed by LOCI, AM, and NNAM. MLR partly even degrades error characteristics, which is probably related to non-linear relations between predictors and local daily precipitation as well as to non-normally distributed and heteroscedastic residuals (compare Wilks, 1995). However, with the simple extension of MLR to MLRR this deficiency can be removed due to the incorporation of error residuals (Equation (4.2)). MLRT corrects the median difference to nearly zero, but shifts the error distribution to negative values, whereas all other statistical approaches show nearly equally distributed differences around the median. Though only two sub-regions are presented here in detail, all DECMs show similar performances in all sub-regions shown in Figure 4.2.
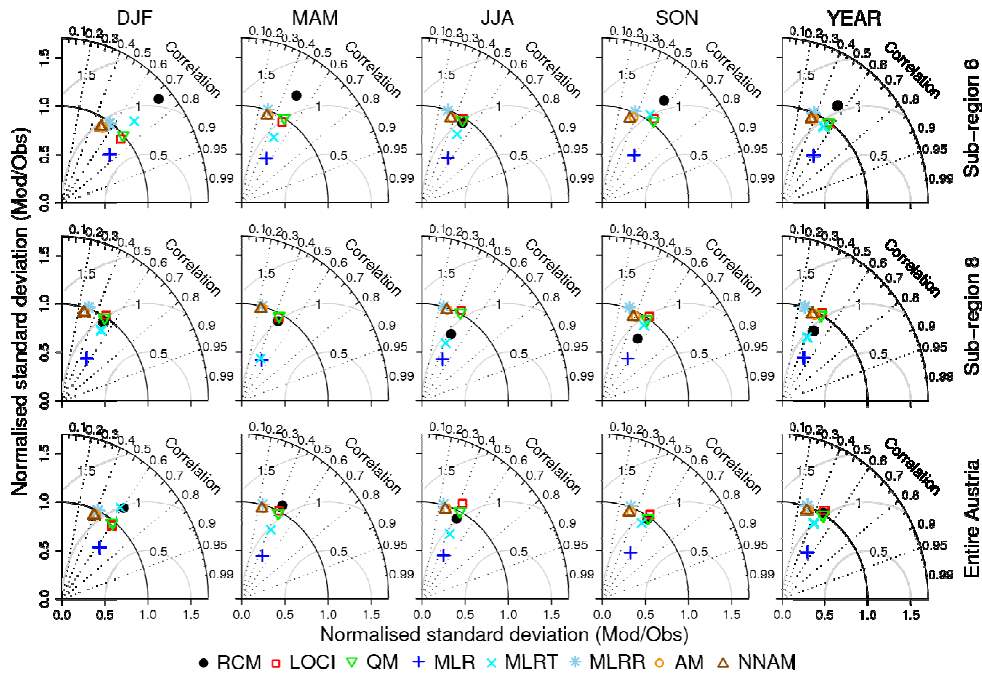
Figure 4.8 Seasonal and annual Taylor diagrams comparing the uncorrected RCM and the considered DECMs with observations in sub-region 6 (upper panels), sub-region 8 (middle panels), and for entire Austria (lower panels). The skill scores (normalized centered RMS, correlation as well as normalized variance ratio) are obtained equally as described in Figure 4.7.

The effect of DECMs on variability is displayed in Figure 4.8. In general, the RCM tends to overestimate day-to-day variability, but also shows pronounced underestimation in sub-region 8. These deficiencies are removed by most DECMs. Major problems remain for MLR which strongly underestimates variability and MLRT which shows non-systematic errors in variability with the tendency to underestimation. However, by adding error residuals (MLRR) the variability is modeled adequately. Minor problems are shown for LOCI, where especially for entire Austria a tendency to variability overestimation is indicated. Additionally, LOCI was more sensitive to a reduced window size than QM concerning the variance ratio (not shown). None of the DECMs is able to increase correlation. This is expected for direct DECMs as they solely rely on temporal characteristics of climate model precipitation. AM, NNAM, and MLRR even degrade correlation. In the case of MLRR this is caused by the random resampling of residuals, whereas concerning the conditional resampling methods this might be

41

an indication that the mesoscale fields, used as predictors, do not fully explain local precipitation. Furthermore, with the exception of MLR, DECMs show no systematic reductions of the RMS, but even sometimes enlarge it. However, an increasing RMS does not indicate a worse model skill, as at low correlation levels an underestimated variance ratio lowers the RMS (compare MLR in Figure 4.8). In summary, most DECMs drastically reduce seasonal precipitation biases, some strongly improve the temporal variability, but any improves temporal correlation on a daily basis. However, since this study focuses on climate applications, the improvement of temporal correlation is not the objective.
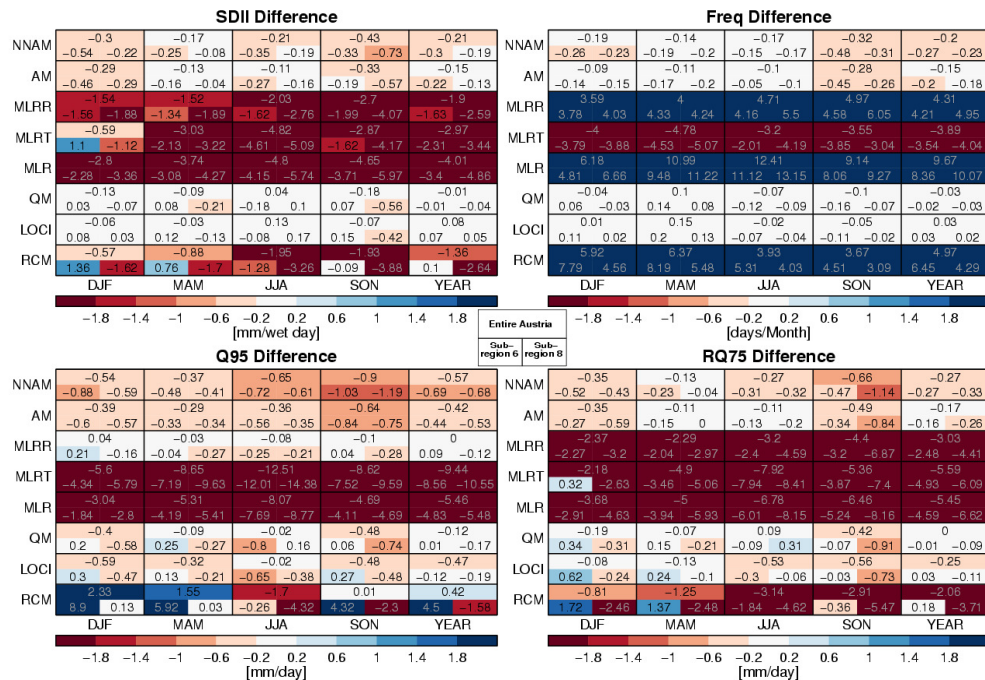


Figure 4.9 Seasonal and annual error portraits comparing the uncorrected RCM and the considered DECMs with observations. For each method and season, the results are given for entire Austria in the upper part, for sub-region 6 (left) and sub-region 8 (right) in the lower part of the respective box. SDII: precipitation intensity, Freq: wet-day frequency, Q95: 95$^{th}$ percentile on all days, RQ75: 75$^{th}$ percentile only on wet days. The skill scores are obtained equally as described in Figure 4.7.

Figure 4.9 depicts several further performance indices: The uncorrected RCM overestimates wet-day frequency (Freq), as already demonstrated. Daily precipitation intensity (SDII), in contrast, shows regional variations, but the tendency to be underestimated by the RCM. This RCM Freq and SDII behavior is characteristic of the "drizzle" problem in climate models (e.g., Gutowski et al., 2003; Fowler and Kilsby, 2007). LOCI and QM correct these errors to virtually zero. Resampling approaches, particularly NNAM, show significant skill, but slight systematic underestimation of the analyzed indicators. Although MLRR improves MLR, both regression approaches fail in reproducing intensity and frequency, with drastic intensity underestimation (up to ─4.6 mm per wet day) and overestimation of frequency (up to about 12 days per month). MLRT shows similar results for intensity, but underestimation of frequency.

Towards extreme precipitation (Q95, RQ75), the uncorrected RCM shows an inhomogeneous picture with overestimation in sub-region 6 and underestimation in sub-region 8. Only in summer all regions agree in underestimation of higher precipitation amounts. QM and LOCI but also AM as well as NNAM systematically reduce RCM error characteristics in these moderately extreme precipitation indices, which is also demonstrated by the quantile-quantile plots in Figure 4.10. MLR and MLRT underestimate Q95 and RQ75 significantly, demonstrating their deficiencies in estimating the daily precipitation's distribution. MLRR captures Q95 surprisingly well, whereas RQ75 is heavily biased. This is related to MLRR resampling, which correctly broadens the entire distribution as seen in Figure 4.10, but does not correct the general MLR problem of estimating the right wet-day probability. The latter fact is confirmed by the underestimation of RQ75. The problematic characteristics of MLRT become obvious in Figure 4.10, which shows a significant curvature in the quantile-quantile relation. Figure 4.10 also confirms the superior performance of LOCI, AM, NNAM, and particularly QM for higher quantiles. However, minor deficiencies still remain. E.g., in winter in sub-region 8, LOCI significantly overestimates heavy precipitation events greater or equal 30 mm/day. This is caused by scaling factors which adequately correct for the mean, but fail to correct these extremer precipitation intensities in the RCM where the error characteristics change from under- to overestimation (compare Figure 4.10 upper leftmost panel).
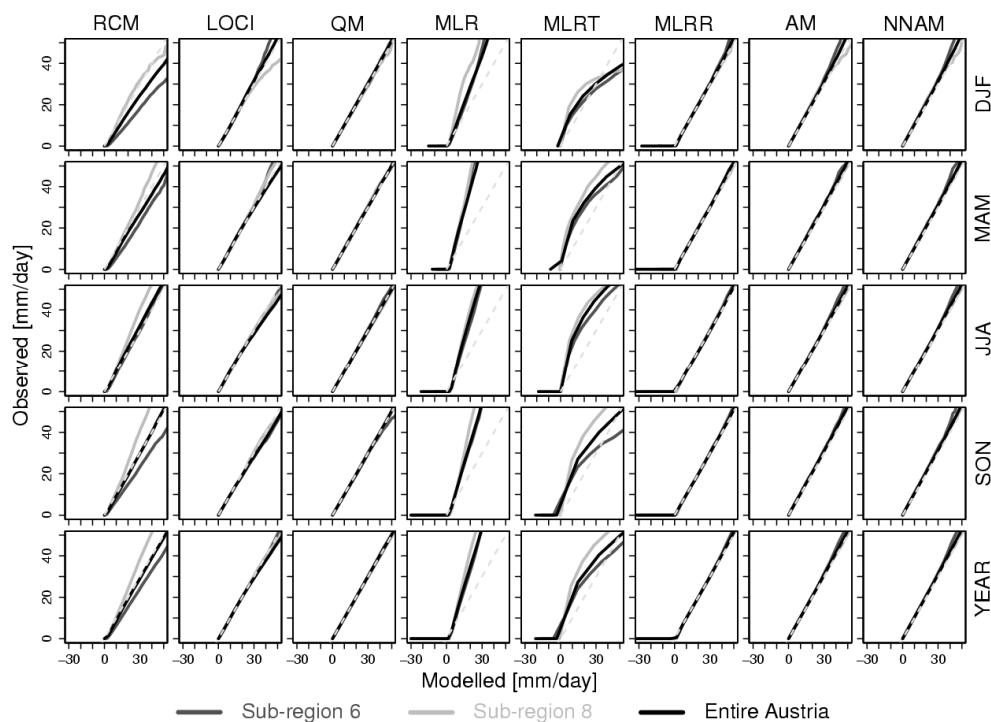
Figure 4.10 Seasonal and annual quantile-quantile (QQ) plots comparing the uncorrected RCM and the considered DECMs with observations. The QQ plots take into account the respective 11 years seasonal or annual data of each individual station within the considered region. If the continuous curves equal the dashed line, the modeled data has the same distribution as the observed data. The limit of 50 mm/day at least represents the 99th percentile in the observed data.

## 4.6 Summary and conclusions

State-of-the-art regional climate models feature significant errors and are therefore often not directly applicable to climate change impact research. This calls on the one hand for further RCM development and improvement and on the other hand for the more pragmatic approach of empirical-statistical post-processing and error correcting RCM results to create user tailored datasets for climate change impact research.

This study evaluated and compared seven different empirical-statistical downscaling and error correction methods (DECMs) for daily precipitation sums from a RCM on the station

scale. For this purpose, a cross validation framework was used, where each evaluated year was post-processed independently from the calibration period.

Most DECMs show enormous potential for reducing RCM error characteristics, which underlines the advantages of combining RCMs and DECMs in climate change impact research. None of the DECMs is able to improve the modeled temporal correlation with observations, but this is of minor importance for climatological studies. Direct DECMs (QM, LOCI) and non-linear indirect resampling methods (AM, NNAM) virtually remove RCM deficiencies in the entire precipitation distribution, i.e. they correct for mean, day-to-day variability, and extremes. The improvements are obtained regardless of season and region. This indicates the transferability of the presented DECMs from the Alpine area in this study to other regions and climates.

QM, LOCI, AM, and NNAM result in nearly similar skills with slight advantages for QM. QM is also favorable due to its simplicity, non-parametric configuration, and consequent applicability to other parameters than daily precipitation. However, instabilities at the highest quantiles of the correction function and the possible extrapolation of the correction function beyond the range of observed values should be further investigated to optimize the applicability of QM to future climate scenarios, particularly with regard to the analysis of trends in extremes. LOCI is of comparable simplicity to QM, also independent of distribution, but features some instabilities in the estimation of temporal variability. Furthermore, as the LOCI scaling factors are calibrated on climatological mean values, LOCI does not adequately correct data which features significantly curved, intensity dependent error characteristics. The dominance of direct DECMs, which only build on model precipitation as predictor, is not surprising as the predictor selection for the indirect DECMs yielded RCM precipitation to be of major relevance for local precipitation. The resampling methods' performances are comparable to QM as well, though they tend to underestimate the highest quantiles as LOCI and slightly degrade temporal correlation. The linear MLR techniques, although optimized by randomization, power law transformation, and objective predictor selection, are defective in estimating non-normally distributed daily precipitation and thus cannot be recommended for RCM precipitation error correction. However, non-linear regression techniques as the support vector regression approach (e.g., Hsieh, 2009) might be worth exploring.

The discussed improvements of DECMs are, in a strict sense, only valid for the MM5 mesoscale climate model used in this study and cannot directly be transferred to other RCMs. However, further works by Piani et al. (2009) or Dobler and Ahrens (2008) and our own experience give confidence in the robustness especially of direct DECMs applied to any RCM, if the general assumption of stationary error-characteristics is not violated.

Considering the application of the presented methods to climate scenarios, it has to be emphasized that this study does not take into account the effect of decadal climate variability on model error characteristics. Thus, the robustness of DECMs applied to long-term RCM simulations remains to be demonstrated in further investigations. Nevertheless, DECMs should be separated into two classes if applied to climate change assessment studies: Firstly, methods that base their calibration on pairs of observed and modeled values are only applicable where the calibration-simulation is correlated with weather. This applies, e.g., to a RCM simulation driven by reanalysis boundary conditions. Such application would correct for the RCM error (assuming perfect boundary conditions), but if used for a future scenario which is driven by a GCM, not correct the GCM error. Secondly, "climatological" methods that base their calibration on climatologies only, like QM and LOCI, do not require calibration-simulations that are correlated to actual weather. They can be, e.g., calibrated on RCM simulations driven by a GCM control run and would, if applied to future scenario simulations, correct for the combined GCM-RCM error, which is a clear advantage in climate applications.

# 5 Paper II: Empirical-statistical downscaling and error correction of regional climate models and its impact on the climate change signal

## 5.1 Abstract

Realizing the error characteristics of regional climate models (RCMs) and the consequent limitations in their direct utilization in climate change impact research, this study analyzes a quantile-based empirical-statistical error correction method (quantile mapping, QM) for RCMs in the context of climate change. In particular the success of QM in mitigating systematic RCM errors, its ability to generate "new extremes" (values outside the calibration range), and its impact on the climate change signal (CCS) is investigated.

In a cross-validation framework based on a RCM control simulation over Europe, QM reduces the bias of daily mean, minimum, and maximum temperature, precipitation amount, and derived indices for extremes by about one order of magnitude and strongly improves the shapes of the related frequency distributions. In addition, a simple extrapolation of the error correction function enables QM to reproduce "new extremes" without deterioration and mostly with improvement of the original RCM quality. QM only moderately modifies the CCS of the corrected parameters. However, if climate variables feature a trend and have magnitude-dependent error characteristics, QM modifies the CCS in a reasonable way. Additionally, QM has large impact on CCSs of non-linearly derived indices of extremes, such as threshold indices.

## 5.2 Introduction

Regional climate models (RCMs; Giorgi and Mearns 1991, 1999; Wang et al., 2004) are widely used tools for providing regional climate information over limited areas. With projects such as ENSEMBLES (van der Linden and Mitchell, 2009) or PRUDENCE (Christensen and Christensen, 2007) the availability and reliability of RCM simulations for Europe has increased rapidly in recent years. However, RCMs still feature considerable systematic errors (e.g., Frei et al., 2003; Suklitsch et al., 2008, 2010) which complicate the application of RCM results in climate change impact research.

One common way to deal with model errors in climate change impact studies is the "delta change approach", also called perturbation method (Déqué, 2007; Fowler and Kilsby, 2007; Graham et al., 2007). This method generates climate scenarios by adding the climate change signal (CCS) from a RCM simulation to daily or monthly observations. CCS is defined as the difference of climatological means (e.g., monthly, seasonal, or annual) between the future (e.g., 2021–2050) and present or past (e.g., 1971–2000) of a climate variable. By taking the difference, systematic model errors are removed as long as they are similar in both periods, but any potential change in temporal variability is removed as well, since variability is inherited from the observations.

Besides the delta approach, more sophisticated RCM post-processing methods have been proposed and evaluated by, e.g., Boé et al. (2007), Graham et al. (2007), Leander and Buishand (2007), Lenderik et al. (2007), Dobler and Ahrens (2008), Piani et al. (2009, 2010), or Themeßl et al. (2010). These approaches belong to the family of Model Output Statistics (MOS; Wilks, 1995; Maraun et al., 2010) and are termed "empirical-statistical downscaling and error correction methods" (DECMs). DECMs are technically identical to empirical-statistical downscaling (ESD; Benestad et al., 2008) but relate modeled instead of observed predictors to observations (predictand). As a consequence, DECMs are only valid for the model they are calibrated on and, in addition to the ESD's traditional purpose of downscaling coarser resolved model results to the local scale, also aim at the reduction of model errors.

Although the success of DECMs for correcting RCM simulations has already been demonstrated (e.g., Dobler and Ahrens, 2008; Terink et al., 2010; Themeßl et al., 2010), these studies focused on the evaluation of DECMs applied on relatively short simulations of the past. Applications to longer climate simulations also exist but primarily in hydrological literature

and rather focus on the results of hydrological models (Dettinger et al., 2004; Wood et al., 2004; Fowler and Kilsby, 2007; Leander et al., 2008; van Pelt et al., 2009) than on the performance of the error correction. Thus, the aim of this study is a) to extend the available studies by analyzing the performance of one particular DECM in the context of climate change, b) to demonstrate the flexibility of the DECM to be successfully applied to different parameters c) to investigate options to reproduce "new extremes" (values outside the calibration range) with DECMs, and d) to analyze the impact of the DECM on the CCS.

The article is structured as follows: Section 5.3 focuses on the data used in this study and on methodological issues. Section 5.4 discusses the performance of the applied DECM for different parameters and its impacts on CCSs, followed by Section 5.5 which summarizes and discusses the key findings of this study.

## 5.3 Data and methods

### *5.3.1 Data*

Daily mean, minimum, and maximum temperature as well as daily precipitation amount from the COSMO model in climate mode (CCLM, version 2.4.6; Böhm et al., 2006) are used in this study. The model data is provided by ETH Zurich within the ENSEMBLES project and covers entire Europe with a gridspacing of about 25 km. The applied simulations are driven by lateral boundary conditions from the ERA-40 reanalysis (Uppala et al., 2005) for the period 1961–2000 (hindcast simulation) and from the coupled ocean-atmosphere general circulation model (GCM) HadCM3 with normal sensitivity (Q0; Gordon et al., 2000) for the period 1951–2050 (control and scenario simulation). The control simulation (1951–2000) is based on observed greenhouse gas concentrations and the scenario simulation (2001–2050) on the SRES greenhouse gas emission scenario A1B (Nakicenovic et al., 2000). Within the ENSEMBLES project 23 RCM simulations driven by 8 different GCMs are available. Compared to other simulations in this ensemble, the CCLM simulation can be regarded as sensitive in terms of temperature change and moderate in terms of precipitation change (Heinrich and Gobiet 2010).

The E-OBS dataset (version 2; Haylock et al., 2008) provides the observational reference for this study. We use the 0.25-degree gridded version of E-OBS, which contains daily mean, minimum, and maximum temperature as well as daily precipitation amount for the period

1950–2006. The data represents spatially averaged values rather than point-scale information and is therefore suited to be compared or related to RCM results (Déqué, 2007). However, E-OBS features known deficiencies, i.e., mean values and temperature parameters are of higher quality than extremes and precipitation amount (Hofstra et al., 2009). As E-OBS lacks some data at the beginning of the 1960s, only grid cells with at least 80 percent data availability between 1961 and 2000 are used.

For the subsequent evaluation, Europe is divided into eight sub-regions, according to Christensen and Christensen (2007) (Table 5.1).

| Sub-region number | Sub-region name | Western Boundary [°E] | Eastern Boundary [°E] | Southern Boundary [°N] | Northern Boundary [°N] |
|---|---|---|---|---|---|
| 1 | British Islands (BI) | −10 | −2 | 50 | 59 |
| 2 | Iberian Peninsula (IP) | −10 | −3 | 36 | 44 |
| 3 | France (FR) | −5 | 5 | 44 | 50 |
| 4 | Mid-Europe (ME) | 2 | 16 | 48 | 55 |
| 5 | Scandinavia (SC) | 5 | 30 | 55 | 70 |
| 6 | Alps (AL) | 5 | 15 | 44 | 48 |
| 7 | Mediterranean (MD) | 3 | 25 | 36 | 44 |
| 8 | Eastern Europe (EA) | 16 | 30 | 44 | 55 |

Table 5.1 European sub-regions.

### *5.3.2 Method*

Basic Quantile Mapping method

Based on the results of a DECM inter-comparison by Themeßl et al. (2010), quantile mapping (QM) is chosen as DECM for this study. Our implementation of QM can be classified as distribution-based (calibrated on climatological distributions rather than on paired data), direct (predictor and predictand are the same parameters), and parameter-free (using empirical cumulative density distributions, *ecdfs*, rather than theoretical cumulative distribution functions). QM is applied on daily basis and for each grid point separately resulting in a corrected time series $Y^{cor}$ at time $t$ and grid cell $i$ via

$$Y_{t,i}^{\text{cor}} = ecdf_{doy,i}^{\text{obs,cal}\,-1}\left(ecdf_{doy,i}^{\text{mod,cal}}\left(X_{t,i}^{\text{raw}}\right)\right) \qquad\qquad (5.1)$$

with $ecdf^{-1}$ being the inverse *ecdf* and doy indicating the respective day of the year in the calibration period (cal). According to Equation (5.1), QM corrects raw climate model output $X^{\text{raw}}$ by mapping the modeled (mod) on the respective observed (obs) *ecdf* of the calibration period (cal) (compare Panofsky and Brier, 1968). For QM calibration, *doy* is centered within a 31 days moving window which is used to construct an *ecdf* for each day of the year.

Frequency adaptation

This study extends the basic QM procedure described in the previous subsection (QMv0, further details in Themeßl et al., 2010) by frequency adaptation (QMv1). Frequency adaptation (FA) is applied in order to account for a methodological problem which occurs if the dry-day frequency in the model results ($ecdf^{\text{mod},cal}$) is greater than in the observations ($ecdf^{\,obs,cal}$). Usually this is not the case, since many RCMs tend to underestimate the dry-day frequency ("drizzling-effect"; e.g., Gutowski et al., 2003), but occurs, e.g., along with the so called summer drying problem of RCMs in south-eastern Europe (Hagemann et al., 2004; Jakob et al., 2007). In such cases, QM without FA results in a systematic wet precipitation bias as any dry day in $X^{\text{raw}}$ is mapped to a precipitation day (compare Figure 5.1). With FA only the fraction $\Delta P_0 = \left(ecdf_{doy,i}^{\text{mod,cal}}(0) - ecdf_{doy,i}^{\text{obs,cal}}(0)\right)/ecdf_{doy,i}^{\text{mod,cal}}(0)$ of such dry-day cases with probability $P_0$ are corrected randomly by linearly interpolating between zero precipitation and the precipitation amount of $ecdf_{doy,i}^{\text{obs,cal}\,-1}\left(ecdf_{doy,i}^{\text{mod,cal}}(0)\right)$ (the first precipitation class in QM without FA). By this means, the wet bias is removed (Figures 5.1a and 5.1b). FA also confirms its applicability and stability in a stricter evaluation setup with differing calibration and evaluation periods (decadal cross validation, see Section 5.4.1) between 1961 and 2000. Applied on the control RCM simulation and the observational reference E-OBS, the summer season precipitation overestimation in Eastern Europe of QM without FA is removed systematically.
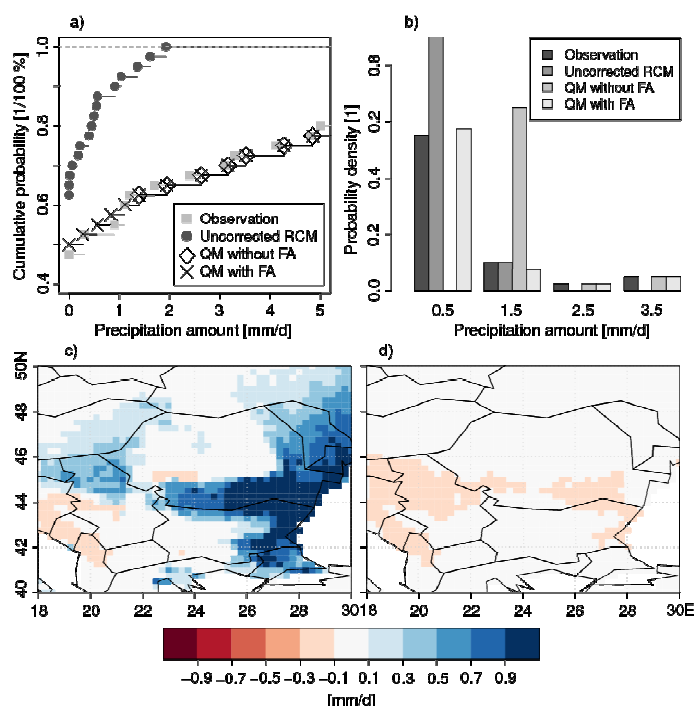
Figure 5.1 The impact of FA on QM's performance for daily precipitation. QMv0 and QMv1 are shown. Both QM versions are calibrated and evaluated for the July 1 1961−2000 at one single grid-cell in Eastern Europe with pronounced summer drying problem (panels a and b). Precipitation in panel b is classified into 1 mm/d bins, where the respective class mid is indicated on the x-axis. Panels c (QMv0) and d (QMv1) show the cross validated summer season precipitation bias maps for Eastern Europe.

New extremes

In climate change applications, the question arises how to treat values outside the range of the calibration period ("new extremes"). Unlike for a wide range of values that are not too close to the outer bounds of the calibration range (where QM clearly reduces model errors, see subsequent sections), DECMs are expected to deteriorate RCM results with regard to new extremes, since they are often per construction not able to reproduce such values. In order to mitigate this problem, two further extensions of QM are proposed here and compared to QMv0. These extensions comprise

a)  a constant extrapolation of the correction value (difference between $ecdf^{obs,cal}$ and $ecdf^{mod,cal}$ ) at the highest and lowest quantiles of the calibration range (QMv1a; compare Boé et al., 2007; Déqué, 2007) and

b)  the same extrapolation, but neglecting the three highest/lowest correction terms (QMv1b). This approach is based on the assumptions that the tails of the correction functions are likely to be very noisy.

Fitting theoretical distributions to the empirical distribution function of the data (e.g., Piani et al., 2009; Dobler and Ahrens, 2008) would resolve the "new extremes" problem as well, but is not part of this study, since this would lead to a loss of information compared to empirical distributions, as well as to a limitation of the flexibility of QM, which can, in our implementation, be applied to any parameter in any region without prior knowledge of its distribution. Also fitting arbitrary functions to the *ecdfs* resolves the "new extremes" problem (see Piani et al., 2010 for an application on global climate models), but still, some a priori information about the shape of the *ecdf* is needed which again limits the flexibility of the method to some degree and potentially introduces biases.

The evaluation of QMv1a and QMv1b is based on the ERA-40-driven CCLM hindcast in order to assure highest possible temporal correlation between the model and observations. This is necessary in order to design an evaluation setup that contains "new extremes" in the observational data of the evaluation period, which should be also represented in a well-performing model. The annual maxima of uncorrected, corrected (3 variants), as well as observed daily time series for the 5 years containing the highest annual maxima in the observation between 1961 and 2000 are evaluated at each grid cell over Europe. The evaluation years are always left out in the calibration process, which guarantees that the evaluated extremes are outside of the calibration range. Quantile-quantile (QQ)-plots in Figure 5.2 show the results for maximum one day precipitation for entire Europe. Comparable results are obtained for the three temperature parameters.
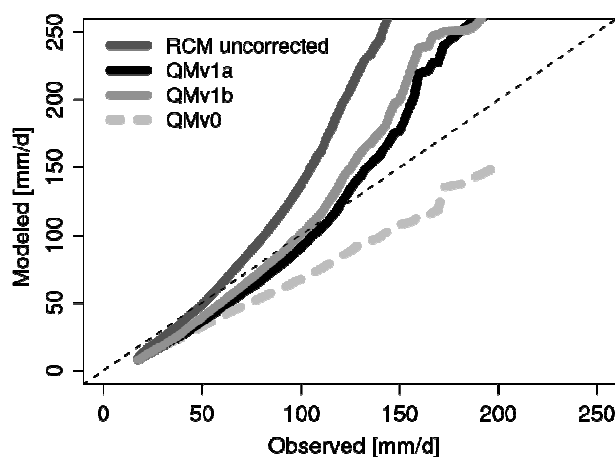
Figure 5.2 Comparison of different QM approaches applied to new extremes of daily precipitation amount. The QQ-plot compares yearly maxima of the 5 years with highest maxima for each grid cell in entire Europe between 1961 and 2000.

The uncorrected RCM overestimates the maxima in this case, while QMv0 significantly underestimates the new extremes due to its methodological constraint of mapping to the maximum of the calibration period. In contrast, QMv1a and QMv1b are both able to generate new extremes. In addition, both extrapolation methods are capable to partly outperform the uncorrected RCM even for new extremes. This is surprising, since the simple extrapolation is based only on weak empirical evidence. The remaining overestimation in the upper tails is due to the fact, that allowing new extremes in QM is accompanied with the loss of the ability to reliably remove outliers from the RCM output (as done by QMv0). Comparing QMv1a to QMv1b, QMv1a performs slightly better in this example due to the relatively smooth tails of the CCLM error correction functions (not shown), but for stability reasons and concerning other parameters and RCMs, it might be still advisable to apply QMv1b in climate change applications. In our subsequent analyses QMv1a is applied. This analysis clearly indicates that such kind of extrapolation of the correction function is favorable, since it removes the major drawback from the QM method without deterioration of the RCM quality.

## 5.4 Results

### *5.4.1 Evaluation of QM with regard to temperature and precipitation*

The evaluation focuses on QM's error correction potential, not on downscaling, as the RCM and E-OBS feature comparable spatial resolutions. In order to assess the applicability of QM to a climate simulation setup (i.e., to a GCM-driven RCM), our evaluation is based on the control simulation between 1961 and 2000 and the observational reference E-OBS. Based on the climate simulation setup, QM corrects for the combined GCM-RCM error. Such application is only possible with distribution-wise DECMs that do not rely on temporal correlation between the model and the observation. The flexibility of QM is assessed by evaluating not only daily precipitation amount, for which the method was originally designed (Themeßl et al., 2010), but also daily mean, minimum, and maximum temperature.

We apply a decadal "leave one out" cross validation approach (e.g., Themeßl et al., 2010), where each decade within 1961—2000 is corrected independently with the remaining 30 years used for calibration. The error characteristics are discussed either averaged over the four validation decades via spatial bias maps or for the entire 40 validation years via probability density functions (*pdfs*) for sub-regions.

Uncorrected daily mean temperature in Figure 5.3 features a seasonally and regionally varying bias between −3.6 K and +2.8 K (sub-regional averages). Smaller scale biases are larger and often associated with orographical features and coastlines. With the exception of the summer season (JJA), CCLM is too cold with strongest negative biases in spring (MAM) in SC as well as throughout the year along the Alpine crest. In summer the model exhibits a strong warm bias in large parts of continental Europe peaking in IP, EA, and MD. Biases of minimum and maximum temperature partly strongly deviate in their patterns as well as in magnitude from these mean temperature characteristics (not shown).

Regardless of the spatially and temporally varying error characteristics, QM corrects the temperature bias to virtually zero throughout Europe (lower panels in Figure 5.3). Remaining absolute biases are ≤ 0.2 K on the sub-regional scale and ≤ 0.3 K on the grid cell scale. Similar results are obtained for daily minimum and maximum temperature (not shown).
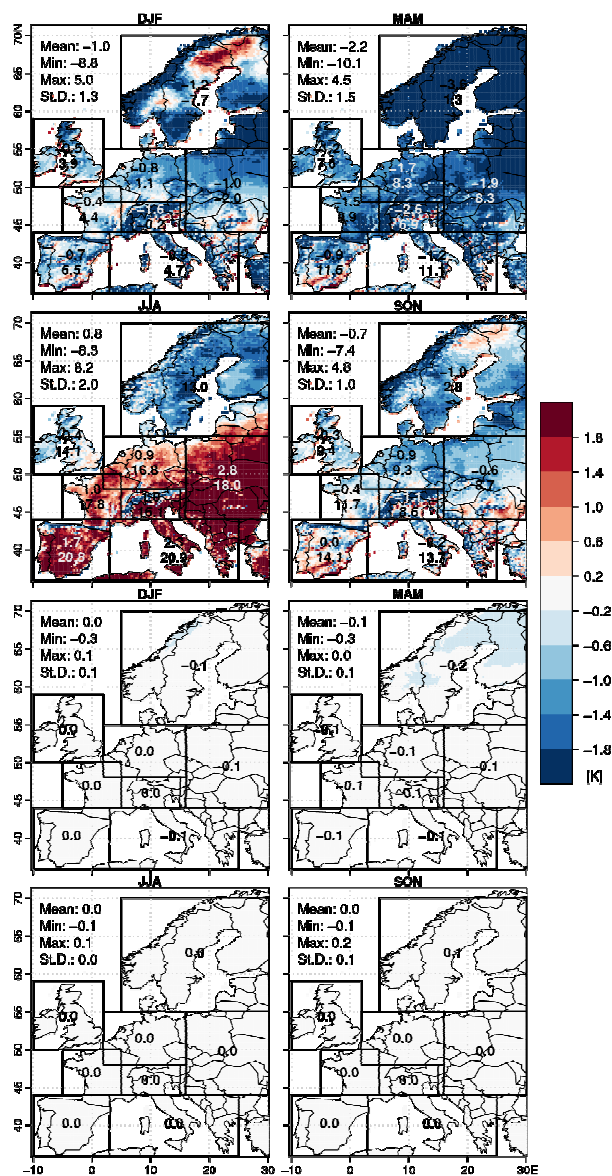
Figure 5.3 Seasonal bias of uncorrected (upper four panels) and corrected (lower four panels) mean temperature from the RCM control simulation (1961–2000) compared to E-OBS. Sub-regional mean biases and the respective mean observations (lower number in upper four panels) are given in the middle of each box. Statistics for entire Europe are given in the left upper corner of each panel.
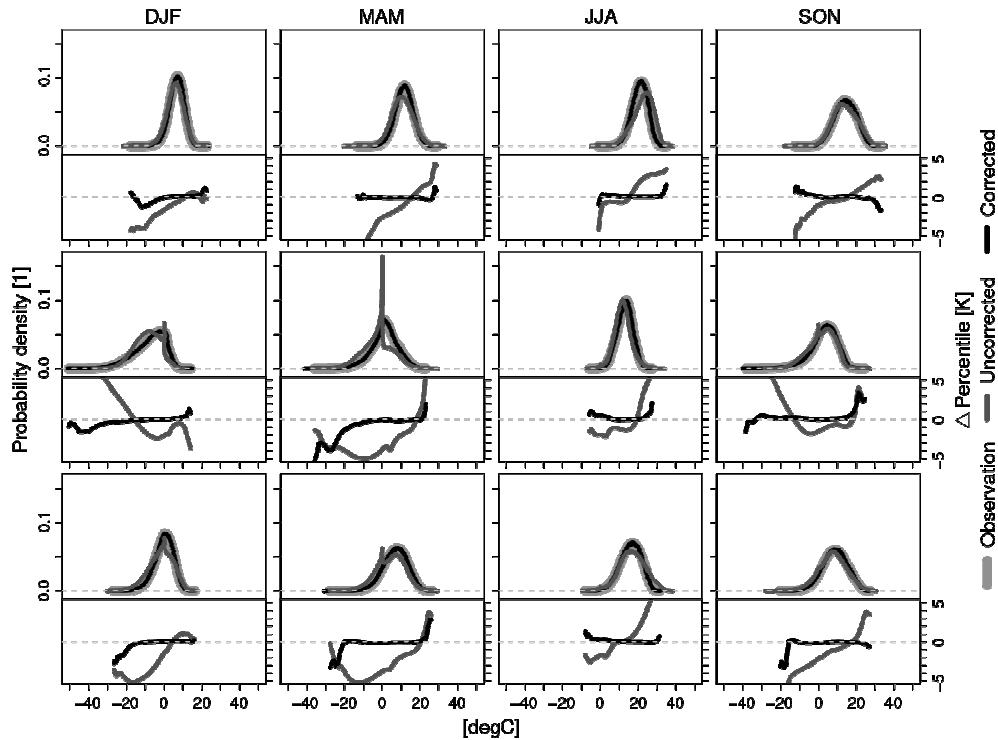
Figure 5.4 Seasonal observed, uncorrected model, and corrected model *pdfs* (upper sub-panels) of mean temperature for IP (top row), SC (middle row) and AL (bottom row). Differences between the uncorrected/corrected and observed data at different percentiles are shown in the lower sub-panels.

The modeled and observed temperature *pdfs* are shown in Figure 5.4 for the sub-regions IP, SC, and AL. These sub-regions have been selected for closer evaluation due to their differing climatic conditions and their differing bias characteristics and will be used subsequently for all further analyzes. The uncorrected RCM *pdfs* roughly capture the seasonal characteristics, but show notable deviations from the observational reference. The errors at different percentiles are mostly in the range of ±3 K and highly magnitude dependent. Remarkable is the frequency peak of the RCM at zero degree in SC. This is most probably related to problems in the representation of melting and freezing processes in the CCLM soil and snow models.

In contrast, all corrected *pdfs* nicely resemble the observations and do not feature any spurious deviations at zero degree. However, the percentile difference plots also show that the tails of the corrected time series still feature considerable errors. Nevertheless, they tend to be

57

smaller than in the uncorrected time series and only concern rare values outside the ±2δ range. Similar results are obtained for daily minimum and daily maximum temperature (not shown).
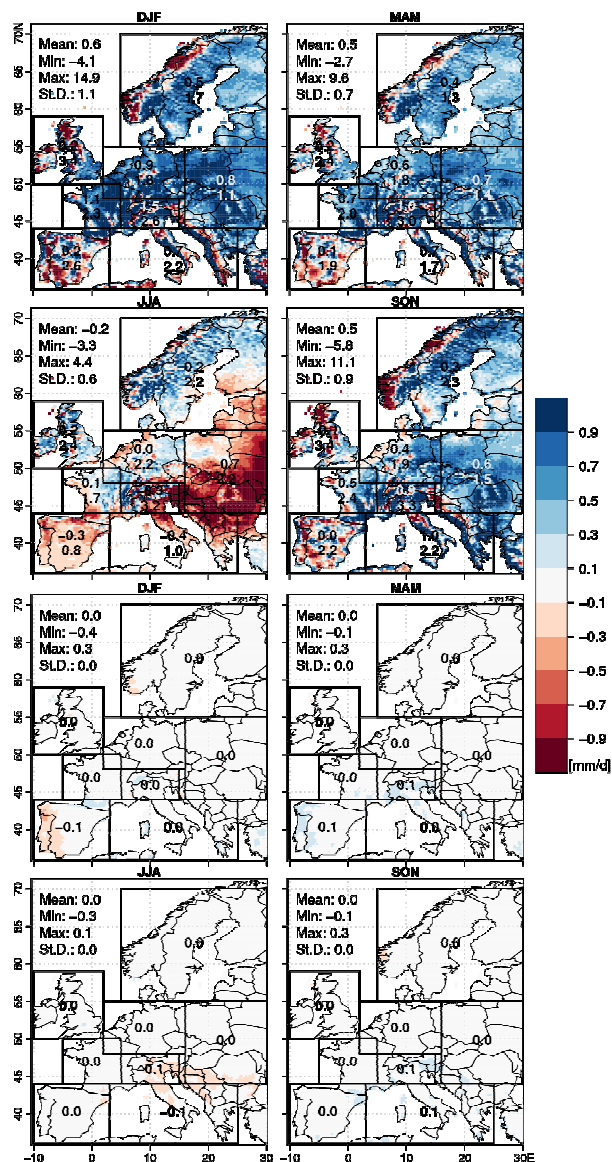


Figure 5.5 The same as in Figure 5.3 but for precipitation amount.

In the case of daily precipitation amount (Figure 5.5) the uncorrected CCLM biases range from −0.7 mm/d to +1.5 mm/d and the relative regional biases vary between about −40 % and +70 % (sub-regional averages). The bias pattern features more pronounced regional/local variation than the temperature bias, and particularly follows orographic structures in winter (DJF), spring and autumn (SON). Referring to Hofstra et al. (2009), who concluded that E-OBS is of limited quality in areas with complex terrain, these biases could be partly due to errors in the reference dataset. Simultaneously, a wet bias covers major parts of continental Europe with widespread areas of more than +0.7 mm/d. In summer, precipitation is strongly underestimated in southern Europe with dry maxima in IP, MD and EA. Together with the strong overestimation of temperature in the same sub-regions, this deviation represents the already mentioned summer drying problem, which is found in several RCMs over Europe.
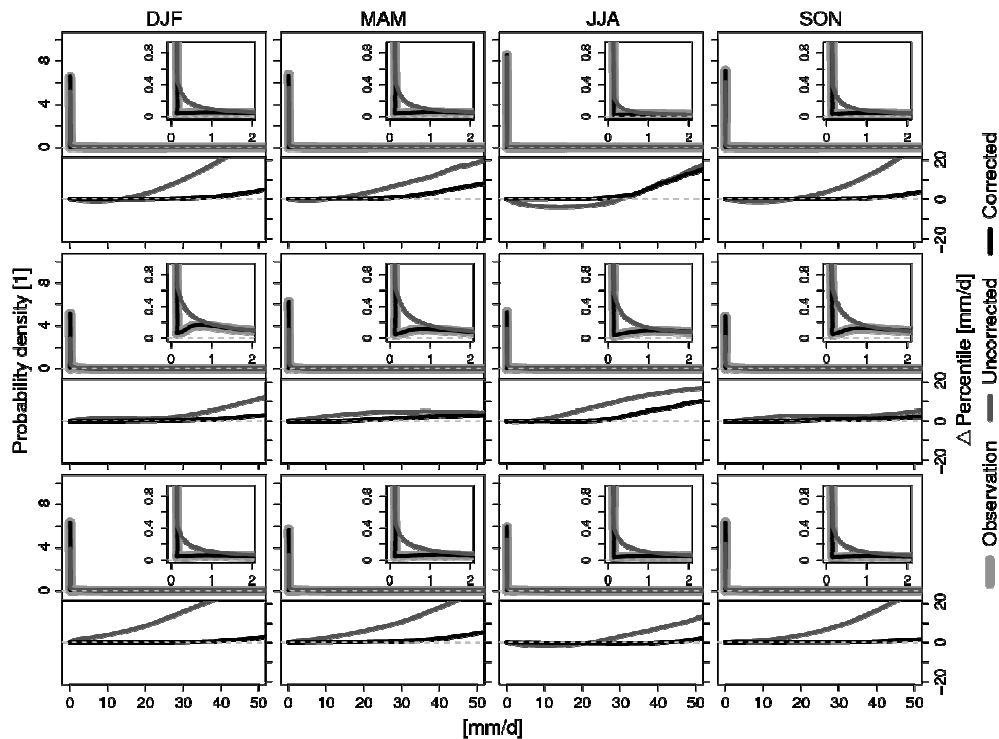


Figure 5.6 The same as in Figure 5.4 but for precipitation amount.

With QM (including frequency adaptation) the bias is almost removed across Europe, independently of region and season. Remaining regional scale absolute biases amount to ≤ 0.1 mm/d at sub-regional scale and ≤ 0.5 mm/d at grid cell scale.

The *pdfs* of daily precipitation amount (Figure 5.6) reveal that the dry-day frequency is systematically underestimated by CCLM, which is compensated by the overestimation of precipitation between 0.1 mm/d and 1 mm/d ("drizzling effect"). The frequency of heavy precipitation events is mostly overestimated by CCLM. Similar to mean temperature, the precipitation error is clearly dependent on the magnitude of precipitation amount.

Quantile mapping corrects the frequency of dry days adequately, also through the implementation of FA, and properly adjusts intensities below 30 to 50 mm/d; the frequency of higher intensities is still often overestimated, depending on season and region. However, these errors are of smaller magnitude than the uncorrected ones.

### 5.4.2 Evaluation of QM with regard to derived parameters

The performance of QM with regard to derived parameters is investigated in this subsection. For this purpose, we consider different indices of extremes including absolute maxima, such as maximum daily mean temperature (tasx) and maximum one day precipitation (px1d), as well as indices related to thresholds, such as number of summer days (days with maximum temperature > 25°C; txn25), number of tropical nights (days with minimum temperature > 20°C; tnn20), precipitation intensity (mean precipitation amount on days ≥ 1 mm/d; pint), and number of heavy precipitation days (days with precipitation amount ≥ 10 mm/d; pn10).

Without correction, all temperature extreme indices in Figure 5.7 feature a positive bias in average and a north-south gradient from negative or nearly no bias in SC and parts of BI to strong positive bias in southern Europe. The positive biases and the respective bias patterns are mainly related to the warm bias of CCLM in summer. For tasx the sub-regional biases vary between −0.6 K and +5.2 K, whereas sub-regional biases for txn25 and tnn20 amount up to about +28 days/year (days/y) and +20 days/y, respectively. In the case of tnn20, the maximum sub-regional biases exceed the respective observations by one order of magnitude.

After error-correction the spatial biases are either completely removed or reduced by roughly one order of magnitude: remaining sub-regional absolute biases are ≤ 0.6 K (≤ 1.6 K on the grid cell scale) for tasx, ≤ 2.2 days/y (≤ 4.4 days/y on grid cell scale) for txn25, and

≤ 0.4 days/y (≤ 2.7 days/y on grid cell scale) for tnn20. QM also adapts the shape and kurtosis of the *pdfs*. Minor discrepancies at the tail of the tasx distribution (< ±3 K) remain, whereas errors for corrected txn25 and tnn20 remain small throughout the entire distribution. Similar results are obtained for IP and SC (not shown).
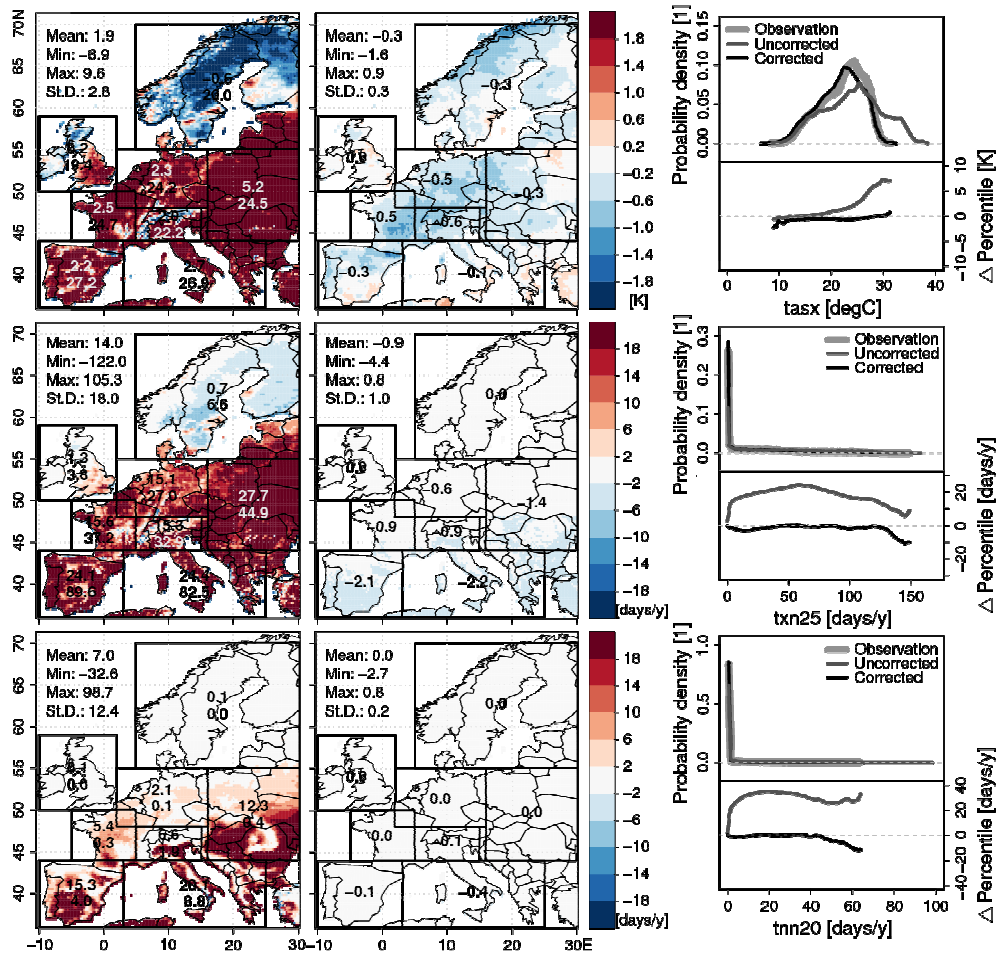


Figure 5.7 Annual bias characteristics of tasx (top row), txn25 (middle row) and tnn20 (bottom row). The left column shows the uncorrected RCM bias, the middle column the corrected RCM bias and the right column the *pdfs* (upper sub-panels) and differences between the uncorrected/corrected and observed data at different percentiles (lower sub-panels) for AL.

61
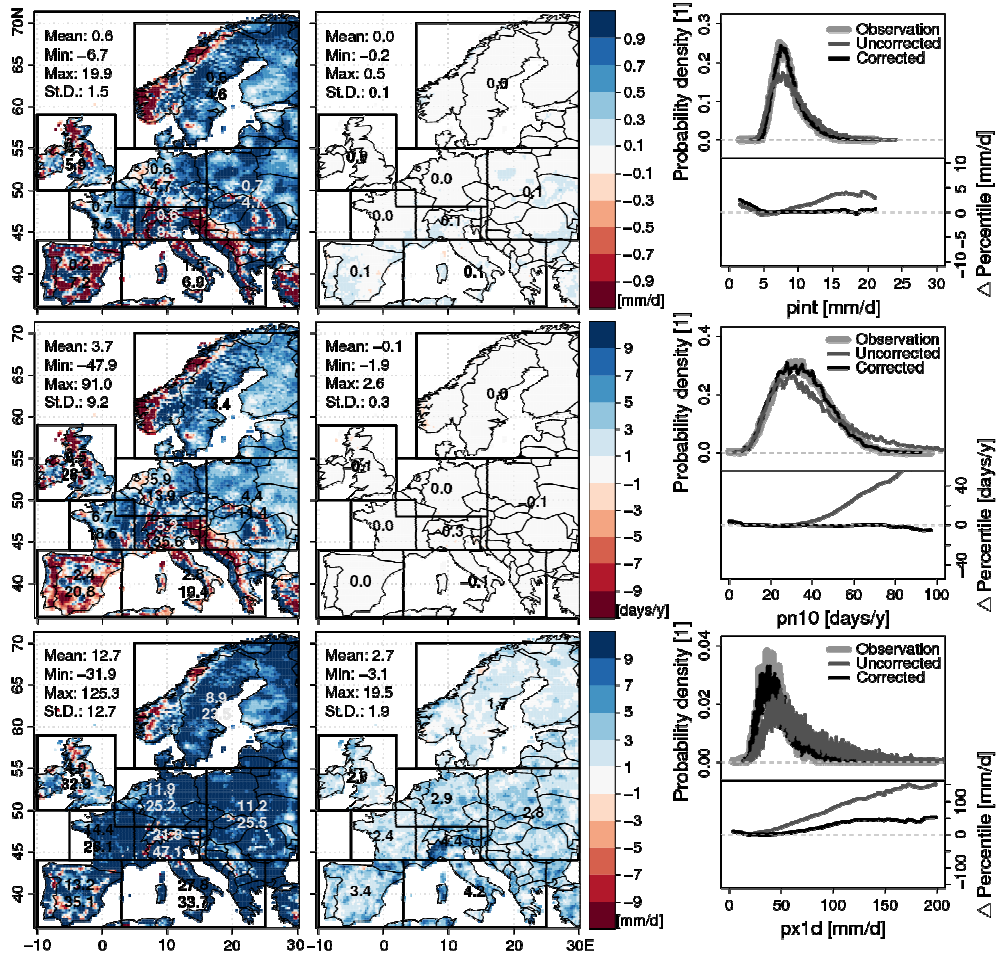
Figure 5.8 The same as in Figure 5.7 but for pint (top row), pn10 (middle row) and px1d (bottom row).

Biases of precipitation-related indices in Figure 5.8 reveal that the uncorrected model overestimates precipitation extremes in most parts of Europe, with the already described orographically induced structures. The overestimation patterns for pint and pn10 essentially cover the same regions with large areas ≥ +0.7 mm/d and ≥ +7 days/y, respectively. px1d is overestimated even more widespread by more than +9 mm/d. The relative biases for all three parameters only exceed +50 % in rare cases for px1d. QM reduces the sub-regional absolute biases to ≤ 0.1 mm/d (≤ 0.5 mm/d on grid cell scale) for pint, ≤ 0.3 days/y (≤ 2.6 days/y on

grid cell scale) for pn10, and ≤ 4.4 mm/d (≤ 19.5 mm/d on grid cell scale) for px1d. Similar to the temperature extremes the errors in the shape as well as in the kurtosis of the *pdfs* are strongly reduced. All three corrected indices of precipitation extremes feature a slight wet bias at the lowest values (similar to the uncorrected indices). Furthermore, a reduced overestimation of px1d remains above about 75 mm/d. Similar results are obtained for IP and SC (not shown).

### 5.4.3 Impact of QM on the climate change signal

In order to evaluate the impact of QM on the CCS, QM calibrated on the control simulation between 1961 and 2000 is applied to the RCM control and scenario simulation until 2050. CCSs are then calculated on monthly basis between the periods 1971─2000 and 2021─2050 for the uncorrected as well as for the corrected time series and compared with regard to the mean CCS, spatial CCS pattern, annual cycle of the CCS, and the significance of the monthly CCS. Differences between corrected and uncorrected CCSs are given in absolute and relative numbers. In cases with small absolute CCSs, numbers are given with higher accuracy, as the small rounded absolute differences then may result in misleading relative differences. Significance is determined by the Wilcoxon rank sum test (Wilks, 1995) on the 95 % significance level.

For the purpose of correcting future scenarios, it is assumed that the error correction function (the statistical model) from the calibration period remains valid for the future period (stationarity assumption; Benestad et al., 2008; Maraun et al., 2010). A consistent evidence for the validity of this assumption is beyond the scope of this study, but it can be expected that in our application the stationary assumption is better met than in most classical statistical downscaling approaches, since we apply a direct method and do not involve large scale-gaps. By this means, the effects of climate change are already represented in the predictor to a large degree and the transfer function can be expected to be largely unaltered by climate change within the calibration range. Given that, and the already demonstrated skill of QM within the calibration range and partly beyond, we are confident in the application of QM to future scenarios. However, a consistent quantification of the limitations arising from the stationary assumption is one of the most important issues for further research in our field.

Figure 5.9 displays the uncorrected and corrected CCS of mean temperature. According to the uncorrected RCM, Europe will experience a mean temperature increase of +2.0 K on average until the mid of the 21st century. Northern Europe features the highest increase (+2.4 K in SC) whereas continental Europe warms between +1.7 K and +2.0 K on the sub-regional scale. The smallest temperature increase is found in BI with +1.5 K. The sub-regional absolute impact of QM on the annual temperature CCS is ≤ 0.3 K, thus ≤ ~20 %.
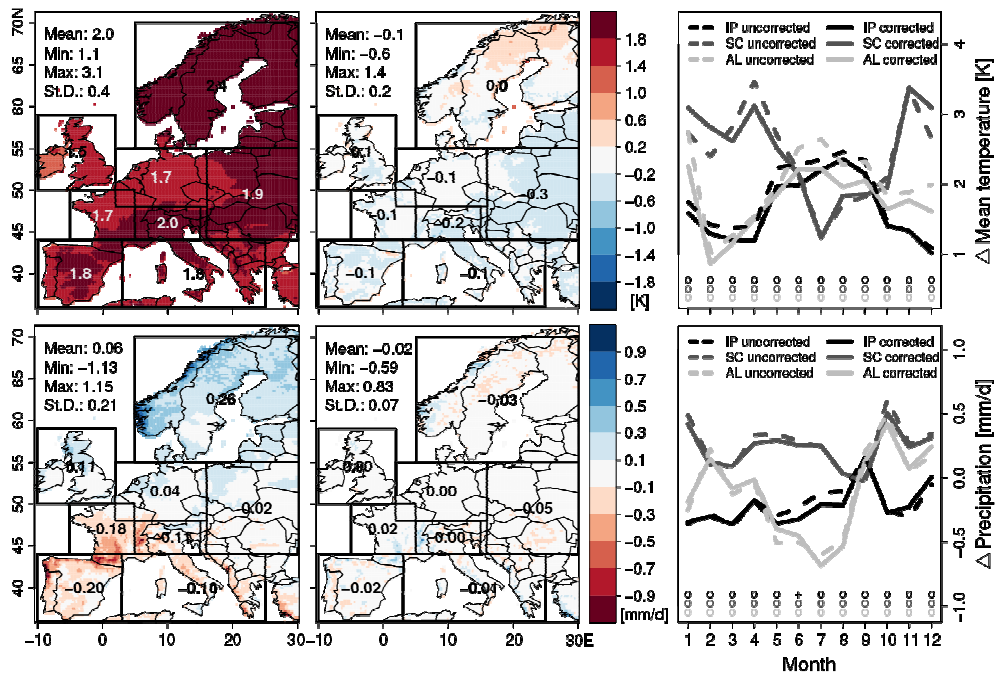


Figure 5.9 Annual mean maps of the uncorrected monthly CCS (left column), the difference between the uncorrected and the corrected CCS (middle column), and the respective annual cycles of the CCS for three sub-regions. Top row: temperature; bottom row: precipitation amount. In the lower part of the annual cycle plots change of significance is indicated with "o" (unchanged significance), "-" (loss of significance after correction), and "+" (significance established after correction).

The impact pattern features a north-south gradient from local increase in northern SC to more widespread decreases in EA, AL, MD, and IP. The corresponding annual cycles for IP, SC, and AL reveal regionally and seasonally differing characteristics, with CCSs between +0.9 K and +3.5 K. The overall characteristics of the annual cycles are not changed by QM.

Nevertheless, QM impact peaks at ─0.5 K in January in AL (~20 % of the uncorrected CCS). This can be related to the shape of the error correction function in combination with a general positive trend (Figure 5.10): E.g., the DJF temperature *pdf* is shifted towards higher temperatures in the scenario period, which are corrected by smaller positive correction terms than the lower temperatures in the control period. Thus, the CCS is reduced compared to the uncorrected scenario. Similar reasoning applies to other regions and seasons. The significance of the monthly CCS remains unchanged.

The precipitation CCS shows a north-south gradient over Europe. The uncorrected precipitation CCS ranges from an increase of +0.3 mm/d in SC to a decrease of ─0.2 mm/d in IP. This pattern is already known from various RCM simulations and, e.g., discussed in Giorgi and Coppola (2007) or van der Linden and Mitchell (2009). QM leads to scattered local impact on the CCS mostly around mountain ridges or coastlines and only small impact (about 10 %) on the sub-regional scale. The sub-regional uncorrected annual cycles of CCSs vary between ─0.7 mm/d and +0.6 mm/d depending on region and season. Similar to mean temperature, the overall characteristics of the annual cycles are not changed by QM. Only during summer months systematic differences of 0.1 mm/d (about 30 %) are visible in IP and AL. This increased negative CCS is related to more wet days between 20 mm/d and 50 mm/d in the control period compared to the scenario period and the high correction values at these intensities (see Figure 5.10 for AL). Similar to mean temperature, the significance in the obtained CCSs remains unchanged in almost all cases.

Figure 5.11 shows uncorrected and corrected CCSs of the investigated temperature-related extremes. Uncorrected tasx increases stronger in the north and south (+1.7 K to +2.0 K sub-regionally) than in the rest of Europe (+1.2 K to +1.6 K). The mean tasx CCS (+1.8 K) is smaller than the mean temperature CCS. txn25 and tnn20 show a gradient of increasing summer days and tropical nights from north to south, with no change northern of a certain latitude. The southern boundary of this zero-change area is strongly related to the threshold values used in the calculation of the indices. The sub-regional maximum changes amount to +1.8 days/month (days/m) ($\triangleq$ 21.6 days/y) and +1.7 days/m ($\triangleq$ 20.4 days/y) for txn25 and tnn20, respectively. The impact of QM on the annual CCS is not uniform between indices and regions. While the sub-regional impact of QM on tasx CCS is around 10 % or smaller, QM modifies the CCS of txn25 and tnn20 by up to 80 %.
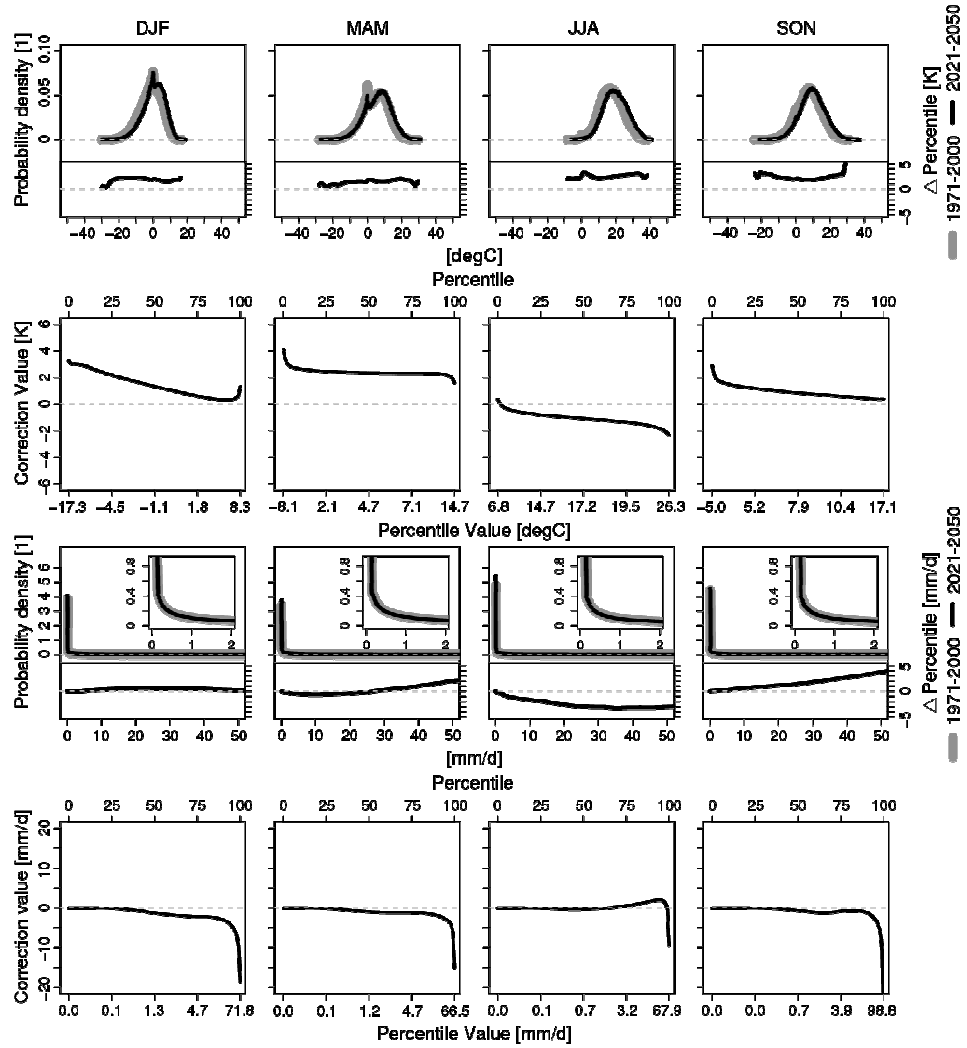
65

Figure 5.10 Seasonal *pdfs* in AL for mean temperature (first row) and precipitation amount (third row) for the period 1971─2000 (light grey) and 2021─2050 (black). The lower part of each panel displays the differences between scenario and control period at different percentiles. The second and fourth rows show the seasonal temperature and precipitation correction functions. Correction terms are derived from differences at different percentiles between observed and modeled *ecdfs*. The regional mean quantities corresponding to these percentiles are indicated on the respective x-axes.

The annual cycle of tasx CCS varies between +0.8 °C and +3.1 °C and is comparable to mean temperature, except that IP does not feature an outstanding tasx warming. Similar to mean temperature, QM does not change the general characteristic of the annual cycle, but shows notable correction in certain months, e.g., in AL peaking in June at +0.7 K (more than 20 %). These differences are related to the same reasons as discussed for mean temperature.
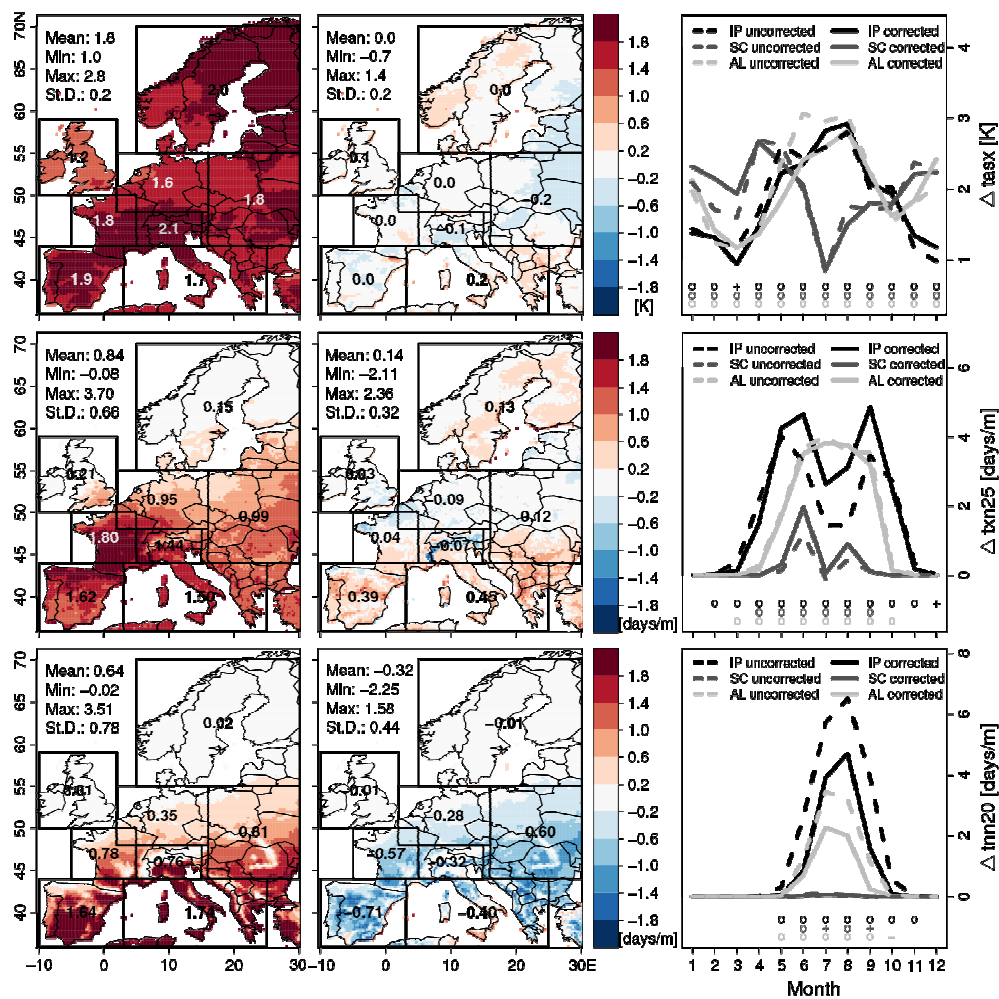


Figure 5.11 Same as Figure 5.9 but for tasx (top row), txn25 (middle row) and tnn20 (bottom row).

The annual cycle of the uncorrected txn25 CCS shows two maxima in IP and SC and one maximum in AL between March and November and ranges up to +5 days/m. Strong impact of error correction is visible for SC and IP, while the scattered local impact cancels out in AL. In IP in August, the CCS is more than doubled by QM and the maximum is shifted from spring to autumn. Here, a positive bias in maximum temperature is removed by QM (not shown) which, in combination with a positive trend in maximum temperature results in far stronger reduction of threshold exceedances in the control than in the scenario period. Regarding tnn20, its annual cycle of the CCS peaks in summer at +3.5 days/m and +6.5 days/m in AL and IP, respectively, and is trivially zero for SC (no tropical nights). With QM the tnn20 CCSs are reduced by up to ─2.3 days/m (about ─60 %). This strong effect of QM can be argued similarly as for txn25. Concerning the changes in the significance of the CCSs for temperature extremes, no systematic impacts are obtained.

Figure 5.12 shows uncorrected and corrected CCSs of precipitation-related indices. pint results in a rather homogenous increase in the uncorrected scenario simulation with a mean annual CCS of +0.3 mm/d, large areas ≥ +0.2 mm/d, and the sub-regional maximum in SC at +0.4 mm/d. This general increase of precipitation intensity is not consistent with the CCS pattern of precipitation amount. The differences can be explained by the increased frequency of dry days in the scenario period as depicted in Figure 5.10 for AL. Contrary to pint, px1d and especially pn10 exhibit north-south gradients from increase to decrease, similar to precipitation amount. The sub-regional CCS for px1d ranges from +1.6 mm/d to ─0.6 mm/d and from +0.3 days/m to ─0.2 days/m for pn10 with maxima in SC and IP.

The impact of QM on the annual CCS of all precipitation related indices is scattered only exceeds 50 % in rare cases on the sub-regional scale. Only in SC and EA a systematic reduction of precipitation extremes due to QM can be found.

The annual cycles of the CCSs of precipitation-related indices show comparable characteristics as the CCS of precipitation amount and range from ─0.6 mm/d to +1.2 mm/d for pint, from ─0.8 days/m to +0.8 days/m for pn10, and from about ─4 mm/d to + 4 mm/d for px1d. pint accords more in its peaks to px1d than to precipitation amount. Similar to precipitation amount, CCSs of all precipitation indices are reduced by QM in summer. E.g., the CCS of pint in IP in July is reduced by about 60 %. By this means also the annual cycle is altered by shifting the maximum to September. Concerning pint and px1d in IP, QM also

strongly increases the autumn CCS (up to 200 % of the uncorrected monthly CCS). As a consequence, single negative months are shifted to positive ones. Again, this can be explained by the shape of the correction function (not shown) and the trend in the parameter.
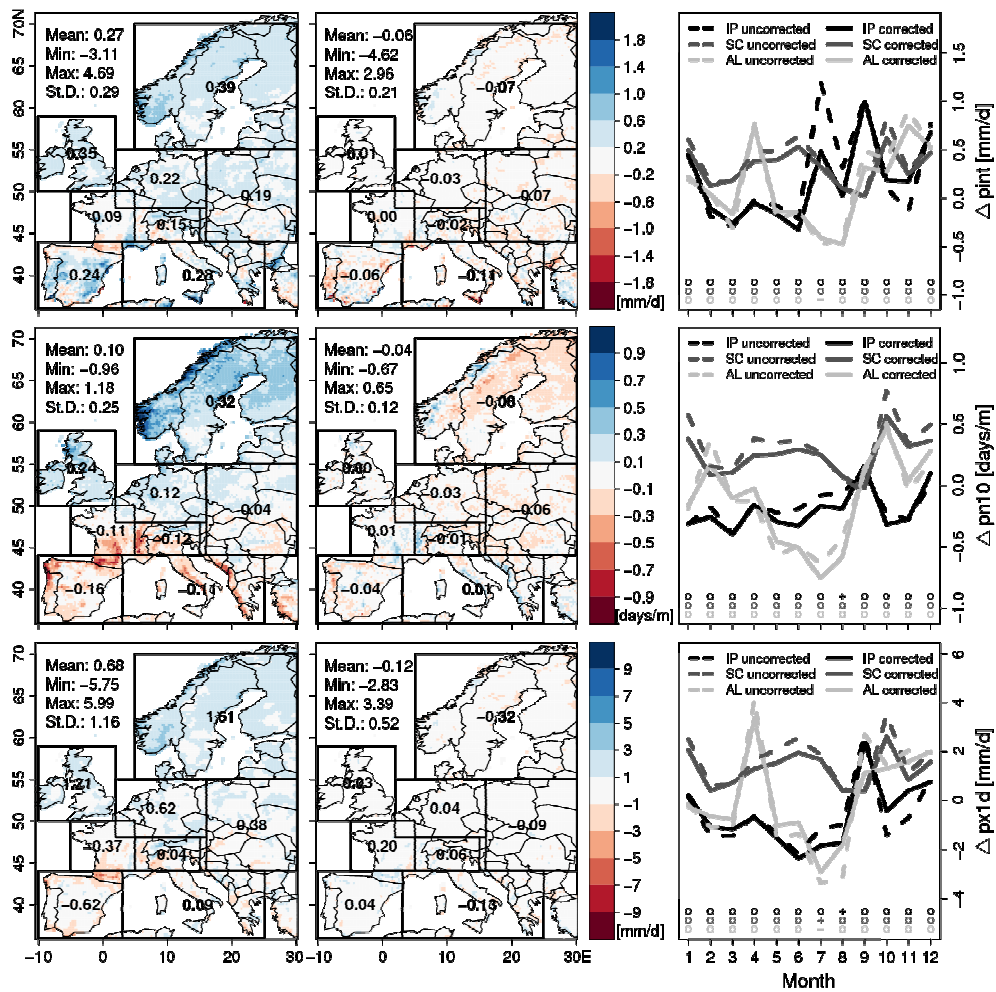


Figure 5.12 As for Figure 5.9 but for pint (top row), pn10 (middle row) and px1d (bottom row).

**5.5 Summary and conclusions**

This study evaluates the performance of an empirical-statistical downscaling and error correction method (DECM), quantile mapping (QM), applied on a RCM climate simulation over Europe regarding daily mean, minimum and maximum temperature, daily precipitation amount, and derived indices for extremes. In addition, two issues related to the climate change contexts are discussed in more detail: A methodological extension of QM which allows "new extremes" (values outside the calibration range) and the impact of QM on the climate change signal (CCS).

In a decadal cross validation of a 40 years RCM (CCLM) control simulation, QM confirms its applicability for longer climate simulations and for several parameters, regardless of spatially and temporally varying error characteristics. This regional transferability strongly suggests a general transferability of QM to any regional climate model.

Quantitatively, we demonstrate that QM reduces biases of daily mean, minimum, maximum temperature, and daily precipitation amount by roughly one order of magnitude. In most cases, the remaining absolute biases are smaller or equal than 0.1 K for temperature and smaller or equal 0.1 mm/d for precipitation amount on the sub-regional scale. For daily precipitation these results are obtained only after frequency adaptation, which assures an adequate performance of QM in situations with more modeled dry days than observed. It has to be noted, that for the application of QM to single decades biases are expected to be larger due to the yet not fully investigated impact of decadal climate variability on the stationarity of DECMs. However, since in our application no considerable scale gaps have to be bridged and since atmospheric parameters are directly mapped, it can be expected, that this restriction affects our application to a lesser extent than most other statistical downscaling approaches. The stationarity of DECMs under climate change is an important issue for further investigation.

Concerning derived indices for extremes, QM shows comparable skill as for daily temperature and precipitation. Particularly indices related to threshold values can feature tremendous biases in uncorrected RCM data, which can be easily removed be QM. Nevertheless, absolute extremes are more prone to biases after error correction than mean parameters. The analyzed indices do not include parameters which are based on temporal correlation, such as maximum number of consecutive dry days. QM corrects distribution-wise

and cannot account for errors in temporal correlation as, e.g., demonstrated by Themeßl et al. (2010).

Errors in the shapes of the daily temperature and precipitation *pdfs* are corrected adequately at least within the $\pm 2\delta$ range around the mean (temperature) or up to the $99^{th}$ quantile (precipitation). Regarding "new extremes", it could be demonstrated that by simple extrapolation of the correction terms, QM successfully produces new extremes without deterioration (and mostly with improvement) of the RCM quality.

Furthermore, this study shows that QM moderately modifies the climate change signal (CCS) of the corrected parameters. This is reasonable considering magnitude-dependent error correction functions and a trend in the underlying data. CCSs of indices that are non-linearly derived from the corrected quantities, such as threshold indices, can be strongly modified by QM. Such modifications are reasonable regarding the biases in uncorrected RCM results, improve the reliability of the CCSs and, obviously, lead to significantly different results in climate change impact investigations based on threshold parameters.

In application to climate change impact research one should keep in mind that QM, as applied here, post-processes each variable separately. As a consequence, the physical consistence between variables and/or autocorrelation structure may be distorted, which could lead to unexpected effects in the impact models (e.g., Boé et al., 2007).

Acknowledgements

# 6 Data demands of the climate impact community

Having described techniques to generate fine scale climate information in the previous chapters, these downscaled data may not be an end itself, but, e.g., function as input for further assessments of the sensitivity of various sectors on climate conditions.
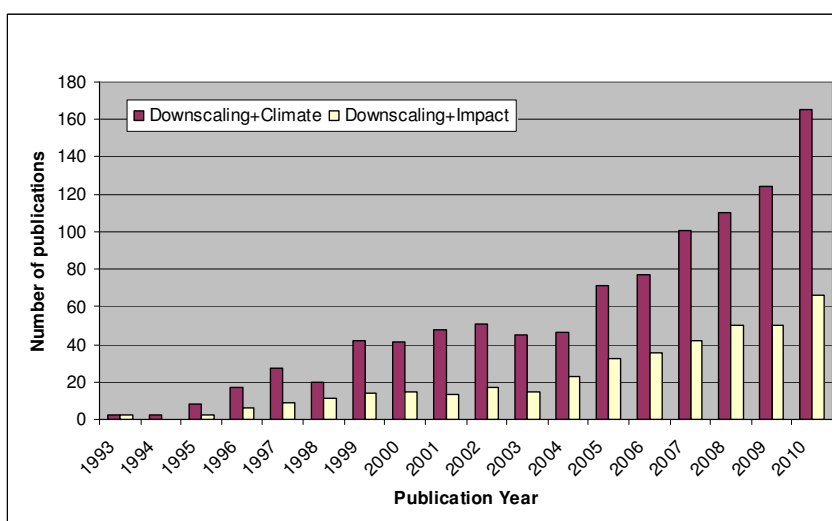


Figure 6.1 Number of scientific publications per year since 1993 in the field of "downscaling" (purple) and "downscaling and impact" (yellow). The data is taken from the ISI Web of Knowledge.

However, regarding Figure 6.1, the ISI Web of knowledge reveals that while the number of publications focusing on pure downscaling issues still strongly increases (search criterion "Downscaling and Climate") only between a third and a half of downscaled data enter further impact assessments (search criterion "Downscaling and Impact"). These numbers underline Cramer et al. (2000) who state "*climate change scenarios for impact assessments are usually an offspring rather than an intended result*". The question thus remains why only such a limited number of downscaling studies go beyond the pure downscaling. Referring to a questionnaire presented at the Infrastructure for the European Network for Earth System

Modelling (IS-ENES) workshop 2011 in Copenhagen entitled "Bridging Climate Research Data and the Needs of the Impact Community" 2011 in Copenhagen (Swart, 2011), impact scientists from different disciplines listed 5 broad categories of problems of available climate data:

1) data format
2) user-friendly data access to data
3) required spatial and temporal resolution
4) reliability and uncertainty
5) specific local needs.

Following this listing, problems are obviously either related to different data management policies and habits in different disciplines (point 1 and 2), which leads to technical difficulties in data exchange or missing communication between producers and user (point 3 to 5), which leads to the problem of unsuitable available data.

## 6.1 Useful data for climate impact studies

Facing the problems of available climate data, the following sections aim at providing a general description of useful data, concepts to produce them, as well as concrete data requirements of different user groups and sectors.

### *6.1.1 General definition of useful data*

McNie (2007) defines useful data in science by salience, credibility, and legitimacy. Salient information represents data that is context-sensitive and tailored to the users demand. I.e., an urban administration is interested in very high resolved precipitation data, temporally as well as spatially, e.g., in order to assess possible discharges for planning canal systems. For such users, monthly climate scenarios from GCMs are obviously insufficient for making appropriate decisions.

Secondly, data users should trust the data they are working with. Credibility in the data may be achieved by various pathways ranging from scientific most rated peer review processes, via strong communication between producers and users to the inclusion of users in

the data production process. Jacobs et al. (2005) stated the ease of interpretation, the clear communication of accuracy, or the possibility to assess the accuracy of the provided data by themselves, e.g., by hands on trainings, to be essential. Furthermore, clear communications about the made assumptions, methodological shortcomings, validation methods as well as statements about uncertainties are important points towards more credibility concerning the applied data (Maraun et al., 2010).

Legitimacy, finally, defines that the provided data are generated free from political suasion or bias and that the interests of the users have been encountered in the generation process (McNie, 2007). Useful data will account for these three points in a balanced way.

### 6.1.2 Generation and evaluation of useful data

Jacobs et al. (2005) describes 6 phases in the direction of useful data generation. Firstly, if, e.g., climate modelers want their data to be implemented in further applications they have to learn what information is needed (intelligence phase). Here for example, workshops are helpful to get the right insights. Then, the data producers enter the promotion phase, where they have to communicate their results, thus make them understandable for further user, but also scientifically sound for peer review processes. This phase is followed by the prescription, implementation and application phase, where data users decide which data to apply. Finally, the termination phase decides if data is no longer useful and the evaluation phase determines if the implemented data was useful in meeting the needs of the users. The lessons learned should then enter via a feedback in the intelligence phase of consequent projects.

The general usefulness of data may be judged in various ways. Possible measurements may be peer-reviewed results in science, or even already measurable reduced losses from climate related impacts such as floods due to adaptation measurements. Furthermore, such measurements could include new and sustainable contacts between research institutions, an open and participatory process, better scientific understanding or increased demand for information (Jacob et al., 2005 and references therein).

Relating to the promotion phase, the problem of scientific boundaries is often visible in interdisciplinary projects. In this context Callahan et al. (1999) and Jacobs et al. (2005) define science integrators as important intermediaries or translators at the boundaries between producers and user. They represent the interface between different science disciplines but also

between science and stakeholders, know and understand both sides and thus ensure that science and its results meet the needs of the users and are understood correctly. Jacobs et al. (2005) define such integrators to have: outside-the-box mentality; interdisciplinary background and willingness to bridge disciplinary gaps; credibility in the science community and knowledge how to translate complex information; expertise in a specific sector; understanding of the institutions and cultures of the partner in the project and the ability to facilitate and to build science-practice networks.

### 6.1.3 Definition of user groups and user sectors

Beyond the former given general definition of useful data, useful data in practice will vary depending on who is using the data and where it is applied. Thus, in the following likely user groups of climate data as well as application sectors are defined. Along with the definition of user sectors, also those there specifically requested meteorological variables are given in order to concretize useful data for climate change impact research and decision making. These definitions are primarily based on a working paper of Themeßl et al. (2011) which was prepared for an Austrian Climate Data Centre as well as on Swart (2011). In general, these sections will only focus on meteorological data, and do not include, e.g., socio-economic or demographic data needs, which are of course equally essential in any integrated assessment (compare IPCC-TGICA, 2007).

Overall, five groups of data users are defined as follows:

1) climate and climate impact researchers
2) experts in environmental and conservation organizations, or in NGOs
3) experts from private sector (consultants, spatial planners, architects,..)
4) people in federal institutions
5) politicians and policy makers.

The order of the user groups is chosen to range from 1) specialist users who need detailed data but little guidance to 5) non-specialist users who mostly need guidance (compare Swart and Pagé, 2010).

| Sector | Required parameters | Temporal Resolution | Spatial Resolution | Periods |
|---|---|---|---|---|
| **Research** | Temperature (mean, min, max), precipitation, humidity, wind (speed, direction), pressure, surface radiation budget, extreme indices | Sub-daily, Daily, Monthly, Seasonal, Annual, Decadal | Point, Grid (50 km−1 km and below) | 1900−2100; Ensemble of scenario periods |
| **Education** | Temperature (mean, min, max), precipitation, wind speed, wind direction, snow (depth cover), indices as frost days, tropical nights | Daily, Monthly, Seasonal, Annual | Point, Grid (50 km), municipality | Climatological normal period (1971−2000); Exemplary scenario period (2021−2050) |
| **Hydrology and Water management** | Temperature (mean, min, max), precipitation, humidity, river discharge, surface runoff, water vapor, global radiation, evapotranspiration, indices of extremes | Sub-daily, Daily, Monthly, Seasonal, Annual | Point, Grid (50 km−1 km and below) | 1900−2011; Ensemble of scenario periods (2021−2050) |
| **Energy** | Temperature (mean, min, max), precipitation, wind (speed direction), global radiation, indices such as heating/cooling degree days | Sub-daily, Daily, Monthly, Seasonal, Annual | Point, Grid (50 km−1 km) | 1961−2011; Ensemble of scenario periods (2021−2050) |
| **Tourism** | Temperature (mean, min, max), precipitation, sunshine duration, snow (depth, cover) | Daily, Monthly, Seasonal, Annual | Point, Grid (25 km), municipality, NUTS regions | 1961−2011; Ensemble of scenario periods (2021−2050) |

Table 6.1 Definition of user sectors in climate change impact research and those there requested data according to the temporal and spatial resolution as well as the needed time span.

According to their diverging working field, Table 6.1 list 10 sectors, which are dependent on or interested in climate data. The listed data requirements should be regarded as an overview and are based on the 50 Essential Climate Variables (ECVs) defined by the Global Climate Observing System (GCOS), as well as on experiences and questionnaires, e.g., from IPCC-TGICA (2007), Swart (2011), the Climate Impacts Group, Center for Science in the Earth System at the University of Washington (Lee and Whitely Binder, 2010), the EU FP6 Climate Change and Variability: Impact on Central and Eastern Europe (CLAVIER) project

(www.clavier-eu.org), the EU FP7 Assessing Climatic Change and Impacts on the Quality and Quantity of Water (ACQWA) project (http://www.acqwa.ch/) as well as Themeßl et al. (2011). For further details, interested readers should consider the mentioned references.

| Sector | Required parameters | Temporal Resolution | Spatial Resolution | Periods |
|---|---|---|---|---|
| **Agriculture-Forestry-Ecosystems** | Temperature (mean, min, max), precipitation, global radiation, wind (speed), snow (depth), humidity, indices as frost, heat waves, drought indices | Daily, Monthly, Seasonal, Annual | Point, Grid (50 km─1 km) | 1900─2011; Ensemble of scenario periods (2021─2050) |
| **Health** | Temperature (mean, min, max), precipitation, humidity, derived indices as human comfort index, heat waves, tropical nights summer days | Daily, Monthly, Seasonal, Annual | Point, Grid (25 km) | 1961─2011; Ensemble of scenario periods (2021─2050) |
| **Infrastructure-spatial planning-transport** | Temperature (mean, min, max), precipitation, snow, wind, river discharges, extremes as maximum one day precipitation, frost days | Sub-daily, Daily, Monthly, Seasonal, Annual | Point, Grid (50 km─1 km and below) | 1961─2011; Ensemble of scenario periods (2021─2050) |
| **Insurance and finances** | Temperature (mean, min, max), precipitation, snow, wind, hail, derived extremes as frost days, heat waves, drought indices, extreme precipitation, wind storms | Sub-daily, Daily, Monthly, Seasonal, Annual (also required frequency or return periods instead of transient time series) | Point, Grid (50 km─1 km), NUTS regions | 1900─2011; Ensemble of scenario periods (2021─2050) |
| **Civil protection** | Temperature (mean, min, max), precipitation, snow, wind, hail, derived extremes as frost days, heat waves, drought indices, extreme precipitation, wind storms | Sub-daily, Daily, Monthly, Seasonal, Annual | Point, Grid (50 km─1 km and below) | 1961─2011; Ensemble of scenario periods (2021─2050) |

Table 6.1 Definition of user sectors in climate change impact research and those there requested data according to the temporal and spatial resolution as well as the needed time spans (continued).

Table 6.1 only contains surface parameters and does not include, e.g., 3 dimensional atmospheric fields, which would be needed as boundary conditions for modeling purposes in the research sector.

In summary, the main meteorological data such as temperature, precipitation, wind, humidity, and radiation as well as derived indices such as extremes, which are in more detail listed in Table 6.2, are required by climate impact researchers and decision makers. Concerning the respective scales, variables are in fact needed on all temporal scales (ranging from sub-daily to annual) and on spatial scales between points of interest to grids (between 50 km and 1 km resolution). Experiences however show that spatially the Europe-wide state-of-the-art ~25 km grid is sufficient for most climate change impact assessments.

Concerning the requested time periods, usually historical data as well as future scenarios are required. Historical records are chosen in this context to represent some kind of base line, which is representative of the present day or recent climate average (e.g., from 1971 to 2000) and also sufficient long to encompass a range of climatic variation (IPCC-TGICA, 2007). For example hydrologists use 100 year time series to estimate a realistic variability for their hydrological models (Salathé et al., 2007). Future scenarios in most cases cover the period until 2050 or until the end of the 21st century. Transient data is demanded if effects such as trends are to be assessed or if the data is needed to initialize subsequent impact models. Often, however, impact modeling is undertaken with equilibrium modeling approaches where impact models are run once for the current climate and then for a future climate. Then, future climates are defined as mean states of time slices extracted from transient runs (Cramer et al., 2000). If climate impacts are assessed until the end of the 21st century, the range of the available scenarios becomes more and more important as from the mid of the century on the scenarios start to diverge (compare IPCC, 2007). In practice, usually not all scenarios are implemented for assessments due to computational limitation. Thus, the spread of differing climate scenarios is then taken into account by applying, e.g., 3 scenarios representing low, mid and high climate changes which are considered to lead to the best, medium, and worst impacts for the focused sector.

Beyond the data requirements of the different user sectors concerning spatial and temporal characteristics, climate impact researchers often also define useful data via attributes such as spatial coherence, temporal persistence or physical consistency (e.g., Kilsby, 2000; Maraun et al., 2010).

### *6.1.4 Categories of useful data*

Concluding the definition of useful data, this section introduces 5 categories of useful data. These categories mainly describe the data's origin and can be classified into

1) metadata
2) observational data
3) direct model output
4) downscaled and error corrected data
5) derived indices.

Metadata represent background information on and descriptions of the data set, i.e., its data sources, its producer, its provided variables, its assumptions and its consequent limitations and uncertainties or other information about its reliability or performance skill. Such meta information is essential for the correct application of the dataset and increases the credibility of the data set (Lee and Whitely Binder, 2010).

Observational data represent the basis of any climate and climate impact study. Observational data enables to assess historic conditions and trends or to calibrate and evaluate models. Besides point-scale time series from federal meteorological services or, e.g., the European Climate Assessment & Dataset (ECA&D) project (http://eca.knmi.nl/) for entire Europe, gridded observational data is, e.g., available globally on monthly basis for temperature, diurnal temperature range, precipitation amount, vapor pressure and cloud cover from 1901─2000 at a spatial resolution of 0.5 degrees (Mitchell et al., 2004). For Europe, the most recent gridded observational dataset ─ E-OBS ─ contains minimum, maximum and mean temperature, as well as precipitation amount and sea level pressure and is available on daily basis between 1961 and 2010 on a 25 km grid (Haylock et al., 2008; van den Besselaar et al., 2011). For precipitation amount a similarly resolved gridded data set on daily basis is available for the Greater Alpine region from 1971─1990 (Frei et al., 2006). For Austria a gridded dataset for mean temperature and precipitation amount at 1 km resolution is also available from 1961─2009 (Beck et al., 2009).

Gridded observational data is important for climate model validation purposes as it does not represent point scale but area average information and thus has the same statistical

properties such as climate model output (e.g., Goodess et al., 2003 and references therein; Déqué, 2007).

Direct model output, especially from RCMs has recently become available for Europe and the USA along with various regional climate modeling projects such as ENSEMBLES or NARCCAP (http://www.narccap.ucar.edu/). Due to their fine resolution between 50 km and 25 km and their strengths listed in Appendix B, such data is often used in complex impact models.

Downscaled and error corrected data represent an important data category for climate change impact studies. Such post-processed data increase the reliability of impact assessments by reducing sources of errors in the modeling chain. In its simplest form an error corrected climate scenario can be obtained via the so called delta approach where the difference between a future period and a reference period from a climate simulation is added as a climate change delta on reference observations (Déqué, 2007; Graham et al., 2007). By this means, it is expected that systematic errors are cancelling out. However, this approach also features shortcomings such as unchanged future variance characteristics. Besides, more sophisticated techniques can be found in literature and are applied and described in Chapter 4 and Chapter 5. Because all DECMs are dependent on observational data, already published studies on downscaled and error corrected data are yet limited on the one hand side to the main meteorological parameters temperature, precipitation, humidity, pressure, wind speed and global radiation (compare Chapter 4 and Chapter 5; Fuchs, 2011; Wilcke et al., 2011) and on the other hand side to industrialized countries with long-term observational networks.

Derived indices are useful information for various sectors, as they describe either sector-relevant thresholds or non-linear sector-relevant effects by combining different variables. Table 6.2 lists commonly demanded indices and provides respective definitions. Some indices are also exemplarily included in the analyses presented in Chapter 4 and Chapter 5.

| Derived indices | Short description or reference |
|---|---|
| **Temperature related indices** | |
| 90th percentile of maximum temperature | 90th percentile of daily maximum temperature |
| 10th percentile of minimum temperature | 10th percentile of daily minimum temperature |
| Heating degree day | Prettenthaler et al. (2008) |
| Cooling degree days | Prettenthaler et al. (2008) |
| Summer days | No. of days with maximum temperature > 25°C |
| Tropical nights | No. of days with minimum temperature > 20°C |
| Frost days | No. of days with minimum temperature < 0°C |
| Growing season length | Nr. of days between first occurrence of at least 6 consecutive day with > 5°C and the first occurrence after 1st July of at least 6 consecutive days with < 5°C |
| Start of growing season | Date due to definition of growing season length |
| **Precipitation related indices** | |
| Precipitation intensity | mean daily precipitation sum on wet days (days where pr_24h exceeds 1 mm/day) |
| Precipitation frequency | No. of wet days |
| 90th percentile of rain days | 90th percentile of daily precipitation sums on wet days |
| Greatest 1-day rainfall | maximum precipitation sum in one day |
| Greatest 5-day rainfall | maximum precipitation sum in 5 consecutive days |
| Intense precipitation | No. of days with precipitation > 10 mm/day |
| Heavy rainfall days | Nr. of events > long-term 90th percentile of wet days |
| **Snow related indices** | |
| Snow days | No. of days with snow height ≥ 1 cm |
| Heavy snow days | No. of days with snow height ≥ 30 cm |
| **Wind related indices** | |
| Strong wind days | No. of days with maximum wind speed > 15 m/s |
| Storm days | No. of days with maximum wind speed > 30 m/s |
| **Aridity indices** | |
| Palmer Drought Severity Index | Heinrich and Gobiet (2011) |
| Standardized Precipitation Index | Heinrich and Gobiet (2011) |
| Aridity Index | Heinrich and Gobiet (2011) |
| **Others** | |
| Potential evapotranspiration | Thornthwaite (1948) |
| Hail frequency | No. of days of hail per hail season |
| Hail intensity | Energy per hail event per area |

Table 6.2 Derived indices for the climate impact research community. The indices are based on STARDEX, CLAVIER and ACQWA project as well as via the ECA&D dictionary of indices (http://eca.knmi.nl/indicesextremes/indicesdictionary.php).

# 7 Discussion and conclusions

This PhD work investigated the applicability and skill of empirical-statistical techniques for downscaling and error correction of daily precipitation, temperature, and derived extremes from regional climate models (RCMs) in Europe. Furthermore, error-corrected, 25 km resolved climate scenarios of the respective parameters were produced for Europe and the impact of the applied empirical-statistical downscaling and error correction methods (DECMs) on the climate change signal (CCS) was investigated. In addition, a definition of useful climate data and their likely users was provided in order to increase the collaboration of climate modeling and climate change impact research or decision making.

In recent years the availability of RCM simulations for Europe tremendously increased due to projects like PRUDENCE and ENSEMBLES. This ensemble of simulations will be further extended by the CORDEX project in the near future. However, the skill of the actually provided regional-scale climate data does not yet fulfill the user demands of the climate change impact community such as generally reproducing the climate conditions of the recent decades. Besides the steady enhancement of RCMs, DECMs, based on the concept of model output statistics (MOS), offer a straightforward option to mitigate such error characteristics and to provide user-tailored climate information.

In a comprehensive inter-comparison study, seven different DECMs were tested for daily precipitation and derived indices. Daily precipitation represents one of the most challenging meteorological parameter due to its stochasticity and its bounded by zero, left-skewed distribution. The evaluation was performed at the station scale in the orographically demanding region of Austria for selected 11 years between 1981 and 1999 and was based on a 10 km resolved MM5 hindcast simulation. The investigated DECMs included Quantile Mapping (QM), Local Intensity Scaling (LOCI), Multiple Linear Regression (MLR), Multiple Linear Regression with transformed predictors (MLRT), Multiple Linear Regression including Randomization (MLRR), the Analogue Method (AM), and the Nearest Neighbor Analogue Method (NNAM). These methods were selected to comprise linear and non-linear, parametric and non-parametric, direct and in-direct, as well as point-wise and spatial techniques. Direct

techniques were defined such that predictor and predictand were the same variable, whereas indirect techniques could combine different predictor variables to estimate the predictand. Spatial techniques used characteristics of coherent atmospheric fields as predictors (e.g., taken from empirical orthogonal functions), whereas point-wise techniques built their statistical model for each grid cell separately. Based on different evaluation skill scores, derived from a yearly temporal cross validation, the following results were obtained:

- all DECMs, except MLR, virtually removed the model bias, independent of region and season considered
- MLRT shifted the error distribution to negative values
- all DECMs, except MLR and MLRT, corrected errors in day-to-day variability
- MLRR, AM, and NNAM degraded RMSE and correlation
- all DECMs, except regression based techniques, improved precipitation intensity and frequency estimation
- QM, LOCI, AM, and NNAM improved parameters related to extremes.

Overall, QM outperformed all other investigated DECMs, although LOCI and the analogue methods resulted in comparable results. However, especially the usage of analogue approaches may be advantageous if multisite, spatially coherent, daily predictand time series are requested, as well as if there is a strong physical linkage between "large scale" circulation patterns and local scale predictands, e.g., comparing extreme weather patterns in Austria and their relation to extreme precipitation (Seibert et al., 2007). In contrast, MLR approaches although optimized by randomization and predictor transformations failed in correcting daily RCM precipitation amounts, as expected.

Based on these findings, QM was selected to produce error-corrected regional climate scenarios for daily mean, minimum, maximum temperature, daily precipitation amount, and derived extreme indices for entire Europe until 2050. The corrected scenarios were based on the RCM CCLM and were corrected according to the E-OBS observational dataset. For this purpose the originally applied QM version was extended in two directions: a) by a frequency adaptation (FA) tool and b) by the constant extrapolation of the extremes' correction terms. The implementation of FA fixed a QM methodological shortcoming in cases when there were more modeled dry days than observed, which usually does not occur due to the RCMs

drizzling problem. The simple extrapolation of the error correction function was proposed as due to global warming new extremes outside the calibration range are expected to occur and the originally implemented QM version would have limited the scenarios to the historical data range. The respective results are summarized in two parts as follows:

Evaluation of QM between 1961 and 2000 in Europe based on a decadal cross-validation:
- QM confirmed its applicability for different meteorological parameters (daily mean, minimum, maximum temperature, and precipitation amount) and reduced the respective biases by roughly one order of magnitude; mostly remaining biases were near zero.
- QM reduced errors in variability and corrected the shape of the *pdfs* to a large degree
- similar results were obtained for derived extremes
- evaluation result were obtained independent of region and season considered
- QM was able to generate new extremes, outside the calibration range without deterioration and mostly with improvement of the original RCM quality

Evaluation of the impacts of QM on the CCS:
- QM moderately modified the CCS of the corrected parameters (mean, minimum, maximum temperature and precipitation amount)
- CCSs of non-linearly derived indices of the corrected quantities, such as threshold indices, partly showed strong modification due to QM

Concerning the impact of QM on the CCS, the obtained effects were related to magnitude-dependent error correction functions as well as trends in the climate scenarios of the uncorrected data.

Based on the described results, several new research questions and topics emerged, especially concerning the application of QM as DECM. In the following these questions are put together. Further development of QM should in particular focus on the correction of extreme events based on, e.g., the theoretical application of extreme value distributions (compare Maraun et al., 2010 and references therein) or the implementation of QM based on weather types (compare Bardossy and Pegram, 2011). In addition, the issue of yet un-

85

improved temporal persistence characteristics such as correlation or wet and dry period spells should be envisaged. Possible pathways in this direction are, e.g., shown by Srikanthan and Pegram (2009) who explicitly correct for autocorrelations. Furthermore, the issue of spatial coherence of daily predictand fields based on point-wise generated QM scenarios is questionable, at least for extreme conditions in smaller sub-region such as river catchments. Another important question, which partly is already investigated in literature, concerns the issue of physical consistency and the question if QM destroys the inherent RCM physical linkages between variables if it is applied separately on several parameters. Concerning the application of QM on future climate conditions, the assumption that the error correction remains stationary is yet not fully answered. Here also the implementation of weather-typing based QM could enhance the robustness of the generated results. Finally, the question of sample uncertainty in the QM procedure as well as the impact of QM on the RCM uncertainty of future climate realizations remains to be answered.

Besides the methodological part of this PhD work, the question of useful climate data frequently re-occurred during the work. Thus, based on already existing literature, experiences and questionnaires from various national and international climate change impact assessment projects, this PhD thesis furthermore dealt with a definition of useful data from the climate impact communities' point of views. Useful data, in general, was defined to be context-sensitive, credible, legitimate, and easily accessible. However, it was shown that the usefulness of climate data is highly dependent on its application purpose. Likely user groups for useful climate data range from specialized users such as researchers that need detailed data via experts in, e.g., private or federal working environments to politician and policy makers that overall need guidance and science translation along with the data. Any user groups, in addition, revealed differing data requirements within different user sectors such as research, energy, tourism, health, insurance, or civil-protection. Overall, questionnaires revealed that the standard meteorological variables such as temperature, precipitation, wind, humidity, and global radiation are mostly required, with resolutions ranging from the sub-daily to the annual scale and from spatial grid information of about 25 km resolution down to the point scale. For certain application also data on political sub-regions such as counties or NUTS regions are requested. NUTS is a nomenclature of regional units for statistics, partitioning Europe in socio-economic regions. Furthermore, the importance of derived indices such as extremes was

underlined during this PhD work as they describe either sector-relevant thresholds or non-linear sector-relevant effects. Dependent on the application, these parameters may be provided by different data sources such as observations, direct model output or post-processed, e.g., error corrected, climate model simulations. Which data in the end is relevant, in fact, can only be decided via strong communication between data producers and users. This communication should also deal with meta-information such as the accuracy and uncertainties of the data in order to increase the user's credibility in the data as well as the suitability of the data. Credibility in the data can furthermore be achieved by an early integration of the users in the data generation process, which additionally eases the accessibility and interpretation of the data, and overall fosters an integrated assessment beyond disciplinary borders.

Finally, this PhD work also revealed the success of science integrators in the generation of useful climate data. Science integrators are scientists that aim at bringing together different science disciplines or different climate data user groups under one research topic. For this purpose, they have to bridge disciplinary borders, build interfaces, translate knowledge; all in all they guide producers as well as users to achieve scientific as well as project-related added values.

# Appendix A

## A.1 Objective predictor selection

Objective predictor selection is based on linear regression models and selects those predictors that contribute significantly to the predictands variability using statistical criteria. In the following, such screening procedures are briefly discussed.

Based on Equation (4.1), a predictand at time $i$ ($y_i^{\text{MLR}}$) is estimated via a linear regression model. The difference between the estimation of $y_i^{\text{MLR}}$ and $y_i$ observed is described by the residual, or noise

$$e_i = y_i - y_i^{\text{MLR}} .$$
(A.1)

Thus, minimizing the residual sum of squares in Equation (A.2)

$$RSS = \sum_{i=1}^{n} e_i^2$$
(A.2)

obviously improves the estimation of the predictand, where $n$ denotes the sample size.

In general, four objective model selection procedures can be found in literature. The following methodological description is based on von Storch and Navarra (1999), Helsel and Hirsch (2002), and Sachs and Hedderich (2006).

Firstly, backward elimination starts by fitting the full regression model using all available predictors $x_p$. Then, iteratively those predictor variables that result in the lowest $\hat{F} -$ statistic according to Equation (A.3)

$$\hat{F} = \frac{RSS_{(p-1)} - RSS_p}{\dfrac{RSS_p}{n - (p+1)}} < F_{out} \qquad\qquad (A.3)$$

are removed from the model with $p$ predictors if the respective $\hat{F}$ − statistic is smaller than the critical value $F_{out}$ (null hypothesis $H_0 : \beta_j = 0$). As the distribution of this statistics varies with the sample size and those in the model included predictors, often a conservative $F_{out}$ threshold of 4 is implemented in order to stop the selection at an adequate step.

Secondly, forward selection starts with an initial model of one predictor variable and then iteratively adds further predictors to the model if they significantly influence the predictand (under the same null hypothesis than given for backward elimination). Included are those predictors with the greatest $\hat{F}$ − statistic according to Equation (A.4).

$$\hat{F} = \frac{RSS_{(p)} - RSS_{p+1}}{\dfrac{RSS_{p+1}}{n - (p+2)}} > F_{in} \qquad\qquad (A.4)$$

For similar reasons as given along with Equation (A.3) a threshold of 4 can be also found here for $F_{in}$.

Thirdly, stepwise regression combines both forward selection and backward elimination. After each forward selection a backward elimination is performed as selected predictors of previous iterations may become redundant later on.

Fourthly, all subset regression calibrates all $2^p$ possible predictor-combinations and selects the best performing predictors according to screening statistics. This procedure has become feasible along with the increasing computational power over the last years. Well-established screening statistics, e.g., are Mallow's Cp, the PRESS statistic, the adjusted coefficient of determination (adjusted $R^2$) or Akaike's Information Criterion. Akaike's Information Criterion (AIC) is often found in literature as it aims at a parsimonious model, thus penalizes models with too many predictors. In the special case of a least squares calibrated regression model, assuming normally distributed residuals with constant variance, the AIC can directly be calculated by Equation (A.5)

$$AIC = n \log(\hat{\sigma}^2) + 2(p+1) \tag{A.5}$$

with

$$\hat{\sigma}^2 = \frac{RSS}{n-p-1}. \tag{A.6}$$

Figure A.1 shows an example of an all subset screening. Daily winter precipitation at the observational station Fresach in Carinthia/Austria is related to various point-wise predictor variables from a RCM hindcast. The figure should be interpreted like a table, where each row represents one calibrated regression model. Colored rectangles indicate that the predictor variable is included in the respective model. The types of the colors are neglectable. The abbreviated predictors on the x-axis are defined in Table 4.1.
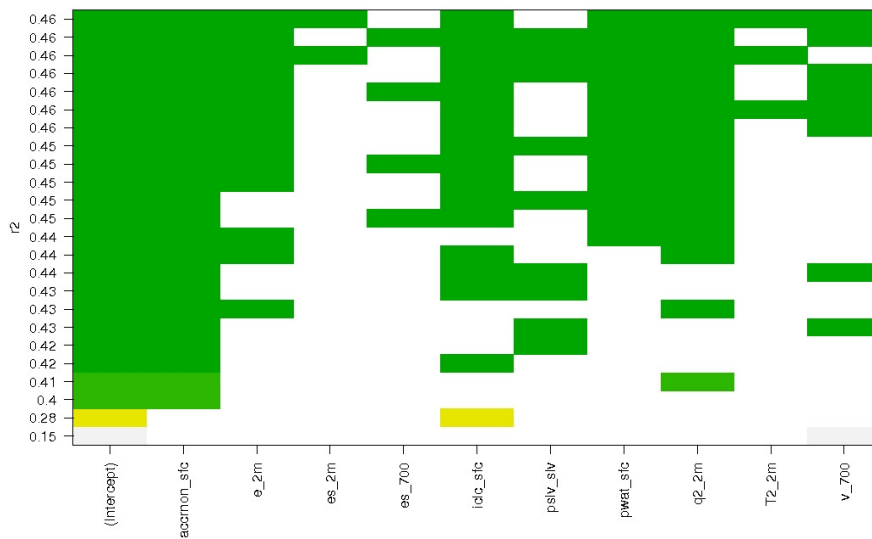


Figure A.1 Example of an all subset screening for daily winter precipitation at the observational station Fresach in Carinthia/Austria. Illustrated are the 3 best performing models with 1 to 8 predictors included. The predictors are listed on the x-axis and defined in Table 4.1. On the y-axis the adjusted coefficient of determination is given as screening statistic.

Thus, in the first model at the bottom of Figure A.1 (grey colors) the intercept as well as the northward wind at 700 hPa (v_700) are included. On the y-axes the screening statistics (adjusted $R^2$) shows that this model explains 15 % of the predictand's variability. For illustrative purposes, only the 3 best performing models including 1, 2, 3, 4, 5, 6, 7, and 8 predictors (neglecting the intercept as predictor) are displayed from the bottom row of Figure A.1 to the top row. Thus, the lowest 3 rows indicate the best three performing models with one predictor variable, the following three rows indicate the respective best performing models with two predictors, and so on.

Overall, the screening reveals that the more predictor variables are included the more variability can be explained, as expected. However, that the inclusion of more than four variables, only adds very little additional information to the model, which is able to describe ~45 % of the entire predictand's variability. In addition, Figure A.1 indicates large scale advective precipitation (accnon_sfc) as the most important variable for this specific location as the predictor reoccurs in all displayed predictor combinations.

# Appendix B

## B.1 Comparison between ESD and DD

|  | Statistical Downscaling | Dynamical Downscaling |
|---|---|---|
| Strength | +Provides station scale information (point values, e.g., for impact studies)<br>+Computationally undemanding and readily transferable<br>+Production of ensembles of climate scenarios for risk assessment/uncertainty analyses readily feasible<br>+Flexibility (e.g., can be applied to GCM/RCM output and also non-meteorological parameters)<br>+Possibility of correcting model biases<br>+Based on standard statistical procedures<br>+Can directly incorporate observations | +Grid resolution of 10 km–50 km<br>+Higher spatial resolution than GCM should reduce some biases (e.g., concerning extremes)<br>+Resolve regional scale atmospheric processes and dynamics such as orographic precipitation<br>+Respond in a physically consistent way to different external forcings (ability of incorporating feedbacks)<br>+Coherent in time, space and between variables<br>+Independent of local observation data (applicable in peripheral regions without observational network)<br>+Readily transferable to other regions |
| Weakness | ─Dependent on realism of dynamical model (affected by their biases)<br>─General assumption are difficult to validate<br>─Do not account for possible systematic changes in regional forcing conditions or feedbacks<br>─Requirement of sufficient long high resolved observation data<br>─Arbitrary choice of domain size<br>─Sensitivity of choice of predictor variables and different transfer functions<br>─By definition limited to the calibrated variability | ─Computationally demanding<br>─Relatively few independent ensembles available<br>─Affected by bias in underlying GCM<br>─Sensitivity of choice of domain size<br>─Parameterization of different processes (e.g., possible influence of different cloud/convection schemes)<br>─Sensitivity of parameterization schemes<br>─Temporal highly resolved GCMs needed as input data (not always available) |
| In betweens | ?Few method provide multi-variate outcomes<br>?Some provide multi-site information |  |

Table B.1 Comparison between dynamical and statistical downscaling (based on Goodess et al., 2001, 2007; Fowler et al., 2007 with own comments). "+" indicates a strength, "─" indicates a methodological weakness.

# Appendix C

## C.1 Downscaling of extremes: issues concerning Quantile Mapping

Accompanied by the expected global warming until the end of the 21st century, new extremes outside the calibration range are likely to occur (IPCC, 2001, 2007; van der Linden and Mitchell, 2009). Dealing with conditions outside the calibration range, any DECM in fact becomes invalid. Thus, the original QM version, applied in Chapter 4, where the extremes are fixed to the historical observational range (compare Figure C.1), had to be questioned and further extended to produce reliable scenarios also for new record breaking extremes. This appendix is intended to give additional information how these extensions were defined. Thus, in the following evaluation results of seasonal error correction functions for single grid cells and European sub-regions are illustrated and discussed.
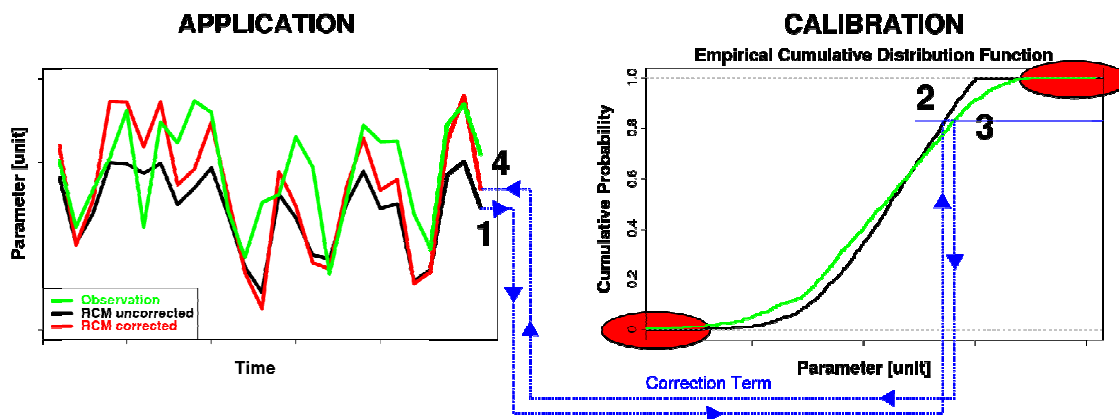


Figure C.1 The scheme of QM. An uncorrected daily value from the climate model (1) is related to its respective empirical cumulative distribution function (*ecdf*, Wilks, 2006) for the same day in the calibration period (2). The value, corresponding to the same probability the calibrated observational *ecdf* (3) is taken as the corrected value in the left panel (4). The question remains, what happens at the tails of the *ecdf*s indicated by the red circle?

Figure C.2 and C.3 illustrate seasonal correction functions for three subjectively chosen grid cells across Europe. The correction functions are derived by comparing the CCLM (Böhm et al., 2006) control run between 1961 and 2000 to the respective E-OBS observational data (Haylock et al., 2008) at all percentiles (compare Figure C.1). The respective correction functions for temperature (Figure C.2) and precipitation (Figure C.3) feature a rather chaotic behavior at the tails compared to the rest of the correction function. In many cases, an obvious break followed by a change of the correction-direction is visible there.
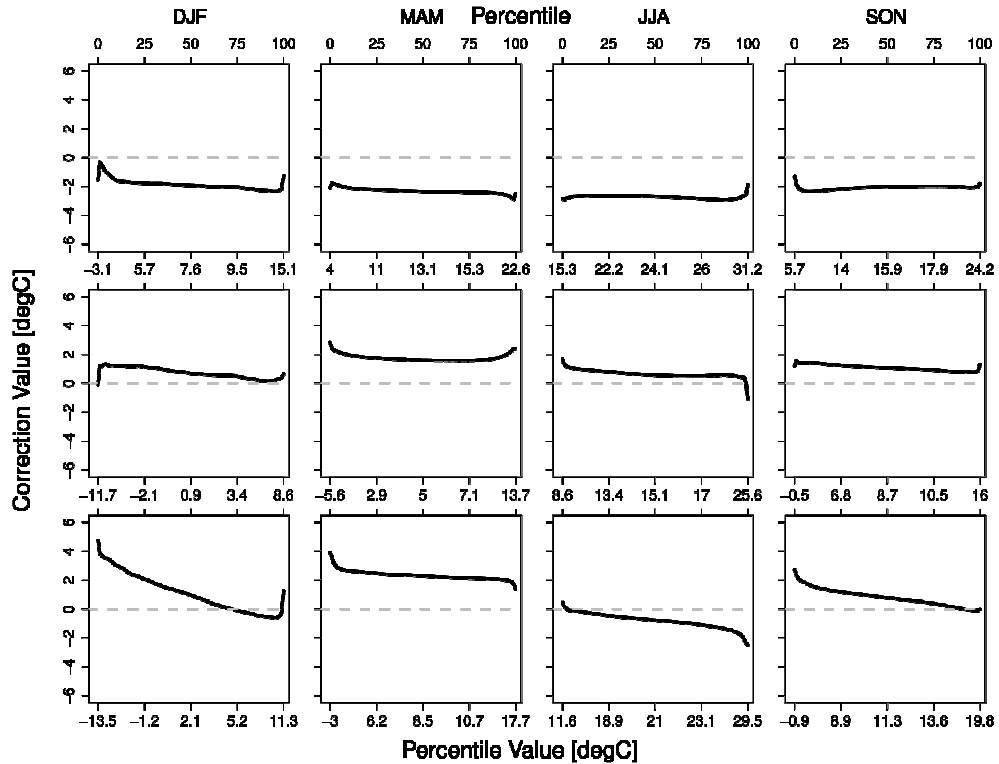


Figure C.2 Seasonal correction functions for daily mean temperature. The correction functions are derived from differences of all percentiles between observed (E-OBS) and modeled (CCLM) *ecdfs* at one selected single grid cell in the Iberian Peninsula (upper panels), in Scandinavia (middle panels) and in the Alpine sub- region between 1961 and 2000. The 0th percentile represents the difference between the *ecdfs'* minima, the 100th percentile represents the difference between the *ecdfs'* maxima. The temperature quantities corresponding to these percentiles are indicated on the x-axes.
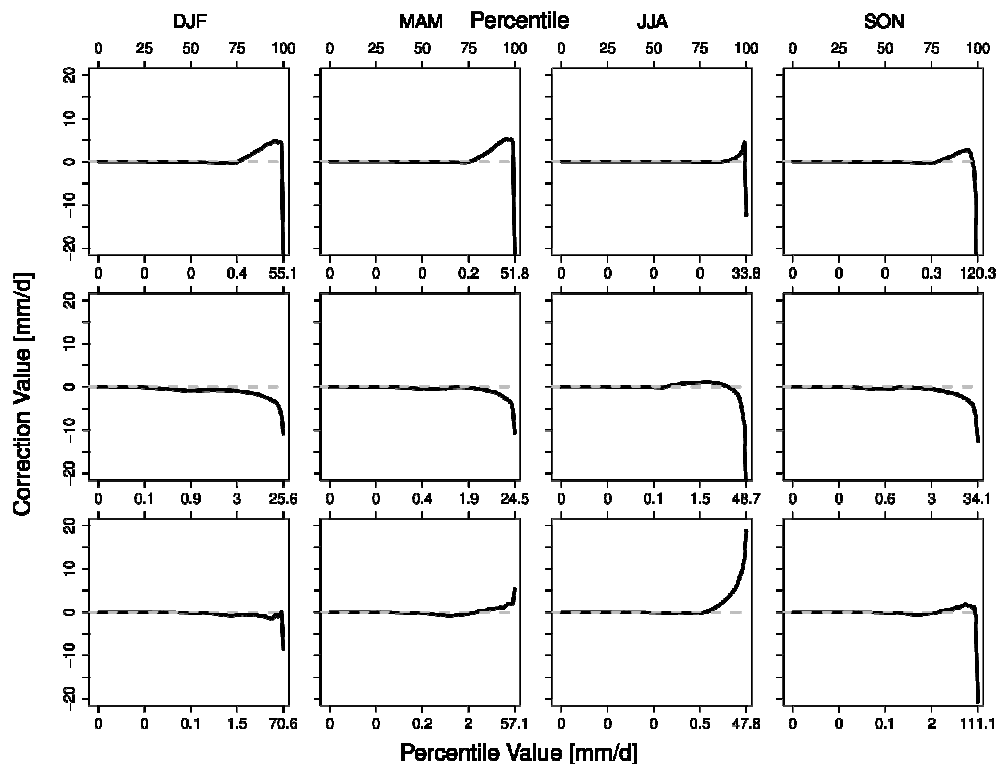
Figure C.3 Seasonal correction function for daily precipitation amount. The respective description is the same as in Figure C.2.

Focusing on the tails of the correction functions, Figure C.4 and C.5 show the lowest and highest ten correction values for mean temperature (Figure C.4) and the highest ten correction terms for precipitation amount (Figure C.5) for three sub-regions in Europe, namely for the Iberian Peninsula, Scandinavia, and the Alpine region. Their exact geographical boundaries are given in Chapter 5. The ten correction values account for about 1 % of the sample size of 1200 values (40 years with a window size of 31 days, compare QM description in Chapter 5). For both parameters, it can be seen that those in Figures C.2 and C.3 illustrated breaks at the 99th percentile for single grid cells are located at the 99.7[th] percentile for the respective sub-regions (compare vertical dashed grey lines).

Based on these error function characteristics, it is firstly concluded that fitting arbitrary functions in the shown correction functions, which is, e.g., done by Piani et al. (2010), would in general lead to an insufficient representation of the specific correction functions at the

extremes. Thus, retaining the empirical error correction functions, Déqué (2007) proposed the constant extrapolation of the correction terms of the extremes for values outside the calibration range and added that "*the application of more sophisticated techniques for correcting new extremes with QM would lack robustness and might introduce unphysical extreme values after correction*". Based on Déqué (2007) and the shown correction function characteristics in Figures C.3 and C.4, two extensions of the original QM version namely QMv1a and QMv1b were proposed. QMv1a constantly extrapolates the lowest/highest correction term on new record breaking extremes, whereas QMv1b assumes the correction functions tails to be noisy and thus disregards the lowest/highest three correction terms for extrapolation. Both QM versions are evaluated in Chapter 5.
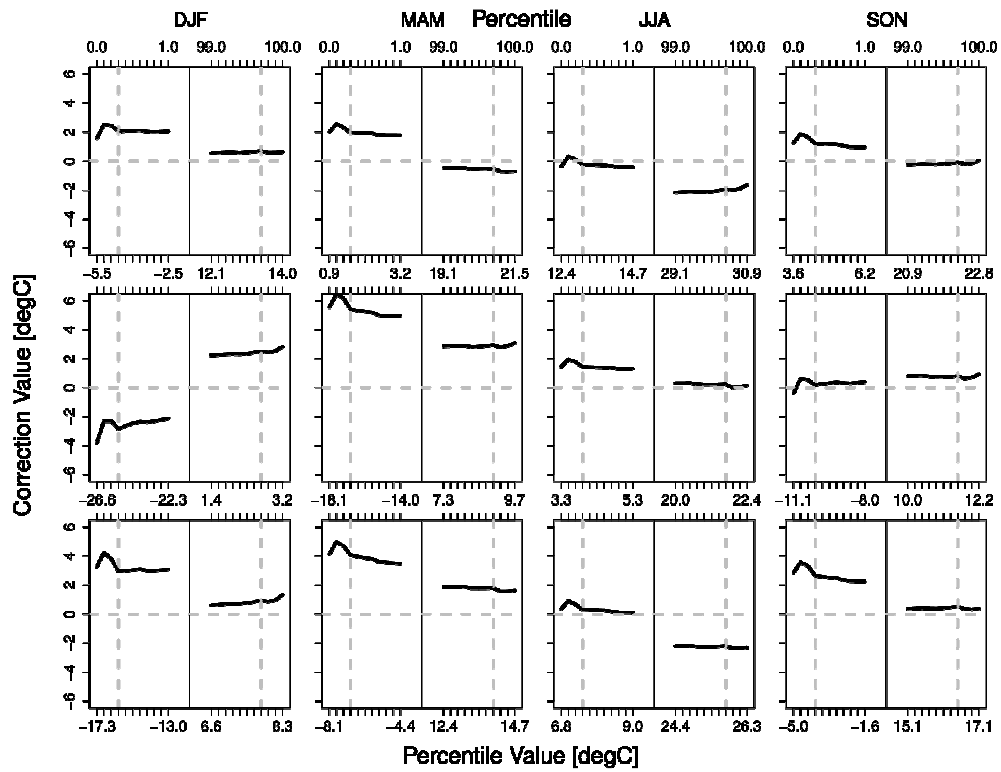


Figure C.4 Seasonal correction functions for daily temperature extremes. The left side in each panel represents the lowest 1 % of correction values in 0.1 % steps. The right side illustrates the respective highest 1 % of the correction functions. The upper row represents the Iberian Peninsula, middle row Scandinavia and the bottom row the Alps sub-region.
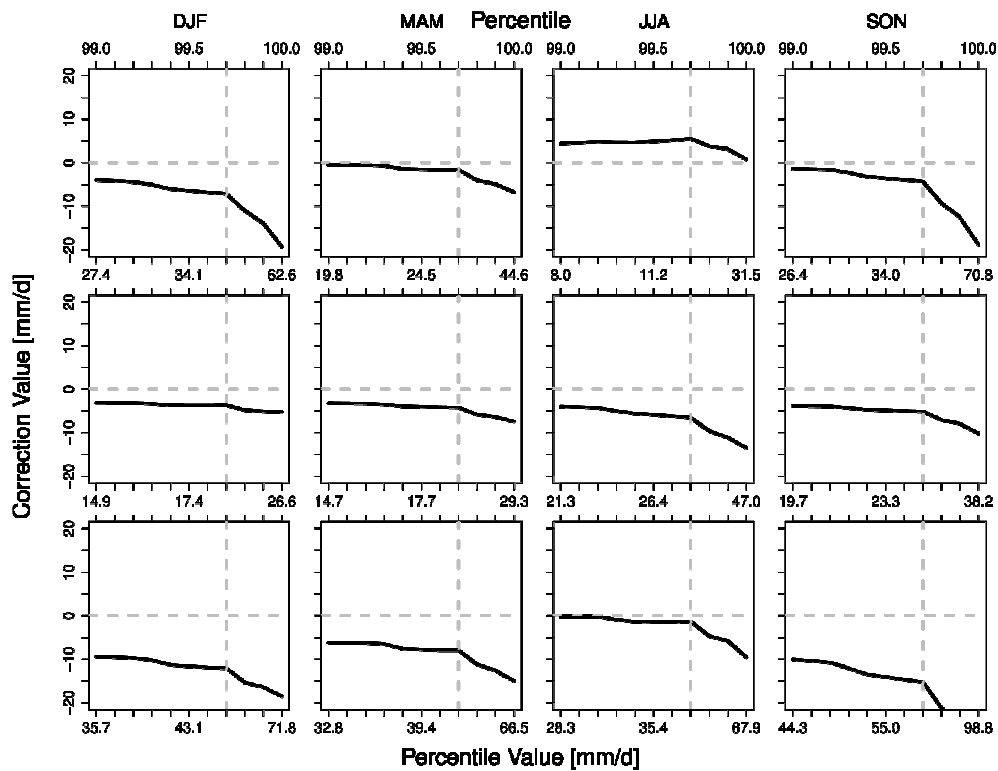
Figure C.5 Seasonal correction functions for daily precipitation extremes. Shown are the highest 1 % of correction values in 0.1 % steps. The respective description is the same as in Figure C.4.

# List of Tables

# List of Figures

# List of Acronyms

**accrnon** advective rain
**accrcon** convective rain
**ACQWA** EU FP7 Assessing Climatic Change and Impacts on the Quality and Quantity of Water
**AIC** Akaike's Information Criterion
**AL** Alps sub-region
**AM** Analogue Method
**ANN** artificial neural network
**AOGCM** Atmosphere-Ocean general circulation model
**Ar** argon

**BI** British Islands sub-region

**cal** calibration period
**CCA** canonical correlation analysis
**CCLM** COSMO model in climate mode
**CCS** climate change signal
**cdf** cumulative density function
**CLAVIER** EU FP6 Climate Change and Variability: Impact on Central and Eastern Europe
**CORDEX** Coordinated Regional climate downscaling experiment
**CO$_2$** carbon dioxide

**DECM** empirical-statistical downscaling and error correction method
**DD** dynamical downscaling
**DJF** winter season

**EA** Eastern Europe sub-region

**ECA&D** European Climate Assessment & Dataset

**ecdf** empirical cumulative density function

**ECV** Essential Climate Variable

**ENSEMBLES** ENSEMBLE-based Predictions of Climate Changes and their Impacts

**ENSO** El Niño- Southern Oscillation

**E-OBS** observation dataset by the ECA&D project

**EOF** empirical orthogonal function

**es** surface vapor pressure

**ESD** empirical-statistical downscaling

**ESDM** empirical-statistical downscaling method

**ETH** Eidgenössische Technische Hochschule

**FA** frequency adaptation

**FR** France sub-region

**Freq** precipitation frequency

**HZB** Austrian Hydrological Service

**H$_2$0** water vapor

**GCM** general circulation model

**GCOS** Global Climate Observing System

**IP** Iberian Peninsula sub-region

**IPCC** Intergovernmental Panel on Climate Change

**IS-ENES** Infrastructure for the European Network for Earth System Modelling

**JJA** summer season

**KKNN** Nearest Neighbor Analogue Method

**LOCI** Local Intensity Scaling

**MAM** spring season

**ME** Mid-Europe sub-region

**MD** Mediterranean sub-region

**MLR** Multiple Linear Regression

**MLRR** Multiple Linear Regression including Randomization

**MLRT** Multiple Linear Regression including predictor Transformation

**MM5** fifth generation of the mesoscale limited area model by NCAR and PSU

**mod** modeled

**MOS** model output statistics

**NAO** North Atlantic Oscillation

**NCAR** National Center of Atmospheric Research

**NUTS** Nomenclature des unités territoriales statistiques

**$N_2$** nitrogen

**obs** observed

**OcCC** Organe Consultatif sur les changements climatiques

**OLS** ordinary least squares

**$O_2$** dioxygen

**PC** principle component

**PCA** principal component analysis

**pdf** probability density function

**pint** precipitation intensity

**pn10** intense precipitation days

**PP** perfect prognosis

**PSU** Penn State University

**pre** precipitation amount

**PRUDENCE** Prediction of regional scenarios and uncertainties for defining European climate change risks and effects

**psfc** pressure at surface

**pslv** pressure at sea level

**pwat** precipitable water
**px1d** maximum one day precipitation

**QM** Quantile Mapping
**QMv0** Quantile mapping, original version (version 0)
**QMv1a** Quantile Mapping, version 1a
**QMv1b** Quantile Mapping, version 1b
**q2** mixing ratio
**Q95** 95th percentile of all modeled days

**RCM** regional climate model
**RMS** root mean square
**RQ75** 75th percentile on wet days

**SC** Scandinavia sub-region
**SDII** precipitation intensity
**SON** autumn season
**STARDEX** EU FP5 projects Statistical and Regional dynamical Downscaling of Extremes for European regions

**tasx** maximum mean temperature
**tnn20** tropical nights
**txn25** summer days

**u** eastward wind

**v** northward wind
**val** validation period

**WT** wet day

**ZAMG** Austrian Meteorological Service

**zg** geopotential height

# Bibliography

Alexander LV, Zhang X, Peterson TC, Caesar J, Gleason B, Klein Tank AMG, Haylock M, Collins D, Trewin B, Rahimzadeh F, Tagipour A, Rupa Kumar K, Revadekar J, Griffiths G, Vincent L, Stephenson DB, Burn J, Aguilar E, Brunet M, Taylor M, New M, Zhai P, Rusticucci M, Vazquez-Aguirre JL. 2006. Global observed changes in daily climate extremes of temperature and precipitation. Journal of Geophysical Research 111. D05109, doi: 10.1029/2005JD006290.

Alexandrov VA, Hoogenboom G. 2000. The impact of climate variability and change on crop yield in Bulgaria. Agricultural and Forest Meteorology 104: 315−327.

Auer I, Böhm R, Jurkovic A, Lipa W, Orlik A, Potzmann R, Schöner W, Ungersböck M, Matulla C, Briffa K, Jones P, Efthymiadis D, Brunetti M, Nanni T, Maugeri M, Mercalli L, Mestre O, Moisselin J-M, Begert M, Müller-Westermeier G, Kveton V, Bochnicek O, Stastny P, Lapin M, Szalai S, Szentimrey T, Cegnar T, Dolinar M, Gajic-Capka M, Zaninovic K, Majstorovicp Z, Nieplova E. 2007. HISTALP – historical instrumental climatological surface time series of the Greater Alpine Region. International Journal of Climatology 27: 17−46.

Baede APM, Ahlonsou E, Ding Y, Schimel D, Bolin B, Pollonais S. 2001. The Climate System: An Overview. In: Climate Change: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change [Hoghton JT, Ding Y, Griggs DJ, Noguer M, van der Linden PJ, Dai X, Maskell K, Johnson CA (eds.)]. Cambridge University Press. Cambridge, New York.

Bardossy A. 2000. Keynote Paper 1: Stochastic Downscaling Methods to Assess the Hydrological Impacts of Climate Change on River Basin Hydrology. In: Beersma J, Agnew M, Viner D, Hulme M. (eds.). ECLAT-2 workshop report No. 2. Downscaling-methods and suitability for producing hydrological impact scenarios.

Bardossy A, Pegram G. 2011. Downscaling precipitation using regional climate models and circulation patterns toward hydrology. Water Resources Research 47: W04505. doi: 10.1029/2010WR009689.

Barry RG, Perry AH. 1973. Synoptic climatology: methods and applications. Methuen & Co. London.

Bauer SJ, Gutdeutsch R, Hantel M, Raith W (editor), Reuter H, Rucker HO, Siedler G, Spohn T, Steinacker R, Zenk W. 2001. Bergmann Schaefer: Lehrbuch der Experimentalphysik, Band 7 Erde und Planeten. Walter de Gruyter. Berlin, New York.

Baur F, Hess P, Nagel H. 1944. Kalender der Grosswetterlagen Europas 1881-1939. Bad Homburg.

Beck A, Hiebl J, Koch E, Potzmann R, Schöner W. 2009. Institutionelle und regulatorische Fragestellungen der Bereitstellung von Wetterdaten. The Economics of Weather and Climate Risks Working Paper Series No. 10. ISSN 2074-9317.

*Bibliography*

Beersma J, Agnew M, Viner D, Hulme M. 2000. ECLAT-2 workshop report No. 2. Downscaling-methods and suitability for producing hydrological impact scenarios. UEA. Norwich.

Beersma JJ, Buishand TA. 2003. Multi-site simulation of daily precipitation and temperature conditional on the atmospheric circulation. Climate Research 25: 121—133.

Benestad RE, Achberger C, Fernandez E. 2005. Empirical-statistical downscaling of distribution functions for daily precipitation, Norwegian Meteorological Institute. met.no report No. 12.

Benestad R. 2007. Novel methods for inferring future changes in extreme rainfall over Northern Europe. Climate Research 34: 195—210.

Benestad RE, Hanssen-Bauer I, Chen D. 2008. Empirical Statistical Downscaling. World Scientific Publishing Company. New Jersey, London.

Bednorz E. 2004. Snow cover in Eastern Europe in relation to temperature, precipitation and circulation. International Journal of Climatology 24: 591—601.

Billet J, DeLisi M, Smith BG. 1997. Use of Regression Techniques to Predict Hail Size and the Probability of Large Hail. Weather and Forecasting 12: 154—164.

Boé J, Terray L, Habets F, Martin E. 2007. Statistical and dynamical downscaling of the Seine basin climate for hydro-meteorological studies. International Journal of Climatology 27: 1643—1655.

Böhm U, Kücken M, Ahrens W, Block A, Hauffe D, Keuler, K Rockel B, Will A. 2006. CLM — the climate version of LM: Brief description and long-term applications. COSMO Newsletter No. 6.

Booij MJ. 2002. Extreme daily precipitation in western Europe with climate change at appropriate spatial scales. International Journal of Climatology 22: 69—85.

Brandsma T, Buishand TA. 1997. Statistical linkage of daily precipitation in Switzerland to atmospheric circulation and temperature. Journal of Hydrology198: 98—123.

Brandsma T, Buishand TA. 1998. Simulation of extreme precipitation in the Rhine basin by nearest-neighbour resampling. Hydrology and Earth System Sciences 2: 195—209.

Brovkin V. 2006. Biosphere Feedbacks. In: Encyclopedia of Quaternary Science, Volume One [Elias SA (eds.)]. Elsevier Science.

Buonomo E, Jones RG, Huntingford C, Hannaford J. 2007. On the robustness of changes in extreme precipitation over Europe from two high resolution climate change simulations. Quarterly Journal or the Royal Meteorological Society 133: 66—81.

Camuffo D. 1978. Cumulated frequency distribution of daily global solar radiation at Venice, Italy. Archiv für Meteorologie Geophysik und Bioklimatologie 26: 45—50.

114

IPCC-TGICA. 2007. General guidelines on the use of scenario data for climate impact and adaptation assessment, Version 2. Prepared by Carter TR on behalf of the IPCC, Task Group on Data and Scenario Support for Impact and Climate Assessment.

Cebon P, Dahinden U, Davies HC, Imboden D, Jaeger CC (eds). 1998. Views from the Alps; Regional Perspectives on Climate Change. The MIT Press. Cambridge.

Christensen JH, Christensen OB. 2007. A summary of the PRUDENCE model projections of changes in European climate by the end of this century. Climatic Change 81: 7–30. doi: 10.1007/s10584-006-9210-7.

Conway D, Wilby RL, Jones PD. 1996. Precipitation and air flow indices over the British Isles. Climate Research 7: 169–183.

Cramer W. 2000. Climate Scenarios for Agricultural, Forest and Ecosystem Impacts - Workshop Summary. In: Cramer W, Doherty R, Hulme M, Viner D (eds.). ECLAT-2 workshop report. Climate scenarios for agricultural, forest and ecosystem impacts. UEA. Norwich.

Cramer W, Doherty R, Hulme M, Viner D. 2000. ECLAT-2 workshop report No. 2. Climate scenarios for agricultural, forest and ecosystem impacts. UEA. Norwich.

Cubasch U, von Storch H, Waszkewitz J, Zorita E. 1996. Estimates of climate change in Southern Europe derived from dynamical climate model output. Climate Research 7: 129–149.

Dehn M, Buma J. 1999. Modelling future landslide activity based on general circulation models. Geomorphology 30: 175–187.

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: 16–26.

Dettinger MD, Cayan DR, Meyer MK, Jeton AE. 2004. Simulated hydrologic responses to climate variations and change in the Merced, Carson, and American river basins, Sierra Nevada, California, 1900–2099. Climatic Change 62: 283–317.

Dobler A, Ahrens B. 2008. Precipitation by a regional climate model and bias correction in Europe and South Asia. Meteorologische Zeitschrift 17: 499–509.

Dudhia J, Gill D, Manning K, Wang W, Bruyere C. 2005. PSU/NCAR Mesoscale Modeling System Tutorial–Class Notes and Users' Guide (MM5 Modeling System Version 3). <http://www.mmm.ucar.edu/mm5/documents/tutorial-v3-notes.html>. Accessed 2009 Aug 12.

Fernández J, Sáenz J. 2003. Improved field reconstruction with the analog method: searching the CCA space. Climate Research 24: 199–213.

Feser F. 2006. Enhanced Detectability of Added Value in Limited-Area Model Results Separated into Different Spatial Scales. Monthly Weather Review 134: 2180–2190.

Fowler HJ, Blenkinsop S, Tebaldi C. 2007. Linking climate change modelling to impact studies: recent advances in downscaling techniques for hydrological modelling. International Journal of Climatology 27: 1547−1578. doi: 10.1002/joc.1556.

Frei C, Schär C. 1998. A precipitation climatology of the Alps from high-resolution rain-gauge observations. International Journal of Climatology 18: 873−900.

Frei C, Christensen JH, Déqué M, Jacob D, Jones RG, Vidale PL. 2003. Daily precipitation statistics in regional climate models: Evaluation and intercomparison for the European Alps. Journal of Geophysical Research 108 (D3), 4124. doi: 10.1029/2002JD002287.

Frei C, Schoell R, Fukutome S, Schmidli J, Vidale PL. 2006. Future change of precipitation extremes in Europe: an intercomparison of scenarios from regional climate models. Journal of Geophysical Research Atmospheres 111, D06105.doi: 10.1029/2005JD005965.

Fuchs R. 2011. Empirical-Statistical Post-Processing of Radiation and Humidity Parameters in Regional Climate Simulations. Wegener Center Verlag. Wegener Center for Climate and Global Change. University of Graz. Graz. ISBN 978-3-9502940-9-5.

Giorgi F, Mearns LO. 1991. Approaches to the simulation of regional climate change: a review. Reviews of Geophysics 29: 191−216.

Giorgi F, Mearns LO. 1999. Introduction to special section: Regional climate modelling revisited. Journal of Geophysical Research 104: 6335−6352.

Giorgi F, Hewitson B, Christensen J, Hulme M, Von Storch H, Whetton P, Jones R, Mearns R, Fu C. 2001. Regional Climate Information − Evaluation and Prjection. In: Climate Change: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change [Hoghton JT, Ding Y, Griggs DJ, Noguer M, van der Linden PJ, Dai X, Maskell K, Johnson CA (eds.)]. Cambridge University Press. Cambridge. UK and New York.

Giorgi F, Coppola E. 2007. European climate-change oscillation (ECO). Geophysical Research Letters 34, L21703. doi: 10.1029/2007GL031223.

Gobiet, A, Truhetz H, Riegler A. 2006. A climate scenario for the Alpine region, reclip:more project year 3. WegCenter progress report, Wegener Center for Climate and Global Change. University of Graz. Graz.

Goodess CM., Palutikof JP. 1998. Development of daily rainfall scenarios for southeast Spain using circulation type approach to downscaling. International Journal of. Climatology 10: 1051−1083.

Goodess CM, Hulme M, Osborne T. 2001. The identification and evaluation of suitable scenario development methods for the estimation of future probabilities of extreme weather events. Tyndall Centre for Climate Change Research. Working paper No.6.

Goodess CM, Hanson C, Hulme M, Osborn TJ. 2003. Representing Climate and Extreme Weather Events in Integrated Assessment Models: A Review of Existing Methods and Options for Development. Integrated Assessment 4: 145─177.

Goodess CM, Anagnostopoulo C, Bardossy A, Frei C, Harpham C, Haylock MR, Hundecha Y, Maheras P, Ribalaygua J, Schmidli J, Schmith T, Tolika K, Tomozeiu R, Wilby RL. 2007. An intercomparison of statistical downscaling methods for Europe and European regions—assessing their performance with respect to extreme temperature and precipitation events. Climatic Change. In press.

Gordon C, Cooper C, Senior CA, Banks H, Gregory JM, Johns TC, Mitchell JFB, Wood RA. 2000. The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments. Climate Dynamics 16: 147─168.

Graham LP, Andréasson J, Carlsson B. 2007. Assessing climate change impacts on hydrology from an ensemble of regional climate models, model scales and linking methods–a case study on the Lule River basin. Climatic Change 81: 293─307.

Gutowski WJ, Decker SG, Donavon RA, Pan Z, Arritt RW, Takle ES. 2003. Temporal-spatial scales of observed and simulated precipitation in central U.S. climate. Journal of Climate 16: 3841─3847.

Häckel H. 2008. Meteorologie, 6. Auflage. Verlag Eugen Ulmer Stuttgart. Stuttgart.

Hagemann S, Machenhauer B, Jones R, Christensen OB, Déqué M, Jacob D, Vidale PL. 2004. Evaluation of Water and Energy Budgets in Regional Climate Models Applied Over Europe. Climate Dynamics 23: 547─567.

Hay LE, Clark MP. 2003. Use of statistically and dynamically downscaled atmospheric model output for hydrologic simulations in three mountainous basins in the western United States. Journal of Hydrology 282: 56─75.

Haylock MR, Goodess CM. 2005. Interannual variability of European Extreme Winter Rainfall and Links with Mean Large-Scale Circulation. International Journal of Climatology 24: 759–776.

Haylock MR Cawley GC, Harpham C, Wilby RL, Goodess CM. 2006. Downscaling heavy precipitation over the United Kingdom: A comparison of dynamical and statistical methods and their future scenarios. International Journal of Climatology 26: 1397─1415.

Haylock, MR., Hofstra N, Klein Tank AMG, Klok EJ, Jones PD, New M. 2008. A European daily high-resolution gridded dataset of surface temperature and precipitation. Journal of Geophysical Research 113: 1–12, D20119. doi: 10.1029/2008JD10201.

Helsel DR, Hirsch RM. 2002. Statistical Methods in Water Resources Techniques of Water Resources Investigations. U.S. Geological Survey. <http://pubs.usgs.gov/twri/twri4a3/html/pdf_new.html>. Accessed 2009 Aug 12.

Heinrich G, Gobiet A. 2010. The future of dry and wet spells in Europe: A comprehensive study based on the ENSEMBLES regional climate models. International Journal of Climatology. Submitted.

*Bibliography*

---

Hess P, Breuowsky H. 1977: Katalog der Grosswetterlagen Europas (1881-1976). Berichte des Deutschen Wetterdienstes Bd 15. Selbstverlag des Deutschen Wetterdienstes. Offenbach am Main.

Hewitson BC, Crane RG. 1996. Climate downscaling: techniques and application. Climate Research 7: 85–95.

Hofstra N, Haylock M, New M, Jones PD.2009. Testing E-OBS European high-resolution gridded data set of daily precipitation and surface temperature. Journal of Geophysical Research 114, D21101. doi: 10.1029/2009JD011799.

Hsieh WW. 2009. Machine Learning Methods in the Environmental Sciences. Cambridge University Press. Cambridge.

Huth R. 1999. Statistical downscaling in central Europe: evaluation of methods and potential predictors. Climate Research 13: 91─101.

Ines AVM, Hansen JW. 2006: Bias correction of daily GCM rainfall for crop simulation studies. Agricultural and Forest Meteorology 138: 44─53.

Imbert A, Benestad RE. 2005. An improvement of analog model strategy for more reliable local climate change scenarios. Theoretical and Applied Climatology 82: 245─255. doi: 10.1007/s00704-005-0133-4.

IPCC. 2001. Climate Change 2001. The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change [Hoghton JT, Ding Y, Griggs DJ, Noguer M, van der Linden PJ, Dai X, Maskell K, Johnson CA (eds.)]. Cambridge University Press. Cambridge, New York.

IPCC. 2007. Climate Change. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, 2007 [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller (eds.)] Cambridge University Press. Cambridge, New York.

Jacob D, Bärring L, Christensen OB, Christensen JH, de Castro M, Déqué M, Giorgi F, Hagemann S, Hirschi M, Jones R, Kjellström E, Lenderink G, Rockel B, Sánchez E, Schär C, Seneviratne SI, Somot S, van Ulden A, van den Hurk B. 2007. An inter-comparison of regional climate models for Europe: design of the experiments and model performance. Climatic Change 81: 31─52. doi: 10.1007/s10584-006-9213-4.

Jacobs K, Garfin G, Lenart M. 2005. More than Just Talk…Connecting Science and Decisionmaking. Environment 47: 6─21.

Johnston GI, Hudson R. 1989. Technologies for Monitoring Global Change. paper no. AIAA 89-0254. AIAA 27th Aerospace Sciences Meeting, January 9-12. 1989.

Jones PD, Hulme M, Briffa KR. 1993. A comparison of Lamb circulation types with an objective classification scheme, Royal Meteorological Society, International Journal of Climatology 13: 655─663.

Kaas E, Li TS, Schmith T. 1996. Statistical hindcast of wind climatology in the North Atlantic and northwestern European region. Climate Research 7: 97─110.

Kapper KL, Gobiet A, Kapper L. 2010. Determination of the Effective Resolution of Regional Climate Models. Geophysical Research Abstracts Vol. 12. EGU2010-10479. EGU General Assembly 2010.

Katz RW, Parlange MB, Naveau P. 2002. Statistics of extremes in hydrology. Advances in Water Resources 25: 1287—1304.

Kidson JW, Thompson CS. 1998. A Comparison of Statistical and Model-Based Downscaling Techniques for Estimating Local Climate Variations. Journal of Climate 11: 735—753.

Kilsby CG, Cowpertwait PSP, O'Conell PE, Jones PD. 1998. Predicting rainfall statistics in England and Wales using atmospheric circulation variables. International Journal of Climatology 18: 523—539.

Kilsby C. 2000. Downscaling - methods and suitability for producing hydrological impact scenarios. In: Beersma J, Agnew M, Viner D, Hulme M. (eds.). ECLAT-2 workshop report No. 2. Downscaling-methods and suitability for producing hydrological impact scenarios.

Klein Tank AMG, Können GP. 2003. Trends in Indices of daily temperature and precipitation extremes in Europe, 1946-1999. Journal of Climate 16: 3665—3680.

Köppen W, Geiger R. 1961. Die Klimate der Erde. Wandkarte. Pertes. Darmstadt.

Lall U, Sharma A, 1996. A Nearest Neighbor Bootstrap For Resampling Hydrologic Time Series. Water Resources Research 32: 679—693.

Lamb HH. 1972. British Isles weather types and a register of daily sequences of circulation patterns, 1861-1971. Geophysical Memoir 116.

Laprise R. 2008. Regional climate modelling. Journal of Computational Physics 227: 3641—3666.

Leander R, Buishand TA. 2007. Resampling of regional climate model output for the simulation of extreme river flows. Journal of Hydrology 332: 487—496.

Leander R, Buishand TA, van den Hurk BJJM, de Wit MJM. 2008. Estimated changes in flood quantiles of the river Meuse from resampling of regional climate models. Journal of Hydrology 351: 331—343.

Lee C, Whitely Binder LC. 2010. Assessing Pacific Northwest water resources stakeholder data needs. In: Final Report for the Columbia Basin Climate Change Scenarios Project. Climate Impacts Group. Center for Science in the Earth System, Joint Institute for the Study of the Atmosphere and Ocean. University of Washington. Seattle.

Lenderik G, Buishand A, van Deursen W. 2007. Estimates of future discharges of the river Rhine using two scenario methodologies: direct versus delta approach. Hydrology and Earth System Sciences 11: 1145—1159.

*Bibliography*

Loibl, W, Beck A, Dorninger M, Formayer H, Gobiet A, Schöner W (eds). 2007. Kwiss-Programm reclip:more: research for climate protection:model run evaluation. Final Report, ARC-sys-0123. Austrian Research Centers–systems research. Vienna. <http://systemsresearch.arcs.ac.at/SE/projects/reclip/>. Accessed 2009 Aug 12.

Maraun D, Ireson A, Wetterhall F, Bachner S, Kendon E, Rust HW, Venema VKC, Widmann M, Chandler RE, Onof CJ, Osborn TJ, Sautner T, Themeßl M, Thiele-Eich I. 2010. Statistical downscaling and modelling of precipitation. Bridging the gap between dynamical models and the end users. Reviews of Geophysics 48,RG3003. doi: 10.1029/2009RG000314.

Matulla C, Penlap EK, Haas P, Formayer H. 2003. Comparative analysis of spatial and seasonal variability: Austrian precipitation during the 20th century. International Journal of Climatology 23: 1577–1588.

Matulla C, Zhang X, Wang XL, Wang J, Zorita E, Wagner S, von Storch H. 2008. Influence of similarity measures on the performance of the analog method for downscaling daily precipitation. Climate Dynamics 30: 133–144. doi: 10.1007/s00382-007-0277-2.

McNie EC. 2007. Reconciling the supply of scientific information with user demands: an analysis of the problem and review of the literature. Environmental Science and Policy 10: 17–38.

Mehrotra A, Sharma A. 2006. Conditional resampling of hydrologic time series using multiple predictor variables: A K-nearest neighbour approach. Advances in Water Resources 29: 987–999.

Mitchell TD, Carter TR, Jones PD. Hulme M. New M. 2004. A comprehensive set of high-resolution grids of monthly climate for Europe and the globe: the observed record (1901-2000) and 16 scenarios (2001-2100). Geography: 55. Tyndall Centre for Climate Change Research. Norwich. UK.

Moron V, Robertson AW, Ward MN, Ndiaye O. 2008: Weather Types and Rainfall over Senegal. Part II: Downscaling of GCM simulations. Journal of Climate 21: 288–307.

Munich Re. 2011. Topics Geo, Natural catastrophes 2010 Analyses, assessments, positions. Munich Re. Munich.

Murphy J. 1999. An Evaluation of Statistical and Dynamical Techniques for Downscaling Local Climate. Journal of Climate 12: 2256–2284.

Murphy J. 2000. Prediction of climate change over Europe using statistical and dynamical downscaling techniques. International Journal of Climatology 20: 489–501.

Nakicenovic, N, Alcamo J, Davis G, de Vries B, Fenhann J, Gaffin S, Gregory K, Grübler A, Jung TY, Kram T, La Rovere EL, Michaelis L, Mori S, Morita T, Pepper W, Pitcher H, Price L, Raihi K, Roehrl A, Rogner H-H, Sankovski A, Schlesinger M, Shukla P, Smith S, Swart R, van Rooijen S, Victor N, Dadi Z. 2000. IPCC Special Report on Emissions Scenarios. Cambridge University Press. Cambridge, New York.

Organe Consultatif sur les changements climatiques (OcCC). 2003: Extremereignisse und Klimaänderung, presented at media conference: Extremereignisse: Die Gefährdung ändert sich. OcCC. Bern.

120

Panofsky HW, Brier GW. 1968. Some Applications of Statistics to Meteorology. The Pennsylvania State University Press. Philadelphia.

Piani C, Haerter JO, Coppola E. 2009. Statistical bias correction for daily precipitation in regional climate models over Europe. Theoretical and Applied Climatology 99 (1-2): 187‒192. doi: 10.1007/s00704-009-0134-9.

Piani C, Weedon GP, Best M, Gomes SM, Viterbo P, Hagemann S, Haerter JO. 2010. Statistical bias correction of global simulated daily precipitation and temperature for the application of hydrological models. Journal of Hydrology. doi: 10.1016/j.jhydrol.2010.10.024.

Poumadere M, Mays C, Le Mer S, Blong R. 2005. The heat wave 2003 in France: dangerous climate here and now. Risk Analysis 25: 1483‒1495.

Prein A, Truhetz H, Suklitsch M, Gobiet A. 2011. NHCM – Non hydrostatic climate modelling, Part III, Evaluation of the Local Climate Model Intercomparison Project (LocMIP) Simulations. Wegener Center Verlag. Wegener Center for Climate and Global Change. University of Graz. Graz. ISBN 978-3-9502940-8-8.

Prettenthaler FE, Töglhofer C, Habsburg-Lothringen C, Türk A. 2008. Klimabedingte Änderung des Heiz- und Kühlenergiebedarfes. In Prettenthaler F, Gobiet A (eds.). Heizen und Kühlen im Klimawandel. Wien. Verlag der Österreichischen Akademie der Wissenschaften.

Prudhomme C, Reynard N, Crooks S. 2002. Downscaling of global climate models for flood frequency analysis: where are we now. Hydrological Processes. 16: 1137‒1150. doi: 10.1002/hyp.1054.

Pryor SC, Schoof JT, Barthelmie RJ. 2005. Climate change impacts on wind speeds and wind energy density in northern Europe: empirical downscaling of multiple AOGCMs, Climate Research. 23: 183‒198.

Räisänen J, Hansson U, Ullerstig A, Döscher R, Graham LP, Jones C, Meier HEM, Samuelsson P, Willen U. 2004. European climate in the late twenty-first century: regional simulations with two driving global models and two forcing scenarios. Climate Dynamics 22: 13‒31.

Randall DA, Wood RA, Bony S, Colman R, Fichefet T, Fyfe J, Kattsov V,.Pitman A, Shukla J, Srinivasan J, Stouffer RJ, Sumi A, Taylor KE. 2007. Climate Models and Their Evaluation. In: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M,. Miller HL (eds.)]. Cambridge University Press. Cambridge, New York.

Rajagopalan B, Lall U, Tarboton TG, Bowles DS. 1997. Multivariate nonparametric resampling scheme for generation of daily weather variables. Stochastic Hydrology and Hydraulics 11: 65‒93.

Rummukainen, M. 1997. Methods of statistical downscaling of GCM simulations, RMK No. 80, Swedish Meteorological and Hydrological. Institute. Norrköping.

Sachs L, Hedderich J. 2006. Angewandte Statistik. Methodensammlung mit R. Springer. Heidelberg, Berlin.

*Bibliography*

Salathé EP, Mote PW, Wiley MW. 2007. Review of scenario selection and downscaling methods for the assessment of climate change impacts on hydrology in the United States pacific northwest. International Journal of Climatology 27: 1611─1621 doi: 10.1002/joc.1540.

Salathé EP. 2003. Comparison of various precipitation downscaling methods for the simulation of streamflow in a rainshadow river basin. International Journal of Climatology 23: 887─901.

Schär C, Vidale PL, Lüthi D, Frei C, Häberli C, Linger MA, Appenzeller C. 2004. The role of increasing temperature variability in European summer heat waves. Nature 427: 332─336.

Schmidli J, Frei C, Vidale PL. 2006. Downscaling from GCM precipitation: A benchmark for dynamical and statistical downscaling methods. International Journal of Climatology 26: 679─689.

Schmidli J, Goodess CM, Frei C, Haylock MR, Hundecha Y, Ribalaygua J, Schmith T. 2007. Statistical and dynamical downscaling of precipitation: An evaluation and comparison of scenarios for the European Alps. Journal of Geophysical Research 112, D04105. doi: 10.1029/2005JD007026.

Schoof JT, Pryor SC. 2001. Downscaling temperature and precipitation: a comparison of regression-based methods and artificial neural networks. International Journal of Climatology 21: 773─790.

Seibert P, Frank A, Formayer H. 2007. Synoptic and regional patterns of heavy precipitation in Austria. Theoretical and Applied Climatology 87: 139-153.

Seneviratne SI, Corti T, Davin EL, Hirschi M, Jaeger EB, Lehner I, Orlowsky B, Teuling AJ. 2010. Investigating soil moisture-climate interactions in a changing climate: A review. Earth-Science Reviews 99(3-4): 125─161. doi: 10.1016/j.earscirev.2010.02.004.

Smith RL. 2001. Environmental statistics. <http://www.stat.unc.edu/postscript/rs/envnotes.ps>. Accessed 2009 March 3.

Srikanthan R, Pegram GGS. 2009. A nested multisite daily rainfall stochastic generation model. Journal of Hydrology 371: 142─153.

Steininger K, Steinschreiber C, Ritz C. 2005. Extreme Wetterereignisse und ihre wirtschaftlichen Folgen. Springer. Berlin, Heidelberg.

Stocker TF, Clarke GKC, Le Treut H, Lindzen RS, Meleshko VP, Mugara RK, Palmer TN, Pierrehumbert RT, Seller PJ, Trenberth KE, Willebrand J. 2001. Physical Climate Processes and Feedbacks. In: Climate Change: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change [Hoghton JT, Ding Y, Griggs DJ, Noguer M, van der Linden PJ, Dai X, Maskell K, Johnson CA (eds.)]. Cambridge University Press. Cambridge, New York.

Suklitsch M, Gobiet A, Leuprecht A, Frei C. 2008. High Resolution Sensitivity Studies with the Regional Climate Model CCLM in the Alpine Region. Meteorologische Zeitschrift 17: 467─476.

Suklitsch M, Gobiet A, Truhetz H, Awan NK, Göttel H, Jacob D. 2010. Error Characteristics of High Resolution Regional Climate Models over the Alpine Area. Climate Dynamics. doi: 10.1007/s00382-010-0848-5.

Sun L, Li H, Ward ML. 2007. Climate Variability and Corn Yields in Semiarid Ceará, Brazil. Journal of Applied Meteorology and Climatology 46: 226─250.

Swart R, Pagé C. 2010. Defining climate modeling user needs: which data are actually required to support impact analysis and adaptation policy development. American Geophysical Union, Fall Meeting. abstract # GC31E-04.

Swart R. 2011. Bridging Climate Research Data and the Needs of the Impact Community. Discussion paper of the IS-ENES/EEA/CIRCLE workshop, 11-12 January 2012. Copenhagen.

Taylor KE. 2001. Summarizing multiple aspects of model performance in a single diagram. Journal of Geophysical Research 106: 7183─7192.

Terink W, Hurkmans RTWL, Torf PJJF, Uijlenhoet R. 2010. Evaluation of a bias correction method applied to downscaled precipitation and temperature reanalysis data for the Rhine basin. Hydrology and Earth System Sciences 14: 687─703.

Themeßl M. 2006. Downscaling of Temperature and Precipitation in the Alpine Region Hohe Tauer. Wegener Center Verlag. Wegener Center for Climate and Global Change. University of Graz. Graz. ISBN 3-9502126-2-0.

Themeßl M, Gobiet A, Leuprecht A. 2010. Empirical-statistical downscaling and error correction of daily precipitation from regional climate models. International Journal of Climatology. doi: 10.1002/joc.2168.

Themeßl M, Gobiet A, Truhetz H. 2011. Definition von Nutzern und die Abschätzung der benötigten Speicherkapazitäten eines Klimadatenzentrums in Österreich. Technical report to the ACRP project KlimDAtZ: Concept for an „Austrian Climate Data Centre". Project no: AC0K00016.

Thornthwaite CW. 1948. An approach toward a rational classification of climate. Geographical Review 38: 55–94. doi: 10.2307/210739.

Truhetz H, Gobiet A, Kirchengast G. 2007. Evaluation of a dynamic-diagnostic modelling approach to generate highly resolved wind fields in the Alpine region. Meteorologische Zeitschrift 16: 191–201.

Uppala SM, Kållberg PW, Simmons AJ, Andrae U, Da Costa Bechtold V, Fiorino M, Gibson JK, Haseler J, Hernandez A, Kelly GA, Li X, Onogi K, Saarinen S, Sokka N, Allan RP, Andersson E, Arpe K, Balmaseda MA, Beljaars ACM, van De Berg L, Bidlot J, Bormann N, Caires S, Chevallier F, Dethof A, Dragosavac M, Fisher M, Fuentes M, Hagemann S, Hólm E, Hoskins BJ, Isaksen L, Janssen PAEM, Jenne R, Mcnally AP, Mahfouf JF., Morcrette JJ, Rayner NA, Saunders RW, Simon P, Sterl A, Trenberth KE, Untch A, Vasiljevic D, Viterbo P, Woollen J. 2005. The ERA-40 re-analysis. Quarterly Journal of the Royal Meteorological Society 131 (612): 2961–3012.

*Bibliography*

van den Besselaar EJM, Haylock MR, van der Schrier G, Klein Tank AMG. 2011. A European Daily High-resolution Observational Gridded Data Set of Sea Level Pressure. Journal of Geopyhsical Research. doi: 10.1029/2010JD015468. In press.

van der Linden P, Mitchell JFB. 2009. ENSEMBLES: Climate Change and its Impacts: Summary of research and results from the ENSEMBLES project. Met Office Hadley Centre. Exeter.

van Pelt SC, Kabat P, ter Maat HW, van den Hurk BJJM, Weerts AH. 2009. Discharge simulation performed with a hydrological model using bias corrected regional climate model input. Hydrology and Earth System Sciences 13: 2387─2397.

von Storch H, Zwiers FW. 1999. Statistical Analysis in Climate Research. Cambridge University Press. Cambridge.

von Storch H, Navarra A (eds). 1999. Analysis of Climate Variability: Applications of Statistical Techniques. Springer. Berlin.

von Storch H. 1999. On the Use of "Inflation" in Statistical Downscaling. Journal of Climate 12: 3505─3506.

von Storch H, Güss S, Heimann M. 1999. Das Klimasystem und seine Modellierung. Springer. Berlin, Heidelberg, New York.

von Storch H, Hewitson B, Mearns L. 2000. Review of Empirical Statistical Downscaling Techniques. In: Iverson T Hoiskar BAK (eds.). Regional climate development under global warming. General Technical Report No. 4 Conf. Proceedings RegClim Spring Meeting Jevnaker, Norway, 8-9 May 2000.

Wang Y, Leung LR, McGregor JL, Lee DK, Wang WC, Ding Y, Kimura F. 2004. Regional Climate Modeling: Progress, Challenges, and Prospects. Journal of the Meteorological Society of Japan 82: 1599─1628.

Wang J, Zhang X. 2008. Downscaling and Projection of Winter Extreme Daily Precipitation over North America. Journal of Climate 21: 923─937.

Warner TT. 2011. Numerical Weather and Climate Prediction. Cambridge University Press. Cambridge.

Wetterhall F, Halldin S, Xu C. 2005. Statistical precipitation downscaling in central Sweden with the analogue method. Journal of Hydrology 306: 174─190.

Widmann M, Bretherton CS, Salathé E. 2003. Statistical Precipitation Downscaling over the Northwestern United States Using Numerically Simulated Precipitation as Predictor. Journal of Climate 16: 799─816.

Wilcke R, Mendlik T, Gobiet A, Themeßl M. 2011. Downscaling and error correction of temperature, precipitation, wind speed, relative humidity and global radiation and its impact for the physical consistency. In preparation.

Wilby R, Wigley T. 1997. Downscaling general circulation model output: a review of methods and limitations. Progress in Physical Geography 21: 530─548.

Wilby R. 1997. Non-stationarity in daily precipitation series: implications for GCM down-scaling using atmospheric circulation indices. International Journal of Climatology 17: 439─454.

Wilby R, Wigley T. 2000. Precipitation predictors for downscaling: observed and general circulation model relationships. International Journal of Climatology 20: 641─661.

Wilby RL, Tomlinson OJ, Dawson CW. 2003. Multi-site simulation of precipitation by conditional Resampling. Climate Research 23: 183-194.

Wilby RL, Charles SP, Zorita E, Timbal B, Whetton P, Mearns LO. 2004: Guidelines for the Use of Climate Scenarios Developed from Statistical Downscaling Methods, Supporting material for the IPCC.

Wilks DS, Wilby RL. 1999. The weather generation game: a review of stochastic weather models, Progress in Physical Geography 23: 329─357.

Wilks DS. 1995. Statistical Methods in Atmospheric Science, Volume 59 of International Geophysics Series. Academic Press. San Diego, London.

Wilks DS. 2006. Statistical Methods in Atmospheric Science, Volume 91 of International Geophysics Series. Academic Press. San Diego, London.

Wood AW, Leung LR, Sridhar V, Lettenmaier DP. 2004. Hydrologic Implications of Dynamical and Statistical Approaches to Downscale Climate Model Outputs. Climatic Change 62: 189─216.

Wu J. 1999. Hierarchy and Scaling: Extrapolating information along a scaling ladder. Canadian Journal of Remote Sensing 25: 367─380.

Yarnal B, Comrie AC, Frakes B, Brown DP. 2001. Developments and Prospects in Synoptic Climatology. International Journal of Climatology 21: 1923─1950.

Zorita E, von Storch H. 1999. The Analog Method as a Simple Statistical Downscaling Technique: Comparison with More Complicated Methods. Journal of Climate 12: 2474─2489.

# Abstract

Although RCMs have already proven their capability to simulate regional climate and its variability, they still feature systematic errors compared to observations. Besides their steady enhancement, empirical-statistical post-processing, based on the concept of model output statistics (MOS), provides a ready opportunity to mitigate RCM error characteristics and to further downscale climate model data to the point-scale. In the course of this PhD work, seven empirical-statistical downscaling and error correction methods (DECMs) are inter compared for their applicability to and error correction potential for daily precipitation, temperature, and derived extreme indices from RCMs in Europe.

Furthermore, error corrected climate scenarios for the respective parameters are generated for Europe and the impact of DECMs on the climate change signal (CCS) is investigated.

Overall, the findings of this PhD work strongly suggest the combination of RCMs and DECMs to provide suitable climate data for climate impact assessments and decision making. DECMs drastically reduce the error characteristics of RCMs regarding mean, variability, and extremes. Particularly, Quantile Mapping (QM) has an outstanding error correction potential and can be considered as highly recommendable due to its simplicity and flexibility. In application to future climate scenarios QM only moderately modifies the CCSs of mean,minimum, maximum temperature, and precipitation amount. In contrast, QM strongly changes the CCSs of non-linearly derived indices of extremes such as threshold indices in some cases. These modifications are caused by magnitude-dependent error characteristics of the respective uncorrected parameters  and trends in the future scenarios. Besides the analysis of DECMs, this PhD work also defines useful climate data from the point of view of the climate impact community and decision makers.