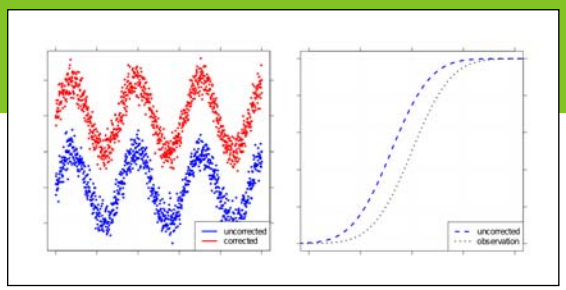
Wegener Center for Climate and Global Change
University of Graz

Scientific Report No. 42-2011



Empirical-Statistical Post-Processing of Radiation and Humidity Parameters in Regional Climate Simulations

Reinhard Fuchs

March 2011

kindly supported also by

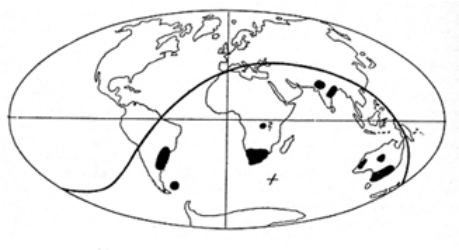
FWF
Der Wissenschaftsfonds.

Wegener Center
www.wegcenter.at

UNI GRAZ

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 information at www.wegcenter.at)

The present report is the result of a Master thesis work completed in December 2010. It was financially supported by the Austrian Science Fund (FWF).



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).

Wegener Center Verlag • Graz, Austria
© 2011 All Rights Reserved.
Selected use of individual figures, tables or parts of text is permitted for non-commercial purposes, provided this report is correctly and clearly cited as the source. Publisher contact for any interests beyond such use: wegcenter@uni-graz.at.

ISBN 978-3-9502940-9-5

March 2011

Contact: Reinhard Fuchs
reinhard.fuchs@alumni.uni-graz.at

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

Reinhard Fuchs, Bakk.rer.nat.

Empirical-Statistical Post-Processing of Radiation and Humidity Parameters in Regional Climate Simulations

Master Thesis
to obtain the degree
Magister Rerum Naturalium
at the Department of Natural Sciences
University of Graz

Supervisor:
Ass.-Prof. Mag. Dr.rer.nat. Andreas Gobiet
Wegener Center for Climate and Global Change
Institute for Geophysics, Astrophysics and Meteorology
University of Graz

2010

Contents

Acknowledgments vii

Abstract ix

Zusammenfassung xi

1 Introduction 1
- 1.1 General Overview of Climate Modeling . . . 1
- 1.2 Climate Model Uncertainties . . . 4
- 1.3 Bias in Regional Climate Models . . . 7
- 1.4 Classifications of Empirical-Statistical Downscaling Methods . . . 9
- 1.5 Criteria for Empirical-Statistical Downscaling and Error Correction . . . 10
- 1.6 The DEC Method: Quantile Mapping . . . 11

2 Data and Methods 13
- 2.1 Data . . . 13
  - 2.1.1 Synthetical Distributions . . . 13
  - 2.1.2 Observations and Regional Climate Simulation Data . . . 14
- 2.2 Quantile Mapping Methods . . . 17
  - 2.2.1 Quantile Mapping Based on Empirical Data . . . 19
  - 2.2.2 Quantile Mapping Based on Fitted Distributions . . . 19
  - 2.2.3 Quantile-Quantile Mapping . . . 22
- 2.3 Data Analysis . . . 23
  - 2.3.1 Statistical Methods . . . 23
  - 2.3.2 Fitting of Empirical Distributions . . . 30

3 Results 37
- 3.1 Evaluation of the Quantile Mapping Methods with Synthetical Distributions . . . 37
  - 3.1.1 Normally Distributed Data . . . 38
  - 3.1.2 Non-normally Distributed Data . . . 40
  - 3.1.3 Normally Distributed Data under Changed Conditions . . . 44
- 3.2 Post-Processing of Regional Climate Model Output . . . 46
  - 3.2.1 Temperature . . . 47

3.2.2 Precipitation 51
3.2.3 Relative Humidity 59
3.2.4 Global Radiation 62
3.3 Influence of Post-Processing on the Climate Change Signal of a Future Period 69

4 Discussion and Conclusion **77**

A Additional Figures **83**

Acronyms **93**

List of Figures **97**

List of Tables **101**

Bibliography **103**

Acknowledgments

I would like to thank my supervisor Ass.-Prof. Mag. Dr.rer.nat. Andreas Gobiet for the opportunity to write my master thesis in the ReLoClim group at the WEGCENTER, the good advices and the motivation. Univ.-Prof. Mag. Dr.rer.nat. Gottfried Kirchengast and the WEGCENTER members for their support. The whole ReLoClim group, especially, Matthias and Georg for the interesting discussions. Armin and Martin for their help with programming problems. Barbara and Martin for the proof-reading of my master thesis. Florian and Michael for the help with LaTeX.

I would also like to thank my colleagues in the office for the nice coffee breaks, discussions and friendly atmosphere. My family and friends for their support during my studies. Especially, Bianca for the very helpful comments.

Finally, I am grateful to the Austrian Student Grant Authority and the WEGCENTER for their financial aid during my studies.

Abstract

Regional climate models (RCMs) provide highly resolved information about different meteorological parameters. However, error characteristics are still strongly pronounced in some regions. Post-processing methods are a possibility to get error corrected climate model data.

In this thesis, the empirical-statistical Downscaling and Error Correction (DEC) method called Quantile Mapping (QM) is used. It has so far mostly been applied to the climate parameters temperature and precipitation, whether on monthly or daily basis. Here, the focus is on evaluating the results of QM applied to daily temperature, precipitation, relative humidity and global radiation from an RCM simulation related to two point-scale observations located in Vienna and Innsbruck.

In many studies, fitted distributions are used in the calibration period of the DEC method, but not any used climate parameter in this thesis has consistently matched to a theoretical distribution. Therefore, the empirical data of a regional climate hindcast simulation (REMO57-ERA40) and observations are taken instead of fitted data. Three different validation approaches are used to analyze the error characteristics of the uncorrected and post-processed data. Furthermore, a scenario run (REMO57-A1B) is post-processed to analyze the Climate Change Signal (CCS).

To test the DEC method itself, QM is applied to synthetic data extracted from different distributions. These results are promising. The findings for the first two validation experiments ("calibration identical validation", "cross-validation") show an improvement of the whole probability density function (PDF). The split-sample test displays advanced results for temperature, precipitation and relative humidity, but almost no improvement for the parameter global radiation. The CCS of the corrected data does not differ too much from the uncorrected CCS and it does not show a better agreement with the observed CCS.

The results demonstrate that it is useful and necessary to post-process temperature, precipitation and relative humidity, but global radiation is a more demanding parameter. Long time periods and high quality observations would be required to capture the whole variability within one climate parameter. In general, QM used with empirical data is applicable to any climate parameter and can provide point-scale data.

Zusammenfassung

Regionale Klimamodelle (RCMs) liefern hoch aufgelöste Informationen über unterschiedlichste meteorologische Parameter. Jedoch sind Fehler in bestimmten Regionen noch deutlich zu erkennen. Verschiedene Nachbearbeitungsmethoden sind in der Lage korrigierte Klimamodelldaten bereitzustellen.

In dieser Arbeit wird die empirisch-statistische Downscaling und Fehlerkorrektur-Methode (DEC) Quantile Mapping (QM) verwendet. Diese ist bis jetzt meist auf die Klimaparameter Temperatur und Niederschlag entweder auf Tages- oder Monatsbasis angewendet worden. Der Schwerpunkt der Arbeit liegt in der Bewertung der nachbearbeiteten Daten für Temperatur, Niederschlag, relative Feuchte und Globalstrahlung auf Tagesbasis für eine RCM Simulation bezogen auf zwei Messstationen in Wien und Innsbruck.

In vielen Studien werden gefittete Verteilungen in den Kalibrierungsperioden der DEC Methode verwendet, aber keiner in dieser Arbeit benutzte Klimaparameter hat konsistent mit einer theoretischen Verteilung übereingestimmt. Daher werden die empirischen Daten des regionalen Hindcast Klimamodells (REMO57-ERA40) und die empirischen Messdaten anstatt der gefitteten Daten benutzt. Drei unterschiedliche Validierungsansätze finden für die Analyse der Fehlercharakteristiken der unkorrigierten und nachbearbeiteten Daten Verwendung. Weiters wird ein Szenario Lauf (REMO57-A1B) nachbearbeitet, um das Klimawandelsignal (CCS) zu untersuchen.

Zuerst ist die DEC Methode mittels synthetisch generierten Daten, die einer bestimmten theoretischen Verteilungsform folgen, getestet worden. Diese Ergebnisse sind vielversprechend gewesen. Die Resultate der ersten beiden Validierungsexperimenten („Kalibrierung identisch Validierung“, „Cross-Validierung“) geben eine Verbesserung der gesamten Wahrscheinlichkeitsdichtefunktion (PDF) wieder. Der Split-Sample Test zeigt Verbesserungen bei den Parametern Temperatur, Niederschlag und relativer Feuchte, aber kaum Verbesserungen bei Globalstrahlung. Das CCS hat sich zwischen den korrigierten und unkorrigierten Daten wenig geändert und zeigt keine eindeutig bessere Übereinstimmung mit dem gemessenen CCS.

Die Ergebnisse zeigen, dass es sinnvoll und notwendig ist Temperatur, Niederschlag und relative Feuchte nachzubearbeiten, aber Globalstrahlung scheint ein anspruchsvollerer Parameter zu sein. Lange Zeitreihen und qualitativ hochwertige Messdaten sind notwendig um die gesamte Variabilität innerhalb eines Parameters abzudecken. QM kann generell auf die empirischen Daten jeglicher Klimaparameter angewendet werden und dabei auch punkt-skalige Daten liefern.

1 Introduction

1.1 General Overview of Climate Modeling

Climate models are simplified reconstructions of the climate system, which are defined through mathematical formulas. The climate system includes the main spheres: the atmosphere, hydrosphere, cryosphere, pedosphere (land surface) and the biosphere as well as all the interactions between these spheres. Furthermore, external forcings (e.g., volcanic activity, changes in solar activity, etc.) and anthropogenic forcings (e.g., change of the composition of the atmosphere, change of land use, etc.) are also contained in these equations. Physical laws (e.g., the conservation of mass, conservation of momentum, conservation of energy) and parameterizations are used to model the climate system. The calculations are done on discrete spatial and temporal scales (numerical grid). Figure 1.1 shows the development of climate models in the last few decades. In the early years of climate modeling, just a few components and processes were taken into account. The models have obviously improved through the years of further development. Step by step more and more components and processes have been added. Almost the same progress can be seen in the development of a finer numerical grid. Figure 1.2 depicts, how the spatial grid for Europe had been improved to a finer one for General Circulation Models (**GCMs**) used in the four Intergovernmental Panel on Climate Change (**IPCC**) assessment reports. In the First Assessment Report (**FAR**) of the IPCC, the GCMs simplified the land mass of Europe extremely as well as the mountain ranges. For example, the Alps were not visible at all, due to the resolution of around 500 km. The spatial resolutions in the GCMs had been clearly improved, which were used in the Fourth Assessment Report (**AR4**) (Figure 1.2). The spatial resolution has been developed through the four IPCC reports from around 500 km to around 110 km (IPCC, 2007, pp. 93–127, 589–662).

GCMs have the skill to represent mean climate parameters such as air temperature, precipitation, radiation, wind and of sea temperature, currents and ice cover. However, GCMs are able to represent mean climate parameters only on a large-scale. While important information of the global climate system behavior can be evaluated, it is not possible to give answers for a smaller scale. Thus, to get information below the grid scale of the GCM a process called downscaling is introduced. Dynamical (with Regional Climate Models (**RCMs**)) and Empirical-Statistical Downscaling (**ESD**) are the two main approaches in this context (IPCC, 2007, pp. 96–97, 918–925).

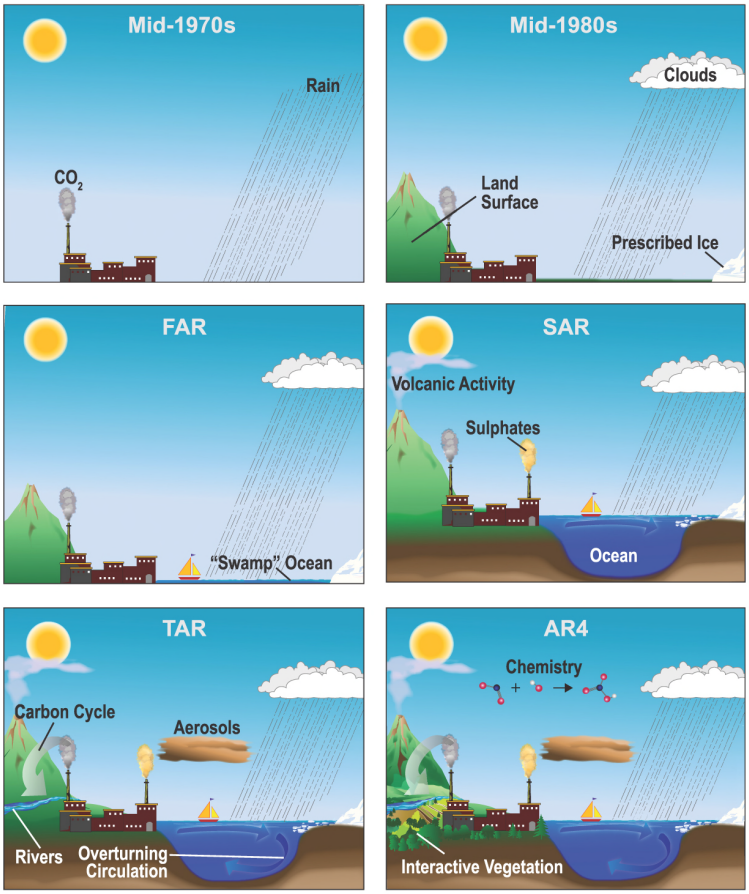


Figure 1.1: Scheme of the development of the processes in climate modeling since mid 1970s. First Assessment Report (FAR) – 1990; Second Assessment Report (SAR) – 1996; Third Assessment Report (TAR) – 2001, Fourth Assessment Report (AR4) – 2007 (IPCC, 2007, Figure 1.2).

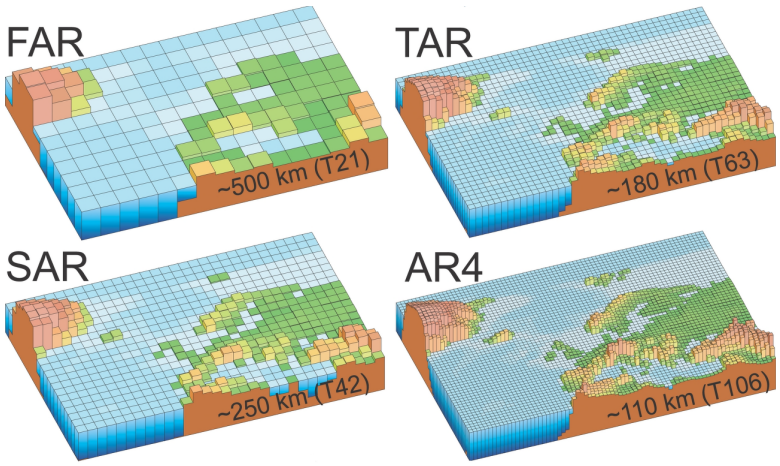


Figure 1.2: Scheme of the development of the spatial grid in the last few decades on the example of Europe. First Assessment Report (FAR) – 1990; Second Assessment Report (SAR) – 1996; Third Assessment Report (TAR) – 2001, Fourth Assessment Report (AR4) – 2007 (IPCC, 2007, Figure 1.4).

RCMs are high resolution models which are formulated using physical principles. The physical processes and parameterizations, the orography, land-sea contrast and land use are modeled on a finer scale. GCMs or observations are used as input at the boundaries. The horizontal grid space is nowadays up to 25 km (e.g., ENSEMBLES, 2009). Drawbacks of RCMs are the computational costs and parameterization schemes, which represent the climate of today and might not be in the range of a future climate (IPCC, 2007, pp. 918–919).

ESD methods are further oppertunities to get physical information on a finer scale. A statistical relationship between predictors (e.g., GCM, large-scale) and predictands (e.g., observations, local/point-scale) has to be derived in a calibration period and then applied to the climate model data in a scenario simulation. The advantages of ESD are the ability of accessing finer scales than an RCM and of being computationally cheap. A disadvantage is the requirement of observation data. Furthermore, a few criteria have to be fulfilled, which are given in Section 1.5. Especially, the relationship between predictor and predictand has to stay stable over time, which is difficult to verify (IPCC, 2007, pp. 918–921). A benefit of ESD is the potential to supply climate data at local/point-scale (ENSEMBLES, 2009, p. 67). The ESD method can be applied between RCM output and a point-scale observation to get downscaled and corrected model data. Themeßl et al. (2010) use the term "empirical-statistical Downscaling and Error Correction (**DEC**)" for a few ESD methods to highlight the error correction part.

1.2 Climate Model Uncertainties

Uncertainty analysis is very important in climate research, hence, it will be explained here in the introduction, even though, it will not be addressed specifically in the thesis. Uncertainties in climate scenarios arise from several sources. The major sources of uncertainties from the first calculations to the final output are (e.g., Hawkins and Sutton, 2009; ENSEMBLES, 2009; Prein, 2009; Rowell, 2006; Déqué et al., 2007):

- Uncertainty in radiative forcing (Special Report on Emissions Scenarios (**SRES**))
- Uncertainty in internal variability
- Uncertainty in GCM formulation (Boundary conditions for RCM)
- Uncertainty in RCM formulation
- Uncertainty due to DEC

Uncertainty in Radiative Forcing (SRES)

A source of uncertainty occurs due to the unknown future greenhouse gas emissions and the resulting radiative forcing (Hawkins and Sutton, 2009). Generally, six groups of emission scenarios are available with different assumptions about the global population, economic growth, energy use and other relevant parameters. In this thesis the A1B scenario is used as an input for the GCM. It represents a future world with very quick economic growth, rapid implementation of new and more efficient technologies and a peaking global population in mid-century with a decline afterwards. The technology usage is balanced across all sources (instead of the fossil intensive A1FI scenario or the non-fossil A1T scenario) (Nakicenovic and Swart, 2000). The sources of uncertainties in surface temperature of the CMIP3 project (for further information see: PCMDI, 2009) are displayed in Figure 1.3. The uncertainty of the emission scenarios is rather small for 2021–2050 (5.56 %) and becomes larger for 2071–2100 (35.9 %)(Prein, 2009). In this thesis, one of the used time periods is from 2021 to 2050. Therefore, the emission scenario uncertainty does not play a too decisive role (ENSEMBLES, 2009, p. 51).

Uncertainty in Internal Variability (Choice of Initial Conditions)

Internal variability of the climate system is another source of uncertainty in climate modeling. Natural fluctuations are due to chaotic behavior of the climate system. Generally, the uncertainty in internal variability decreases with larger spatial and longer time scales (Hawkins and Sutton, 2009). Déqué et al. (2007) discover for the PRUDENCE project (more information in: PRUDENCE, 2004) that the uncertainty in internal variability is less important compared to other sources of uncertainty towards the end of

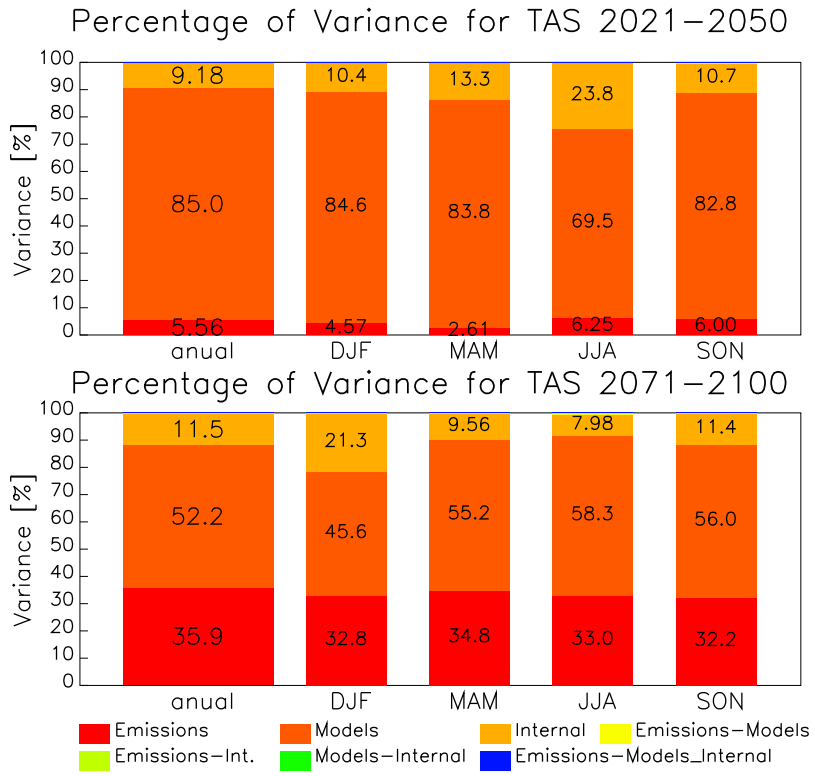


Figure 1.3: Uncertainty in surface temperature of the CMIP3 GCMs for a specific domain size over Europe (N 32° – N 60°, W 15° – E 35°)(Prein, 2009).

the 21st century, but relatively large in the first decades as the climate change signal is not that large for this period (ENSEMBLES, 2009, pp. 51–52). The uncertainty in internal variability is 9.18 % for 2021–2050 (see Figure 1.3). This is only slightly larger than the uncertainty of the emission scenario for 2021–2050 and stays approximately at the same percentage level for 2071–2100.

Uncertainty in GCM Formulation (Driving Data for RCMs)

The formulation of GCMs is an additional source of uncertainty. Different climate change responses are simulated from different GCMs but with the same radiative forcing (Hawkins and Sutton, 2009). These different responses are due to uncertainties in the general structure, used parameterizations of sub-grid-scale processes (e.g., cloud physics) and coupling between the unresolved variability of the sub-grid scale and the resolved grid-scale flow (ENSEMBLES, 2009, p. 11). As can be seen in Figure 1.3, the uncertainty in GCM formulation is by far the largest source of uncertainty. GCMs are used as lateral boundary conditions for RCMs (driving data). Prein and Gobiet (2010) investigate in the quality of driving data in six upper air and three surface parameters. They conclude that "[...]there is no perfect GCM but some GCMs perform better than others [...]". Furthermore, they identify five best performing GCMs within the CMIP3 project. These are the UKMO-HadCM3, UKMO-HadGEM1, ECHAM5/MPI-OM, MIROC3.2(hires) and INGV-ECHAM4.

Uncertainty in RCM Formulation

So far, the explained sources of uncertainty are taken all into account for RCMs through the GCM, which provides the lateral boundary conditions for the RCM. Sources of uncertainty in RCM formulation are similar to the uncertainties in GCM formulation. The RCM formulations differ among each other. "[...] one particular RCM may amplify climate change signals in some GCM, another RCM may weaken the same" (ENSEMBLES, 2009, p. 51). Further sources of uncertainty are the resolution of the grid, the domain settings (Prein, 2009) and the choice of the GCM/RCM pair (ENSEMBLES, 2009, p. 11). Rowell (2006) states that "[...] uncertainty in the formulation of the RCM has a relatively small, but non-negligible, impact on the range of possible outcomes of future UK seasonal mean climate [...]" and that the "[...] largest source of uncertainty, for both variables and in all seasons, is the formulation of the global coupled model [...]".

Uncertainty due to DEC

The application of DEC adds another source of uncertainty. Especially, the relationship between predictor and predictand and the stationarity are the main sources of uncertainty (ENSEMBLES, 2009, p. 11). This will be discussed in more detail in Section 1.5.

In the ENSEMBLES summary, DEC uncertainties are first divided into three groups: "uncertainties 'previous' to downscaling", "uncertainties related to downscaling itself" and "'downstream' uncertainties". This is a quite basic classification, where the "uncertainties 'previous' to downscaling" is the overall uncertainty in the large-scale forcing. The "uncertainties related to downscaling itself" are linked to "stationarity, overfitting, range of applicapility, overall underlying skill and spatial resolution of predictands". The final uncertainty in this definition is the "'downstream' uncertainties" which identifies the uncertainty on, for example, how the projected changes impact human and natural systems. Furthermore, observations are a further source of uncertainty as they are used in the calibrating process (ENSEMBLES, 2009, p. 63).

1.3 Bias in Regional Climate Models

Section 1.2 provided a background about the uncertainty in climate research. Another important issue is the bias in climate models since one of the main parts of this thesis is the analysis of the performance of a DEC method with a hindcast simulation and observations. Different studies are available on how well RCMs perform (e.g., Déqué et al., 2007; Jacob et al., 2007; Suklitsch et al., 2010). In this section general results of a recent study by Suklitsch et al. (2010) are presented.

Suklitsch et al. (2010) examine four different RCMs (CCLM, MM5, MM5, REMO) and 62 one-year hindcast simulations regarding the error characteristic of high resolution regional climate models over a mountainous orography (the Alps). Figure 1.4 depicts the mean air temperature bias and the mean precipitation bias of these model simulations in 1999. Six different subregions within the greater Alpine region are defined for the analysis of the results. Two RCMs (CCLM and MM5) show nearly exclusively cold biases for air temperature, while MM5 and REMO show cold and warm biases depending on the model settings. The seasonal error range varies between -3.2 K and $+2.0$ K for the subregions. The right panel in Figure 1.4 shows the relative bias of mean precipitation. The two RCMs, WRF and MM5 simulate too wet conditions in nearly all subregions. The seasonal biases for mean precipitation show an overestimation in winter precipitation in all models. REMO and CCLM show more variations in the bias depending on the subregion and season. The seasonal error range varies between -45.7 % and $+94.7$ % for the subregions (Suklitsch et al., 2010). These results show that some RCMs perform quite well in these regions, but there still exists large biases. A possibility to reduce the error is to use DEC methods to get bias corrected data.

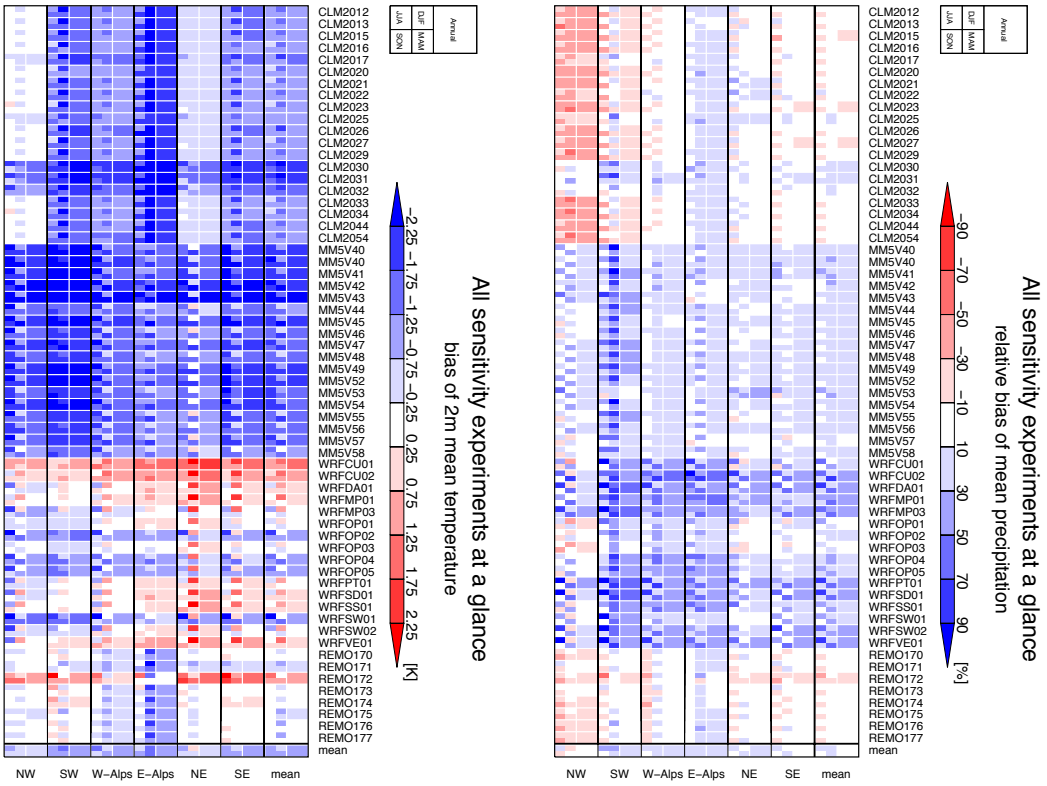


Figure 1.4: Bias of mean air temperature (left) and mean precipitation (right) of 62 one-year hindcast experiments with four different RCMs (CCLM, MM5, MM5, REMO) for the greater Alpine region. This region is divided into North-West (NW), South-West (SW), Western-Alps (W-Alps), Eastern-Alps (E-Alps), North-East (NE) and South-East (SE) (Suklitsch et al., 2010).

1.4 Classifications of Empirical-Statistical Downscaling Methods

The bias in climate models for a complex area was explained in the former section and introduced the need of DEC to get corrected point-scale model data. In this section, two opportunities to classify DEC methods are presented in a basic way.

Maraun et al. (2010) combine the ideas of Wilby and Wigley (1997) and Rummukainen (1997) and separate the DEC methods into:

- Perfect prognosis
- Model output statistics
- Weather generators

Perfect prognosis and model output statistics usually do not indicate different DEC methods but different procedures. The correction function for DEC methods used in perfect prognosis are calculated between large-scale and local-scale observations. This statistical relationship is then applied to model simulations (GCM or RCM) to get finer scales. Hence, perfect prognosis assumes that the simulated large-scale predictors represent the "perfect" climate ($\rightarrow$ no error correction). An advantage is that the derived function can be applied to different models. Model output statistic methods use a different approach. The calibration is done between simulated (large-scale) and observed (local-scale) data and then applied to the model which is used in the calibration period (Maraun et al., 2010). Weather generators return random numbers which are realistically looking sequences depending on the large-scale state (Storch et al., 1999*b*).

Déqué (2007) suggests a more impact orientated classification of DEC methods. The impact is the difference between model (scenario) output ($Out(S)$) and reference output ($Out(R)$).

$$Impact = Out(S) - Out(R) \tag{1.1}$$

Post-processed data, marked with a $'$, is used in the impact equation. Déqué (2007) proposes five ways to achieve the impact; most methods need a (daily) observation dataset ($O(d)$):

- The confident method: $R'(d) = R(d)$, $S'(d) = S(d)$
- The delta method: $R'(d) = O(d)$, $S'(d) = O(d) + \langle S - R \rangle$
- The unbiasing method: $R'(d) = R(d) + \langle O - R \rangle$, $S'(d) = S(d) + \langle O - R \rangle$
- The variable correction method: $R'(d) = f(R(d)|O)$, $S'(d) = f(S(d)|O)$
- The regime method: $R'(d) = O(d1)$, $S'(d) = O(d2)$

where $\langle\rangle$ stands for a multiannual average and f for a function. Since no post-processing is applied in the confident method, the impact is derived from the difference between the model and the reference output. The post-processed scenario used in the delta method is derived from the difference between a scenario and reference run which is added to the daily observations. The delta method is simple and robust, but neglects a changing climate variability. Within the unbiasing method the difference of a multiannual average between observations and reference output is added to the daily reference and scenario output. The advantage of this method is that no daily observations are needed. The variable correction method uses a function built on an observation dataset to receive the post-processed datasets. In the regime method, observations at different dates represent the reference and scenario data, which allows a changed frequency but not new extremes (Déqué, 2007).

The approach used in this thesis will be calibrated between RCM output (large-scale, predictor) and observations (point-scale, predictand) and can be classified as model output statistics method as defined by Maraun et al. (2010) and as a variable correction method as specified by Déqué (2007).

1.5 Criteria for Empirical-Statistical Downscaling and Error Correction

This section gives an overview on general criteria for applying DEC. Storch et al. (1999*b*) defines three criteria:

1. "The predictors are variables of relevance and are realistically modeled by the AOGCM."

2. "The transfer function is valid also under altered climatic conditions. This is an assumption that in principle cannot be proven in advance. The observational record should cover a wide range of variations in the past; ideally, all expected future realizations of the predictors should be contained in the observational record."

3. "The predictors employed fully represent the climate change signal."

Benestad et al. (2008, pp. 27–32) summarize four criteria for the application of DEC:

1. Strong relationship

2. Model representation

3. Description of change

4. Stationarity

A *strong relationship* between the predictor (large-scale, here RCM) and the predictand (point-scale, here observation) is necessary. This is given as long as the predictor and predictand have a large degree of covariance and similar time structures. The predictor has to be represented by the model in a realistical way. Furthermore, changes have to be captured which can be achieved as long as the predictor responds to given perturbations in a similar way as the predictand. Finally, the statistical relationship between predictor and predictand must not change over time (*stationarity*) (Wilby, 1997; Benestad et al., 2008, pp. 27–32).

The criteria of *strong relationship* and *model representation* are fulfilled due to the used DEC method. The same parameter (e.g., mean air temperature) is used as predictor and predictand. The predictor is given by the RCM (large-scale), the predictand by observations (point-scale). The criteria of *stationarity* and the *description of change* are crucial. It is not possible to give an ad hoc answer on this. It will be analyzed within this thesis. In the ENSEMBLES summary (ENSEMBLES, 2009, p. 68) it is noted:

> "The problem of stationarity – i.e., the assessment of whether the statistical relationships between large-scale GCM fields (predictors) and local observations (predictands) derived from present climate data are still valid under future climate conditions – remains a concern for statistical downscaling."

A way to analyze whether this criterion is fulfilled or not is to calibrate the DEC method in the coldest (driest) period and apply it to the warmest (wettest) period. If the results are still satisfying the stationarity criterion seems to be fulfilled (Maraun et al., 2010).

1.6 The DEC Method: Quantile Mapping

In the previous sections, an introduction to the field of DEC was given. In this section, the used method is explained in more detail and the research question will be posed. The DEC method has different names in the literature and will be called Quantile Mapping (**QM**) and the advanced version Quantile-Quantile Mapping (**Q-QM**) in this thesis. It relies on the empirical transformation of Panofsky and Brier (1968). In general, a correction function is calculated by matching two distributions (Boé et al., 2007; Ines and Hansen, 2006). QM was already used for different purposes. For example, the DEC method was used in the context of downscaling to modify the climate time series of the Parallel Climate Model (**PCM**) (Dettinger et al., 2004; Wood et al., 2004). In many studies, it was successfully applied to reduce the bias in (mean, minimum or maximum) temperature and precipitation amount on monthly or daily scale (e.g., Piani et al., 2010; Benestad, 2007; Dobler and Ahrens, 2008; Ines and Hansen, 2006). QM was even used to correct soil moisture by matching satellite and model data (Reichle and Koster, 2004).

In this thesis the main research question is:

- What are the error characteristics of the climate parameters daily temperature, precipitation, relative humidity and global radiation in a regional climate simulation after applying QM?

Additional questions are:

- How are the daily mean climate parameters temperature, precipitation, relative humidity and global radiation distributed for a specific observation and model grid cell? Is it useful to fit those data for the post-processing or is it proper to take the empirical data?
- How well does QM perform with synthetically generated (random) distributions?
- What is the influence of the method on the climate change signal?
- How does the added feature of Quantile Mapping called Quantile-Quantile Mapping perform?
- Are the criteria for empirical-statistical Downscaling and Error Correction fulfilled?

2 Data and Methods

2.1 Data

In this section, the used data for the thesis is described. The DEC methods (QM and Q-QM) are tested with randomly generated data and with observed and simulated data. Therefore, the data chapter is divided into a "synthetical distributions" and an "observations and regional climate simulation data" subsection.

2.1.1 Synthetical Distributions

The Probability Density Functions (**PDFs**) of the used synthetical distributions are presented in this subsection. For the analysis (hypothesis testing, see Subsection 2.3.2) of the empirical data, mainly the normal, gamma and beta distribution are taken. Moreover, different synthetical distribution datasets are created with random numbers with the programming language Interactive Data Language (**IDL**) to study how well the DEC method performs (see Section 3.1). The normal, gamma, log-normal and bimodal distributions are chosen for these experiments.

The normal (or Gaussian) distribution is defined with the two parameters *mean* $\mu \epsilon \mathbb{R}$ and the *variance* ($\sigma^2 \geq 0$) and with $x \epsilon \mathbb{R}$:

$$f(x|\mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \tag{2.1}$$

The gamma distribution is specified by the shape parameter $\alpha > 0$ and the scale parameter $\beta > 0$ and with $x \epsilon (0, \infty]$:

$$f(x|\alpha, \beta) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{(\alpha-1)} \exp\left(-\frac{x}{\beta}\right) \tag{2.2}$$

The PDF of the beta distribution with the two shape parameters $p > 0$ and $q > 0$ is defined between $0 \leq x \leq 1$:

$$f(x|p, q) = \left[\frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)}\right] x^{p-1}(1-x)^{q-1} \tag{2.3}$$

The log-normal distribution is characterized like the normal distribution, but the value

x is logarithmized and therefore, the limitation of $x\epsilon(0,\infty]$:

$$f_{ln}(x|\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right) \tag{2.4}$$

2.1.2 Observations and Regional Climate Simulation Data

Regions and Time Period

The validation of the DEC method is done on daily basis for point-scale data. The observation data is provided by the Central Institute for Meteorology and Geodynamics (**ZAMG**). Two criteria have to be satisfied for choosing the stations. The first is to have a long enough time series on daily basis for all four relevant parameters: temperature, precipitation, relative humidity and global radiation. A thirty year period from 1971 to 2000 is supposed to be a long enough time period. Long time series are quite common for temperature, precipitation and relative humidity, but not every weather station covers long time series of measurements for global radiation. The second criterion is that one station should be within the mountainous region and one in the more flat parts of Austria.

Under these conditions, two locations or rather two weather stations are chosen. Figure 2.1 shows the locations of all weather and climate stations from ZAMG in Austria. These are around 250 stations, most of them are semi-automatic weather recording systems called TAWES. The station *Hohe Warte in Vienna* and *University in Innsbruck* are chosen and their locations are marked with circles in Figure 2.1. This figure also pictures in a simplified way the orography, where brown colors indicate mountainous regions with higher altitude and green colors represent lower altitude. Vienna is located in a basin (Vienna basin) at the easternmost extension of the Alps. The weather station *Vienna, Hohe Warte* is at an altitude of 198 m. Innsbruck is situated in a valley between high mountains (to the north, the *Nordkette* with the highest peak *Hafelekar 2334 m* and to the south, *Patscherkofel 2246 m* and *Serles 2718 m*). The weather station *Innsbruck, University* is 578 m above sea level. More details about these weather stations are given in Table 2.1, where in the last row the time of the beginning of the operations is stated. Both stations were moved or rebuilt once a few meters within the period 1971–2000.

Observation Data

The observation data is provided by the ZAMG, where in general measurements are taken every 10 minutes. A quality check by the ZAMG is carried out before the data is released. Since 1971 meter-reading has been done at 6 am, 1 pm and 6 pm and before 1971 at 6 am, 1 pm and 8 pm. All time specifications are given in Coordinated Universal Time (**UTC**), if it is not explicitly indicated. The calculations in the thesis are done with daily mean values, which are also provided by the ZAMG. The daily mean

Table 2.1: Detailed information on the two chosen weather stations.

Station	Latitude / Longitude	Altitude	since
Vienna, Hohe Warte	N 48.25000° / E 16.35778°	198 m	1872
	N 48.24861° / E 16.35639°	198 m	1991
Innsbruck, University	N 47.26056° / E 11.38500°	577 m	1917
	N 47.26000° / E 11.38417°	578 m	1986

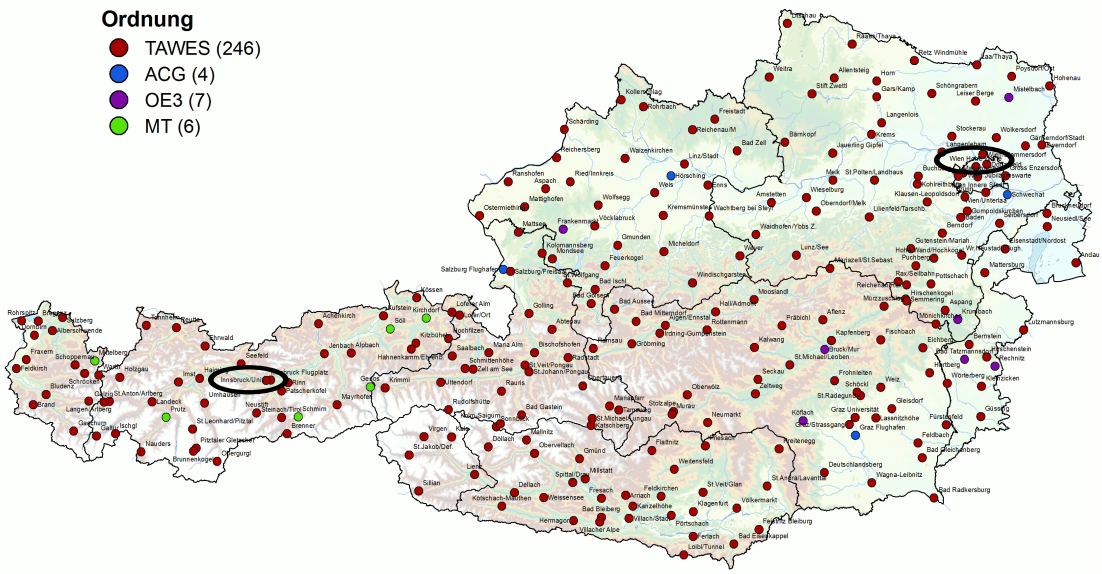


Figure 2.1: Locations of the weather and climate stations in Austria (ZAMG, 2010, accessed Jan., 2010).

values are calculated differently from parameter to parameter. A short description for the four parameters will be given here.

Daily mean temperature $\bar{T}$ (unit: °C) is the mean of the daily minimum temperature and the daily maximum temperature from 6 pm (day before) until 6 pm:

$$\bar{T} = \frac{T_{\min} + T_{\max}}{2} \tag{2.5}$$

Daily precipitation amount $\bar{P}$ (unit: mm) is added up from 6 am until 6 am of the next day.

$$\bar{P} = \sum_{i=1}^{n} P_i \tag{2.6}$$

Daily mean relative humidity $\bar{RH}$ (unit: %) is calculated as the mean of the three meter-readings:

$$\bar{RH} = \frac{RH_{7\text{am}} + RH_{2\text{pm}} + RH_{7\text{pm}}}{3} \tag{2.7}$$

The ZAMG provides daily mean global radiation in J/cm^2/d. This is the added up value from the hourly mean global radiation $\hat{GR}$ and is calculated from 0 am until 12 pm (CET). The SI-unit W/m^{2}will be used as this is also what the climate models provide. So, the daily mean global radiation is recalculated into W/m^2(J/cm^2/d→W/m^2).

$$\hat{GR} = \sum_{i=1}^{24} (\hat{GR}_i) \text{ in J/cm}^2\text{/d} \tag{2.8}$$

$$\bar{GR} = \hat{GR} \cdot 10000/24/3600 \text{ in W/m}^2 \tag{2.9}$$

Regional Climate Model

The output of the RCM REgional MOdel (**REMO**) in the version 5.7 from the Max Planck Institute for Meteorology (**MPI-OM**) is used in this thesis (Jacob, 2001). The simulation data is provided on a 0.22° grid (~25 km) over entire Europe. The hindcast simulation with ECMWF Re-Analysis (ERA40) (REMO57-ERA40, 1951–2000) (UPPALA et al., 2005) is used to test QM and Q-QM against observation data. REMO57-ERA40 should reproduce the past weather as closely as possible and therefore, being comparable with observation data. Furthermore, the Climate Change Signal (**CCS**) is analyzed with the simulation data from REMO driven by the GCM ECHAM5 (REMO57-A1B, 1951–2051) (Roeckner et al., 2003). It represents the climate (such as mean and variance), but it does not correlated to the weather (Themeßl and Gobiet, 2010). The ECHAM5 simulation

is based on A1B SRES (Nakicenovic and Swart, 2000). These simulations are also used in the ENSEMBLES project and the data are available online (`http://ensembles-eu.metoffice.com/`, ENSEMBLES, 2009).

The daily mean values for temperature, relative humidity and global radiation are calculated from 0 am till 12 pm (UTC). Daily precipitation amount is calculated from 6 am till 6 am (next day).

The DEC method will be applied on a point-scale with simulated data and observations. The model data used for the two point-scales are the mean of the four closest model grid points. This is done to provide smooth data and to get the effective resolution of RCMs into the data (Kapper, 2009). For example, *Vienna, Hohe Warte* is located at N 48.25000° / E 16.35778°. The data of the four next grid points in REMO are averaged for this geographical location (see Table 2.2 and Figure 2.2).

Table 2.2: Detailed information on the four closest grid points to the chosen weather stations.

Station	Latitude / Longitude	Altitude
Vienna, Hohe Warte	N 48.3192° / E 16.4374°	187.3100 m
	N 48.0993° / E 16.4444°	175.7740 m
	N 48.3141° / E 16.1070°	239.5830 m
	N 48.0941° / E 16.1154°	407.0470 m
Innsbruck, University	N 47.2432° / E 11.2861°	1402.8900 m
	N 47.2627° / E 11.6084°	1549.0100 m
	N 47.4623° / E 11.2565°	1334.7000 m
	N 47.0241° / E 11.3153°	2008.0400 m

2.2 Quantile Mapping Methods

An overview of the DEC method QM was already given in Section 1.6. For the application of QM, first a calibration and validation period have to be defined. In the calibration period, the correction function for a given parameter is generated by matching the Cumulative Distribution Function (**CDF**) of the simulated data with the CDF of the observed data depending on the quantile. This correction function is then used to correct the data in the validation period (Boé et al., 2007). Different ways are feasible to determine the correction function between simulated and observed data in the calibration period. Here, three possibilities of calculating the correction function will be mentioned (Ines and Hansen, 2006).

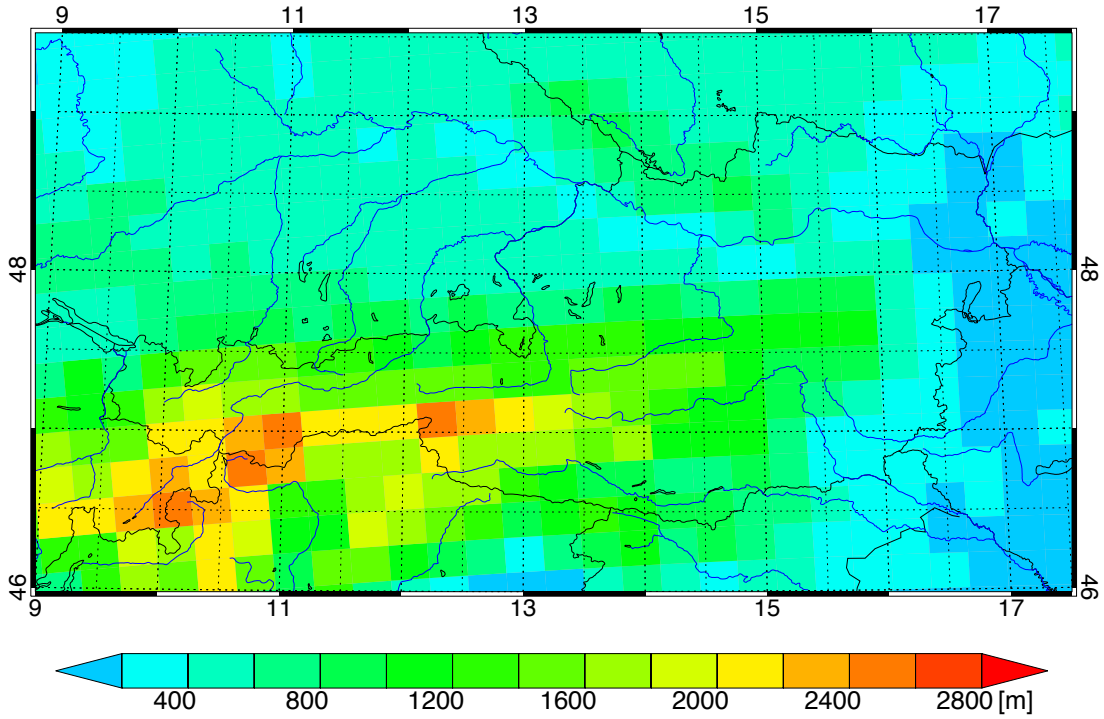


Figure 2.2: Orography of the REgional MOdel (REMO) v5.7.

1. The Empirical Cumulative Distribution Functions (**ECDFs**) of the simulated and observed data are used in the calibration period (see Subsection 2.2.1).

2. It is also possible to fit the empirical data to a theoretical distribution and then use the fitted CDFs in the correction process instead of the empirical data (see Subsection 2.2.2). But just as long as the meteorological parameter follows a specific theoretical distribution.

3. In some cases, it might be useful to take ECDFs for simulated data and fitted CDFs for observed data (Ines and Hansen, 2006).

In this study the focus is put on alternative 1. ECDFs are created with a 31-day moving window centered on the day which should be corrected for all years in the calibration period (e.g., March 15th ± 15 days from 1971–2000). This allows a more sensitive correction of annual cycles (Themeßl and Gobiet, 2010). Other possibilities are to use the whole data (e.g., Piani et al., 2010) or to split the data seasonally (e.g., Boé et al., 2007). This might lead to inconsistencies at the boundaries of each seasons. Hence, the moving window is preferred here.

2.2.1 Quantile Mapping Based on Empirical Data

Figure 2.3 illustrates the functionality of QM. Empirical CDFs are used in the correction process. Therefore, the QM method can generally be applied to any meteorological parameter (Themeßl and Gobiet, 2010).

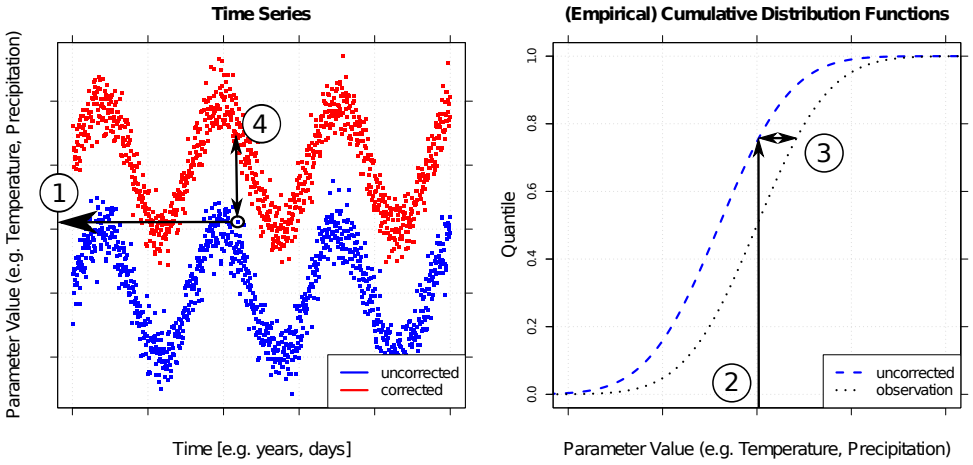
A given value at a time (x_t, e.g., March 15th) at a location (i, e.g., one grid cell center or latitude/longitude of a weather station, ...) of the simulated data is first taken out of the validation period. The next step is to find the quantile of this value in the calibration period ($ECDF_t^{\text{sim,cal}}(x_t^{\text{sim,val}})$). Finally, this quantile is used to calculate the value of the observed data (obs) ($x_t^{\text{obs,cal}}$). The difference between the simulated data value and this observed data value is then the correction value (Δx_t). Additionally, linear interpolations are applied between two quantiles. If some values are higher (or lower) than the highest (or lowest) values in the calibration period, the correction value of the highest (or lowest) quantile is taken (Boé et al., 2007).

$$x_t^{\text{obs,cal}} = ECDF_t^{\text{obs,cal}^{-1}}(ECDF_t^{\text{sim,cal}}(x_t^{\text{sim,val}})) \tag{2.10}$$

$$\Delta x_t = x_t^{\text{sim,val}} - x_t^{\text{obs,cal}} \tag{2.11}$$

2.2.2 Quantile Mapping Based on Fitted Distributions

Another possibility is to use QM with fitted distributions. But an advantage of the method in Subsection 2.2.1 is that it can be applied, in general, to any meteorological



(a) Validation Period (b) Calibration Period

Figure 2.3: Overview of the DEC method: Quantile Mapping (QM).

parameter. This is important since the distributions of only a few meteorological parameters follow a theoretical distribution. A reasonable fit can be found for daily mean temperature and daily precipitation amount. The normal distribution is considered for daily mean temperature. Many different theoretical distributions are proposed for daily precipitation amount, for example, an exponential distribution in Benestad (2007), a three parameter hyper-exponential distribution in Ines and Hansen (2006) or the S_B distribution in Swift and Schreuder (1981). But the most common used distribution to fit daily precipitation amounts is the gamma distribution (Vlček and Huth, 2009; Wilks, 2006; Piani et al., 2010; Dobler and Ahrens, 2008)

The slight difference of this procedure to the method in Subsection 2.2.1 is that at first the parameters of the theoretical distribution which suit the data best, have to be found. The correction function is then calculated with CDFs of these theoretical distributions.

Daily Mean Temperature

A normal distribution is assumed to fit the simulated and observed data of daily mean temperature. The equation for the PDF and CDF of the normal distribution is:

$$PDF_{\text{normal}}(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \tag{2.12}$$

$$CDF_{\text{normal}}(x) = \int_{-\infty}^{x} PDF_{\text{normal}}(x')\,\mathrm{d}x' = \frac{1}{2} + \frac{1}{2}\text{erf}\left(-\frac{(x-\mu)^2}{\sigma\sqrt{2}}\right) \tag{2.13}$$

The parameters μ and σ in (2.12) have to be identified. Different procedures are feasible for this. The Maximum Likelihood Estimation (**MLE**) method or just the "moments estimator" are able to calculate these parameters (Wilks, 2006, pp. 114–116). The solved parameters can be used within the CDF in eq. (2.13) to calculate the correction value (Δx_t).

$$x_t^{\text{obs,cal}} = CDF_t^{\text{obs,cal}^{-1}}(CDF_t^{\text{sim,cal}}(x_t^{\text{sim,val}})) \tag{2.14}$$

$$\Delta x_t = x_t^{\text{sim,val}} - x_t^{\text{obs,cal}} \tag{2.15}$$

Daily Precipitation Amount

For daily precipitation, the methodology differs compared to daily mean temperature. Daily precipitation fits in general a gamma distribution, but the gamma distribution is not defined at $x = 0$. A common way to arrange this is to do separated frequency and intensity correction (e.g., Ines and Hansen, 2006; Piani et al., 2010). The approach by Ines and Hansen (2006) is presented in the next two paragraphs.

Frequency Correction For the frequency correction, a threshold value $\tilde{x}^{\text{sim}}$ for the simulated data has to be calculated from the observed and simulated ECDFs. Generally, climate models tend to drizzle too much. Therefore, a minimum observed precipitation amount $\tilde{x}^{\text{obs}}$ needs to be defined (e.g., $\tilde{x}^{\text{obs}} = 0.1\ mm$), which considers a day to be wet. The threshold value $\tilde{x}^{\text{sim}}$ for the simulated data is calculated by the inverse ECDF of the simulated data ($CDF^{\text{sim,cal}^{-1}}$) from the ECDF of the observed data with the input value $\tilde{x}^{\text{obs}}$. In other words, the quantile of $\tilde{x}^{\text{obs}}$ in the observed data is first calculated and then the corresponding value $\tilde{x}^{\text{sim}}$ of the simulated data is determined as the threshold value.

$$\tilde{x}^{sim} = CDF^{\text{sim,cal}^{-1}}(CDF^{\text{obs,cal}^{-1}}(\tilde{x}^{obs})) \tag{2.16}$$

Simulated data in the validation period will be set to 0, if it is less than this threshold value.

$$x_t^{\text{sim,val}} = 0 \; \forall \; x_t^{\text{sim,val}} < \tilde{x}^{\text{sim}} \tag{2.17}$$

The simulated data, which is higher than or equal to the threshold value will be used in the intensity correction.

Intensity Correction At first, the parameters of the gamma distribution have to be estimated from the data, which is not used in the frequency correction. The PDF and CDF for gamma distributed data is defined as:

$$PDF_{\text{gamma}}(x) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)} x^{(\alpha-1)} \exp\left(-\frac{x}{\beta}\right) \tag{2.18}$$

$$CDF_{\text{gamma}}(x) = \int_{-\infty}^{x} PDF_{\text{gamma}}(x')\,\mathrm{d}x' \tag{2.19}$$

After having estimated the necessary parameters (α, β), the correction value for $x_t^{\text{sim,val}} \geq \tilde{x}^{\text{sim}}$ is:

$$x_t^{\text{obs,cal}} = CDF_t^{\text{obs,cal}^{-1}}(CDF_t^{\text{sim,cal}}(x_t^{\text{sim,val}})) \tag{2.20}$$

$$\Delta x_t = x_t^{\text{sim,val}} - x_t^{\text{obs,cal}} \tag{2.21}$$

Pros and Cons of Fitted Distributions

The pros of fitted distributions are primarily that the fitted data is smooth and well defined. Outliers or discontinuity do not exist in the fitted dataset. Therefore, the correction function has also a smooth form without extreme outliers. The cons of this method are mainly that not every climate parameter follows a specific theoretical distribution. A hypothesis test will be carried out in Subsection 2.3.2 for the used data in this thesis.

2.2.3 Quantile-Quantile Mapping

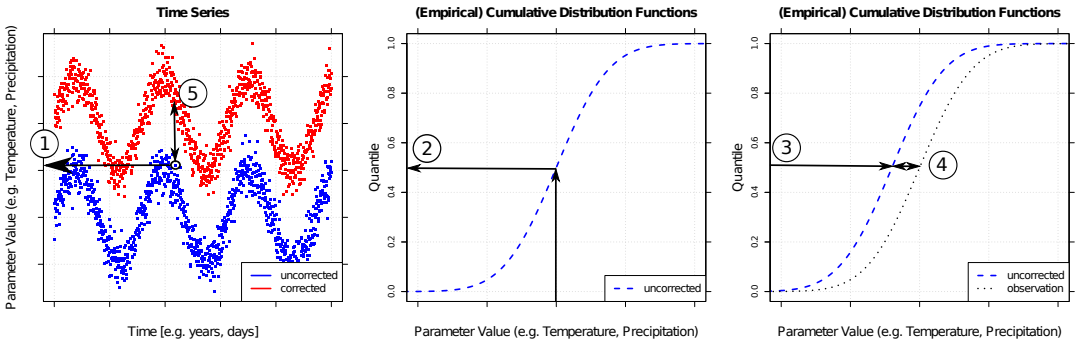
So far, a given value in the validation period was used to find the correction value in the calibration period. In this version the correction value at the same quantile as in the validation period is taken (see Figure 2.4). First the quantile at the given value in the validation period is calculated (e.g., March 15th, 2032 → March 15th ± 15 days from 2021–2050). The difference between the observed and simulated data at this quantile

in the calibration period is used as correction value:

$$x_t^{\text{sim,cal}} = ECDF_t^{\text{sim,cal}^{-1}}(ECDF_t^{\text{sim,val}}(x_t^{\text{sim,val}})) \quad (2.22)$$

$$x_t^{\text{obs,cal}} = ECDF_t^{\text{obs,cal}^{-1}}(ECDF_t^{\text{sim,val}}(x_t^{\text{sim,val}})) \quad (2.23)$$

$$\Delta x_t = x_t^{\text{sim,cal}} - x_t^{\text{obs,cal}} \quad (2.24)$$



(a) Validation Period (b) Validation Period (c) Calibration Period

Figure 2.4: Overview of the DEC method: Quantile-Quantile Mapping (Q-QM).

One assumption needs to be made for this additional version of QM:

- The model error depends on the shape of the distribution. A value which is in the validation period the median will be corrected like the median in the calibration period.

This method is not limited to empirical data, and can also be used with fitted distributions as already explained in Subsection 2.2.2.

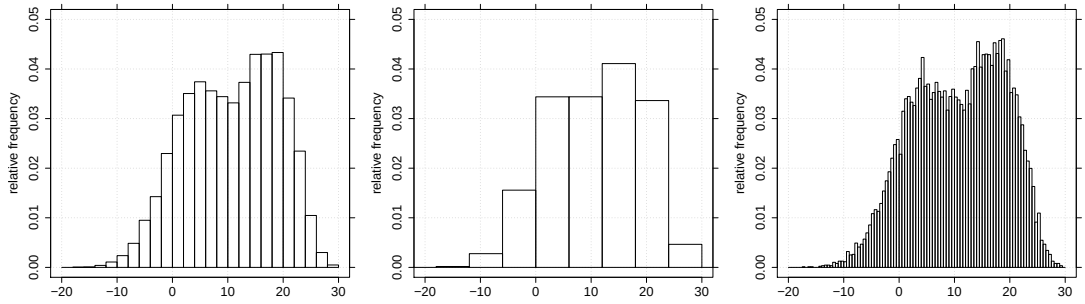
2.3 Data Analysis

2.3.1 Statistical Methods

Different kinds of statistical methods are used in this thesis to present the data. Most of them are basic knowledge like arithmetic mean, standard deviation and median, and will not be explained in more detail here. But the used types of graphics and the hypothesis tests will be presented in the following subsections. The primarily sources are Wilks (2006) and Sachs and Hedderich (2006). The statistical analyses are done with the programming language R (R Development Core Team, 2009). Therefore, the available documentation is also a possible reference for further information.

Histogram

Histograms are used in Subsection 2.3.2 to display the empirical data of the observed and simulated data. The data is split up in bins or class intervals. The height of a bin represents the added up number of values falling into this interval. Some characteristics of the data are quickly uncovered such as location, spread, symmetry, and if the dataset is multimodally distributed. It is important to choose a bin width, which allows the histogram to be easily read. In Figure 2.5, three histograms represent the same data, but with different bin widths. Meaningful results might be masked if the intervals are



(a) Bin width as suggested in Equation 2.25

(b) Three times the bin width as in Figure 2.5a

(c) One quarter of the bin width as in Figure 2.5a

Figure 2.5: Example of histograms showing some random data with relative frequency on the y-axis and different bin widths.

too wide as in Figure 2.5b. If the intervals are too narrow, the histogram might look inhomogeneous and is difficult to construe as in Figure 2.5c. A rule of thumb for the width of the bins h, is

$$h \cong \frac{c\,IQR}{n^{1/3}} \tag{2.25}$$

where c is a constant between 2.0 and 2.6, Interquartile Range (**IQR**) and n is the number of data values. Note that all histograms in Subsection 2.3.2 are comparable within each parameter. The bins have the same widths and are first estimated with Equation 2.25 and then adapted for each parameter (Wilks, 2006, pp. 33–35).

Probability Density Function Plot (PDF-plot)

A histogram can be seen as a nonparametric estimator of the underlying probability distribution. But the choice of the bin widths is arbitrary, and a further essential

requirement for constructing a histogram is that each data value is "rounded" to the midpoint of the bin. The examples in the Figures 2.5b and 2.5c have already shown how the characteristics of a histogram change with different bin widths. Wilks (2006, p. 85) explains PDFs as the "continuous, and theoretical, analogs of the familiar histogram and of the nonparametric kernel density estimate."

Characteristic shapes called kernels (e.g. Gaussian, Triangular, Quadratic, Quartic) are necessary to smooth the data. The Gaussian kernel is chosen for the PDF-plots. All data values contribute to the smoothing process (Wilks, 2006, pp. 35–39). Furthermore, two criteria have to be fulfilled for a PDF. The integral over all values must be equal 1 and all values of $f(x)$ have to be greater or equal 0 (Wilks, 2006, pp. 85–87).

$$\int_x f(x)\mathrm{d}x = 1 \tag{2.26}$$

$$f(x) \geq 0 \; \forall \; x \tag{2.27}$$

In Figure 2.6, PDF and histograms are drawn in one plot. It is not necessary to specify a bin width. Therefore, the PDF-plots in Figure 2.6 look the same, although the histograms differ.

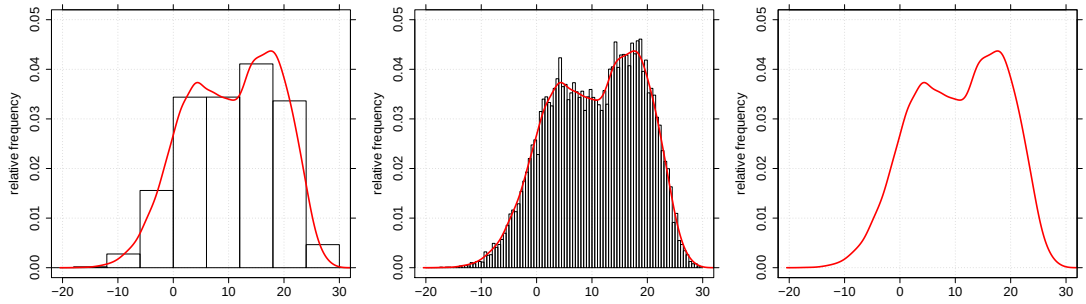


Figure 2.6: Example of PDF-plots and histograms.

Quantile-Quantile Plot (QQ-plot)

Quantile-Quantile (**QQ**)-plots are an easy possibility to present how well one or more datasets (e.g., uncorrected and corrected modeled data) fit a reference dataset (e.g., observed data). A QQ-plot shows where the main differences to the reference dataset occur. The plot just reproduces the sorted datasets as long as the compared datasets have the same data length, otherwise the two datasets are compared at the same quantiles. Figure 2.7 gives an example of a QQ-plot. The straight line shows the ideal case,

where the two datasets would be equal. On the x-axis the values for the "observed data" and on the y-axis the values for the "modeled data" are plotted. The red curve shows how the modeled data differs from the observation on each data value. In this example, the main differences appear at lower values (Wilks, 2006, pp. 113–114).

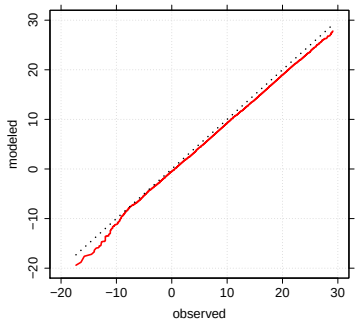


Figure 2.7: Example of a QQ-plot.

Empirical Cumulative Distribution Function Plot (ECDF-plot)

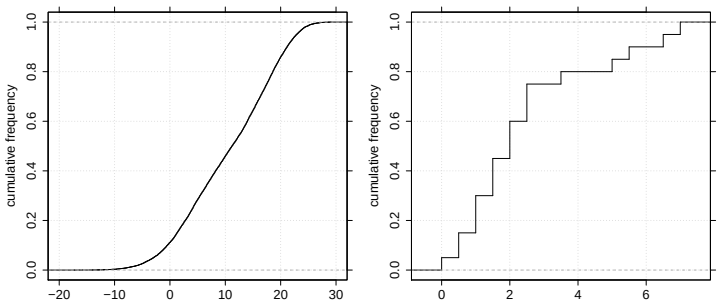
The ECDF-plot reproduces the data with cumulative probabilities on the y-axis and the data values on the x-axis. Figure 2.8a depicts the ECDF of the same data as in the examples above, but without dividing the data into bins. Hence, the steps after each value are not visible. Figure 2.8b shows the results for some random data with clearly visible jumps. To get a continuous CDF, the same procedure of the smoothing can be used as explained in Section 2.3.1. For a discrete CDF (or ECDF) the probability (Pr_{discrete}) is calculated by summation:

$$Pr_{\text{discrete}}\{X \leq x\} = \sum_{X \leq x} Pr\{X \leq x\} \tag{2.28}$$

The probability of a continuous CDF ($Pr_{\text{continuous}}$) is given by the integral:

$$Pr_{\text{continuous}}\{X \leq x\} = \int_{X \leq x} f(x)\mathrm{d}x \tag{2.29}$$

X represents a random quantity which does not exceed a specific value x (Wilks, 2006, pp. 39–42, 86–87).



(a) (b)

Figure 2.8: Example of two ECDF-plots. On the left, the same data as in the previous examples is used and on the right, random values are showing the steps in a clearer way.

Box-and-Whisker Plot (boxplot)

A Box-and-Whisker Plot (**boxplot**) is a graphical tool and was introduced by Tukey (1977). The distribution of an underlying data is delivered in a simple way. Five numbers reflect the data in the basic version of a boxplot: the minimum, the lower quartile $q_{0.25}$, the median $q_{0.5}$, the upper quartile $q_{0.75}$ and the maximum. Here, a more sophisticated form of the boxplot is used. Figure 2.9 shows a characteristic boxplot. The box in the middle indicates the lower quartile, median and upper quartile. The two horizontal lines represent the last data points which are within a distance of 1.5 times the IQR from each quartile ("fence"). Data values outside of this range can be symbolized with dots, while in this thesis only the minimum and maximum value is marked with a dot (Wilks, 2006, pp. 31–33).

$$\text{Upper fence} = q_{0.75} + 1.5 \cdot IQR \tag{2.30}$$

$$\text{Lower fence} = q_{0.25} - 1.5 \cdot IQR \tag{2.31}$$

Taylor Diagram

The Taylor diagram was introduced by Taylor (2001) and provides a precise statistical summary of how well patterns fit each other. Three statistical parameters are plotted in one diagram: the correlation, the Root-Mean-Square (**RMS**) difference, and the ratio of the variance (see Figure 2.10). The correlation coefficient R is usually applied in

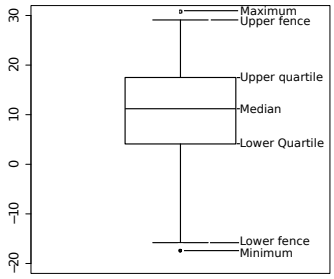


Figure 2.9: Example and overview of a boxplot.

statistics to verify pattern similarities. The centered RMS difference E' is commonly used to quantify differences in two fields.

$$R = \frac{\frac{1}{N}\sum_{n=1}^{N}(f_n - \bar{f})(r_n - \bar{r})}{\sigma_f \sigma_r} \tag{2.32}$$

$$E' = \left\{ \frac{1}{N} \sum_{n=1}^{N} [(f_n - \bar{f}) - (r_n - \bar{r})]^2 \right\}^{1/2} \tag{2.33}$$

f_n and r_n represent two variables with N discrete points. The mean values are shown with a bar over each variable ($\bar{f}$, $\bar{r}$), σ_f and σ_r are the corresponding standard deviations of f and r. Taylor (2001) used the relationship of the four statistical parameters (R, E', σ_f, σ_r) and the law of cosine,

$$E'^2 = \sigma_f^2 + \sigma_r^2 - 2\sigma_f \sigma_r R \tag{2.34}$$

$$c^2 = a^2 + b^2 - 2ab\cos\phi \tag{2.35}$$

where a, b and c are the lengths of the triangle and ϕ is the angle between a and b. Finally, each variable has to be normalized to get comparable results, even when the units of measure are different ($\hat{E}' = E'/\sigma_r$,$\hat{\sigma}_f = \sigma_f/\sigma_r$, $\hat{\sigma}_r = 1$). In Figure 2.10, the black dot symbolizes the reference data (f_n) (compared with the reference data (f_n)) and the blue square the test data (f_r) (compared with the reference data (f_n)). In this example, the correlation between the test data and the reference data is about 0.94, the centered RMS difference is roughly 0.3, and the normalized standard deviation is close to 1.

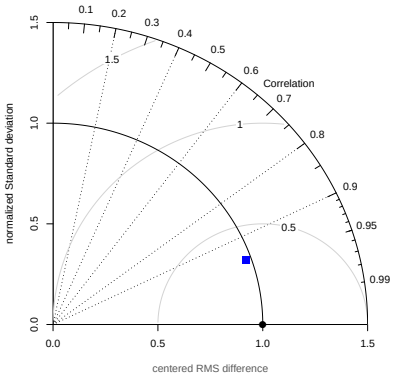


Figure 2.10: Example of a Taylor diagram.

Hypothesis Testing

Different hypothesis tests are used in this thesis. The Lilliefors test and the Cramér-von Mises test (**CvM-test**) are chosen as a parametric and nonparametric test, respectively. The test level was set to 5 %. The PDF- and QQ-plots already show qualitatively how the data is distributed and how well two distributions fit each other. But to get a quantitative information, a hypothesis test is required.

A common one-sample test (parametric test), where the distribution parameters are estimated from the test data, is the Lilliefors test which is based on the same test statistic D as the one-sample Kolmogorov Smirnov test (**KS-test**) (where the distribution parameters must not be estimated from the test data):

$$D_n = \max_x |F_n(x) - F(x)| \tag{2.36}$$

The test statistic D_n represents the largest difference between the empirical $F_n(x)$ and the theoretical (fitted) cumulative distribution function $F(x)$ (Wilks, 2006, p. 148). The null hypothesis (H_0) where the empirical distribution is drawn from the same population, is rejected, if the test statistic is larger than the critical value D_α (available in Crutcher, 1975).

The CvM-test is used as the two-sample test (nonparametric test) and as an alternative to the KS-test. Sachs and Hedderich (2006) mention that the CvM-test shows a very high "statistical power". The test statistic (C) is:

$$C = \frac{n_1 \cdot n_2}{(n_1 + n_2)^2} \sum_{i=1}^{n_1+n_2} D_i^2 \tag{2.37}$$

D is the difference between the two ECDFs and n_1 the number of values of the first dataset and n_2 of the second dataset. The test statistic is the sum of the squared differences between the ECDFs (Sachs and Hedderich, 2006, p. 408). The critical values for the CvM-test can be found in tables by Anderson and Darling (1952) and Csörgö and Faraway (1996).

Another possibility to present the information, whether rejecting the null hypothesis or not, is the p-value. Wilks (2006, p. 133) explains the p-value as

> "[...] the specific probability that the observed value of the test statistic, together with all other possible values of the test statistic that are at least as unfavorable to the null hypothesis, will occur (according to the null distribution). Thus, the null hypothesis is rejected if the p-value is less than or equal to the test level, and is not rejected otherwise."

In Table 2.3, the interpretation of the p-value by Sachs and Hedderich (2006, p. 324) is given.

Table 2.3: Interpretation of the p-value (Sachs and Hedderich, 2006, p. 324).

p-value	against H_0 speaks
>15 %	nothing
10 % to <15 %	hardly anything
5 % to <10 %	something
1 % to <5 %	much
<1 %	very much

2.3.2 Fitting of Empirical Distributions

The histograms of the four parameters – daily mean temperature, precipitation amount, relative humidity, and global radiation – are analyzed in this subsection. In Subsection 2.1.1, different theoretical distributions were presented which are now used to fit the empirical datasets. The observed and simulated data are given for the period of 1971 to 2000 and each month is analyzed separately. The results are important for the further work, while different approaches of QM are discussed in Section 2.2. It would be reasonable to use QM with fitting (see Subsection 2.2.2) as long as the empirical data can be approximated with a theoretical distribution. Otherwise, QM with empirical data, presented in Subsection 2.2.1, would be preferred.

The histograms are comparable within one parameter. The same bin width and length of the axis is used. For each month, histograms with fitted distributions are

produced, while in the following sections, just one month for each season is shown (Winter: January, Spring: April, Summer: July, Fall: October). Nevertheless, these months show the behaviors of the corresponding season. Additional information is printed at the bottom of the plot like the fitting parameters, the median, the minimum, the maximum, the used bin width and if the null hypothesis is rejected (TRUE) or not (FALSE).

Temperature Fitted with a Normal Distribution

For daily mean temperature, a normal distribution is commonly used to fit the data (e.g., Storch and Zwiers, 2002; Dettinger et al., 2004; Wilks, 2006). For example, Storch and Zwiers (2002, pp. 55–56) discovers that daily mean temperature can vary quite extensively from region to region and might feature more complicated distributions. The results of the observed (top row) and simulated data (bottom row) for *Vienna, Hohe Warte* are shown in Figure 2.11. The normal distributions seem to fit quite well the histograms of the observations (ZAMG). But the Lilliefors test points out that seven of the twelve months are not normally distributed. The null hypothesis ("The data is normally distributed.") is rejected (marked with "TRUE" in the bottom line of the plots). The simulation data of REMO57-ERA40 shows that the warmer months (June, July, August and September) are normally distributed, while the other months are not. One reason which is clearly visible in the plots is that REMO is overestimating the influence of the phase change at zero-degree. The same analyses are done with REMO57-A1B data and the results, which are not specifically shown here, are that not even one month is normally distributed.

Daily Precipitation Amount (> 0.1 mm) Fitted with a Gamma Distribution

The distribution of daily precipitation amount is usually not normally distributed, because for most days of a month no precipitation or just a little occurs. The distribution is strongly skewed to the right. It is possible to fit a normal distribution for longer time periods of precipitation, for example, seasonal precipitation (Wilks, 2006, pp. 88–89). The two-parameter gamma distribution is applied for daily precipitation as it is presented in many studies (e.g., Vlček and Huth, 2009; Wilks, 2006; Aksoy, 2000). It has to be mentioned that the gamma distribution is not defined at $x = 0$ (see Equation 2.2). Therefore, a threshold has to be defined which separates the frequency and intensity of daily precipitation amount. Here, the threshold is set at 0.1 mm. The Lilliefors test has been applied as proposed in Crutcher (1975). Vlček and Huth (2009) did already research on European station data (seasonal) with the result that especially in winter more than 40% of the stations are not gamma distributed. In this thesis the data is tested for each month separately on daily basis. The result of the Lilliefors test is that observed and simulated daily precipitation is not gamma distributed. Figure 2.12 shows

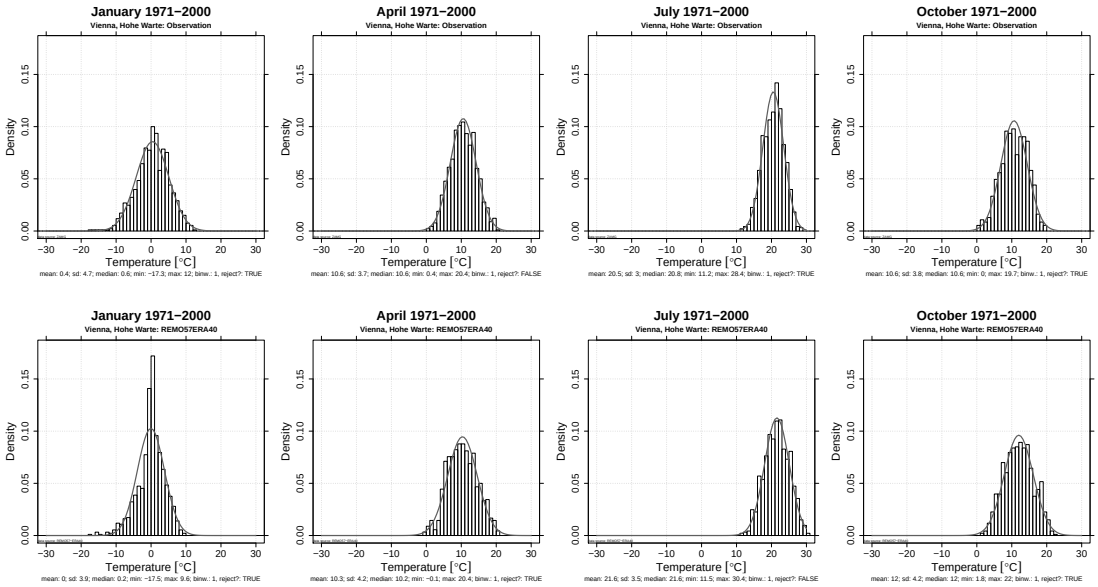


Figure 2.11: Histogram and fitted normal distribution plots of observed (top row) and REMO57-ERA40 simulated (bottom row) temperature [°C] for *Vienna, Hohe Warte* from 1971 to 2000. For each season one representative month is shown (for winter: January, spring: April, summer: July, fall: October).

qualitatively how well the fit matches with the histogram.

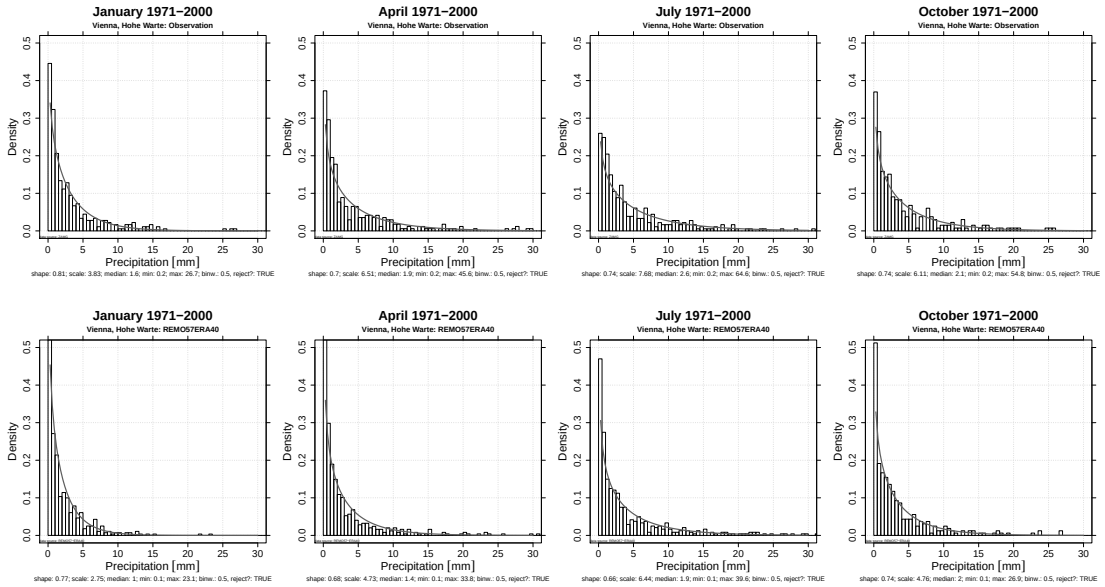


Figure 2.12: Same as Figure 2.11, but for *Vienna, Hohe Warte*, precipitation (>0.1) [mm/d] and a fitted gamma distribution.

Daily Relative Humidity Fitted with a Beta Distribution

Relative humidity is bounded on two sides (0 and 1 or 0 % and 100 %) as it is the case with some other meteorological parameters such as cloud amount. For this specific type, the beta distribution is a common choice (e.g., Yao, 1974; Wilks, 2006, pp. 102–104). No critical values of the Lilliefors test are available in Crutcher (1975) for the beta distribution. The common KS-test is carried out instead. It has to be kept in mind that the type II error (the null hypothesis is not rejected, when it is false) is higher for the KS-test, if the parameters are estimated from the data which should be tested (Vlček and Huth, 2009). The results for relative humidity are shown in Figure 2.13 for the selected months. The KS-test depicts that none of the months for the observed data are beta distributed and for REMO57-ERA40 output seven months are rejected, and eight months for REMO57-A1B respectively. Furthermore, a normal distribution and a "reflected" log-normal distribution were also fitted, with the result that both do not fit the data (these results are not shown in the figures).

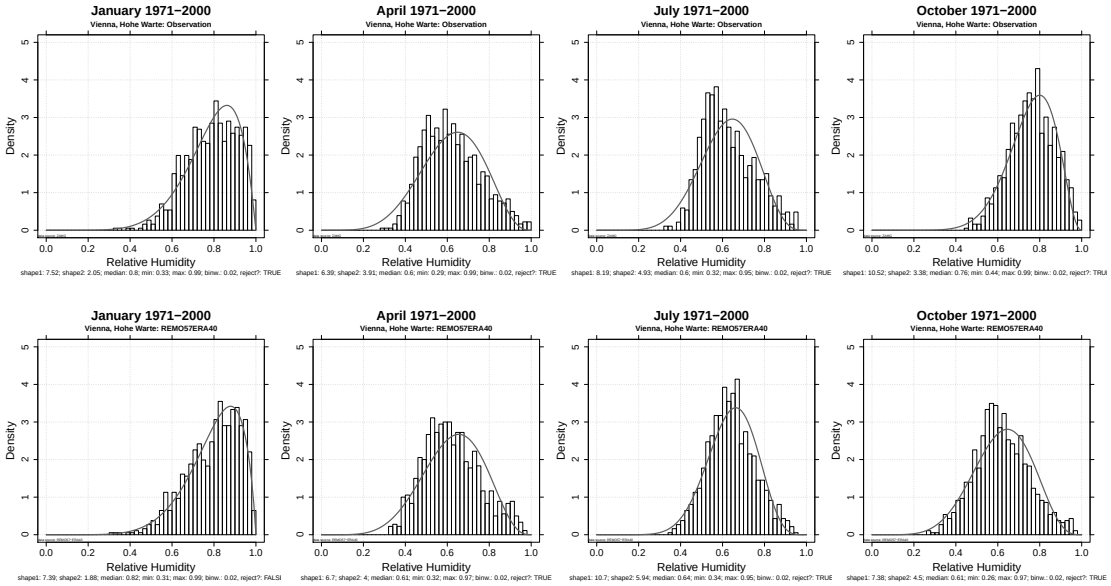


Figure 2.13: Same as Figure 2.11, but for *Vienna, Hohe Warte*, relative humidity [%] and a fitted beta distribution.

Daily Global Radiation Fitted with a Normal Distribution

The distribution of global radiation varies strongly throughout the year as it can be seen in Figure 2.14. There is no theoretical distribution known which could fit daily mean global radiation in a simple way. There are research studies available, where global radiation is estimated from other climate variables or through empirical formulae (e.g., Bindi and Miglietta, 1991; Donatelli et al., 2003; Matyasovszky and Bogardi, 1996; Trnka et al., 2005). Camuffo (1978) reveals for Venice, Italy in a six year period that the low and moderate intensities might be drawn from a log-normal distribution, and high intensities can be approximated with a log-normal or normal distribution. Such an approach can be useful, but is unfeasible for this thesis. The data should be taken as a whole and not divided into pieces. Daily global radiation is tested, if it is normally distributed, just to show that it is a more complex climatic parameter. The results in all tests show that the normal distribution does not fit daily global radiation.

Conclusion

In many studies fitted datasets for the climate parameters (e.g., daily precipitation amount) are applied in the post-processing (see the method in Subsection 2.2.2). But the results here of how well a theoretical distribution fits a specific climate parameter

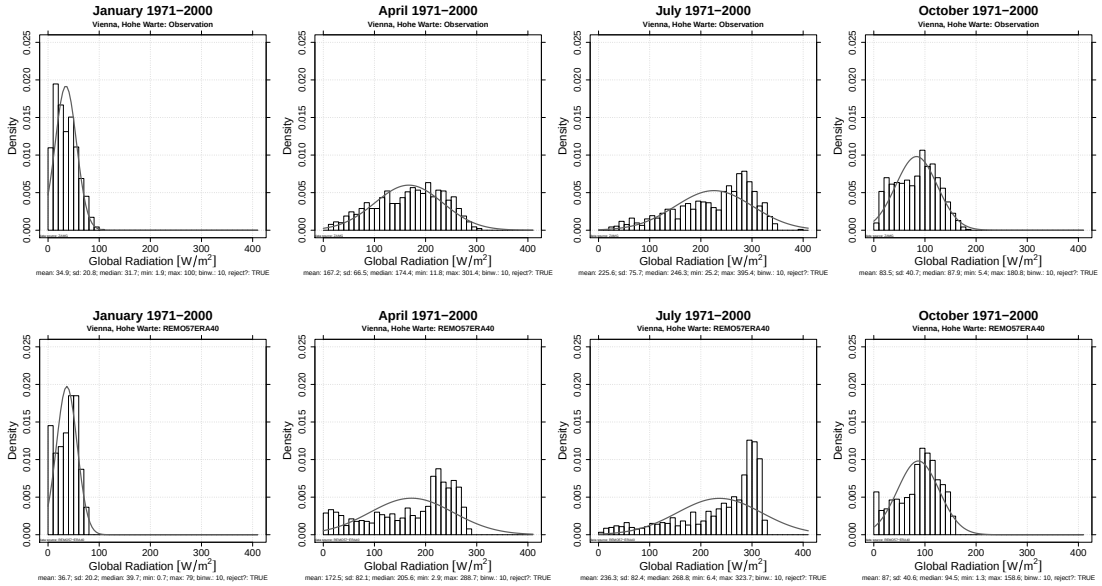


Figure 2.14: Same as Figure 2.11, but for *Vienna, Hohe Warte*, global radiation [W/m^2] and a fitted normal distribution.

have shown that in most cases the null hypothesis is rejected. Therefore, the DEC method will be applied with empirical datasets.

3 Results

This chapter is divided into a section, where the QM methods are first tested with synthetically generated data and a section, where the four climate parameters from the REMO model are post-processed with observational data of two specific locations mentioned in the data section.

3.1 Evaluation of the Quantile Mapping Methods with Synthetical Distributions

Synthetical distributions (compare Subsection 2.1.1) are now used to verify the performance of the QM methods. Beside normally distributed, gamma distributed and log-normally distributed, also bimodally distributed data is analyzed. The data is synthetically generated by (5000) random numbers with the programming language IDL. The terms calibration and validation period are also used for the following examples. PDF-, QQ- and boxplot display the findings. The (synthetical) reference data is always shown in black-dotted, the (synthetical) uncorrected data in blue-dashed and the (synthetical) corrected data in red-full. The blue boxplot stands for the (synthetical) uncorrected minus the (synthetical) reference data, and the red boxplot for the (synthetical) corrected minus the (synthetical) reference. Two different plots for the correction function are produced, one with the quantiles on the x-axis and the other one with the absolute value on the x-axis. The QQ-plots show the (synthetical) uncorrected and the (synthetical) corrected data compared with the reference data. The provided p-values are calculated with a two-sample CvM-test hypothesis test. The data which is corrected by the method Quantile Mapping will be called QM-corrected and the term Q-QM-corrected will be used for the data corrected with the method Quantile-Quantile Mapping, respectively. The aim of the two methods is to correct the synthetically introduced errors. The corrected data should be as close as possible to the reference data in the validation period. The corrected data is a result of the uncorrected data in the validation period plus the correction value which is calculated within the calibration period.

3.1.1 Normally Distributed Data

Normally Distributed Data

Both QM methods are tested with normally distributed data differing in mean and standard deviation. These parameters are listed in Table 3.1 for one specific case. The "seed" stands for the used random number sample within the programming language IDL. The "seed" of the reference data is the same in the calibration and validation period in order to compare the results easier. The "seed" of the uncorrected datasets differs between calibration and validation period to have some fluctuations in the data. In Figure 3.1 the PDF-plot and the boxplot display these datasets for the calibration

Table 3.1: Description of the chosen random normally distributed data.

Distribution	Period	Data	Seed	Mean	SD
Normal	Calibration	Reference	8	0	1
		Uncorrected	1	2	2
	Validation	Reference	8	0	1
		Uncorrected	30	2	2

period. The differences between the two distributions are around 2 in mean and 1 for the standard deviation. Different plots in Figure 3.2 reveal the results after applying

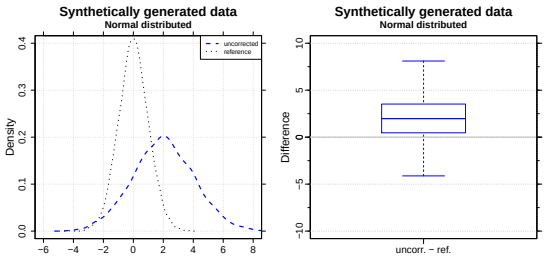


Figure 3.1: PDF-plot and boxplot for synthetically generated data in the calibration period. The uncorrected (blue-dashed) and reference (black-dotted) data are normally distributed and differ in the first (mean) and second parameter (standard deviation).

QM in the validation period. The QM-corrected data is close to the reference data in the PDF-plot, which is also visible through the p-value of 0.87 in Table 3.2. Mean and median are corrected to nearly 0, and the norm-sd is almost 1 (see Table 3.2). A few

outliers are only slightly corrected, since the correction function does not supply a high correction for the lowest values. These relicts are due to the chosen set of random numbers. The findings for Q-QM are shown in Figure 3.3 and Table 3.2 in the same

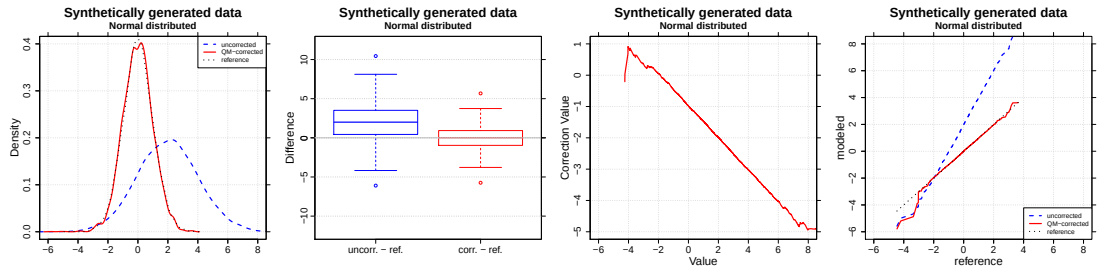


Figure 3.2: PDF-plot, boxplot, correction function, and QQ-plot for synthetically generated data in the validation period. The uncorrected (blue-dashed) and reference (black-dotted) data are normally distributed and differ in mean and standard deviation. The DEC method Quantile Mapping is applied. The line of the corrected data is displayed in red-full.

way as for QM. The results are similar, but small differences occur due to the different procedures of choosing the correction value. Q-QM has also corrected the mean and median roughly from 2 to 0, and the normalized standard deviation to close to 1. The p-value shows that the corrected data and the reference data are statistically drawn from the same distribution. The anomaly around the peak can be explained by the imperfect normally distributed datasets. The QQ-plot for the Q-QM-corrected data is at lower levels a bit smoother than the QM-corrected data, due to the fact that more values are positively corrected below -3 (compare y-axis in Figure 3.2 and 3.3). QM corrects all data values which are lower (higher) than the lowest (highest) uncorrected data value in the calibration period with this lowest correction value.

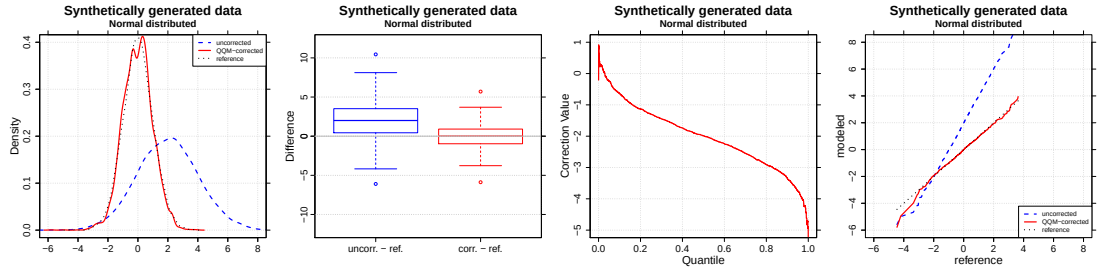


Figure 3.3: Same as Figure 3.2, but with Q-QM.

Table 3.2: Results for synthetically generated data: Normal distributions.

Comparison	Mean	Median	Normalized SD	p-value
Uncorrected – Reference	1.98	2.00	2.01	0.00
QM-corrected – Reference	−0.01	0.02	0.99	0.87
Q-QM-corrected – Reference	–0.02	0.01	0.98	0.23

3.1.2 Non-normally Distributed Data

Both QM methods are also applied on non-normally distributed data, which are described in more detail in Table 3.3. Some of these distributions are based on the results in Subsection 2.3.2. Parameters 1 and 2 are called scale and shape parameter for the gamma distribution, meanlog and meansd for the log-normal distribution. It is not specifically defined for the bimodal distribution, since the bimodal distribution is in this thesis a combination of two normally distributed datasets.

Table 3.3: Description of the chosen random non-normally distributed data.

Distribution	Period	Data	Seed	Parameter 1	Parameter 2
Gamma	Calibration	Reference	8	1	1
		Uncorrected	1	2	0.5
	Validation	Reference	8	1	1
		Uncorrected	30	2	0.5
Log-normal	Calibration	Reference	8	0	0.5
		Uncorrected	1	1	0.3
	Validation	Reference	8	0	0.5
		Uncorrected	30	1	0.3
Bimodal	Calibration	Reference	8	-	-
		Uncorrected	1	-	-
	Validation	Reference	8	-	-
		Uncorrected	30	-	-

Gamma Distributed Data

The results for gamma distributed data are presented in Figure 3.4 and Table 3.4. The differences between the corrected and reference data are almost 0 and the normalized standard deviation is close to 1 for QM and Q-QM. The QM-corrected data provides a p-value of 0.87 compared with the reference data while the Q-QM-corrected data and the reference data do not statistically come from the same distribution (p-value of 0.02).

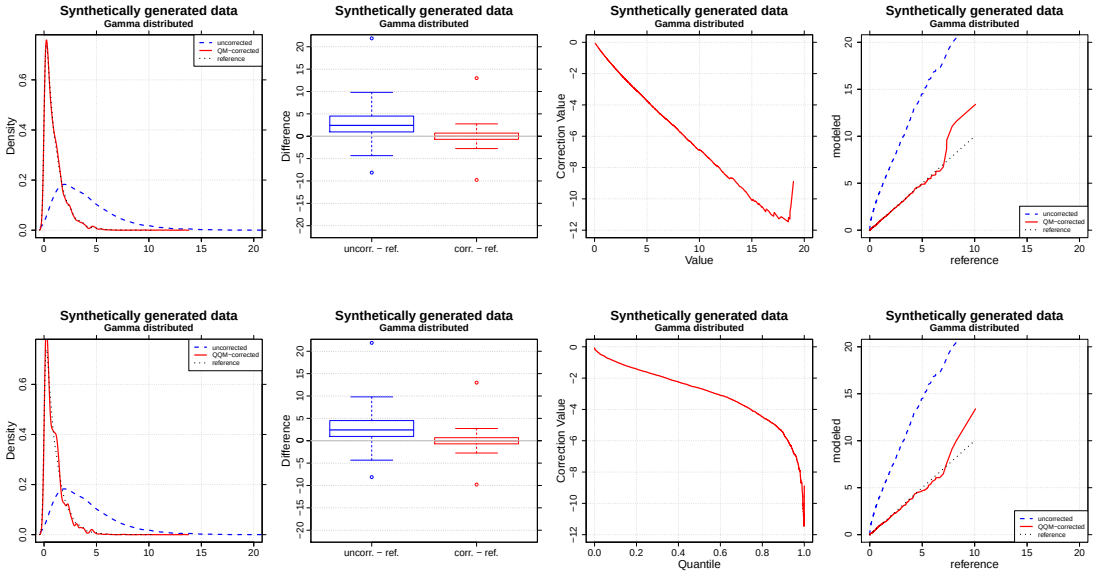


Figure 3.4: Same as Figure 3.2 for QM (top) and Q-QM (bottom), but for gamma distributed data.

Table 3.4: Results for synthetically generated data: Gamma distributions.

Comparison	Mean	Median	Normalized SD	p-value
Uncorrected – Reference	2.99	2.42	2.84	0.00
QM-corrected – Reference	-0.01	-0.01	1.00	0.87
Q-QM-corrected – Reference	-0.04	-0.02	0.96	0.02

Log-normally Distributed Data

Figure 3.5 and Table 3.5 show the findings for log-normally distributed data. Both QM methods have corrected the data closely to 0 and the normalized standard deviation is close to 1 after the correction process. The hypothesis test (corrected and reference data are equal) yields a p-value of 0.87 for the QM-corrected data and 0.39 for the Q-QM-corrected data.

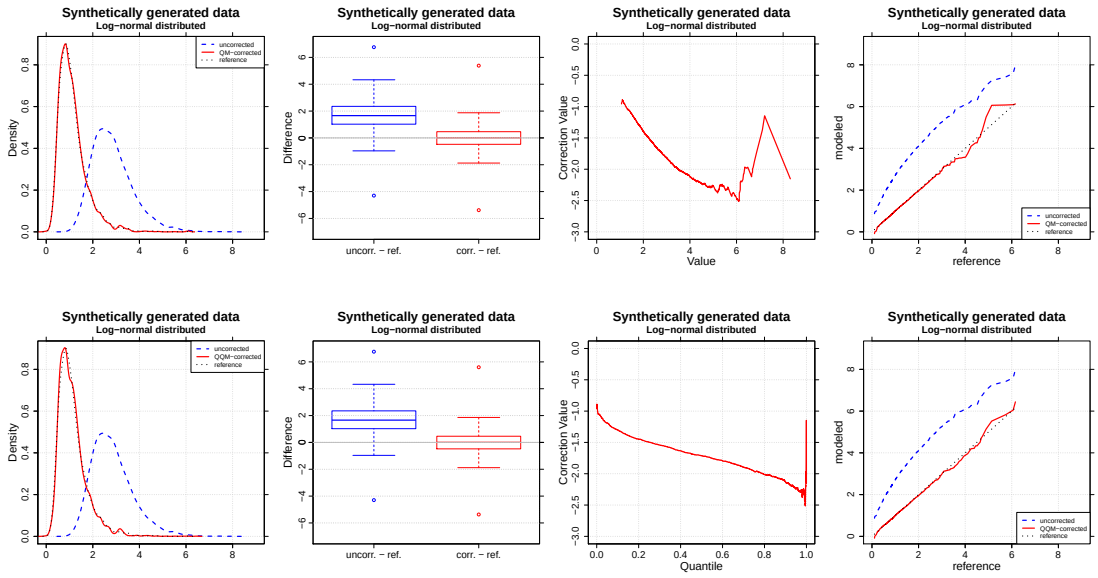


Figure 3.5: Same as Figure 3.2 for QM (top) and Q-QM (bottom), but for log-normally distributed data.

Table 3.5: Results for synthetically generated data: Log-normal distributions.

Comparison	Mean	Median	Normalized SD	p-value
Uncorrected – Reference	1.70	1.66	1.44	0.00
QM-corrected – Reference	-0.01	0.01	0.98	0.87
Q-QM-corrected – Reference	-0.01	0.01	0.98	0.39

Bimodally Distributed Data

The used REMO simulation overestimates the zero-degree peak for daily air temperature in winter due to the phase transformation (see Figure 2.11). A bimodally distributed dataset, which is a combination of two normal distributions, is used to verify how QM and Q-QM perform in this case. The synthetical simulation data is a bimodally distributed data and the reference data does not have this second peak (normally distributed). The results show that QM and Q-QM reduce the error (see Figure 3.6 and Table 3.6). The QM-corrected data reaches a p-value of 0.95 and the Q-QM-corrected data a p-value of 0.93.

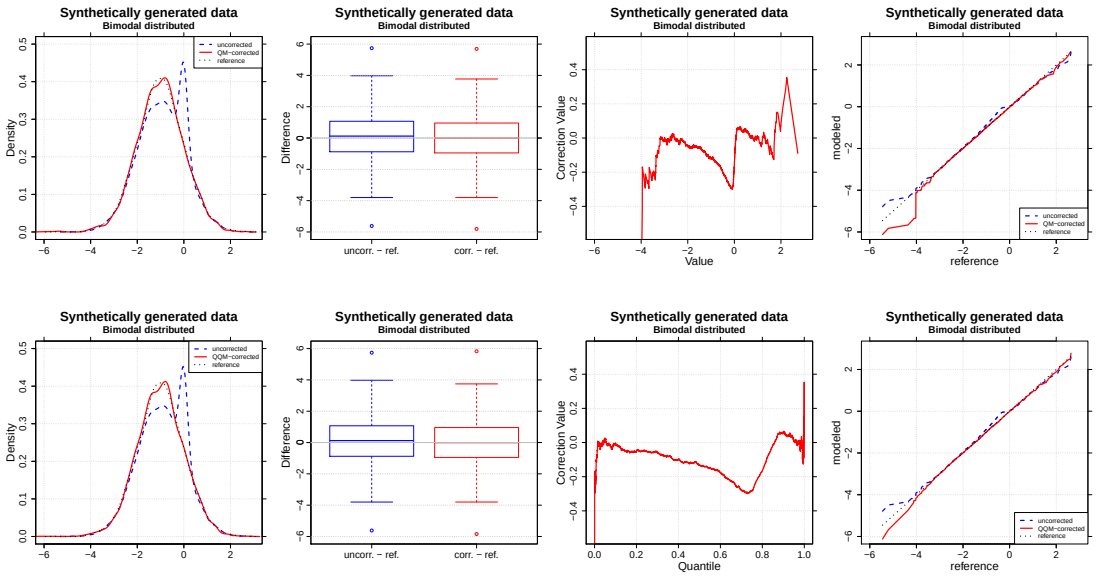


Figure 3.6: Same as Figure 3.2 for QM (top) and Q-QM (bottom), but for bimodally distributed data.

Table 3.6: Results for synthetically generated data: Bimodal distributions.

Comparison	Mean	Median	Normalized SD	p-value
Uncorrected – Reference	0.09	0.12	1.00	0.00
QM-corrected – Reference	-0.01	-0.01	1.00	0.95
Q-QM-corrected – Reference	-0.01	-0.01	0.99	0.93

3.1.3 Normally Distributed Data under Changed Conditions

Until now, the synthetical distributions have been tested under the same conditions. The same parameters 1 and 2 were taken for the calibration and validation period. But what happens, if the uncorrected and reference data in the validation period is shifted by a fixed value? Normally distributed data which differs in mean and sd is shifted by 4 between the calibration and validation period (see Table 3.7 and Figure 3.7). QM and Q-QM shall now be tested, how they can handle this "delta approached change". The

Table 3.7: Description of the chosen random normally distributed data.

Distribution	Period	Data	Seed	Mean	SD
Normal	Calibration	Reference	8	-2.5	5
		Uncorrected	1	-7.5	6
	Validation	Reference	8	1.5	5
		Uncorrected	30	-3.5	6

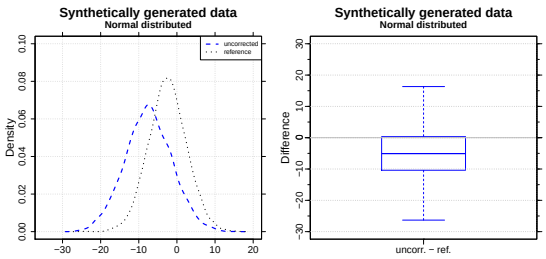


Figure 3.7: Same as Figure 3.1, but for normally distributed data under changed conditions.

results are given in Figure 3.8 and Table 3.8. The differences are quite large between QM and Q-QM. The difference in mean to the reference data is for the Q-QM-corrected data almost 0, and for the QM-corrected -0.77. Furthermore, the p-value indicates that the QM-corrected data is not of the form of the reference data, while the Q-QM-corrected data shows good agreement. The reasons for the differences in the outcome can be explained by the shifted distributions. Especially, the correction value for QM is between -10 and $+10$ too small. Another reason is the last correction value which is used for all uncorrected data values higher than the last one. This correction value is quite low as well. In contrast, the Q-QM-corrected data shows good agreement with

the reference data due to the method itself, which uses the quantiles in the validation period to get the correction value of the calibration period (see Subsection 2.2.3).

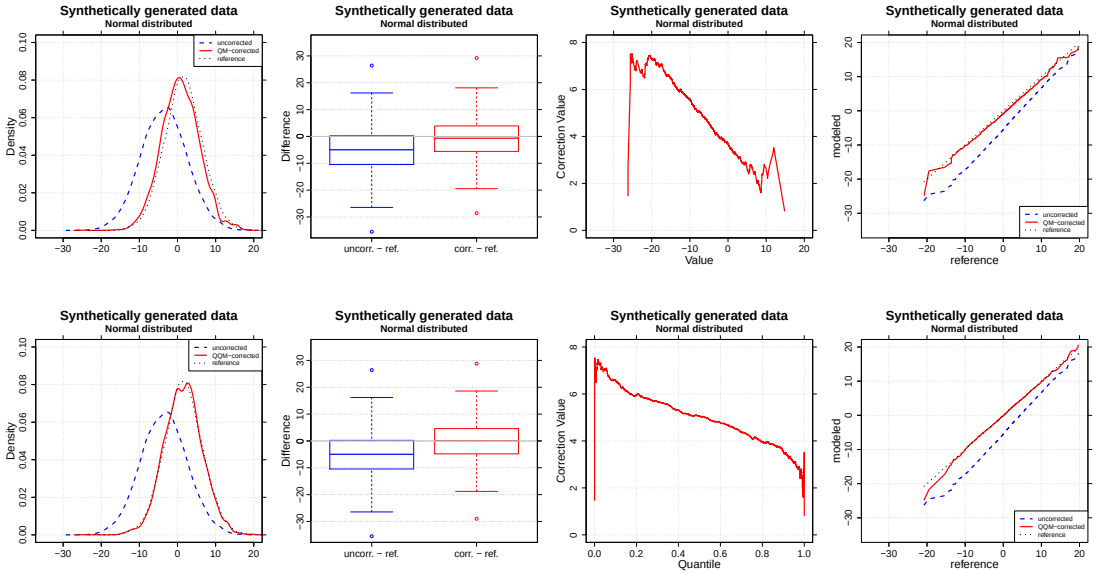


Figure 3.8: Same as Figure 3.2 for QM (top) and Q-QM (bottom), but for normally distributed data under changed conditions.

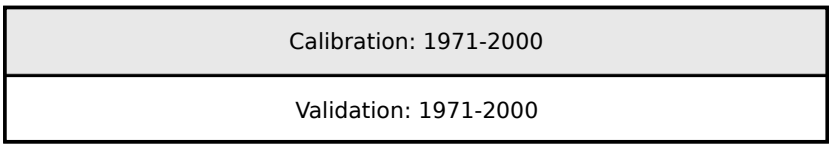
Table 3.8: Results for synthetically generated data: Normally distributed data under changed conditions.

Comparison	Mean	Median	Normalized SD	p-value
Uncorrected – Reference	-5.06	-4.97	1.21	0.00
QM-corrected – Reference	-0.77	-0.66	0.99	0.00
Q-QM-corrected – Reference	-0.06	0.11	0.99	0.68

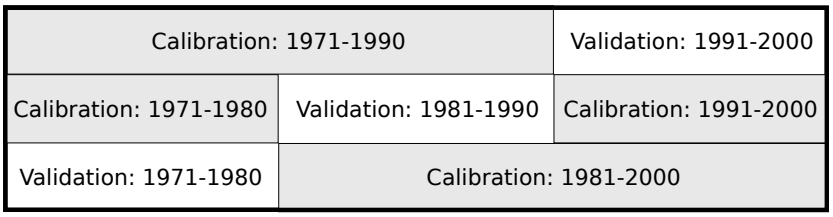
This simplified example shall highlight, when and where differences between QM and Q-QM appear. The good results for Q-QM are only valid due to no changes in variability. But climate change refers beside changes in mean, also to changes in the variability (IPCC, 2007, p. 667), therefore, the results for such cases might be different.

3.2 Post-Processing of Regional Climate Model Output

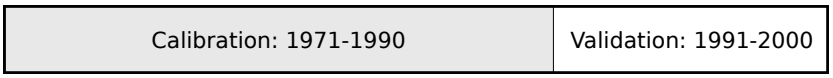
In the previous section, the two QM methods were tested on synthetically generated data. Now, QM and Q-QM are applied on REMO data related to point-scale ZAMG-observations. Three different calibration and validation period combinations (calibration (**cal**)/validation (**val**) approaches) are chosen to get comparable performance results for the two stations in Vienna (Hohe Warte) and Innsbruck (University) and the four climate parameters (daily: temperature, precipitation, relative humidity, global radiation). Matulla et al. (2003) use a similar procedure. A scheme is shown in Figure 3.9. In the first, the calibration and validation period is identical from 1971 to 2000 (see Figure 3.9a). A so-called "leave one out cross-validation" is used as the second cal/val approach (see Figure 3.9b). Within the thirty year period, twenty years are used for the calibration and ten years are corrected. The whole period is corrected after three cycles, without taking the same period for the validation as in the calibration. The third validation experiment is a split-sample test. The three decades which are tested in the cross-validation, are now validated independently with the attention to the third period (calibration from 1971 to 1990 and validation from 1991 to 2000, see Figure 3.9c).



(a) Validation period is identical with the calibration period.



(b) Leave one out cross-validation.



(c) Split-sample test.

Figure 3.9: Schematic graphic of the three different calibration-validation approaches. The gray fields indicate the calibration periods and the white fields the validation periods.

Different types of figures are used in this section for the evaluation. PDF- and QQ-

plots display the data in the same way as in Section 3.1; corrected: red-full, uncorrected: blue-dashed, observed: black-dotted. These figures provide the information of how the data is distributed (PDF-plot) and how well the corrected and uncorrected fit the observations (QQ-plot). Boxplots present statistical values of the differences between the corrected (in red) and uncorrected (in blue) data and the observed data. Taylor diagrams point out the correlation, the normalized standard deviation and the centered root-mean-square (RMS) difference between the corrected (red-triangle) and uncorrected (blue-square) data with the observations as reference. The mean error (bias) from the corrected and the uncorrected data to the observations and the p-value (reference: observations) are shown in tables. Furthermore, correction function plots reveal the error of the uncorrected data and therefore, how the correction values are selected. They are calculated as the differences between observed and uncorrected data on each quantile. For each decade and season a correction function plot is created. Note that the correction function plots are seasonally averaged while in the post-processing a specific correction function is calculated for each day with a moving window. Finally, a plot for the CCS represents the difference between the period 1991–2000 and 1971–1990 for each month.

In this section, the figures show the results for *Vienna, Hohe Warte* (with Quantile Mapping). The figures for *Innsbruck, University* (QM) are presented in the appendix (see Appendix A). The QM results of both stations are explained in this section, while the results for Q-QM are not explicated separately. The plots for the whole year are presented here, and the seasonal results for both QM methods are in the electronic appendix. The CCS plots are displayed for both stations in this section for QM but not for Q-QM. In this section, the term "uncorrected data" is used for the raw REMO simulation data and the term "corrected data" for the REMO57-ERA40 simulation data post-processed with QM.

3.2.1 Temperature

The results for daily mean temperature of the three different cal/val approaches are shown in Figure 3.10 and Figure 3.11 for *Vienna, Hohe Warte*. The uncorrected data and observations are not statistically drawn from the same distribution (p-value = 0.00), which is visible in the PDF- and QQ-plot (Figure 3.10, blue-dashed). The main differences occur at the very low absolute values and due to an overvaluation of the temperature at zero-degree Celsius of REMO. The bias of the uncorrected data is small for the year (Δmean = 0.24 °C), but the seasonal differences are between −0.88 °C and 1.10 °C (see Table 3.9). The correlation coefficient is close to 0.93 which indicates a high correlation to the observations (see Figure 3.11).

The corrected data and the observations come significantly from the same distribution in all three cal/val approaches. The p-value is for all seasons higher than 0.88 and the bias of the corrrected data is lower than 0.09 °C for the "calibration identical validation"-

case. The "Cross Validation"-case leads to p-values higher than 0.42 and a bias of lower than 0.11 °C in all seasons. The split-sample test results in p-values higher than 0.26 for all seasons except summer (p-value = 0.01). The bias between corrected and observed data is not higher than 0.34 °C in all seasons. The correlation coefficient has not changed visibly. The normalized standard deviations are closer to 1 after the correction.

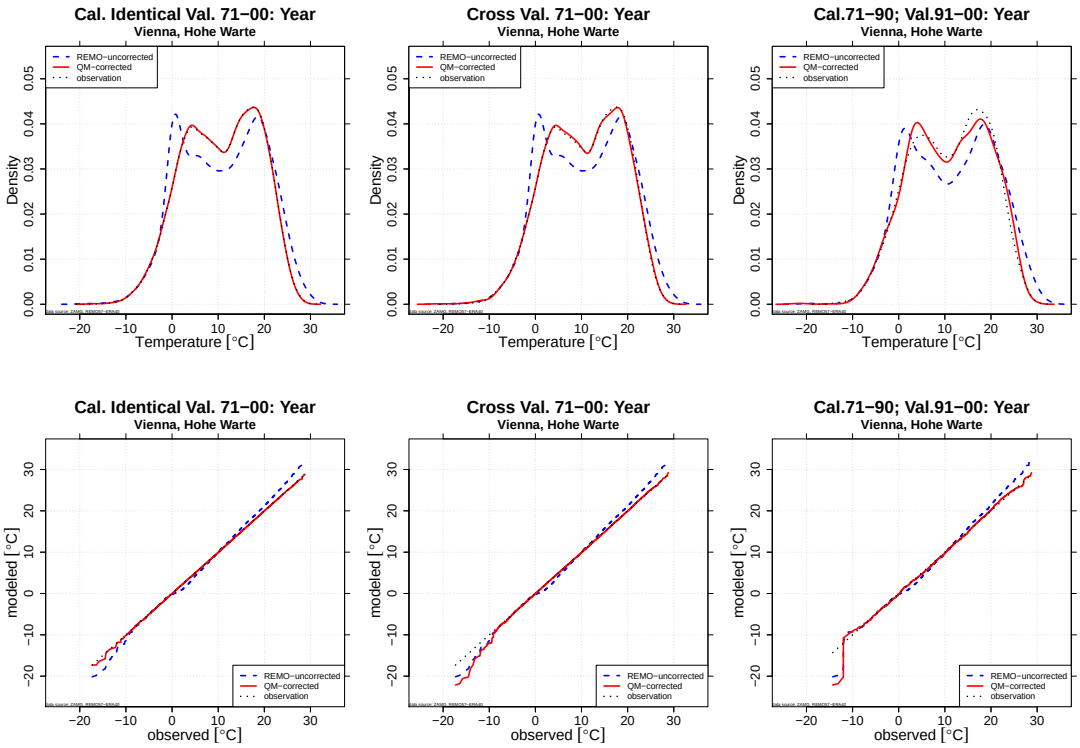


Figure 3.10: PDF- and QQ-plots of temperature [°C] for *Vienna, Hohe Warte* of the three different calibration/validation approaches. Blue-dashed line represents the uncorrected, red-full line the corrected and the black-dotted line the observed data. The DEC method QM is applied.

The results for the station *Innsbruck, University* are summarized in Table 3.10 and in Figure A.1. The bias of the uncorrected data is between −5.50 °C and −8.14 °C for the seasons. This is rather high due to the different altitudes of the ZAMG station (578 m) and REMO grid (between 1334 m and 2008 m). Therefore, the downscaling effect of the DEC method is larger. The biases of the corrected data for the "Calibration identical Validation" and "Cross Validation" cases are below 0.10 °C. The corrected data show a significant agreement to the observations. The bias in the split-sample test is below

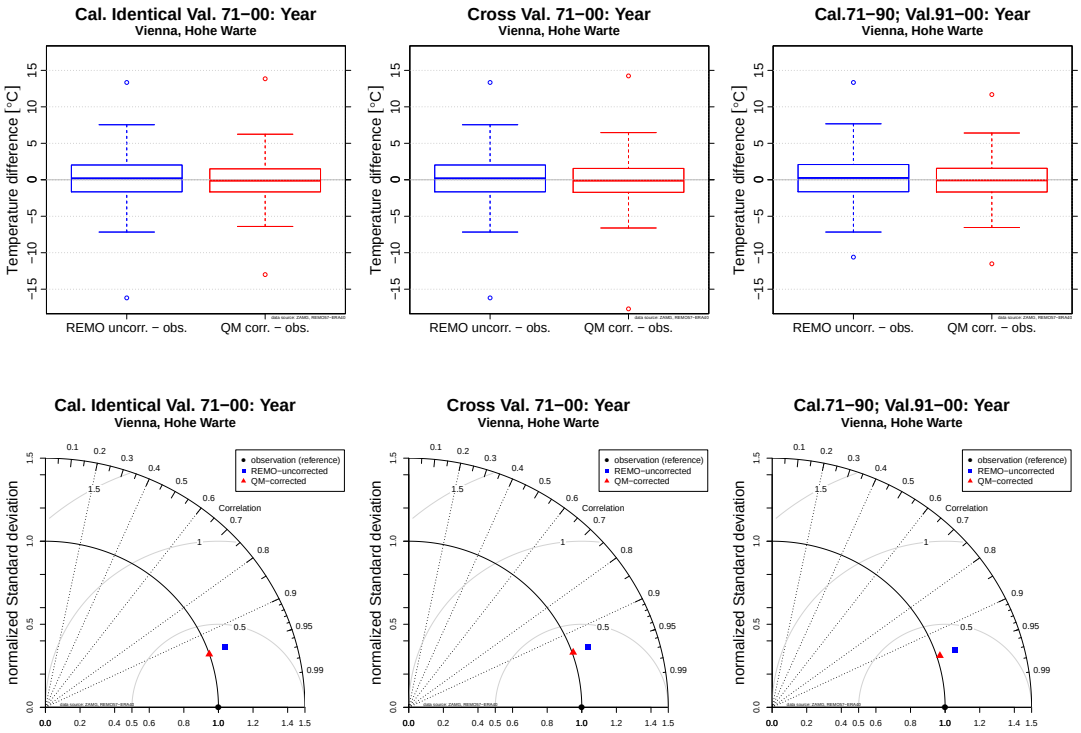


Figure 3.11: Boxplots and Taylor diagrams of temperature [°C] for *Vienna, Hohe Warte* of the three different calibration/validation approaches. Blue-dashed line represents the uncorrected, red-full line the corrected and the black-dotted line the observed data. The DEC method QM is applied.

Table 3.9: Mean error (ΔMean) and p-value (pVal) for each season and the whole year for temperature [°C] for *Vienna, Hohe Warte*. The uncorrected (a) and the three different cases for the corrected are compared with the observed data. (b) Calibration identical validation (1971–2000); (c) Cross-validation; (d) Split-sample test (calibration: 1971–1990, validation: 1991–2000) and in parenthesis the appropriated mean error for uncorrected-observed data for the period 1991–2000. The DEC method QM is applied.

	(a)		(b)		(c)		(d)	
Season	ΔMean	pVal	ΔMean	pVal	ΔMean	pVal	ΔMean	pVal
DJF	-0.88	0.00	-0.06	0.99	-0.10	0.99	-0.26 (-0.97)	0.29
MAM	-0.28	0.00	-0.04	0.99	-0.03	0.99	-0.10 (-0.32)	0.26
JJA	1.01	0.00	-0.08	0.88	-0.09	0.42	0.34 (1.47)	0.01
SON	1.10	0.00	-0.01	0.99	-0.03	0.93	-0.09 (1.10)	0.41
Year	0.24	0.00	-0.05	0.99	-0.06	0.98	-0.02 (0.32)	0.47

0.50 °C for all seasons and the whole year. The p-values are higher than 0.14 for all seasons.

Table 3.10: Same as Table 3.9, but for *Innsbruck, University*, temperature [°C], QM.

	(a)		(b)		(c)		(d)	
Season	ΔMean	pVal	ΔMean	pVal	ΔMean	pVal	ΔMean	pVal
DJF	-6.26	0.00	0.01	0.96	0.02	0.94	-0.32 (-6.43)	0.21
MAM	-8.14	0.00	-0.08	0.99	-0.09	0.96	-0.42 (-8.32)	0.15
JJA	-6.44	0.00	-0.06	0.98	-0.05	0.95	-0.11 (-6.51)	0.76
SON	-5.50	0.00	-0.08	0.99	-0.08	0.99	-0.59 (-5.86)	0.14
Year	-6.59	0.00	-0.05	0.99	-0.05	0.99	-0.36 (-6.78)	0.21

Correction Function

Figure 3.12 presents the correction functions for each decade, the whole period and each season for temperature in *Vienna, Hohe Warte*. They are calculated as the difference between observed and uncorrected data on each quantile. The first row shows the correction function for the period 1971–1980, the second row for 1981–1990, the third row for 1991–2000 and the fourth for the whole period 1971–2000 for each season (columns).

In general, the uncorrected simulation data seems to be too cold in winter (positive

correction values are more frequent) and the summer and fall months are too warm (negative correction values are more often). The correction functions look different between the decades. The first decade 1971–1980 shows a lower correction value for spring, summer and fall than the other decades. The decades from 1981–2000 are more similar to each other while the correction values around the median show the most similarities. The correction function for extreme values differ much more. The range is from −3.3 °C in summer to 1.6 °C in winter.

The resulting correction functions for *Innsbruck, University* are presented in Figure A.2. The error based on the different altitude between the observation and REMO grid is crucial. The correction function is slightly u-shaped with higher correction values for lower and higher quantiles in winter and fall. Spring and summer offers a more straight correction function, while the correction values for spring decrease the higher the values are.

Influence of Post-Processing on the Climate Change Signal

Figure 3.13 displays the CCS for *Vienna, Hohe Warte* (left panel) and for *Innsbruck, University* (right panel). The CCS is calculated as the difference between the periods 1991–2000 and 1971–1990 (reference period). These rather short periods are chosen to be able to compare the outcome with observations.

The observed CCS for temperature for *Vienna, Hohe Warte* varies between 0 and 1 with one exception. The signal is negative in December. The uncorrected CCS is similar to the observed one, but shows a much higher change in August. The CCS shows slight changes between the uncorrected and corrected data. Especially, in the summer months the corrected CCS is a bit lower than the uncorrected. This is due to a greater frequency of higher values in summer and therefore, a higher negative correction (compare Figure 3.12). The corrected CCS is for some months a bit closer to the observed CCS.

The observed CCS for *Innsbruck, University* is similar to the CCS of *Vienna, Hohe Warte*, but has slightly higher values. The lowest CCS of the uncorrected data is instead of December in November. The differences between the uncorrected and corrected CCS are even smaller as for *Vienna, Hohe Warte*.

3.2.2 Precipitation

Figures 3.14 and 3.15 display the results for daily precipitation. The uncorrected and observed data are not drawn from the same distribution (p-value = 0.00). This is also visible in the QQ-plot (blue dashed line).

Differences between uncorrected and observed data can be found mainly at low values (<3 mm/d) and high values (>3 mm/d), but with opposite signs (see Figure 3.14). At the low values an overestimation exists (area of the PDF), while at the high values an

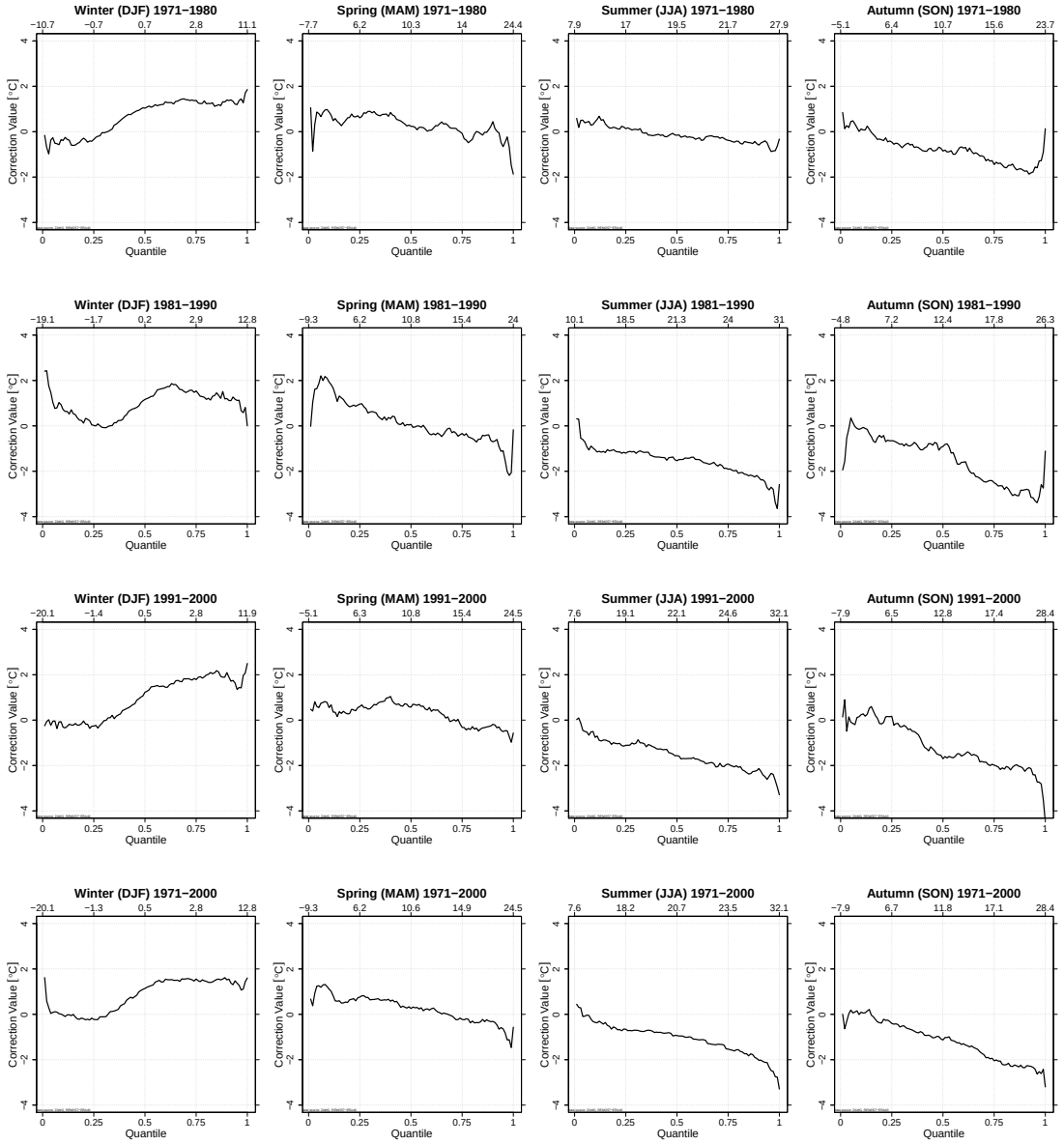


Figure 3.12: Correction function for *Vienna, Hohe Warte* for temperature [°C]. Correction functions based on the differences between observed and uncorrected data on each quantile. At the bottom x-axis the quantiles are displayed and at the top x-axis the absolute value for minimum, lower quartile, median, upper quartile and the maximum. Top row shows the correction function for the period 1971–1980, second row for 1981–1990, the third row for 1991–2000 and the bottom row for the whole period 1971–2000 for each season (colums).

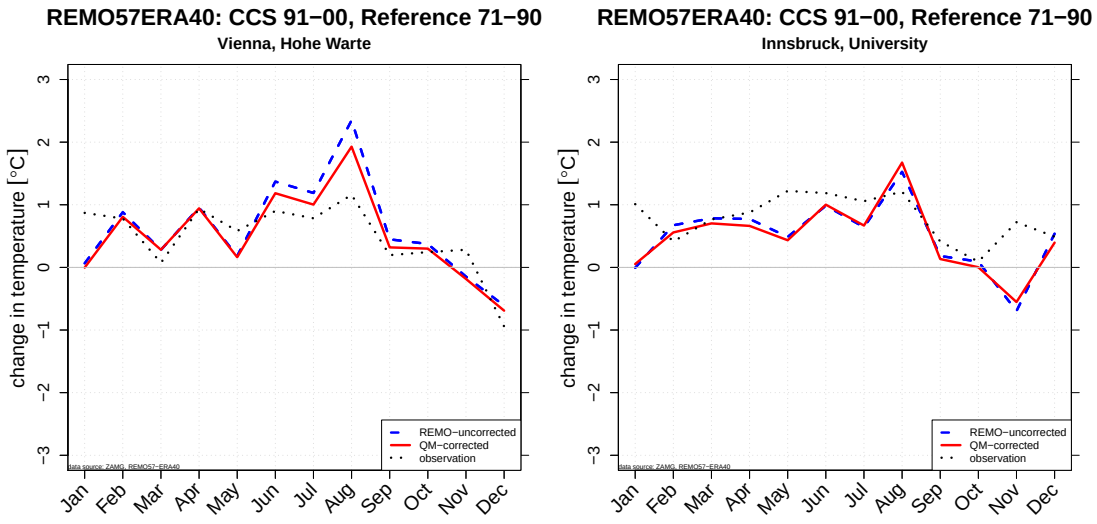


Figure 3.13: CCS for temperature [°C] for the period 1991–2000 compared with 1971–1990 (reference) for each month from REMO57-ERA40 and ZAMG-observations. CCS for *Vienna, Hohe Warte* is on the left and *Innsbruck, University* on the right.

underestimation appears. In general, the higher values of the uncorrected data are lower than the higher values of the observed data, but with one outlier in the uncorrected data (maximum). The bias of the uncorrected data fluctuates between −0.21 mm/d in fall (SON) and 0.15 mm/d in winter (DJF), and is over the year at −0.05 mm/d (see Table 3.11). A low correlation coefficient is given between uncorrected and observed data.

The corrected and observed data are significantly drawn from the same distribution in the first two cal/val approaches (see Table 3.11 (b) and (c)). The biases are below 0.06 mm/d for the "Calibration identical Validation" and below 0.10 mm/d for the "Cross Validation" in all seasons. The bias of the whole year is −0.02 mm/d and 0.05 mm/d, respectively. The bias has changed in the split-sample test from −0.26 mm/d to −0.33 mm/d. The p-value is over the 0.05 significance level for spring and summer. The normalized standard deviations are closer to 1 after the correction, but the correlation coefficients have not changed.

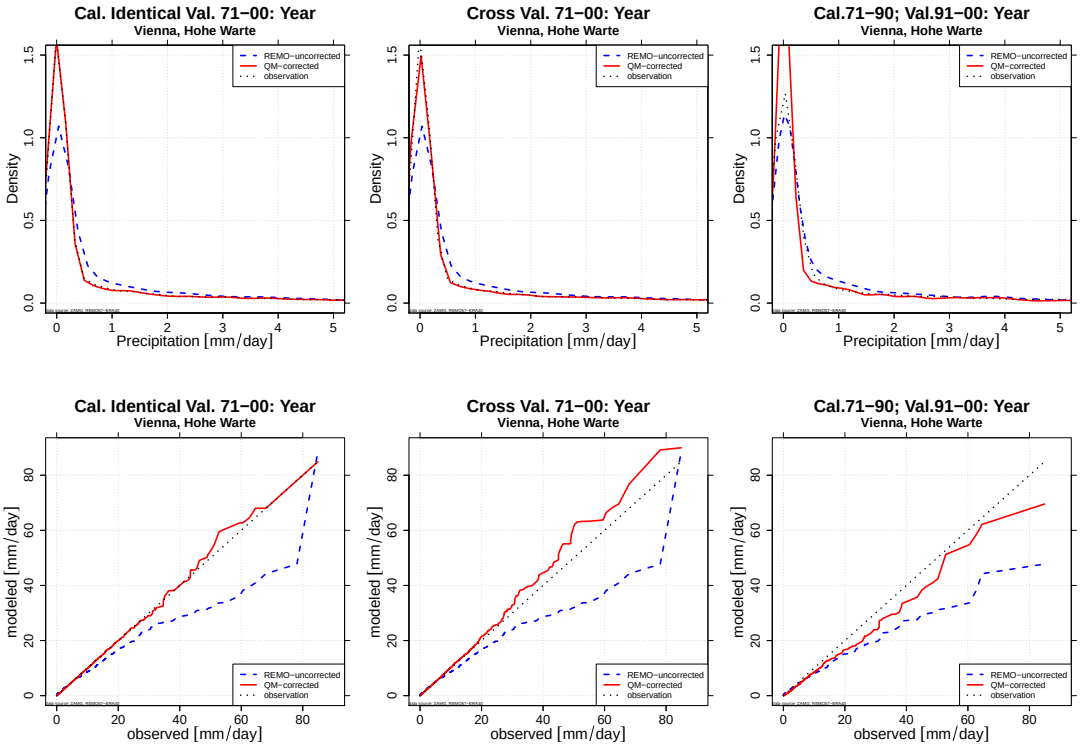


Figure 3.14: Same as Figure 3.10, but for *Vienna, Hohe Warte*, precipitation [mm/d], QM.

The results for *Innsbruck, University* are summarized in Table 3.12 and in Figure A.3.

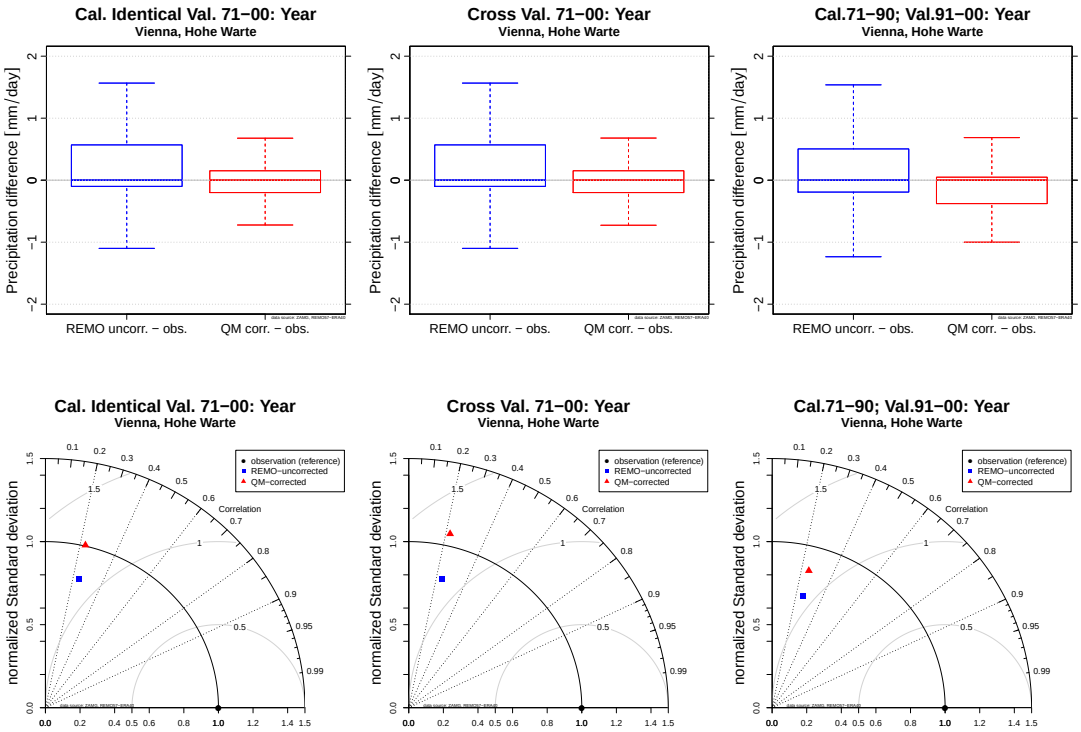


Figure 3.15: Same as Figure 3.11, but for *Vienna, Hohe Warte*, precipitation [mm/d], QM.

Table 3.11: Same as Table 3.9, but for *Vienna, Hohe Warte*, precipitation [mm/d], QM.

	(a)		(b)		(c)		(d)	
Season	ΔMean	pVal	ΔMean	pVal	ΔMean	pVal	ΔMean	pVal
DJF	0.15	0.00	-0.01	0.89	0.02	0.79	-0.15 (0.04)	0.00
MAM	-0.06	0.00	0.00	0.99	0.06	0.99	-0.57 (-0.43)	0.12
JJA	-0.07	0.00	-0.01	0.95	0.09	0.99	-0.46 (-0.39)	0.63
SON	-0.21	0.00	-0.05	0.76	0.01	0.87	-0.15 (-0.27)	0.00
Year	-0.05	0.00	-0.02	0.65	0.05	0.63	-0.33 (-0.26)	0.00

The bias of the uncorrected data is quite high for precipitation amount and lies between 0.89 mm/d and 2.03 mm/d for the seasons, and 1.67 mm/d for the whole year. The error characteristic is different to *Vienna, Hohe Warte*. At lower and higher values a general overestimation occurs. The bias is corrected in all three cal/val combinations to less than 0.04 mm/d for the whole year, and with more variations within the seasons in the split-sample test. The p-values are in all experiments (except for fall in the split-sample test) higher than the significance level of 0.05.

Table 3.12: Same as Table 3.9, but for *Innsbruck, University*, precipitation [mm/d], QM.

	(a)		(b)		(c)		(d)	
Season	ΔMean	pVal	ΔMean	pVal	ΔMean	pVal	ΔMean	pVal
DJF	1.86	0.00	0.04	1.00	0.06	0.99	0.20 (1.85)	0.82
MAM	1.90	0.00	-0.08	0.99	-0.05	0.99	-0.22 (1.70)	0.18
JJA	2.03	0.00	0.00	0.99	0.12	0.99	-0.28 (1.70)	0.08
SON	0.89	0.00	-0.07	0.85	-0.01	0.99	0.25 (1.24)	0.01
Year	1.67	0.00	-0.03	0.85	0.03	0.93	-0.01 (1.62)	0.75

Correction Function

The correction function plots for daily precipitation amount are shown in Figure 3.16 for *Vienna, Hohe Warte*. All plots show similar results with one exception. REMO overestimates lower values which leads to a negative correction until roughly 3 mm (or 0.8 Quantile). The model underestimates higher precipitation events, and therefore, the correction function is positive at higher quantiles. The exception is the summer of the period 1971–1980 (hence the whole period 1971–2000 as well), where the model has produced an extreme precipitation event of 88.4 mm/d. The highest observed precipitation amount in this period is around 10 mm less. Especially, these large differences at the highest quantiles are "problematic" for QM. Future simulated precipitation, which is higher than the highest simulated one in the calibration period, will be corrected with the correction value at the highest quantile. Another obvious problem is that the correction function shows rapid changes at the highest quantiles in some seasons (like in winter 1981–1990).

The correction functions for *Innsbruck, University* are provided in Figure A.4. The correction values are mainly negative due to the general overestimation of precipitation (see Table 3.12). Only at higher quantiles in the seasons spring, summer and fall positive correction values are produced due to the underestimated high precipitation events.

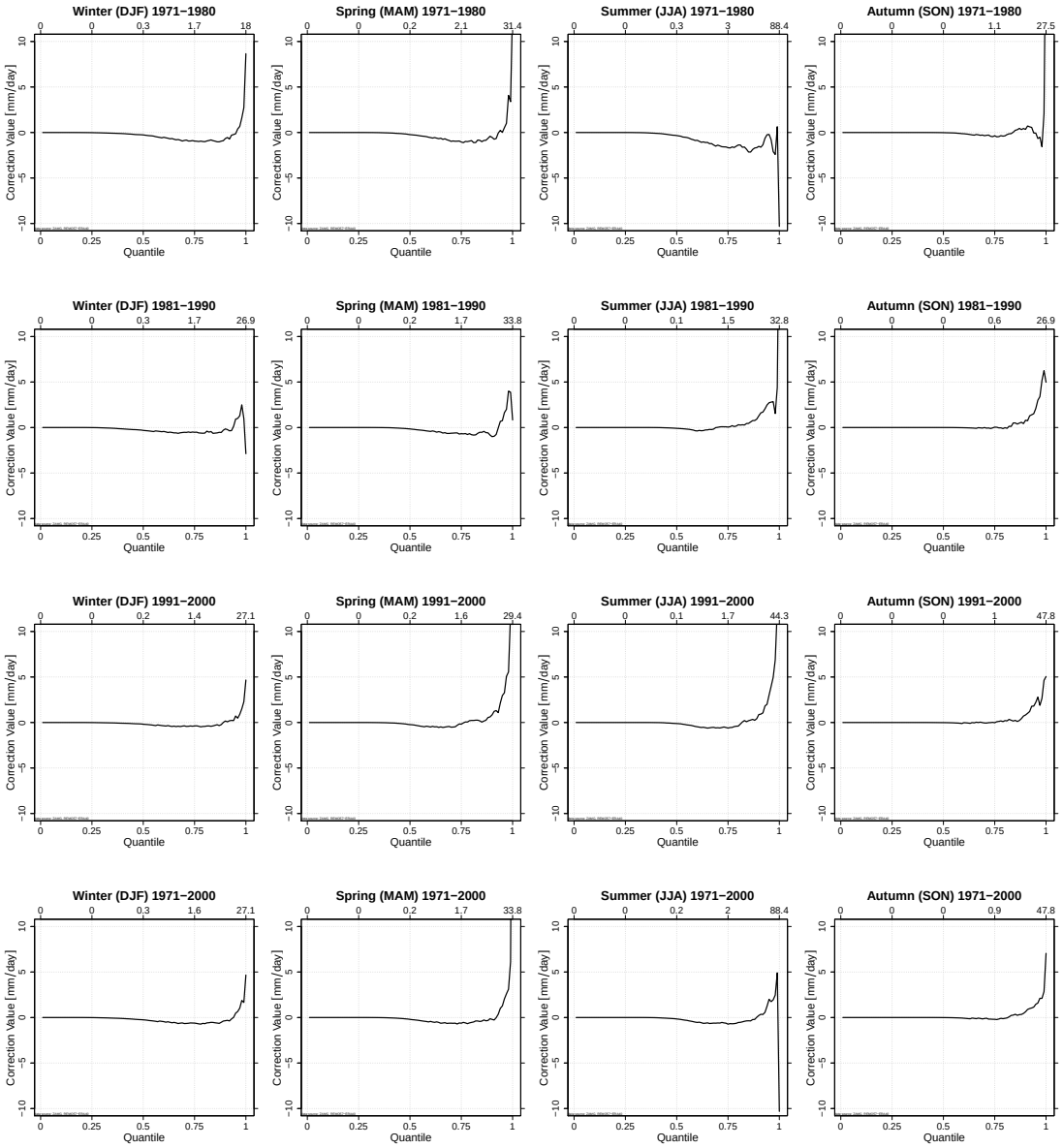


Figure 3.16: Same as Figure 3.12, but for *Vienna, Hohe Warte*, precipitation [mm/d].

Influence of Post-Processing on the Climate Change Signal

The observed CCS for monthly precipitation shows for *Vienna, Hohe Warte* large fluctuations with the highest amplitudes of about −15 mm/month in June and roughly 28 mm/month in July (see Figure 3.17). The CCS of the uncorrected data also displays large fluctuations, but generally a lower signal.

The CCS for precipitation shows some moderate changes after the correction too. These differences between uncorrected and corrected CCS are more difficult to interpret due to the complex form of the correction function given in Figure 3.16. The differences occur, when more values in the (scenario) period (1991–2000) are within the lower quantiles, where negative corrections are dominant, or in the higher quantiles, where the values are mostly positively corrected. For example, the positive change of the CCS for July is due to the greater frequency of higher values in the period 1991–2000, but still not too high that the correction might become negative, as it is at approximately 88.4 mm/d. Nonetheless, the corrected CCS does not get closer to the observed CCS.

The range of the observed CCS for *Innsbruck, University* is similar to *Vienna, Hohe Warte*. In August, the CCS is −25 mm/month, while it is 32 mm/month in November. The differences between uncorrected and corrected CCS are larger for *Innsbruck, University* than for *Vienna, Hohe Warte*. In the seasons winter, spring and summer the corrected CCS are higher than the uncorrected and lower in fall.

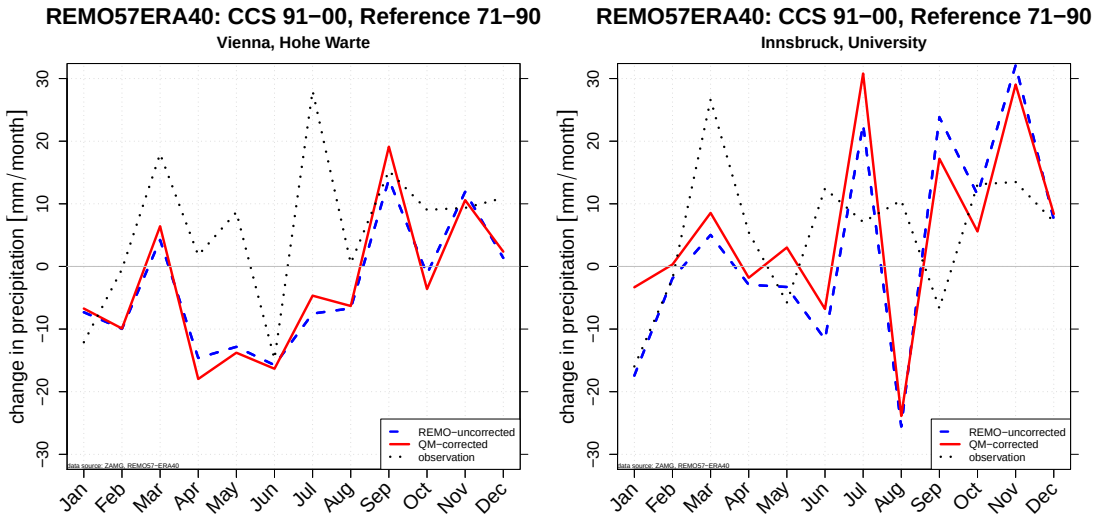


Figure 3.17: Same as Figure 3.13, but for precipitation [mm/month], QM.

3.2.3 Relative Humidity

The results for the parameter relative humidity (*Vienna, Hohe Warte*) are presented in Figure 3.18, Figure 3.19 and in Table 3.13.

The uncorrected and observed data do not come statistically from the same distribution. The bias of the uncorrected data is −2.12 % for the year. The largest differences occur in fall with −11.52 %, and is less than 1.14 % for the other seasons. The correlation coefficient is close to 0.4 for the whole year.

The bias after the correction for the first two cal/val approaches is less than 0.50 % (year), and is also corrected closely to 0.50 % in fall. The p-values for the first two cases are all higher than the significance level of 0.05 with two exception (summer and year in the "Cross-Validation"). The split-sample test shows a bias of −4.37 % before and −3.91 % after the correction for the whole year. For the seasons winter, spring and summer the bias is before the correction less than afterwards (see Table 3.13), but the error for fall has been corrected from −12.77 % to −2.61 %. The p-values do not show any agreement in the distributions of the corrected and the observed data.

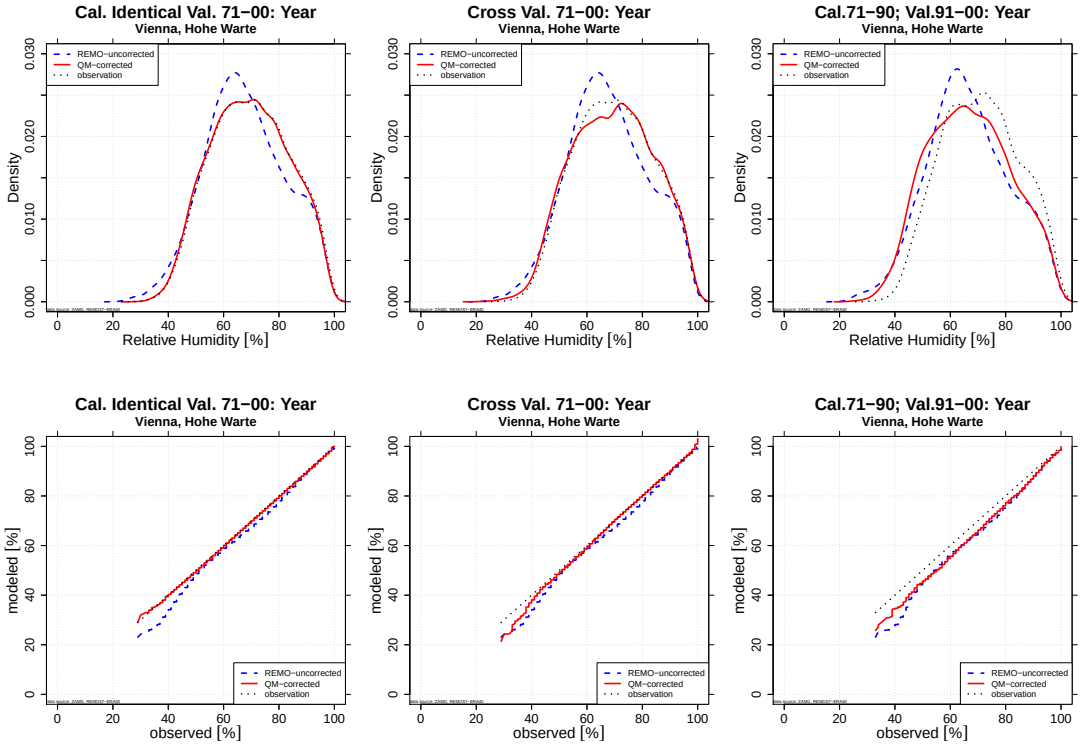


Figure 3.18: Same as Figure 3.10, but for *Vienna, Hohe Warte*, relative humidity [%], QM.

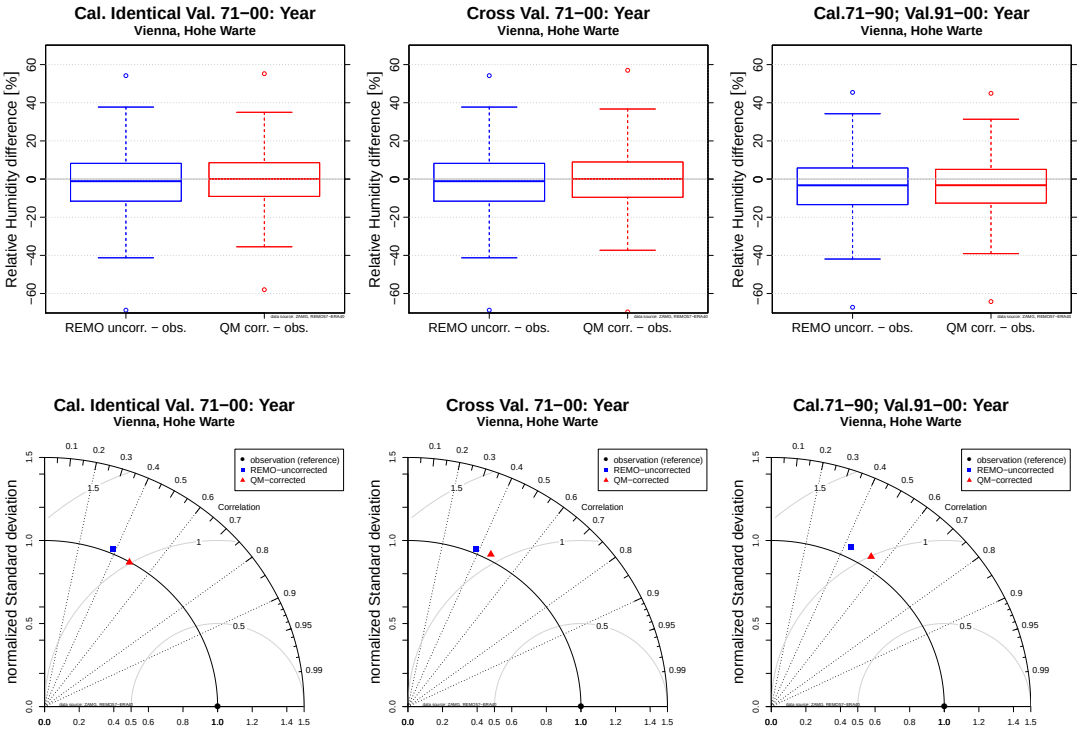


Figure 3.19: Same as Figure 3.11, but for *Vienna, Hohe Warte*, relative humidity [%], QM.

The results for the station *Innsbruck, University* are summarized in Table 3.14 and in Figure A.6. The seasonal results show improvements after applying QM. The bias has been corrected in all seasons, but the p-values are almost all at 0.

Correction Function

The correction function plots for relative humidity are shown in Figure 3.20 for *Vienna, Hohe Warte*. The seasons winter, spring and summer do not differ that much from each other while a few differences appear between the decades. The REMO-produced relative humidity in summer is too low for the period 1971–1980 and too high for the period 1991–2000 compared with the observations. These differences are even more obvious for fall, where the largest correction values (positive correction values due to underestimated relative humidity) are produced. The correction value for the period 1971–1981 is mainly at a level of 8 %, and for the period 1981–1990 at a level of 15 %. Relative humidity is bounded at 0 and 100, therefore, the absolute correction values at the extremes cannot be that high, while the lower boundary is not reached, anyway.

Table 3.13: Same as Table 3.9, but for *Vienna, Hohe Warte*, relative humidity [%], QM.

	(a)		(b)		(c)		(d)	
Season	ΔMean	pVal	ΔMean	pVal	ΔMean	pVal	ΔMean	pVal
DJF	1.13	0.00	-0.30	0.55	-0.50	0.41	-2.21 (-0.07)	0.00
MAM	1.09	0.00	-0.58	0.46	-0.55	0.39	-4.77 (-1.69)	0.00
JJA	0.80	0.00	-0.37	0.50	-0.08	0.01	-6.02 (-2.98)	0.00
SON	-11.52	0.00	-0.62	0.19	-0.55	0.20	-2.61 (-12.77)	0.00
Year	-2.12	0.00	-0.47	0.12	-0.42	0.02	-3.91 (-4.37)	0.00

Table 3.14: Same as Table 3.9, but for *Innsbruck, University*, relative humidity [%], QM.

	(a)		(b)		(c)		(d)	
Season	ΔMean	pVal	ΔMean	pVal	ΔMean	pVal	ΔMean	pVal
DJF	-2.60	0.00	-0.57	0.32	-0.65	0.31	-2.32 (-3.89)	0.00
MAM	10.97	0.00	-0.30	0.66	-0.29	0.46	-3.62 (8.90)	0.00
JJA	5.81	0.00	-0.39	0.57	-0.35	0.44	-2.08 (4.81)	0.01
SON	-8.01	0.00	-0.56	0.20	-0.67	0.14	1.57 (-5.67)	0.00
Year	1.59	0.00	-0.46	0.10	-0.49	0.07	-1.62 (1.08)	0.00

The correction functions for *Innsbruck, University* are given in Figure A.6. The seasons vary more from each other than it is the case for *Vienna, Hohe Warte*. In winter smaller values are positively corrected, while the correction value above the quartile is close to zero. The correction function for spring is u-shaped with high negative correction around the median. The correction value decreases with smaller values for summer. In fall, the correction function starts again at low quantiles with high correction, but decreases with higher values.

Influence of Post-Processing on the Climate Change Signal

The observed CCS for relative humidity lies between -2 % and 4 %, with a negative signal in January, February and August, and a positive one in the other months (see Figure 3.21). The uncorrected CCS looks different with only positive CCSs in November and December. The corrected CCS for *Vienna, Hohe Warte* does not show large differences compared with the CCS of the uncorrected data. The largest negative differences occur in June, July and November, while in August and September the largest positive differences appear.

The observed and uncorrected CCS for *Innsbruck, University* shows a bit less divergences, but still no correlation. The CCS between uncorrected and corrected data displays larger deviations. The main differences occur in October, November and December with a lower CCS signal after the correction and in January and February with a higher CCS signal. The reason for the larger differences for *Innsbruck, University* are mainly the higher gradients in the correction functions.

For both stations, the CCSs do not show a better agreement with the observed CCS after the correction.

3.2.4 Global Radiation

The results for the parameter global radiation are shown in Figure 3.22, Figure 3.23 and in Table 3.15 for *Vienna, Hohe Warte*.

The uncorrected and observed data are not drawn from the same distribution (see PDF- and QQ-plots). The bias is 5.19 $\mathrm{W/m^2}$ for the year, and varies within the seasons between -0.46 $\mathrm{W/m^2}$ and 9.40 $\mathrm{W/m^2}$. The correlation coefficient is around 0.78 for the whole year.

The biases after the QM-application for the first two cal/val approaches are less than 0.50 $\mathrm{W/m^2}$ for the whole year. The maximum seasonal bias is around 1 $\mathrm{W/m^2}$ in spring. The p-values are above 0.75. The results of the split-sample test show a bias of -0.88 $\mathrm{W/m^2}$ before and -9.81 $\mathrm{W/m^2}$ after the correction for the whole year. Even higher discrepancies can be found within the seasons (see Table 3.15).

The results for the station *Innsbruck, University* are very similar to *Vienna, Hohe Warte* for the first two cal/val combinations (see Table 3.16). The split-sample test

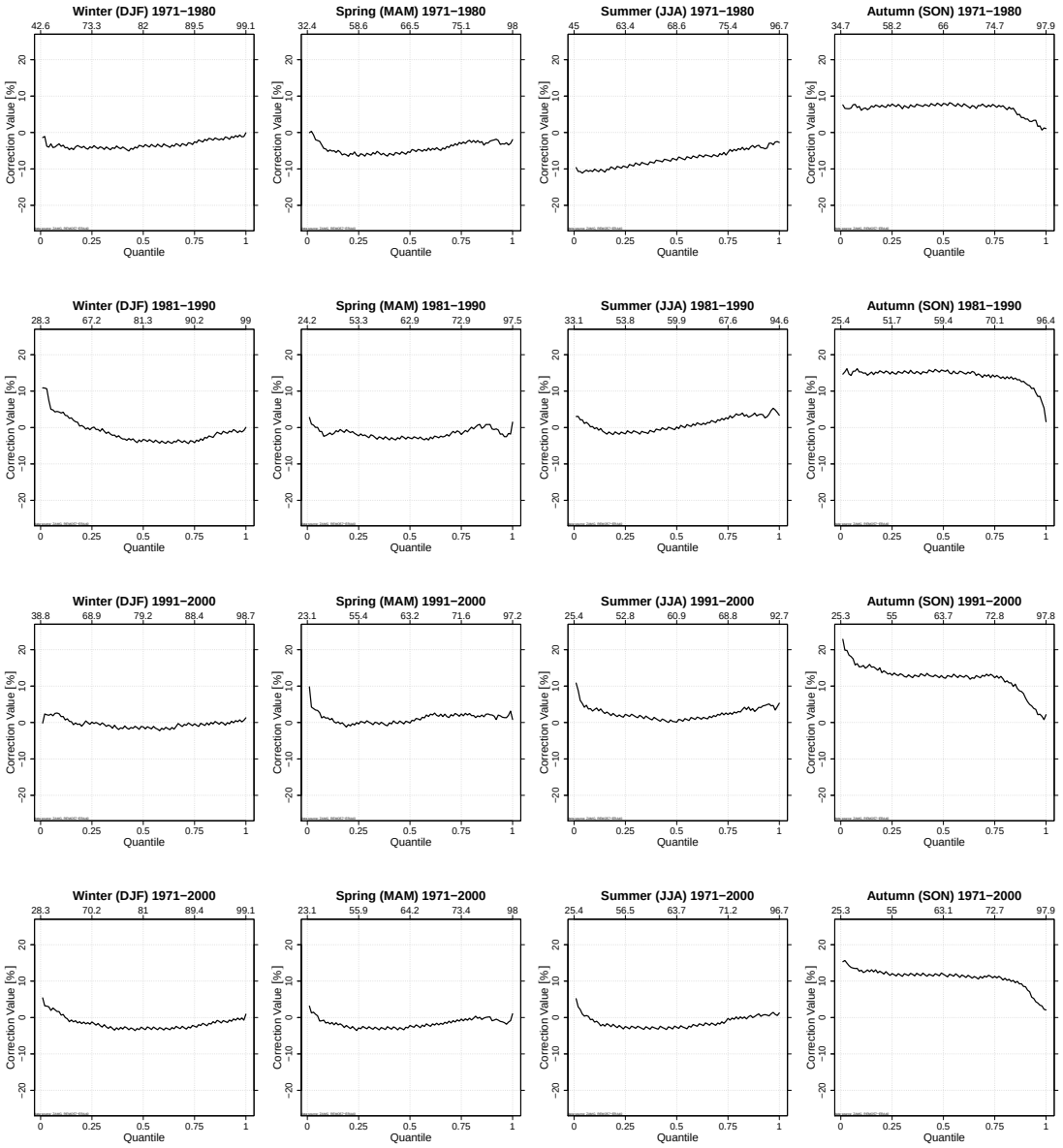


Figure 3.20: Same as Figure 3.12, but for *Vienna, Hohe Warte*, relative humidity [%].

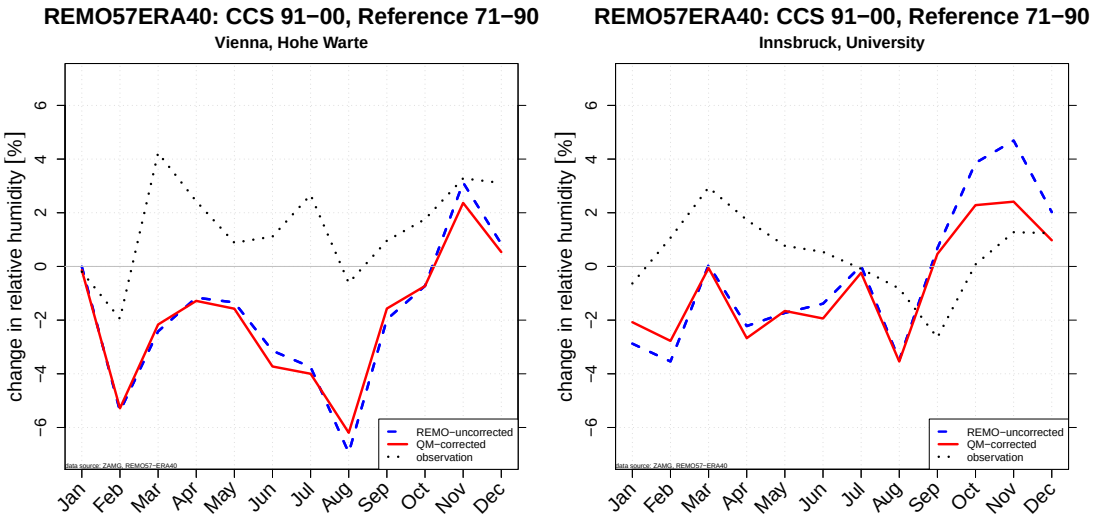


Figure 3.21: Same as Figure 3.13, but for relative humidity [%], QM.

Table 3.15: Same as Table 3.9, but for *Vienna, Hohe Warte*, global radiation [W/m^2], QM.

	(a)		(b)		(c)		(d)	
Season	ΔMean	pVal	ΔMean	pVal	ΔMean	pVal	ΔMean	pVal
DJF	-0.46	0.00	-0.47	0.77	-0.49	0.80	-6.30 (-4.33)	0.00
MAM	5.65	0.00	-1.04	0.97	-1.00	0.95	-15.91 (-4.49)	0.01
JJA	9.40	0.00	0.24	0.93	-0.07	0.91	-10.48 (3.39)	0.00
SON	6.08	0.00	-0.39	0.96	-0.40	0.93	-6.45 (1.86)	0.16
Year	5.19	0.00	-0.41	0.99	-0.49	0.99	-9.81 (-0.88)	0.01

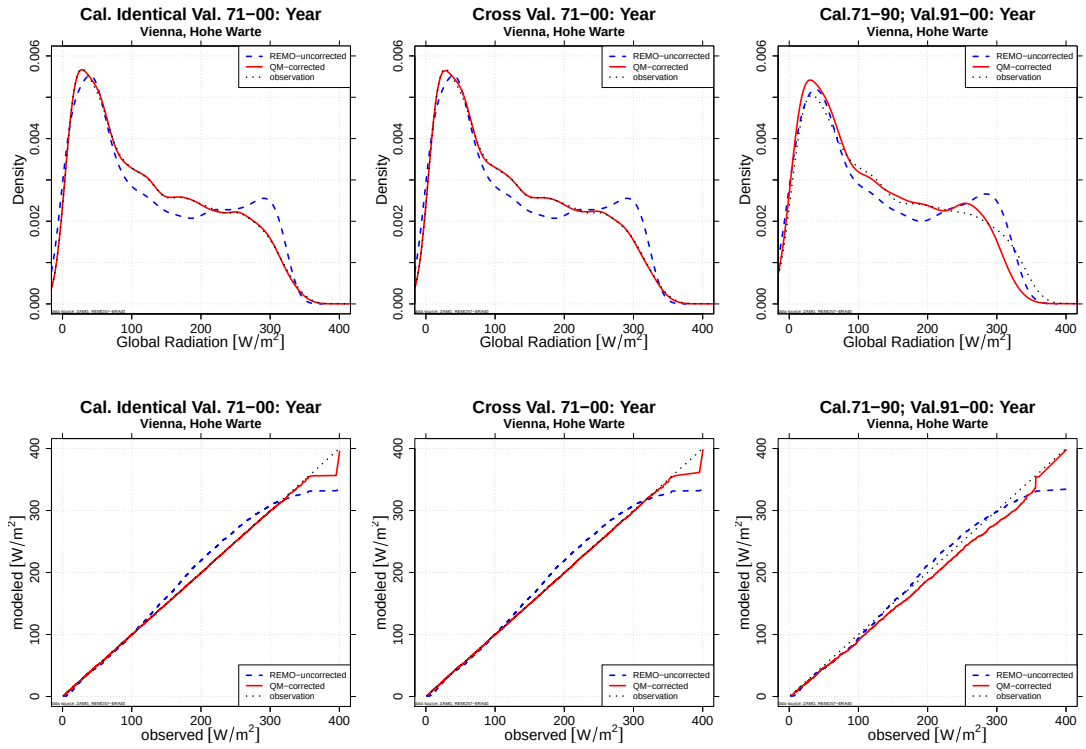


Figure 3.22: Same as Figure 3.10, but for *Vienna, Hohe Warte*, global radiation [W/m^2], QM.

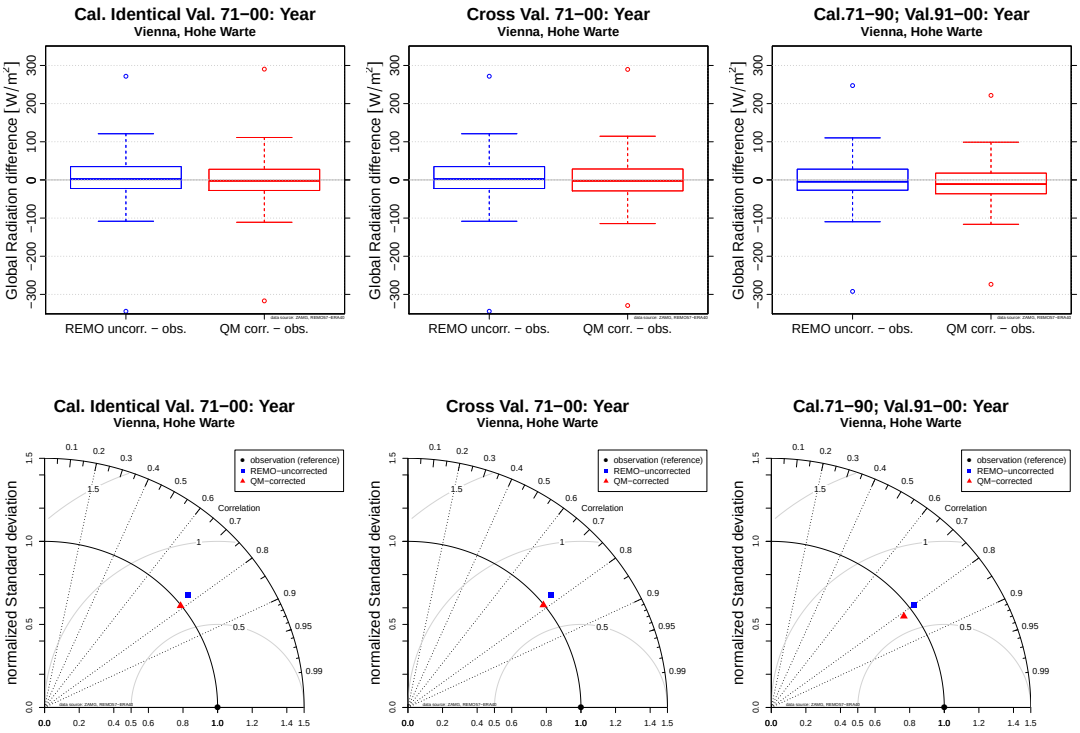


Figure 3.23: Same as Figure 3.11, but for *Vienna, Hohe Warte*, global radiation [W/m^2], QM.

shows slightly different results compared with the results of *Vienna, Hohe Warte*. The biases after the correction are lower for the seasons winter, summer and fall and higher for summer and the whole year. The p-value indicates only for the seasons winter and fall a significant agreement between corrected and observed distributions.

Table 3.16: Same as Table 3.9, but for *Innsbruck, University*, global radiation [W/m^2], QM.

	(a)		(b)		(c)		(d)	
Season	ΔMean	pVal	ΔMean	pVal	ΔMean	pVal	ΔMean	pVal
DJF	-15.99	0.00	-0.14	0.99	-0.16	0.99	1.95 (-14.10)	0.70
MAM	-16.85	0.00	-1.22	0.99	-1.52	0.99	13.82 (-6.26)	0.00
JJA	13.01	0.00	0.45	0.77	0.18	0.63	15.28 (23.94)	0.00
SON	-12.05	0.00	-0.45	1.00	-0.54	0.99	-1.72 (-13.50)	0.33
Year	-7.95	0.00	-0.34	1.00	-0.51	0.99	7.30 (-2.54)	0.00

Correction Function

The correction function plots for daily global radiation are displayed in Figure 3.24 for *Vienna, Hohe Warte*. They have a similar form for the seasons winter and fall, and also for spring and summer. The correction function for winter is within a range of ± 5 W/m^2 apart from the upper extremes, which are up to 25 W/m^2. The figure for the fall months look similar. In general, it shows lower correction values and higher ones at the extremes. The correction functions for the seasons spring and summer show a wide range from the lowest to the highest quantile. Between the 0.1 and 0.9 quantiles, the range of the correction function is from 30 W/m^2 to -25 W/m^2. The exceptions at the upper extremes are due to very high values in the observational data. Generally, the correction function is at lower and higher quantiles positive and around the median negative. Furthermore, the correction functions vary enormously throughout the decades (especially, summer and fall). For example, the variation of the correction around the median in spring is up to 25 W/m^2.

The correction functions for *Innsbruck, University* are given in Figure A.8 and show a very similar behavior to the correction functions for *Vienna, Hohe Warte*.

Influence of Post-Processing on the Climate Change Signal

The CCSs for global radiation are presented in Figure 3.25. The observed CCS is for all months positive for *Vienna, Hohe Warte* and negative (apart from January) for *Innsbruck, University*. The signals are higher in midyear. The uncorrected CCS for

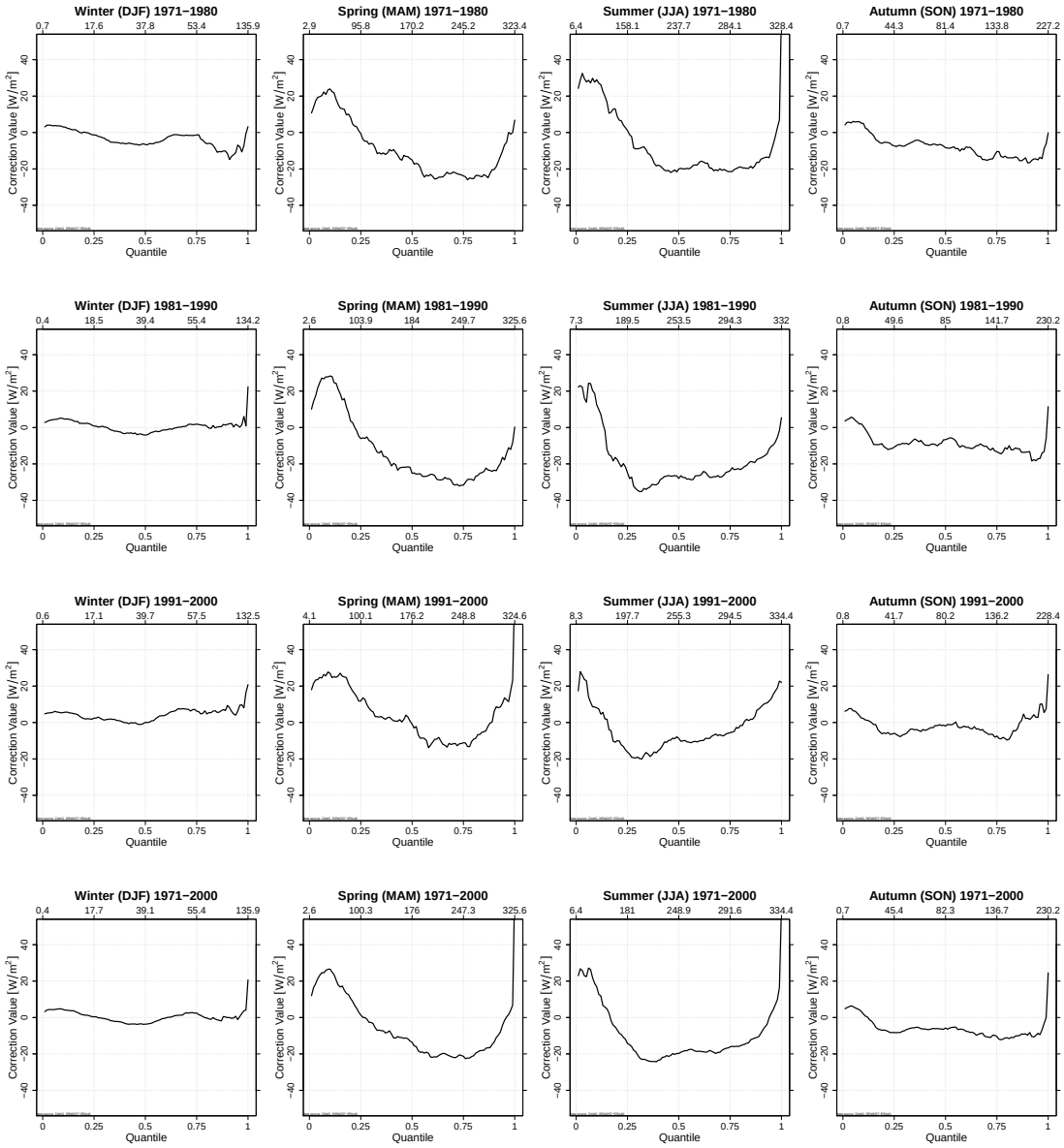


Figure 3.24: Same as Figure 3.12, but for *Vienna, Hohe Warte*, global radiation [W/m^2].

Innsbruck, University does not match to the observed CCS. It shows similarities to the simulated CCS of *Vienna, Hohe Warte*. The uncorrected CCS for *Vienna, Hohe Warte* looks like the observed CCS, but with a lower positive signal. Only small differences are visible in the warmer periods of the year after the correction. The corrected CCS signal is lower for these months (up to 3 W/m^2).

The CCSs do not show a better agreement with the observed CCS after the correction.

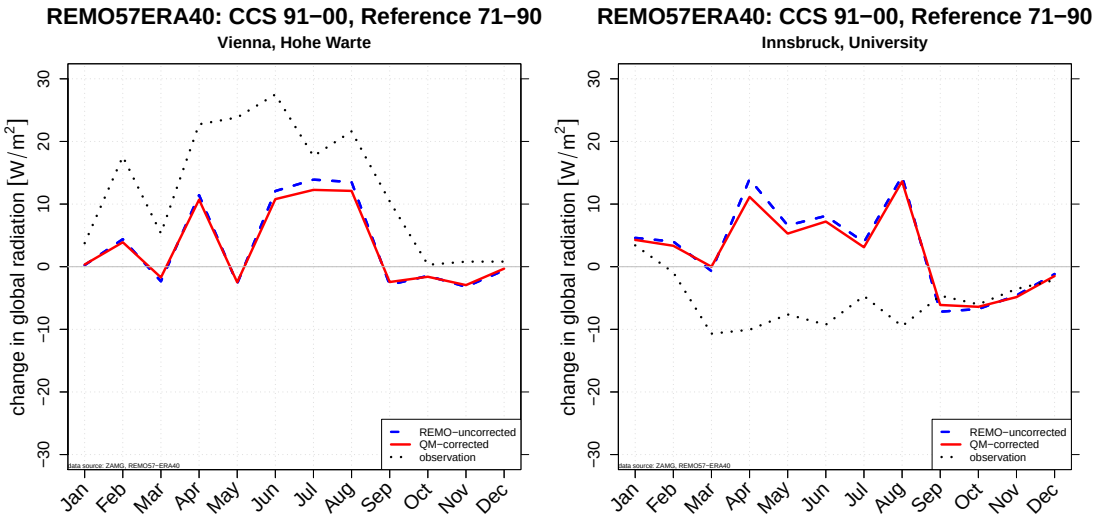


Figure 3.25: Same as Figure 3.13, but for global radiation [W/m^2], QM.

3.3 Influence of Post-Processing on the Climate Change Signal of a Future Period

In the previous sections, the DEC methods have been applied to historical datasets in order to evaluate and compare the findings with observations. The chosen calibration and validation periods were rather short. Therefore, the outcome was partly difficult to interpret. The next step is to analyze how post-processing affects a future CCS with extended time periods. The REMO model driven by the GCM ECHAM5 with a A1B-SRES Scenario is used. The calibration period (1971–2000) and the validation period (2021–2050) have been extended. The results of QM are presented in this section. The CCS is calculated for each parameter as the difference between 2021–2050 and 1971–2000 averages. Each figure displays the seasonal correction functions for *Vienna, Hohe Warte* (top row) and *Innsbruck, University* (mid row), and the CCS for *Vienna, Hohe Warte* (bottom left) and *Innsbruck, University* (bottom right). These plots are useful

to evaluate the criteria *description of change* (see Section 1.5). The displayed correction functions are seasonal averages, while in the post-processing a 31-day moving window is implemented.

Temperature

The correction function and CCS plots are shown in Figure 3.26 for the parameter temperature. The uncorrected CCS lies between 0.4 °C and 1.8 °C. The corrected CCS is slightly different after the correction. The corrected CCSs for January and February are higher than the uncorrected due to higher temperature values in the scenario period and the shape of the correction function, where at higher quantiles the correction values are positive. In contrast, the corrected CCS is lower for August due to a (higher) negative correction value for higher quantiles. The CCS for *Innsbruck, University* is similar to *Vienna, Hohe Warte*. The differences between the corrected and the uncorrected CCS are all slightly positive and difficult to interpret due to the different altitudes between the model and the observation.

Precipitation

Figure 3.27 displays the results for the parameter precipitation. The variations of the uncorrected CCS are between −10 mm/month and 22 mm/month. The corrected CCS differs positively in almost all months with the largest change in June (more than 10 mm/month). The reasons are the same as for the parameter temperature. The uncorrected data shows already a high CCS for this month which indicates a shift or a change in the shape of the distribution. The correction value at higher (or especially the highest) quantile is also very large. This value is more frequently used for the future scenario and therefore, the corrected CCS shows a higher signal. The largest correction value is available in spring, but the difference between uncorrected and corrected CCS is very low in these months. This is due to a low CCS of the uncorrected data. This indicates that the form and position of the distribution has not changed much. Therefore, the high correction values are not taken too often in the correction process. The corrected CCS for *Innsbruck, University* also shows slight changes to the uncorrected CCS.

Relative Humidity

The uncorrected CCS for the parameter relative humidity (see Figure 3.28) is between −2 % and 2 % for *Vienna, Hohe Warte* and shows just very small differences to the corrected CCS. The corrected signal is higher than the uncorrected one in June due to the positive signal and higher positive correction values for higher quantiles. The uncorrected CCS for *Innsbruck, University* is a bit higher, therefore, the corrected CCS differs more. For example, the uncorrected CCS is negative in November, therefore,

lower relative humidity values are more frequent. The correction values at lower values are more often used, which are more positive. Hence, the corrected CCS is higher than the uncorrected one.

Global Radiation

The CCS and correction functions for global radiation are shown in Figure 3.29. The uncorrected CCS shows fluctuations between -10 W/m^2 and 5 W/m^2. The corrected CCS differs just slightly to the uncorrected CCS for both stations. The correction functions vary strongly within the quantiles, and offer high correction values at the extremes. Nonetheless, the corrected signal differs just slightly from the uncorrected one due to the small CCS.

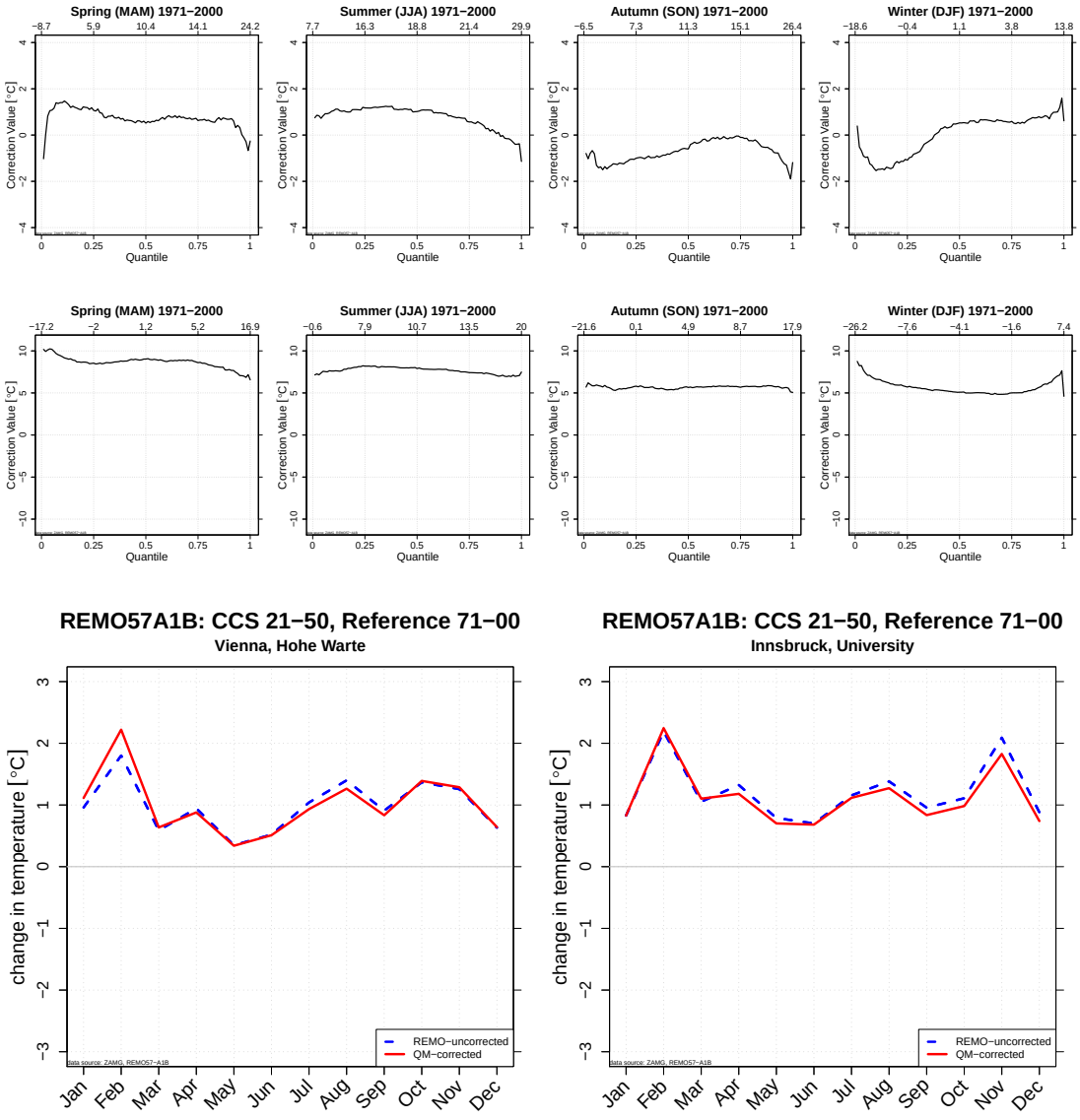


Figure 3.26: Correction function plots for *Vienna, Hohe Warte* (top row) and *Innsbruck, University* (mid row), and CCS (bottom row) for temperature [°C] for the period 2021–2050 compared with 1971-2000 (reference) for each month from REMO57-A1B. CCS for *Vienna, Hohe Warte* is on the left and *Innsbruck, University* on the right.

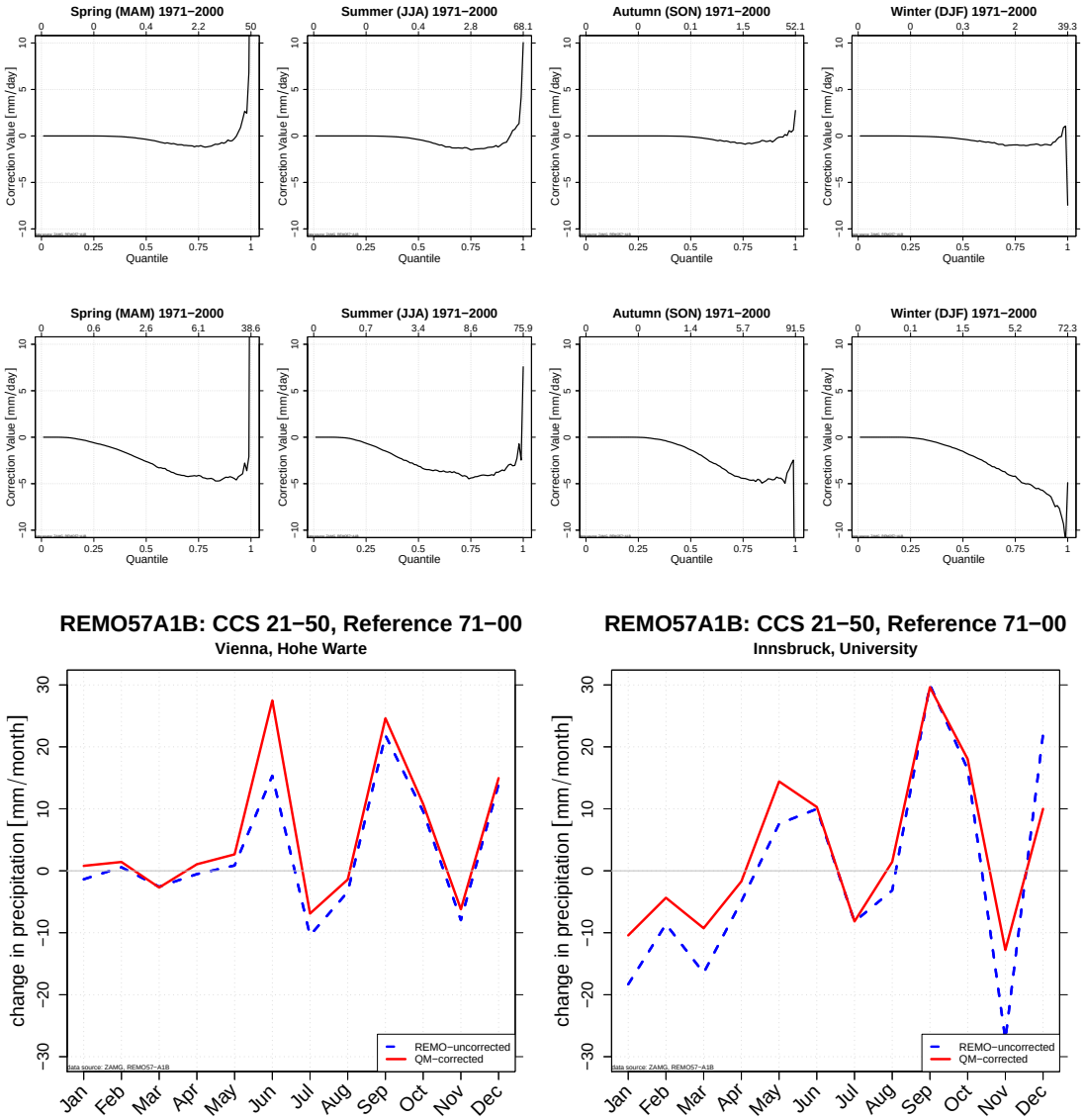


Figure 3.27: Same as Figure 3.26, but for precipitation [mm/month], QM.

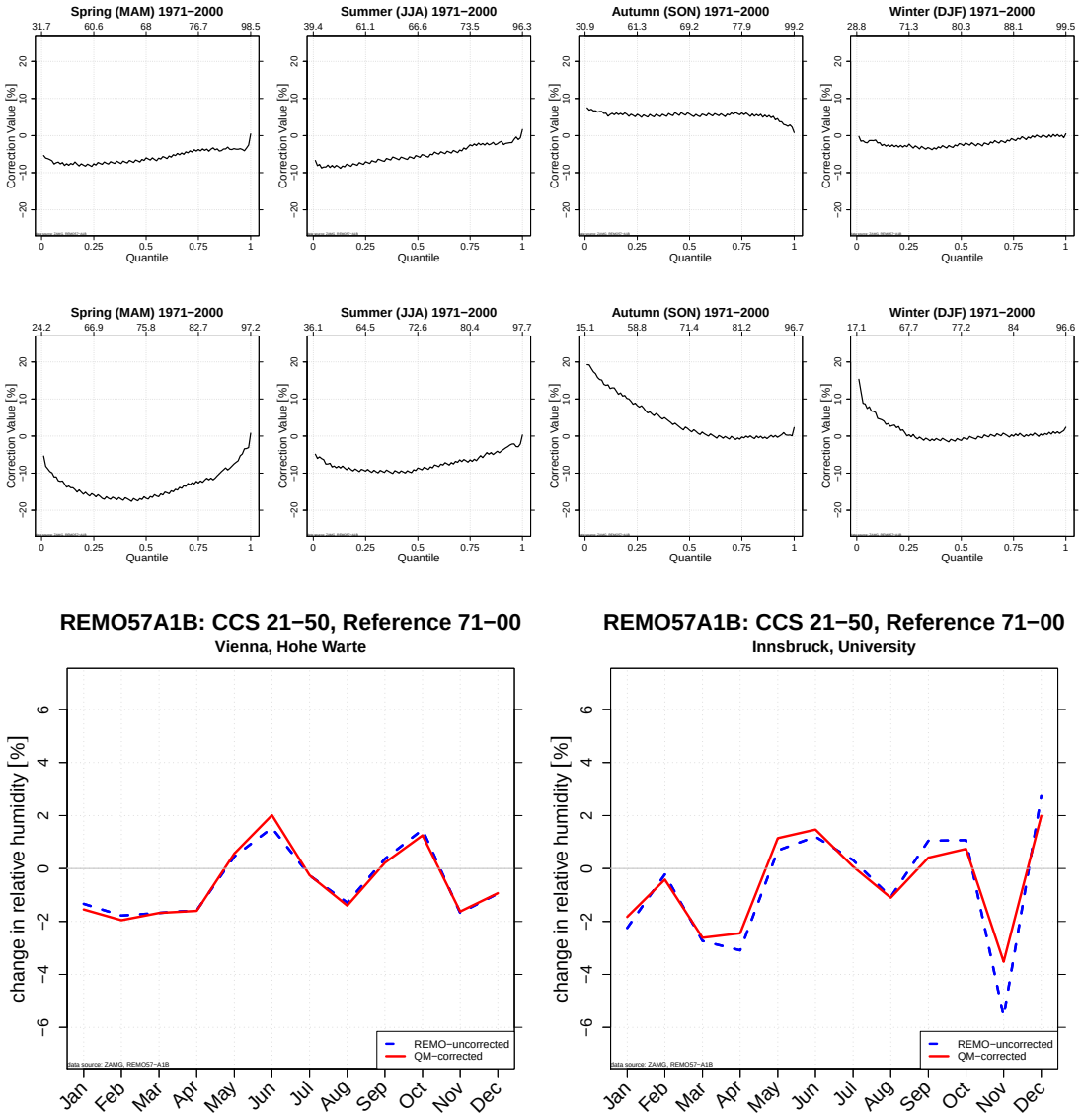


Figure 3.28: Same as Figure 3.26, but for relative humidity [%], QM.

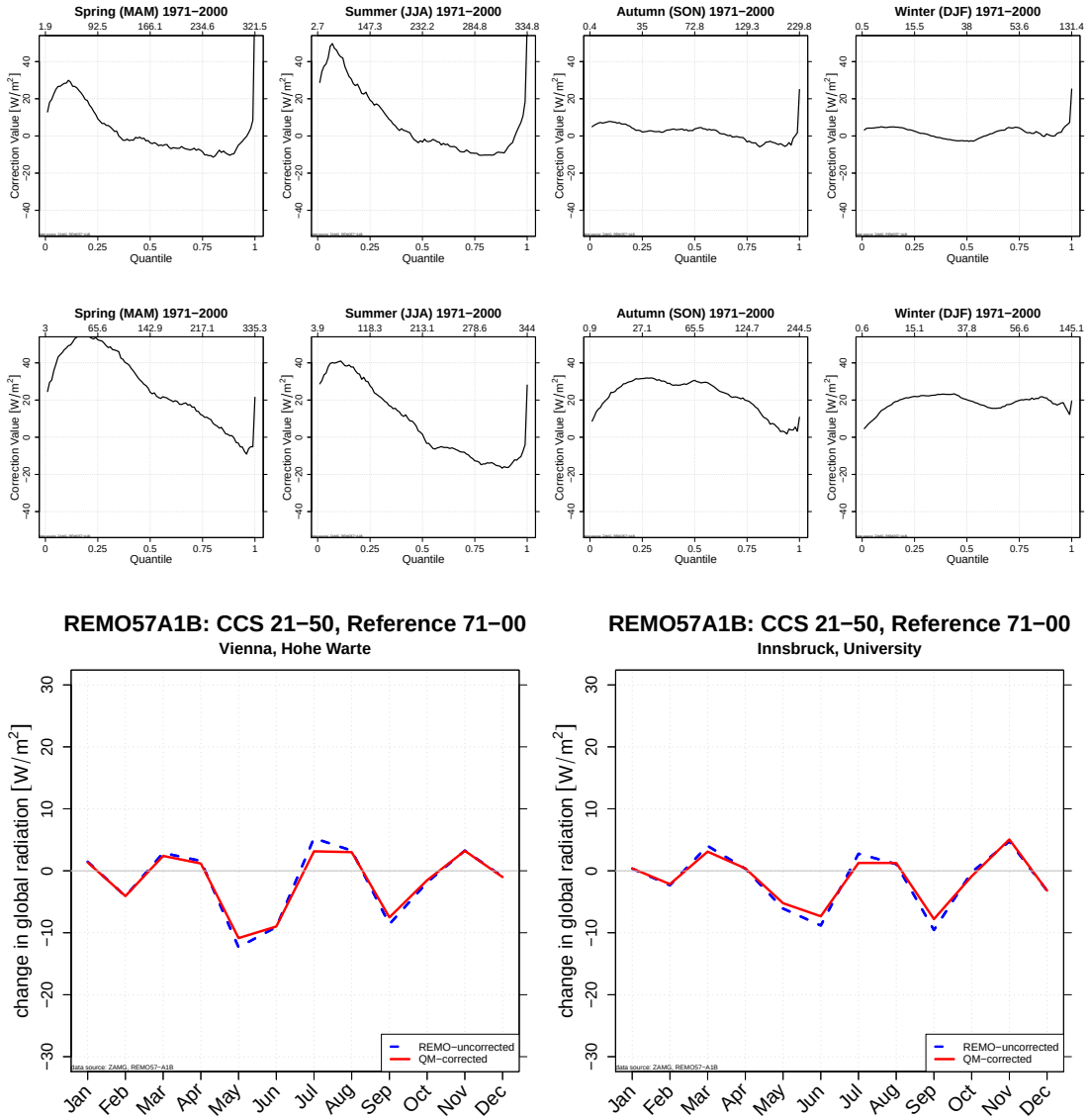


Figure 3.29: Same as Figure 3.26, but for global radiation [W/m^2], QM.

4 Discussion and Conclusion

So far, the DEC method Quantile Mapping has mostly been applied to climate variables like temperature and precipitation, either on monthly or daily basis. Furthermore, in many studies the method is used to correct and/or downscale GCM output. In this thesis the main research question was:

- What are the error characteristics of the climate parameters daily temperature, precipitation, relative humidity and global radiation in a regional climate simulation after applying QM?

Additional questions are:

- How are the daily mean climate parameters temperature, precipitation, relative humidity and global radiation distributed for a specific observation and model grid cell? Is it useful to fit those data for the post-processing or is it proper to take the empirical data?
- How well does QM perform with synthetically generated (random) distributions?
- What is the influence of the method on the climate change signal?
- How does the added feature of Quantile Mapping called Quantile-Quantile Mapping perform?
- Are the criteria for empirical-statistical Downscaling and Error Correction fulfilled?

Discussion of the Results

In many studies, fitted datasets for the climate parameters are used in the post-processing with QM. The normal distribution is mostly applied to fit daily temperature (e.g., Storch and Zwiers, 2002; Dettinger et al., 2004; Wilks, 2006) and the gamma distribution is commonly taken for daily precipitation amount (e.g., Ines and Hansen, 2006; Piani et al., 2010). But the hypothesis tests of the observed and simulated data do not show an overall significant agreement with theoretical distributions. This result is also verified in a study of Vlček and Huth (2009). They test daily precipitation for each season, with the result that the gamma distribution is not as valid for daily precipitation as commonly believed. It is even more difficult to find a reasonable theoretical

distribution for relative humidity and global radiation. These findings lead to use the DEC method with empirical data.

At first, QM is tested with synthetically generated distributions. The findings reveal that it is not limited to a specific distribution. QM is able to correct normally, gamma, log-normally and bimodally distributed data. After all, the results of the application on synthetically generated data are promising.

The next step is to apply QM on RCM output related to point-scale observations. Three different calibration/validation approaches are carried out. In the first, the calibration and validation period are identical, the second experiment is a cross-validation, and the third a split-sample test (Matulla et al., 2003). The "calibration identical validation" approach shows an almost perfect agreement between observed and corrected data. This is expected from the post-processing method. The small deviations of the corrected data to the observed are due to the calculation of the ECDFs. The cross-validation shows very good results for all four parameters, due to the fact that biases are canceled out within the three calibration/validation cycles. Only small differences are left over compared to the "calibration identical validation" approach, which are mainly because of to corrections at the extremes. The strongest informative value concerning the *stationarity* assumption is given by the split-sample test.

The split-sample test for temperature delivers significantly improved data after the correction. Mean difference, norm-sd and the p-value are all better. Similar results can be found in the literature (e.g., Dettinger et al., 2004; Wood et al., 2004). QM corrects the whole PDF. Not only the mean but also the variability is corrected (Déqué, 2007). The downscaling effect is visible for the point-scale *Innsbruck, University*, where the main differences between uncorrected and observed data can be explained by the different altitudes of the observational station and grid cells.

The split-sample test on precipitation delivers significantly improved data after the correction for *Innsbruck, University*. The results for *Vienna, Hohe Warte* also show improvements but not in mean. This is due to a larger negative difference between uncorrected and observed data in the validation period (1991–2000) compared to the calibration period (1971–1990). Many studies show advances in the precipitation data, while in some papers empirical data (e.g., Themeßl et al., 2010; Wood et al., 2004; Dettinger et al., 2004; Boé et al., 2007; Déqué et al., 2007), and in others fitted distributions (e.g., Piani et al., 2010; Dobler and Ahrens, 2008; Ines and Hansen, 2006; Benestad, 2007) are used.

The error for the parameter relative humidity varies strongly from season to season. The split-sample test for *Vienna, Hohe Warte* shows especially in fall an improvement, where the largest biases occur. Furthermore, the large variations in the error within the seasons are corrected as well as the error at low values ($< 40\,\%$). The bias for *Innsbruck, University* is after applying QM less than before in all seasons.

The split-sample test for the parameter global radiation displays improved model data in the seasons winter, summer and fall for *Innsbruck, University*. But for *Vienna,*

Hohe Warte, the bias is larger after applying QM than before. Reasons for that are the large differences of the error between the calibration and validation period. The correction function plots for *Vienna, Hohe Warte* show in the (calibration) period 1971–1990 high negative correction values, while these differences are not as large in the (validation) period 1991–2000. The split-sample test of the other two validation periods 1971–1980 and 1981–1990 returns improved results after the correction. This might indicate that the period 1991–2000 is a particular decade in respect to the differences between observed and simulated data ("interdecadal fluctuations"). A 30-year period is usually recommended in climate research (as defined by the World Meteorological Organization (**WMO**)) in order to eliminate from year to year variations (Met Office, 2005). A further issue is the replacement of the weather stations in 1991 (*Vienna, Hohe Warte*) and in 1986 (*Innsbruck, University*). Storch et al. (1999*a*) mention that changes of the location might lead to problems. Especially, the measurement of global radiation strongly depends on the location.

Two different REMO simulations are used to analyze the CCS of all four parameters. In the hindcast simulation REMO57-ERA40 the period 1991–2000 is compared with 1971–1990. The period 2021–2051 is compared with 1971–2000 from the REMO57-A1B simulation. It is difficult to give a clear statement as long as no specific trend can be seen in the data. Differences between the uncorrected and corrected CCSs occur only because of shifted distributions between the scenario and the reference period and high gradients in the correction functions. Some of the corrected CCS match slightly better with the observed CCS than the uncorrected. But these differences are quite small. So far, not many studies are available on this issue. Themeßl and Gobiet (2010) discover similar results for temperature and precipitation. They also analyze the influence on weather extremes, where larger differences between uncorrected and corrected CCS are visible. Usually, a 10-year-period is too short for the CCS to do a reasonable evaluation. Furthermore, a few years should be between the scenario and reference period. However, in the present thesis, the observation data are only available for a specific but short period. Therefore, CCS of the REMO57-A1B is also studied with longer time periods and a gap between the reference and scenario data. Moderate changes of the CCS appear after the correction.

The additional DEC method Q-QM mainly shows similar results to QM. For the synthetically generated distributions, differences to QM appear when the correction functions offer high gradients and the distributions are shifted between the validation and the calibration period. Therefore, as long as the variability does not change, Q-QM shows promising results. For the post-processing of the REMO57-ERA40 data, the largest differences to QM occur at the extremes, where the Q-QM-corrected data has sometimes shown better agreement with the observations than the QM-corrected data (e.g., precipitation extremes). Nonetheless, other statistical parameters (e.g., mean, p-value) show worse results. The assumption that the model error depends on the shape of the distribution is in some cases problematic. In the Q-QM method, a value which is in the validation

period the median will be corrected like the median in the calibration period. But some errors depend on physics and not on statistics, for example the mentioned zero-degree error.

Benestad et al. (2008) summarize four criteria which have to be fulfilled for applying DEC. The criteria of *strong relationship* and *model representation* are fulfilled due to the used DEC method. The same parameter (e.g., mean air temperature) is used as predictor and predictand. The criterion of *description of change* is fulfilled due to a still existing trend in the data after the correction (visible in the CCS). A possibility to analyze *stationarity* is to calibrate the method in the coldest (driest) period and to apply it to the warmest (wettest) or vice versa (Maraun et al., 2010). If the results are satisfying for this case the issue of *stationarity* can be regarded as fulfilled. In this thesis the split-sample test is used to give a simplified answer. The shown correction functions, which display the difference between observed and uncorrected simulated data depending on the quantile, provide the information. One can compare each decadal correction function with the overall 30-year correction function, but with the knowledge that the used decade is also included in the 30-year period. However, it is still difficult to give a reasonable answer due to the short periods. For temperature and precipitation, less variations throughout the decades appear while larger deviations occur for the parameters relative humidity and global radiation. This leads to the worse results in the split-sample test for the validation period 1991–2000. The issue of *stationarity* is also evaluated on temperature with three different ESD methods in the ENSEMBLES project. There, the results on this are positive (ENSEMBLES, 2009, pp. 68–69).

Discussion of the Method

The performance of the DEC method Quantile Mapping and the additional version Quantile-Quantile Mapping is analyzed with different climate parameters in this thesis. QM has the strength to provide point-scale information. DEC methods are generally easy to implement and relatively cheap in computational costs. A drawback is especially the issue of stationarity between predictor and predictand. Furthermore, high spatially and temporally resolved observation data is required in a high quality (Themeßl and Gobiet, 2010; Maraun et al., 2010; ENSEMBLES, 2009).

Boé et al. (2007) mention three limitations of QM. The temporal autocorrelation characteristics cannot be corrected. Furthermore, a variable might not be independent from another one, but the method corrects each variable independently. Finally, the spatial autocorrelation might be biased, which is given by the RCM.

An already mentioned advantage of QM with empirical data is that it is generally applicable to any climate parameter. However, a few adaptations should be carried out for specific parameters. Daily precipitation amount should be separated into frequency and intensity correction, as it is mostly recommended in the literature, but not done here. A systematic error can be avoided with this approach. Another issue with bounded

parameters (e.g., relative humidity, precipitation) is that they can contain values outside of the boundary after the correction process. Therefore, an adaptation within the method might be reasonable.

Outlook

A few subjects remain unanswered and are mentioned here. QM (and Q-QM) on daily precipitation with a separated frequency and intensity correction might deliver different results with the advantage to overcome systematic introduced errors. The interpretation of the results would be easier with extended time periods (at least 25 years of calibration and 25 years of validation) in a hindcast simulation. Therefore, homogenized observation datasets would provide the high quality predictands. A more detailed analysis of the zero-degree peak could deliver interesting results. A larger gap between the calibration and validation period should be used in this case. Finally, the parameters daily relative humidity and global radiation should be studied in a larger context, e.g., the whole Alpine region. The results of just two specific observational stations might not deliver the whole information.

Conclusion

In this thesis, an empirical-statistical Downscaling and Error Correction method is applied to synthetically generated data and to four point-scale daily climate parameters: temperature, precipitation, relative humidity and global radiation. Many studies are already available, where QM is used to downscale and/or error correct the parameters temperature and precipitation. But it is quite new to apply QM to daily relative humidity and global radiation. The results show that QM works well with different kinds of synthetical generated distributions. All parameters show improved results for the "calibration identical validation" and cross-validation case. The split-sample test returns improved results for temperature, precipitation and relative humidity, but does not display an overall improvement for global radiation. This might also be due to the chosen observation data ("replacement of the observational station") and short time periods ("interdecadal fluctuations"). Therefore, further work should be carried out on longer time periods with high quality observations.

A Additional Figures

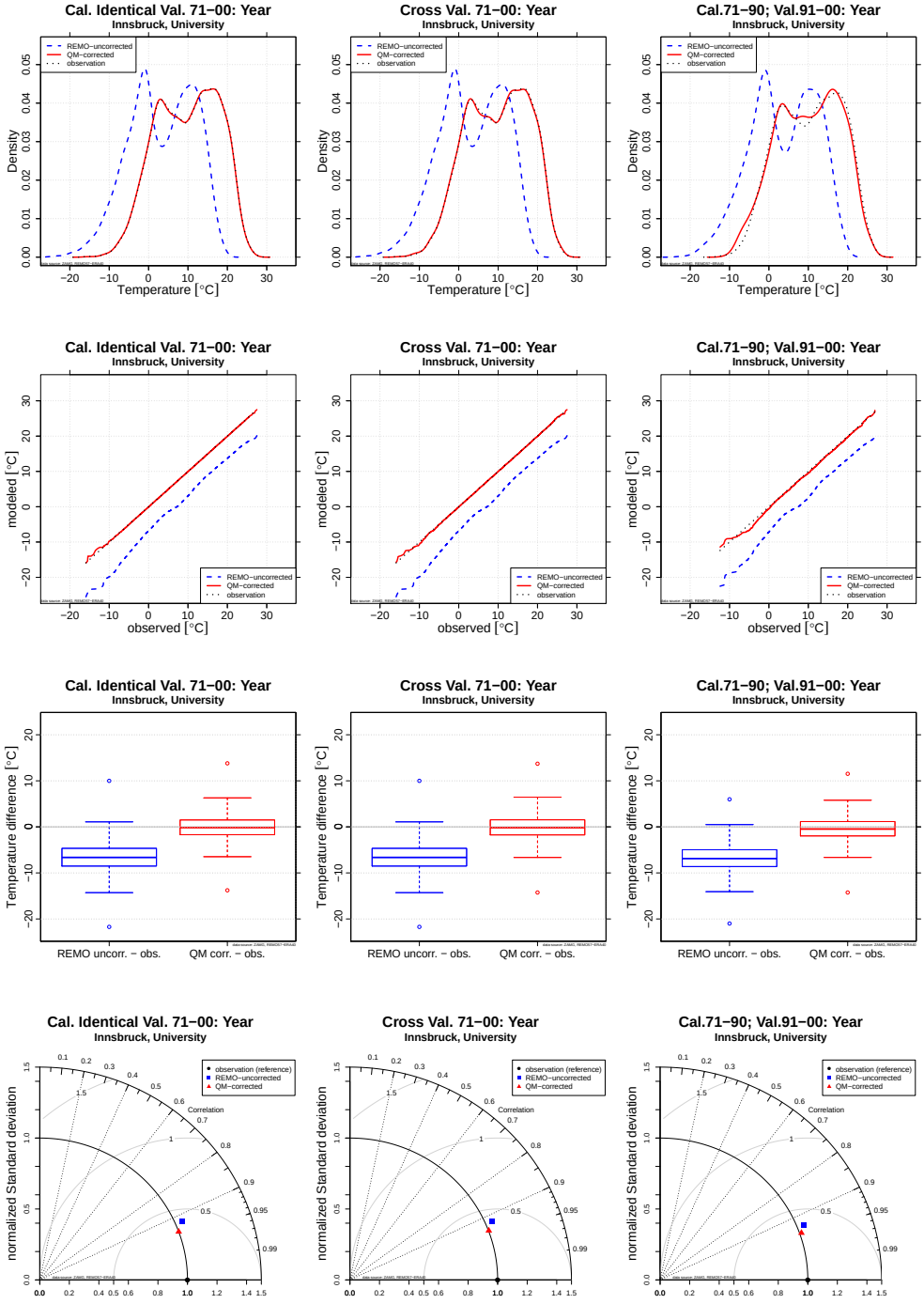


Figure A.1: Same as Figure 3.10 and 3.11, but for *Innsbruck, University*, temperature [°C], QM.

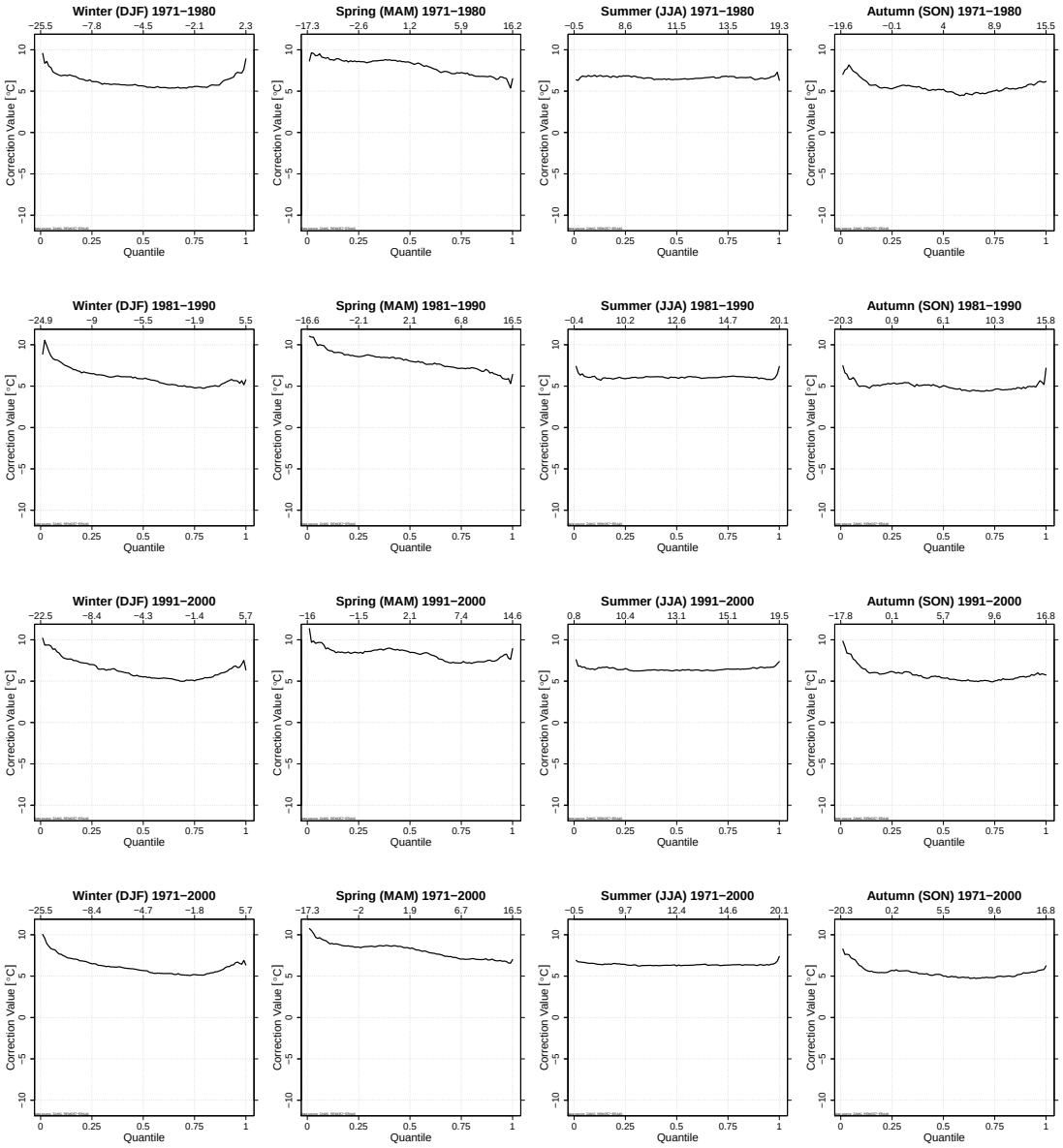


Figure A.2: Same as Figure 3.12, but for *Innsbruck, University*, temperature [°C].

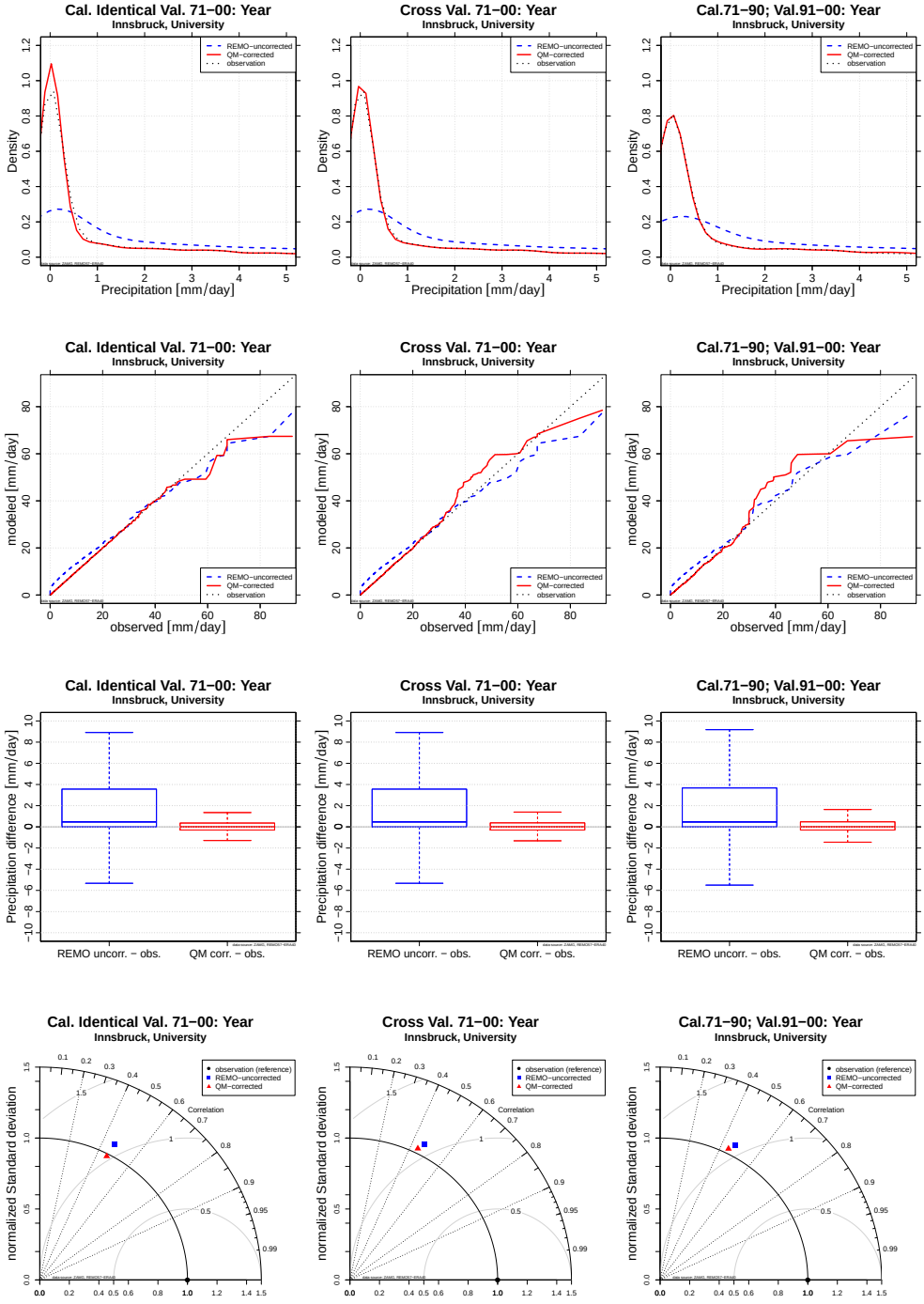


Figure A.3: Same as Figure 3.10 and 3.11, but for *Innsbruck, University*, precipitation [mm/d], QM.

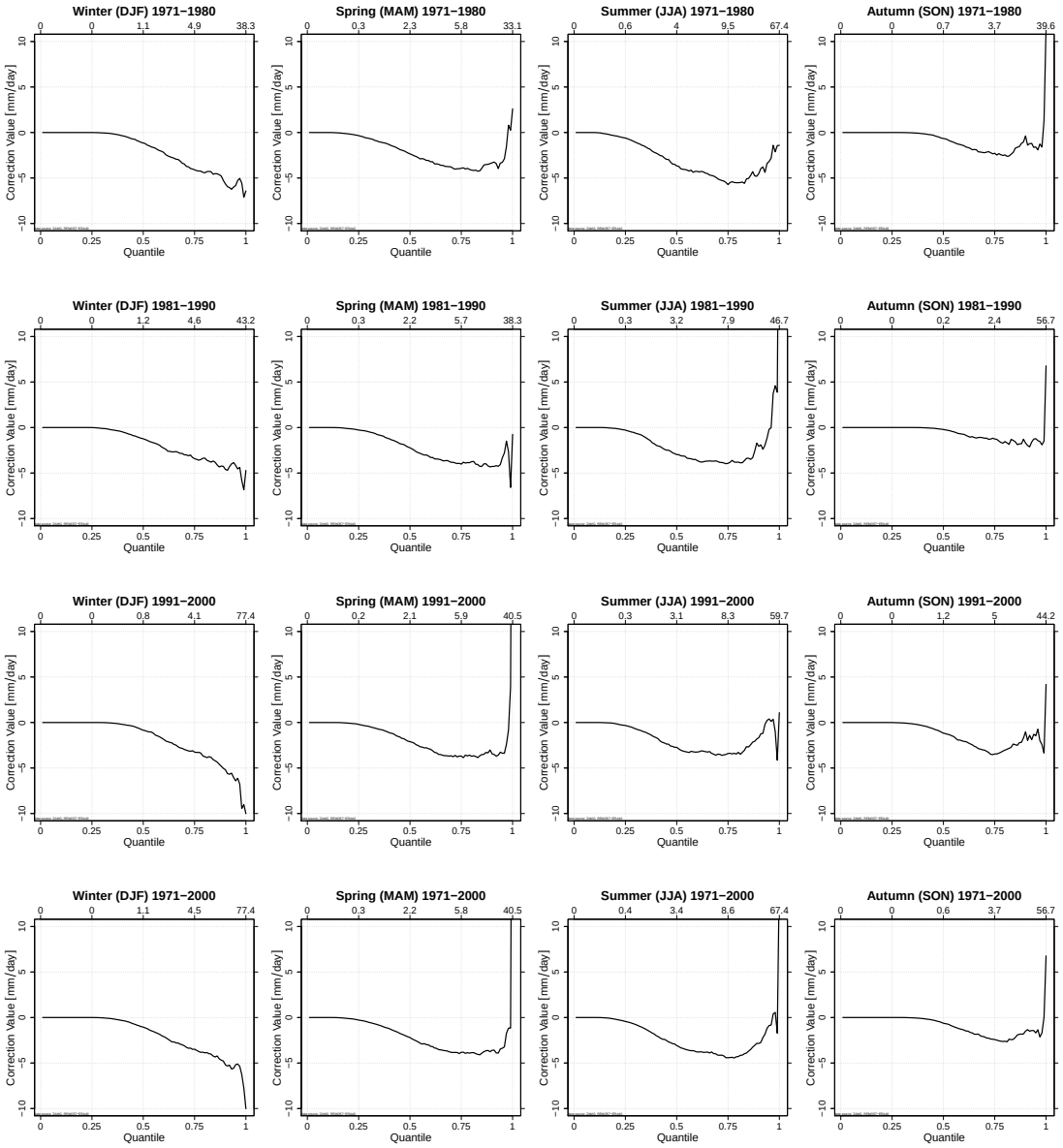


Figure A.4: Same as Figure 3.12, but for *Innsbruck, University*, precipitation [mm/d].

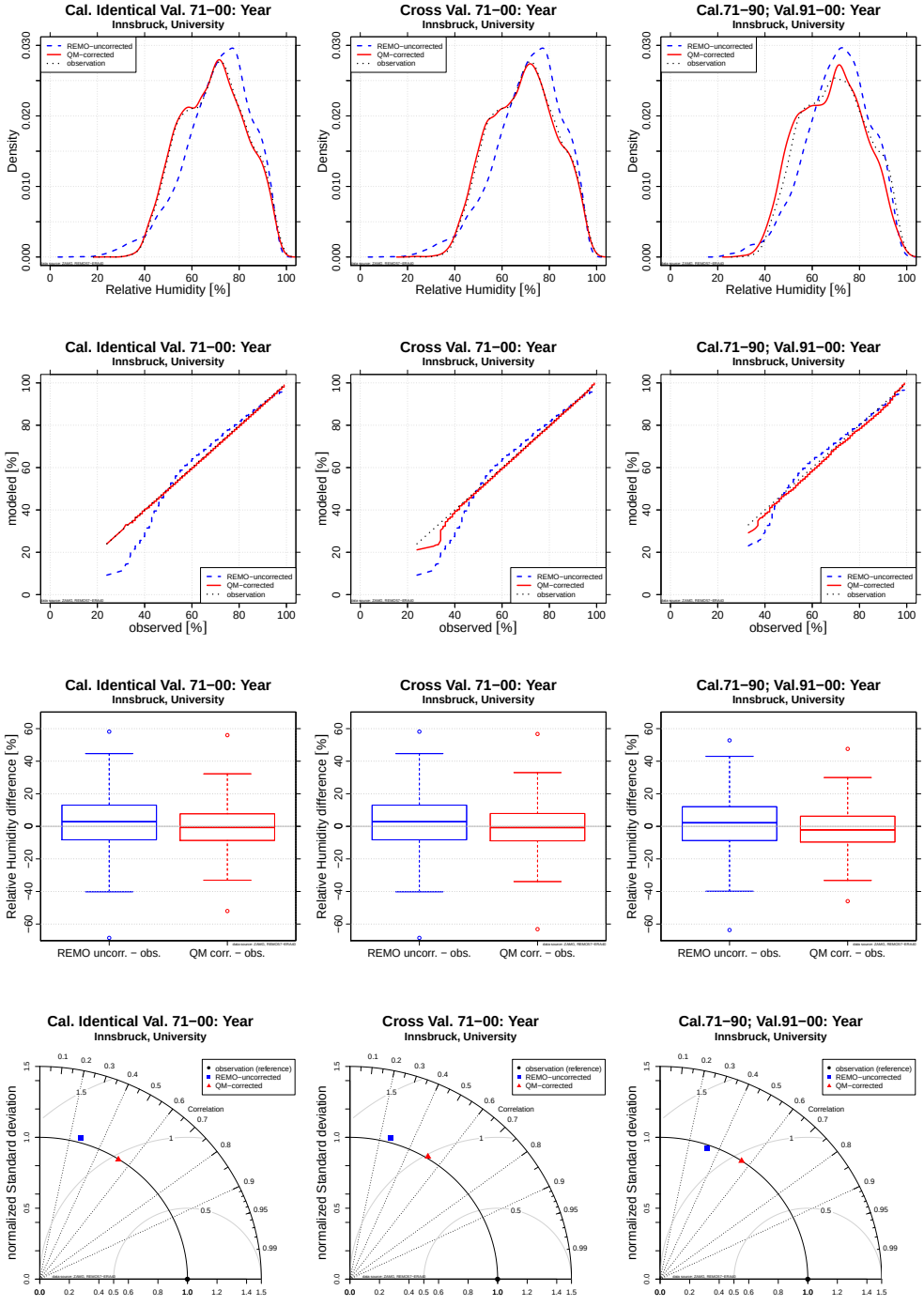


Figure A.5: Same as Figure 3.10 and 3.11, but for *Innsbruck, University*, relative humidity [%], QM.

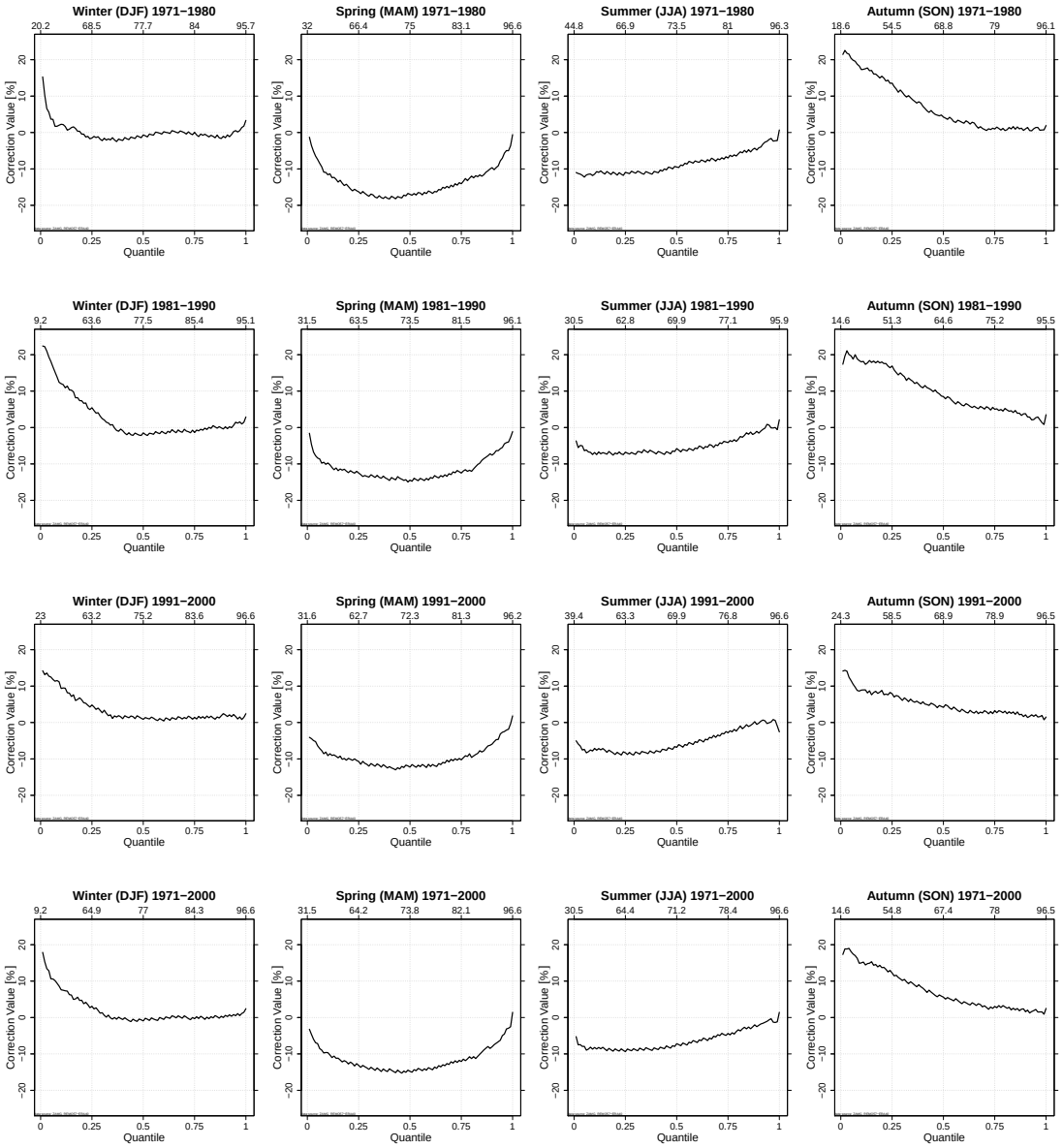


Figure A.6: Same as Figure 3.12, but for *Innsbruck, University*, relative humidity [%].

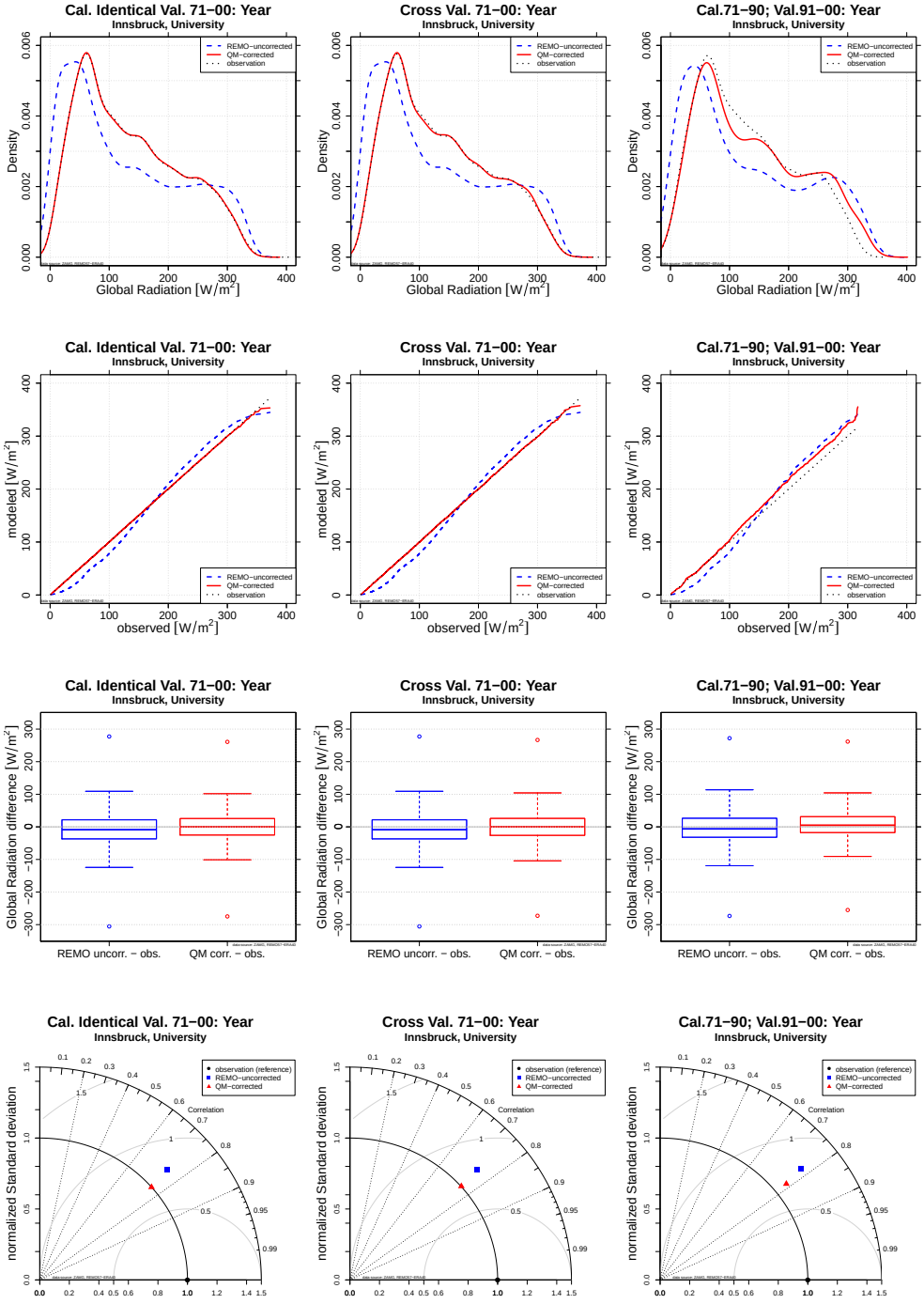


Figure A.7: Same as Figure 3.10 and 3.11, but for *Innsbruck, University*, global radiation [W/m^2], QM.

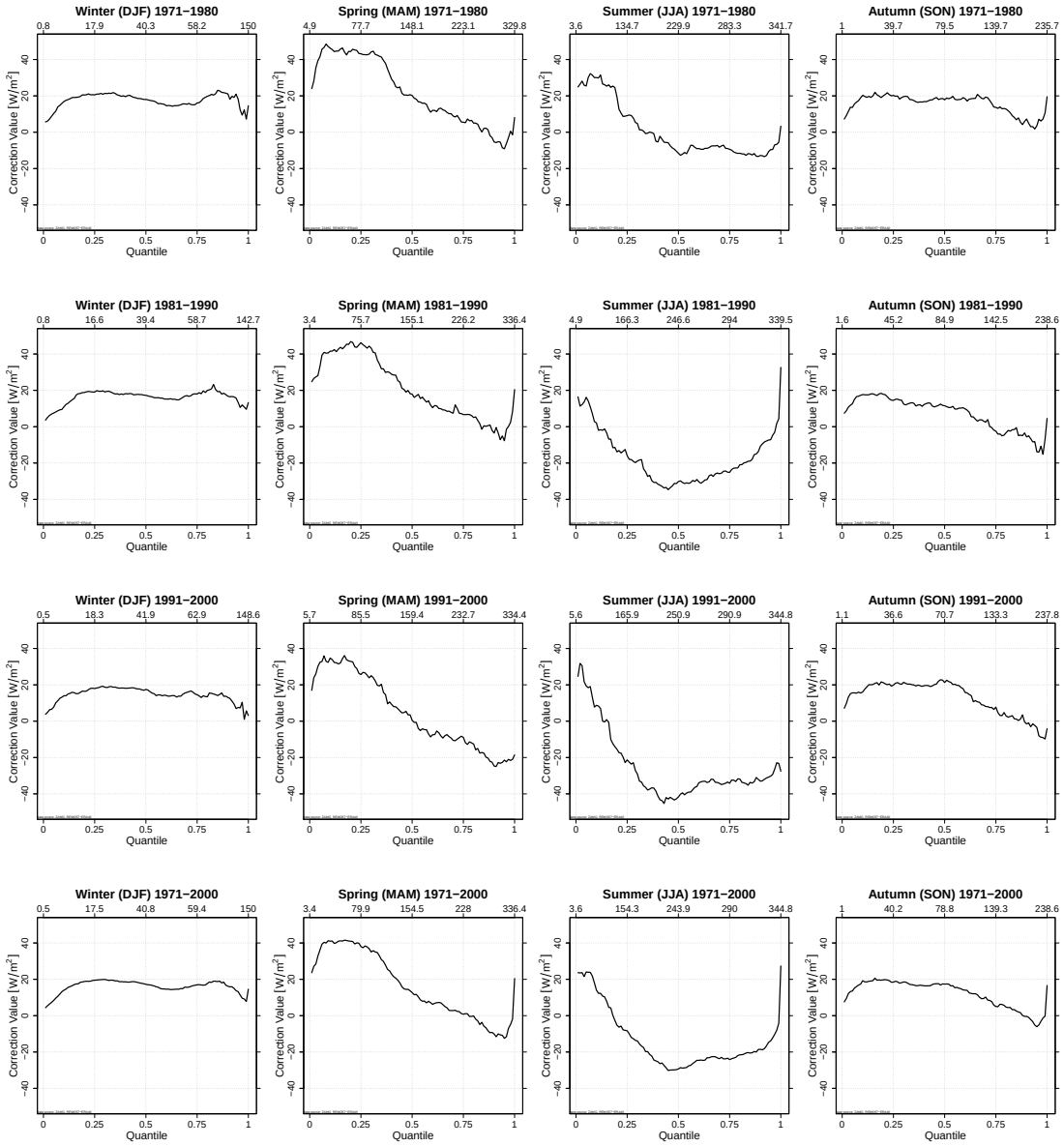


Figure A.8: Same as Figure 3.12, but for *Innsbruck, University*, global radiation [W/m^2].

Acronyms

A

A1B Balanced Scenario. 4, 17, 69

A1FI Fossil Fuel Intensive Scenario. 4

A1T Predominantly Non-Fossil Fuel Scenario. 4

AOGCM Atmosphere–Ocean General Circulation Model. 10

AR4 Fourth Assessment Report. 1–3, 97

B

boxplot Box-and-Whisker Plot. 27, 28, 37–39, 97, 98

C

cal calibration. 19, 21–23, 46, 47, 54, 56, 59, 62

CCLM COSMO Model in Climate Mode. 7, 8, 97

CCS Climate Change Signal. ix, xi, 16, 47, 51, 53, 58, 62, 67, 69–72, 79, 80, 99

CDF Cumulative Distribution Function. 17, 19–22, 26

CMIP3 Coupled Model Intercomparison Project 3. 4–6, 97

CvM-test Cramér-von Mises test. 29, 30, 37

D

DEC empirical-statistical Downscaling and Error Correction. ix, xi, 3, 4, 6, 7, 9–14, 17, 20, 23, 35, 39, 48–50, 69, 77–81, 97–99, 101

E

E-Alps Eastern-Alps. 8, 97

ECDF Empirical Cumulative Distribution Function. 19, 21, 23, 26, 27, 30, 78, 97

ECHAM4 Global Climate Model ECHAM, version 4. 6

ECHAM5 Global Climate Model ECHAM, version 5. 6, 16, 69

ERA40 ECMWF Re-Analysis. 16

ESD Empirical-Statistical Downscaling. 1, 3, 80

F

FAR First Assessment Report. 1–3, 97

G

GCM General Circulation Model. 1, 3–6, 9, 16, 69, 77, 97

H

HadCM3 Hadley Centre Coupled Model, version 3. 6

HadGEM1 Hadley Centre Global Environmental Model, version 1. 6

I

IDL Interactive Data Language. 13, 37, 38

INGV Istituto Nazionale di Geofisica e Vulcanologia, Italy. 6

IPCC Intergovernmental Panel on Climate Change. 1

IQR Interquartile Range. 24, 27

K

KS-test Kolmogorov Smirnov test. 29, 33

M

MIROC3.2(hires) Model for Interdisciplinary Research on Climate. 6

MLE Maximum Likelihood Estimation. 21

MM5 Mesoscale Model 5. 7, 8, 97

MPI-OM Max Planck Institute for Meteorology. 6, 16

N

NE North-East. 8, 97

NW North-West. 8, 97

O

obs observed data. 19, 21–23

P

PCM Parallel Climate Model. 11

PDF Probability Density Function. ix, xi, 13, 21, 22, 24, 25, 29, 37–39, 46–48, 51, 62, 78, 97, 98

PRUDENCE Prediction of Regional scenarios and Uncertainties for Defining EuropeaN Climate change risks and Effects. 4

pVal p-value. 50, 101

Q

Q-QM Quantile-Quantile Mapping. 11–13, 16, 23, 37, 39–47, 77, 79–81, 97, 98

QM Quantile Mapping. ix, xi, 11–13, 16, 17, 19, 20, 23, 30, 37–50, 54–56, 58–62, 64–67, 69, 73–75, 77–81, 84, 86, 88, 90, 97–102

QQ Quantile-Quantile. 25, 26, 29, 37, 39, 46–48, 51, 62, 97, 98

R

RCM Regional Climate Model. ix, xi, 1, 3, 4, 6–11, 16, 17, 78, 80, 97

ReLoClim Regional and Local Climate Modeling and Analysis Research. vii

REMO REgional MOdel. 7, 8, 16–18, 31, 37, 43, 46–48, 51, 56, 60, 69, 79, 97

REMO57-A1B REgional MOdel v5.7 driven by the General Circulation Model ECHAM5 with A1B SRES Scenario. ix, xi, 16, 31, 33, 72, 79, 99

REMO57-ERA40 REgional MOdel v5.7 driven by ECMWF Reanalysis. ix, xi, 16, 31–33, 47, 53, 79, 98, 99

RMS Root-Mean-Square. 27, 28

S

SAR Second Assessment Report. 2, 3, 97

SE South-East. 8, 97

sim simulated data. 19, 21–23

SRES Special Report on Emissions Scenarios. 4, 17, 69

SW South-West. 8, 97

T

TAR Third Assessment Report. 2, 3, 97

TAWES semi automatic weather recording system. 14

U

UK United Kingdom. 6

UKMO United Kingdom, Met Office. 6

UTC Coordinated Universal Time. 14, 17

V

val validation. 19, 21–23, 46, 47, 54, 56, 59, 62

W

W-Alps Western-Alps. 8, 97

WEGCENTER Wegener Center for Climate and Global Change. vii

WMO World Meteorological Organization. 79

WRF Weather Research and Forecasting Model. 7

Z

ZAMG Central Institute for Meteorology and Geodynamics. 14, 16, 31, 46, 48, 53, 99

List of Figures

1.1 Scheme of the development of the processes in climate modeling since mid 1970s. First Assessment Report (FAR) – 1990; Second Assessment Report (SAR) – 1996; Third Assessment Report (TAR) – 2001, Fourth Assessment Report (AR4) – 2007 (IPCC, 2007, Figure 1.2). 2

1.2 Scheme of the development of the spatial grid in the last few decades on the example of Europe. First Assessment Report (FAR) – 1990; Second Assessment Report (SAR) – 1996; Third Assessment Report (TAR) – 2001, Fourth Assessment Report (AR4) – 2007 (IPCC, 2007, Figure 1.4). . . . 3

1.3 Uncertainty in surface temperature of the CMIP3 GCMs for a specific domain size over Europe (N 32° – N 60°, W 15° – E 35°)(Prein, 2009). . 5

1.4 Bias of mean air temperature (left) and mean precipitation (right) of 62 one-year hindcast experiments with four different RCMs (CCLM, MM5, MM5, REMO) for the greater Alpine region. This region is divided into North-West (NW), South-West (SW), Western-Alps (W-Alps), Eastern-Alps (E-Alps), North-East (NE) and South-East (SE) (Suklitsch et al., 2010). 8

2.1 Locations of the weather and climate stations in Austria (ZAMG, 2010, accessed Jan., 2010). 15

2.2 Orography of the REgional MOdel (REMO) v5.7. 18

2.3 Overview of the DEC method: Quantile Mapping (QM). 20

2.4 Overview of the DEC method: Quantile-Quantile Mapping (Q-QM). . . . 23

2.5 Example of histograms showing some random data with relative frequency on the y-axis and different bin widths. 24

2.6 Example of PDF-plots and histograms. 25

2.7 Example of a QQ-plot. 26

2.8 Example of two ECDF-plots. On the left, the same data as in the previous examples is used and on the right, random values are showing the steps in a clearer way. 27

2.9 Example and overview of a boxplot. 28

2.10 Example of a Taylor diagram. 29

2.11 Histogram and fitted normal distribution plots of observed (top row) and REMO57-ERA40 simulated (bottom row) temperature [°C] for *Vienna, Hohe Warte* from 1971 to 2000. For each season one representative month is shown (for winter: January, spring: April, summer: July, fall: October). 32

2.12 Same as Figure 2.11, but for *Vienna, Hohe Warte*, precipitation (>0.1) [mm/d] and a fitted gamma distribution. 33

2.13 Same as Figure 2.11, but for *Vienna, Hohe Warte*, relative humidity [%] and a fitted beta distribution. 34

2.14 Same as Figure 2.11, but for *Vienna, Hohe Warte*, global radiation [W/m^2] and a fitted normal distribution. 35

3.1 PDF-plot and boxplot for synthetically generated data in the calibration period. The uncorrected (blue-dashed) and reference (black-dotted) data are normally distributed and differ in the first (mean) and second parameter (standard deviation). 38

3.2 PDF-plot, boxplot, correction function, and QQ-plot for synthetically generated data in the validation period. The uncorrected (blue-dashed) and reference (black-dotted) data are normally distributed and differ in mean and standard deviation. The DEC method Quantile Mapping is applied. The line of the corrected data is displayed in red-full. 39

3.3 Same as Figure 3.2, but with Q-QM. 39

3.4 Same as Figure 3.2 for QM (top) and Q-QM (bottom), but for gamma distributed data. 41

3.5 Same as Figure 3.2 for QM (top) and Q-QM (bottom), but for log-normally distributed data. 42

3.6 Same as Figure 3.2 for QM (top) and Q-QM (bottom), but for bimodally distributed data. 43

3.7 Same as Figure 3.1, but for normally distributed data under changed conditions. 44

3.8 Same as Figure 3.2 for QM (top) and Q-QM (bottom), but for normally distributed data under changed conditions. 45

3.9 Schematic graphic of the three different calibration-validation approaches. The gray fields indicate the calibration periods and the white fields the validation periods. 46

3.10 PDF- and QQ-plots of temperature [°C] for *Vienna, Hohe Warte* of the three different calibration/validation approaches. Blue-dashed line represents the uncorrected, red-full line the corrected and the black-dotted line the observed data. The DEC method QM is applied. 48

3.11 Boxplots and Taylor diagrams of temperature [°C] for *Vienna, Hohe Warte* of the three different calibration/validation approaches. Blue-dashed line represents the uncorrected, red-full line the corrected and the black-dotted line the observed data. The DEC method QM is applied. 49
3.12 Correction function for *Vienna, Hohe Warte* for temperature [°C]. Correction functions based on the differences between observed and uncorrected data on each quantile. At the bottom x-axis the quantiles are displayed and at the top x-axis the absolute value for minimum, lower quartile, median, upper quartile and the maximum. Top row shows the correction function for the period 1971–1980, second row for 1981–1990, the third row for 1991–2000 and the bottom row for the whole period 1971–2000 for each season (colums). 52
3.13 CCS for temperature [°C] for the period 1991–2000 compared with 1971–1990 (reference) for each month from REMO57-ERA40 and ZAMG-observations. CCS for *Vienna, Hohe Warte* is on the left and *Innsbruck, University* on the right. 53
3.14 Same as Figure 3.10, but for *Vienna, Hohe Warte*, precipitation [mm/d], QM. 54
3.15 Same as Figure 3.11, but for *Vienna, Hohe Warte*, precipitation [mm/d], QM. 55
3.16 Same as Figure 3.12, but for *Vienna, Hohe Warte*, precipitation [mm/d]. 57
3.17 Same as Figure 3.13, but for precipitation [mm/month], QM. 58
3.18 Same as Figure 3.10, but for *Vienna, Hohe Warte*, relative humidity [%], QM. 59
3.19 Same as Figure 3.11, but for *Vienna, Hohe Warte*, relative humidity [%], QM. 60
3.20 Same as Figure 3.12, but for *Vienna, Hohe Warte*, relative humidity [%]. 63
3.21 Same as Figure 3.13, but for relative humidity [%], QM. 64
3.22 Same as Figure 3.10, but for *Vienna, Hohe Warte*, global radiation [W/m^2], QM. 65
3.23 Same as Figure 3.11, but for *Vienna, Hohe Warte*, global radiation [W/m^2], QM. 66
3.24 Same as Figure 3.12, but for *Vienna, Hohe Warte*, global radiation [W/m^2]. 68
3.25 Same as Figure 3.13, but for global radiation [W/m^2], QM. 69
3.26 Correction function plots for *Vienna, Hohe Warte* (top row) and *Innsbruck, University* (mid row), and CCS (bottom row) for temperature [°C] for the period 2021–2050 compared with 1971-2000 (reference) for each month from REMO57-A1B. CCS for *Vienna, Hohe Warte* is on the left and *Innsbruck, University* on the right. 72
3.27 Same as Figure 3.26, but for precipitation [mm/month], QM. 73
3.28 Same as Figure 3.26, but for relative humidity [%], QM. 74

3.29 Same as Figure 3.26, but for global radiation [W/m^2], QM. 75

A.1 Same as Figure 3.10 and 3.11, but for *Innsbruck, University*, temperature [°C], QM. 84
A.2 Same as Figure 3.12, but for *Innsbruck, University*, temperature [°C]. . 85
A.3 Same as Figure 3.10 and 3.11, but for *Innsbruck, University*, precipitation [mm/d], QM. 86
A.4 Same as Figure 3.12, but for *Innsbruck, University*, precipitation [mm/d]. 87
A.5 Same as Figure 3.10 and 3.11, but for *Innsbruck, University*, relative humidity [%], QM. 88
A.6 Same as Figure 3.12, but for *Innsbruck, University*, relative humidity [%]. 89
A.7 Same as Figure 3.10 and 3.11, but for *Innsbruck, University*, global radiation [W/m^2], QM. 90
A.8 Same as Figure 3.12, but for *Innsbruck, University*, global radiation [W/m^2]. 91

List of Tables

2.1 Detailed information on the two chosen weather stations. 15
2.2 Detailed information on the four closest grid points to the chosen weather stations. 17
2.3 Interpretation of the p-value (Sachs and Hedderich, 2006, p. 324). 30

3.1 Description of the chosen random normally distributed data. 38
3.2 Results for synthetically generated data: Normal distributions. 40
3.3 Description of the chosen random non-normally distributed data. . . . 40
3.4 Results for synthetically generated data: Gamma distributions. 41
3.5 Results for synthetically generated data: Log-normal distributions. . . 42
3.6 Results for synthetically generated data: Bimodal distributions. 43
3.7 Description of the chosen random normally distributed data. 44
3.8 Results for synthetically generated data: Normally distributed data under changed conditions. 45
3.9 Mean error (ΔMean) and p-value (pVal) for each season and the whole year for temperature [°C] for *Vienna, Hohe Warte*. The uncorrected (a) and the three different cases for the corrected are compared with the observed data. (b) Calibration identical validation (1971–2000); (c) Cross-validation; (d) Split-sample test (calibration: 1971–1990, validation: 1991–2000) and in parenthesis the appropriated mean error for uncorrected-observed data for the period 1991–2000. The DEC method QM is applied. 50
3.10 Same as Table 3.9, but for *Innsbruck, University*, temperature [°C], QM. 50
3.11 Same as Table 3.9, but for *Vienna, Hohe Warte*, precipitation [mm/d], QM. 55
3.12 Same as Table 3.9, but for *Innsbruck, University*, precipitation [mm/d], QM. 56
3.13 Same as Table 3.9, but for *Vienna, Hohe Warte*, relative humidity [%], QM. 61
3.14 Same as Table 3.9, but for *Innsbruck, University*, relative humidity [%], QM. 61
3.15 Same as Table 3.9, but for *Vienna, Hohe Warte*, global radiation [W/m^2], QM. 64

3.16 Same as Table 3.9, but for *Innsbruck, University*, global radiation [W/m^2], QM. 67

Bibliography

Aksoy, H. (2000). "Use of Gamma Distribution in Hydrological Analysis". In: *Turkish Journal of Engineering and Environmental Sciences* 24, 419–428.

Anderson, T. W. and D. A. Darling (1952). "Asymptotic Theory of Certain "Goodness of Fit" Criteria based on Stachastic Processes". In: *Annals of Mathematical Statistics* 23, 193–212.

Benestad, R. E. (2007). "Novel methods for inferring future changes in extreme rainfall over Northern Europe". In: *Climate Research* 34, 195–210.

Benestad, R. E., I. Hanssen-Bauer, and D. Chen (2008). *Empirical-Statistical Downscaling.* World Scientific Publishing Co. Pte. Ltd. ISBN: 9812819126.

Bindi, M. and F. Miglietta (1991). "Estimating daily global radiation from air temperature and rainfall measurements". In: *Climate Research* 1, 117–124.

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

Camuffo, D. (1978). "Cumulated Frequency Distribution of Daily Global Solar Radiation at Venice, Italy". In: *Archiv für Meteorologie, Geophysik und Bioklimatologie* 26, 45–50.

Crutcher, H. L. (1975). "A Note on the Possible Misuse of the Kolmogorov–Smirnov Test". In: *Journal of Applied Meteorology* 14, 1600–1603.

Csörgö, S. and J. J. Faraway (1996). "The Exact and Asymptotic Distributions of Cramér-von Mises Statistics". In: *Journal of the Royal Statistical Society: Series B* 58, 221–234.

Dettinger, M. D., D. R. Cayan, M. K. Meyer, and A. E. Jeton (2004). "Simulated hydrological responses to climate variations and change in the Merced, Carson, and American river basins, Sierra Nevada, California, 1900-2099". In: *Climatic Change* 62, 283–317.

Dobler, A. and B. Ahrens (2008). "Precipitation by a regional climate model and bias correction in Europe and South Asia". In: *Meteorologische Zeitschrift* 4, 499–509.

Donatelli, M., G. Bellocchi, and F. Fontana (2003). "RadEst3.00: software to estimate daily radiation data from commonly available meteorological variables". In: *European Journal of Agronomy* 18, 363–367.

Déqué, M., D. P. Rowell, D. Lüthi, F. Giorgi, J. H. Christensen, B. Rockel, D. Jacob, E. Kjellström, M. de Castro, and B. van den Hurk (2007). "An intercomparison of regional climate simulations for Europe: assessing uncertainties in model projections". In: *Climatic Change* 81, 53–70.

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

ENSEMBLES (2009). *ENSEMBLES: Climate Change and its Impacts: Summary of research and results from the ENSEMBLES project.* Ed. by P. van der Linden and J. Mitchell. FitzRoy Road, Exeter EX1 3PB, UK: Met Office Hadley Centre, 160pp.

Hawkins, E. and R. Sutton (2009). "The potential to narrow uncertainty in regional climate predictions". In: *American Meteorological Society,* 1095–1107.

IPCC (2007). *Climate Change 2007 - The Physical Science Basis: Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change.* Ed. by S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. Averyt, M. Tignor, and H. Miller. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press, 996pp. ISBN: 0521705967.

Ines, A. V. and J. W. Hansen (2006). "Bias correction of daily GCM rainfall for crop simulation studies". In: *Agricultural and Forest Meteorology* 138, 44–53.

Jacob, D. (2001). "A note to the simulation of the annual and inter-annual variability of the water budget over the Baltic Sea drainage basin". In: *Meteorology and Atmospheric Physics* 77, 61–73.

Jacob, D. et al. (2007). "An inter-comparison of regional climate models for Europe: model performance in present-day climate". In: *Climatic Change* 81, 31–52. DOI: 10.1007/s10584-006-9213-4.

Kapper, K. (2009). *Scale sensitive analysis of regional climate models.* Wegener Center Verlag Graz. ISBN: 978-3-9502615-6-1.

Maraun, D. et al. (2010). "Precipitation Downscaling under Climate Change. Recent Developments to Bridge the Gap between Dynamical Models and the End User". In: *Reviews of Geophysics* (in press).

Matulla, C., H. Scheifinger, A. Menzel, and E. Koch (2003). "Exploring two methods for statistical downscaling of Central European phenological time series". In: *International Journal of Biometeorology* 48, 56–64.

Matyasovszky, I. and I. Bogardi (1996). "A space-time stochastic climatological approach to daily global solar radiation". In: *Climate Research* 7, 11–20.

Met Office (2005). *Climate averages.* [Online; accessed 21-October-2010]. URL: `http://www.metoffice.gov.uk/climate/uk/averages/`.

Nakicenovic, N. and R. Swart (2000). *Special Report on Emissions Scenarios: A Special Report of Working Group III of the Intergovernmental Panel on Climate Change.* 1st ed. Cambridge University Press. ISBN: 0521804930.

PCMDI (2009). *Program for Climate Model Diagnosis and Intercomparison.* [Online; accessed 28-July-2010]. URL: `www-pcmdi.llnl.gov`.

PRUDENCE (2004). *Prediction of Regional scenarios and Uncertainties for Defining EuropeaN Climate change risks and Effects.* [Online; accessed 28-July-2010]. URL: `http://prudence.dmi.dk/`.

Panofsky, H. and G. Brier (1968). "Some Applications of Statistics to Meteorology". In: *The Pennsylvania State University Press: University Park* 224.

Piani, C., J. O. Haerter, and E. Coppola (2010). "Statistical bias correction for daily precipitation in regional climate models over Europe". In: *Theoretical and Applied Climatology* 99, 187–192.

Prein, A. F. (2009). *Uncertainties in the Driving Data of Regional Climate Models in the Alpine Region.* Wegener Center Verlag Graz. ISBN: 978-3-9502615-7-8.

Prein, A. F. and A. Gobiet (2010). "Quality, climate change, and uncertainty in the CMIP3 ensemble over Europe from a regional climate modeler's point of view". In: *Journal Of Climate* (submitted).

R Development Core Team (2009). *R: A Language and Environment for Statistical Computing.* ISBN 3-900051-07-0. R Foundation for Statistical Computing. Vienna, Austria. URL: `http://www.R-project.org`.

Reichle, R. H. and R. D. Koster (2004). "Bias reduction in short records of satellite soil moisture". In: *Geophysical Research Letter* 31.

Roeckner, E. et al. (2003).
The atmospheric general circulation model ECHAM5: Part 1. Tech. rep.
Max Planck Institute for Meteorology.

Rowell, D. P. (2006). "A demonstration of the uncertainty in projections of UK climate change resulting from regional model formulation". In: *Climatic Change* 79, 243–257.

Rummukainen, M. (1997). *Methods of statistical downscaling of GCM simulations.* Reports Meteorology and Climatology 80.
SE-601 76 Norrkping, Sweden: Swedish Meteorological and Hydrological Institute.

Sachs, L. and J. Hedderich (2006). *Angewandte Statistik: Methodensammlung mit R.* 12. Aufl. Springer, Berlin. ISBN: 3540321608.

Storch, H. von and F. W. Zwiers (2002). *Statistical analysis in climate research.* Cambridge University Press, 500pp. ISBN: 9780521012300.

Storch, H. von, A. Navarra, and C. of the European Communities (1999*a*).
Analysis of climate variability. Springer, 366pp. ISBN: 9783540663157.

Storch, H. von, B. Hewitson, and L. Mearns (1999*b*).
Review of Empirical Downscaling Techniques. [Online; accessed 19-April-2010].
URL:
`http://regclim.met.no/rapport_4/presentation02/presentation02.htm`.

Suklitsch, M., A. Gobiet, H. Truhetz, N. K. Awan, H. Göttel, and D. Jacob (2010). "Error characteristics of high resolution regional climate models over the Alpine area". In: *Climate Dynamics* (accepted).

Swift Jr., L. W. and H. T. Schreuder (1981).
"Fitting Daily Precipitation Amounts Using the S_b Distribution".
In: *Monthly Weather Review* 109, 2535–2541.

Taylor, K. E. (2001).
"Summarizing multiple aspects of model performance in a single diagram".
In: *Journal of Geophysical Research* 106, 7183–7192.

Themeßl, M. J., A. Gobiet, and A. Leuprecht (2010).
"Empirical-statistical downscaling and error correction of daily precipitation from regional climate models". In: *International Journal of Climatology* (accepted).

Themeßl, M. and A. Gobiet (2010). "Empirical-statistical downscaling and model error correction of daily temperature and precipitation from regional climate simulations and the effects on climate change signals".
EGU General Assembly, 02 – 07 May, 2010, Vienna.

Trnka, M., Z. Zalud, J. Eitzinger, and M. Dubrovsky (2005). "Global solar radiation in Central European lowlands estimated by various empirical formulae". In: *Agricultural and Forest Meteorology* 131, 54–76.

Tukey, J. W. (1977). *Exploratory Data Analysis.* Addison-Wesley Publishing Co. Inc. ISBN: 0201076160.

UPPALA, S. M. et al. (2005). "The ERA-40 re-analysis". In: *Quarterly Journal of the Royal Meteorological Society* 131, 2961–3012.

Vlček, O. and R. Huth (2009). "Is daily precipitation Gamma-distributed? Adverse effects of an incorrect use of the Kolmogorov–Smirnov test". In: *Atmospheric Research* 93, 759–766.

Wilby, R. L. (1997). "Non-stationarity in daily precipitation series: Implications for GCM down-scaling using atmospheric circulation indices". In: *International Journal of Climatology* 17, 439–454.

Wilby, R. and T. Wigley (1997). "Downscaling general circulation model output: a review of methods and limitations". In: *Progress in Physical Geography* 21.4 (Dec. 1997), 530–548. URL: `http://ppg.sagepub.com/content/21/4/530.abstract`.

Wilks, D. S. (2006). *Statistical Methods in the Atmospheric Sciences.* 2nd ed. Academic Press. ISBN: 0127519661.

Wood, A. W., L. R. Leung, V. Sridhar, and D. P. Lettenmaier (2004). "Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs". In: *Climatic Change* 62, 189–216.

Yao, A. Y. M. (1974). "A Statistical Model for the Surface Relative Humidity". In: *Journal of Applied Meteorology* 13, 17–21.

ZAMG (2010). *Wetter- und Klima-Stationen.* [Online; accessed 18-January-2010]. URL: `http://www.zamg.ac.at/wir_ueber_uns/datenmanagement/stationsnetz/`.

Abstract:
In this thesis, the empirical-statistical Downscaling and Error Correction method called Quantile Mapping (QM) is used to post-process regional climate model (RCM) output. It has so far mostly been applied to the climate parameters temperature and precipitation, whether on monthly or daily basis. Here, the focus is on evaluating the results of QM applied to daily temperature, precipitation, relative humidity and global radiation related to two point-scale observations located in Vienna and Innsbruck. Three different validation approaches are used to analyze the error characteristics of the uncorrected and post-processed data. Furthermore, a scenario run (REMO57-A1B) is post-processed to analyze the Climate Change Signal (CCS). The findings for the first two validation experiments ("calibration identical validation", "cross-validation") show an improvement of the whole probability density function. The split-sample test displays advanced results for temperature, precipitation and relative humidity, but almost no improvement for the parameter global radiation. The CCS of the corrected data does not differ too much from the uncorrected CCS and it does not show a better agreement with the observed CCS. The results demonstrate that it is useful and necessary to post-process temperature, precipitation and relative humidity, but global radiation is a more demanding parameter. Long time periods and high quality observations would be required to capture the whole variability within one climate parameter. In general, QM used with empirical data is applicable to any climate parameter and can provide point-scale data.

Zum Inhalt:
In dieser Arbeit wird die empirisch-statistische Downscaling und Fehlerkorrektur-Methode Quantile Mapping (QM) für die Nachbearbeitung regionaler Klimamodelldaten (RCM) verwendet. Diese ist bis jetzt meist auf die Klimaparameter Temperatur und Niederschlag entweder auf Tages- oder Monatsbasis angewendet worden. Der Schwerpunkt der Arbeit liegt in der Bewertung der nachbearbeiteten Daten für Temperatur, Niederschlag, relative Feuchte und Globalstrahlung auf Tagesbasis bezogen auf zwei Messstationen in Wien und Innsbruck. Drei unterschiedliche Validierungsansätze finden für die Analyse der Fehlercharakteristiken der unkorrigierten und nachbearbeiteten Daten Verwendung. Weiters wird ein Szenario Lauf (REMO57-A1B) nachbearbeitet, um das Klimawandelsignal (CCS) zu untersuchen. Die Resultate der ersten beiden Validierungsexperimenten („Kalibrierung ident Validierung", „Kreuzvalidierung") geben eine Verbesserung der gesamten Wahrscheinlichkeitsdichtefunktion wieder. Der Split-Sample Test zeigt Verbesserungen bei den Parametern Temperatur, Niederschlag und relativer Feuchte, aber kaum Verbesserungen bei Globalstrahlung. Das CCS hat sich zwischen den korrigierten und unkorrigierten Daten wenig geändert und zeigt keine eindeutig bessere Übereinstimmung mit dem gemessenen CCS. Die Ergebnisse zeigen, dass es sinnvoll und notwendig ist Temperatur, Niederschlag und relative Feuchte nachzubearbeiten, aber Globalstrahlung scheint ein anspruchsvollerer Parameter zu sein. Lange Zeitreihen und qualitativ hochwertige Messdaten sind notwendig um die gesamte Variabilität innerhalb eines Parameters abzudecken. QM kann generell auf die empirischen Daten jeglicher Klimaparameter angewendet werden und dabei auch punktskalige Daten liefern.

Wegener Center for Climate and Global Change
University of Graz
Leechgasse 25
A-8010 Graz, Austria
www.wegcenter.at
ISBN 978-3-9502940-9-5