

Oliver Gutjahr

Vom Fachbereich VI
(Raum- und Umweltwissenschaften)
der Universität Trier
zur Verleihung des akademischen Grades
Doktor der Naturwissenschaften (Dr. rer. nat.)
genehmigte Dissertation

**Climate simulations with the regional climate
model COSMO-CLM for the Saar-Lor-Lux region
and Rhineland-Palatinate**

Betreuender: Univ.-Prof. Dr. Günther Heinemann

Berichterstattende:
Univ.-Prof. Dr. Günther Heinemann
Prof. Dr. Markus Casper

Datum der wissenschaftlichen Aussprache

17.12.2013

Trier, 2014

Abstract

High-resolution projections of the future climate are required to assess climate change realistically at a regional scale. This is in particular important for climate change impact studies since global projections are much too coarse to represent local conditions adequately. A major concern is thereby the change of extreme values in a warming climate due to their severe impact on the natural environment, socio-economical systems and the human health. Regional climate models (RCMs) are, however, able to reproduce much of those local features.

Current horizontal resolutions are about 18-25 km, which is still too coarse to directly resolve small-scale processes such as deep-convection. For this reason, projections of a possible future climate were simulated in this study with the regional climate model COSMO-CLM at horizontal resolutions of 4.5 km and 1.3 km for the region of Saarland-Lorraine-Luxemburg and Rhineland-Palatinate for the first time. At a horizontal scale of about 1 km deep-convection is treated explicitly, which is expected to improve particularly the simulation of convective summer precipitation and a better resolved orography is expected to improve near surface fields such as 2 m temperature.

These simulations were performed as 10-year long time-slice experiments for the present climate (1991–2000), the near future (2041–2050) and the end of the century (2091–2100). The climate change signals of the annual and seasonal means and the change of extremes are analysed with respect to precipitation and 2 m temperature and a possible added value due to the increased resolution is investigated. To assess changes in extremes, extreme indices have been applied and 10- and 20-year return levels were estimated by "peak-over-threshold" models.

Since it is generally known that model output of RCMs should not directly be used for climate change impact studies, the precipitation and temperature fields were bias-corrected with several quantile-matching methods. Among them is a new developed parametric method which includes an extension for extreme values and is hence expected to improve the correction. In addition, the impact of the bias-correction on the climate change signals and on the extreme value statistics was investigated.

The results reveal a significant warming of the annual mean by about $+1.7^{\circ}\text{C}$ until 2041–2050 and $+3.7^{\circ}\text{C}$ until 2091–2100, but considerably stronger signals of up to $+5^{\circ}\text{C}$ in summer in the Rhine Valley. Furthermore, the daily variability increases by about $+0.8^{\circ}\text{C}$ in summer but decreases by about -0.8°C in winter. Consequently, hot extremes increase moderately until the mid of the century but strongly thereafter, in particular in the Rhine Valley. Cold extremes warm continuously in the complete domain in the next 100 years but strongest in mountainous areas.

The change signals with regard to annual precipitation are of the order $\pm 10\%$ but not significant. Significant, however, are a predicted increase of $+32\%$ of the seasonal precipitation in autumn until 2041–2050 and a decrease of -28% in summer until 2091–2100. No significant changes were found for days with intensities $> 20\text{ mm day}^{-1}$, but the results indicate that extremes with return periods ≤ 2 years increase as well as the frequency and duration of dry periods.

The bias-corrections amplified positive signals but dampened negative signals and considerably reduced the power of detection. Moreover, absolute values and frequencies of extremes were altered by the correction, but change signals remained approximately constant. The new method outperformed

other parametric methods, in particular with regard to extreme value correction and related extreme indices and return levels. Although the bias correction removed systematic errors, it should be treated as an additional layer of uncertainty in climate change studies.

Finally, the increased resolution of 1.3 km improved predominantly the representation of temperature fields and extremes in terms of spatial heterogeneity. The benefits for summer precipitation were not as clear due to a severe dry-bias in summer, but it could be shown that in principle the onset and intensity of convection improves.

This work demonstrates that climate change will have severe impacts in this investigation area and that in particular extremes may change considerably. An increased resolution provides thereby an added value to the results. These findings encourage further investigations, for other variables as for example near-surface wind, which will be more feasible with growing computing resources. These analyses should, however, be repeated with longer time series, different RCMs and anthropogenic scenarios to determine the robustness and uncertainty of these results more extensively.

Zusammenfassung

Um eine möglichst realistische Erfassung des regionalen Klimawandels zu ermöglichen, werden hochaufgelöste Projektionen des zukünftigen Klimas benötigt. Dies ist insbesondere für Klimafolgestudien von Bedeutung, da globale Projektionen viel zu grob aufgelöst sind, um lokale Bedingungen angemessen abbilden zu können. Ein Hauptinteresse besteht dabei an potentiellen Änderungen von Extremwerten in einem sich erwärmenden Klima aufgrund ihrer teils verheerenden Auswirkungen auf die Umwelt, sozio-ökonomischen Systeme und die menschliche Gesundheit. Regionale Klimamodelle (RCM) sind hingegen in der Lage diese lokalen Besonderheiten zu simulieren.

Gegenwärtige Auflösungen von etwa 18-25 km sind jedoch räumlich immer noch nicht ausreichend aufgelöst, um kleinskalige Prozesse wie Konvektion direkt zu simulieren. Aus diesem Grund werden in dieser Arbeit erstmals Projektionen des zukünftigen Klimas mithilfe des Regionalmodells COSMO-CLM mit Auflösungen von 4.5 km und 1.3 km für die Region Saarland-Lorraine-Luxemburg und Rheinland-Pfalz simuliert. Bei einer Skala von etwa 1 km kann Konvektion aufgelöst werden, wodurch eine bessere Simulationen von konvektivem Niederschlag erwartet wird und durch die feiner aufgelöste Orographie bodennahe Felder wie die 2 m-Temperatur verbessert werden.

Die Simulationensdauer beträgt je 10 Jahre $\frac{1}{4}$ des gegenwärtigen Klima (1991–2000), die nahe Zukunft (2041–2050) und das Ende des Jahrhunderts (2091–2100). Analysiert werden die Klimaänderungssignale der jährlichen und saisonalen Mittelwerte, sowie der Extreme von Niederschlag und 2 m Temperatur. Untersucht wird zudem der Mehrwert einer erhöhten Auflösung. Änderungen von Extremwerten werden anhand von Indizes und 10- und 20-jährigen Wiederkehrereignissen aus *peaks-over-threshold* Modellen analysiert.

Es ist generell bekannt, dass Modelldaten aus regionalen Klimamodellen nicht direkt als Eingangsdaten für Klimafolgenmodelle verwendet werden sollten, weshalb die Niederschlags- und Temperaturfelder zudem mit verschiedenen Korrekturverfahren des *quantile-matchings* korrigiert wurden. Unter diesen befindet sich auch ein neu entwickeltes Verfahren, das eine Erweiterung speziell für Extreme beinhaltet und die Korrektur dieser verbessert. Weiterhin wird die Auswirkung der Bias-Korrektur auf die Klimaänderungssignale und die Extremwertstatistik untersucht.

Die Ergebnisse zeigen eine signifikante Erwärmung im Jahresmittel von etwa $+1.7^\circ\text{C}$ bis 2041–2050 und von $+3.7^\circ\text{C}$ bis 2091–2100. In flachen, tiefliegenden Regionen, wie dem Oberrheingraben, kann diese Erwärmung im Sommer auf $+5^\circ\text{C}$ ansteigen. Die Variabilität der Tagestemperaturen steigt zudem im Sommer um etwa $+0.8^\circ\text{C}$ an, während sie sich im Winter um etwa -0.8°C reduziert. Heiße Extreme nehmen daraufhin bis zur Mitte des Jahrhunderts moderat zu, steigen danach aber stark an, wiederum vor allem im Oberrheingraben. Kalte Extreme erwärmen sich hingegen kontinuierlich in der gesamten Region in den nächsten 100 Jahren, am stärksten in Gebirgen.

Die Klimaänderungssignale für den Jahresmittelniederschlag betragen etwa $\pm 10\%$ und sind nicht signifikant. Signifikant ist hingegen eine prognostizierte Zunahme von $+32\%$ des saisonalen Niederschlags im Herbst bis 2041–2050 sowie eine Abnahme von -28% im Sommer bis 2091–2100. Es konnte keine Zunahme von Niederschlagstagen mit Intensitäten >20 mm pro Tag nachgewiesen werden.

Die Ergebnisse weisen jedoch darauf hin, dass sowohl Extreme mit Jährlichkeiten von < 2 Jahren als auch die Frequenz und Dauer von Trockenperioden zunehmen.

Die Bias-Korrekturen führten zu einer Verstärkung der positiven und einer Abschwächung der negativen Signale. Dabei wurden Absolutwerte und Häufigkeiten von Extremen verändert, die relativen Signale blieben jedoch weitestgehend konstant. Das neue Verfahren war dabei anderen parametrischen Verfahren überlegen, insbesondere bei der Korrektur von Extremwerten. Obwohl die Bias-Korrektur systematische Fehler entfernt, sollte die Anwendung solcher Verfahren als eine weitere Unsicherheitskomponente in Klimawandelstudien betrachten.

Eine erhöhte Auflösung von 1.3 km verbesserte vor allem die Temperatur-Felder und Extreme hinsichtlich der räumlichen Variabilität. Ein Mehrwert für die Simulation des konvektiven Sommerniederschlags konnte aufgrund eines deutlichen Trockenfehlers nicht zufriedenstellend gezeigt werden. Der Beginn und die Intensität der Konvektion verbessern sich jedoch prinzipiell.

Mit dieser Arbeit konnte gezeigt werden, dass der Klimawandel deutliche Auswirkungen im Untersuchungsgebiet haben wird. Insbesondere zeichnet sich eine drastische Änderung in den Temperaturextremen ab, die einer wesentlich erhöhten Erwärmung unterliegen als die Mittelwerte. Durch eine erhöhte räumliche Auflösung von regionalen Klimamodellen kann ein Informationsgewinn in den Ergebnissen erzielt werden. Diese Erkenntnis bekräftigt weiterführende Untersuchungen, auch von anderen Variablen wie z.B. dem bodennahen Wind, die mit wachsenden Computerressourcen auch realisierbarer werden. Diese Analysen sollten jedoch mit längeren Zeiträumen, anderen RCM und anthropogenen Szenarien wiederholt werden, um die Robustheit und Unsicherheit der Ergebnisse weitergehend zu untersuchen.

Contents

Abstract	i
1 Introduction	1
1.1 Climate change and modelling	2
1.1.1 Climate change detection and uncertainty	2
1.1.2 Climate models	4
1.2 Bias-correction	6
1.2.1 Origin of the bias	6
1.2.2 Bias-correction methods	7
1.2.3 Quantile-matching	9
1.3 Extreme value analysis	10
1.3.1 Extreme value analysis under stationarity	10
1.3.2 Extreme value analysis under non-stationarity	11
1.3.3 Multivariate extreme value theory	13
1.4 Aims and outline of this study	13
2 Climate model and data sets	15
2.1 The regional climate model COSMO-CLM	15
2.1.1 Model structure	15
2.1.2 Physical parameterizations	16
2.1.3 Model configuration	19
2.2 Data sets	20
2.2.1 Initial and lateral boundary data	20
2.2.2 CCLM runs	20
2.2.3 Observational data	22
2.2.4 Short evaluation of the forcing data	24
3 Statistical methods and bias correction	27
3.1 Climate change signal detection	27
3.1.1 Statistical tests and the field significance	27
3.1.2 Power analysis of the t-test	28
3.2 Bias-correction	29
3.2.1 Quantile-matching methods	29
3.2.2 Cross-validation	32
3.2.3 MAEx	32
3.3 Extreme value analysis	33
3.3.1 Extreme indices	33

3.3.2	Fitting of a POT model	34
4	Evaluation of the hindcast simulations	41
4.1	Evaluation of the hindcast runs	41
4.1.1	Annual means	41
4.1.2	Monthly means	45
4.1.3	Taylor diagrams	47
4.1.4	Standard deviation	49
4.1.5	Histograms and diurnal cycle	50
4.2	Summary and discussion	57
5	Bias correction of the hindcast and climate simulations	59
5.1	Cross-validation	59
5.2	Correction of precipitation and temperature fields	61
5.2.1	Precipitation	61
5.2.2	2 m temperature	66
5.3	Summary and discussion	68
6	Climate change signal analysis	71
6.1	Climate change signals from uncorrected model data	71
6.1.1	Annual signals	71
6.1.2	Seasonal signals	74
6.2	Climate change signals from bias-corrected model data	81
6.2.1	Effect on precipitation signals	81
6.2.2	Effect on temperature signals	87
6.3	Summary and discussion	87
7	Extreme value analysis of precipitation and temperature fields	90
7.1	Extreme indices	91
7.1.1	Evaluation of the hindcast	93
7.1.2	Extreme indices from the climate simulations	100
7.1.3	Impact of the bias-correction	111
7.2	Return levels from the climate simulations	114
7.2.1	Precipitation	114
7.2.2	Maximum and minimum temperature	120
7.2.3	Impact of the bias-correction on the precipitation return levels	123
7.3	Summary and discussion	125
8	Conclusion and outlook	129
8.1	Future predictions	130
8.1.1	Temperature	130
8.1.2	Precipitation	130
8.2	Resolution effect	131
8.3	Bias-correction effect	131
8.4	Limitations and future suggestions	132
	List of Tables	134
	List of Figures	135

List of Abbreviations and Symbols	138
A CCLM configurations	141
B Additional information for chapter 4	147
C Additional information for chapter 6	149
D Additional information for chapter 7	151
Bibliography	155
Acknowledgements	170
Curriculum Vitae	171

Introduction

Climate change and its implications on the system earth and the human society is one of the major challenges of the 21st century. Since the early 20th century, the mean surface temperature of the Earth is rising. Due to this Global Warming, the mean surface temperature is expected to rise about $+4^{\circ}\text{C}$ in Mid-Europe until the end of the century. In the last report of the Intergovernmental Panel on Climate Change (IPCC 2007) the main source of this Global Warming could be identified to be increasing greenhouse gases resulting from human activities. Related to this Global Warming are also changes in precipitation and the water cycle, which is expected to intensify. Furthermore, the occurrence of extreme weather is expected to increase with rising temperatures and may lead to more frequent and intense heavy precipitation events, flooding, heat-waves or drought-spells. An increase of extreme events may have severe impacts on ecological and socio-economical systems.

Yet these impacts will not occur similarly in every region but will vary in frequency, intensity and spatial extent on a regional to local scale (Sillmann and Roeckner, 2008). Moreover, the impacts of the climate change are afflicted with uncertainties, since the climate change signal is a composite of natural climate variations and the human footprint. It is of fundamental importance to quantify these uncertainties not only for detection or attribution purposes but also for developing adaptation strategies (Deser et al., 2012). However, this is a difficult task because it is unclear how humanity will behave in the next 100 years. The best available tools to study climate change are therefore numerical models. They are based on fundamental thermo-dynamical equations and enable the simulation of a potential future climate with respect to certain assumptions about the development of the human society. Climate models can be set up at global or regional scales with different horizontal resolutions. The methodological philosophy is then to simulate the present climate and a future projection with an assumed greenhouse gas scenario. This procedure allows to study the influences of increased greenhouse gas concentrations caused by humanity on the future climate isolated from the natural variability. However, there are some drawbacks (Deser et al., 2012; Hawkins and Sutton, 2009; Tebaldi and Knutti, 2007): (i) it is unclear how much the greenhouse gas concentrations will increase (forcing uncertainty), (ii) internal variability arises from non-linear dynamical processes intrinsic to the atmosphere (climate noise) and (iii) climate models are afflicted with errors and models react differently on a given forcing (model uncertainty).

A key challenge is to give physically based estimations of the future climate on the one hand but also to quantify the associated uncertainties on the other. However, to capture the whole uncertainty requires the use of many climate models with different forcing data and different scenarios. This task can only be accomplished by so called ensemble runs in large-scale projects like PRUDENCE (Christensen

and Christensen, 2007), ENSEMBLES (van der Linden and Mitchell, 2009) or CORDEX (Giorgi et al., 2009). The expected gain of such a multi-model combination of general circulation models (GCMs) and dynamically nested regional climate models (RCMs) is that the ensemble outperforms single models (Buser et al., 2009), because individual models are overconfident (Weigel et al., 2008). By the use of many individual models the ensemble spread is widened while the average ensemble-mean error is reduced.

However, the horizontal resolution of such large-scale projects (approximately 18-25 km) lag behind the currently possible scales. Significantly finer scales of down to 1 km are in the focus of ongoing research in stand-alone high-resolution experiments (Knote et al., 2010; Kendon et al., 2012; Hohenegger et al., 2008; Prein et al., 2013). Such climate simulations on convection permitting scales are very computational cost-intensive, though. Up to now, there are no coordinated ensemble projects in progress at these scales. At the 1 km scale the orography is better resolved and an explicit treatment of sub-scale processes and dynamics (Hohenegger et al., 2008), such as deep-convection, becomes possible. In particular, the simulation of summer precipitation could benefit from an improvement in the convection dynamics. A more accurate representation of orography may not only improve the triggering of convection and associated wind- and precipitation fields, but will most likely affect the near-surface temperature. In coarser climate simulations the convection process is not resolved explicitly but parametrized in simple terms, which constitutes a source for errors (Hohenegger et al., 2008).

Within this thesis, the regional climate model COSMO-CLM ("CONsortium of SMall-Scale MOdelling" - "Climate Local-area Modelling") was set up for the state of Rhineland-Palatinate and Saarland in Germany, the department of Lorraine in France and Luxemburg (Saar-Lor-Lux & RLP). The present climate and the future projections for the mid and end of the 21st century were simulated on very high resolutions of 4.5 km and 1.3 km. Changes in the annual and seasonal mean of precipitation and 2 m temperature and furthermore the changes of extremes in terms of frequency and intensity were analysed.

To investigate the potential added value of the 1.3 km resolved climate simulations, all analyses were performed for both resolutions. In addition, an essential part consisted of correcting systematic model errors and to investigate the effect of the bias-correction on the climate change signals of means and extreme values.

The next sections briefly review the usage of climate models in climate change studies (1.1), the procedure of bias-correcting model errors (1.2) and the application of extreme value theory to model output (1.3).

1.1 Climate change and modelling

1.1.1 *Climate change detection and uncertainty*

"'Climate change' [...] [is defined as] a change in the state of the climate that can be identified (e.g., using statistical tests) by changes in the mean and/or variability of its properties, and that persists for an extended period, typically decades or longer" (Hegerl et al., 2007). It is furthermore a combination of responses from internal processes and/or external forcing. External forcing can be of natural origin, e.g. due to changes in solar radiation or because of volcanic eruptions or may be caused by anthropogenic induced processes, e.g. changes in the composition of the atmosphere as a result of increased greenhouse gas concentrations (Hegerl et al., 2007; Deser et al., 2012). In contrast, internal processes or variabilities (also called "climate noise" (e.g. Madden, 1976; Schneider and Iii, 1994)) are always present over all time scales, ranging from instantaneous over weekly to annual or even millennial time scales. Slow reacting climate components are for example oceans or large ice sheets, which dampen or integrate high-frequency atmospheric fluctuations into delayed responses (Hasselmann, 1976) and thus play an

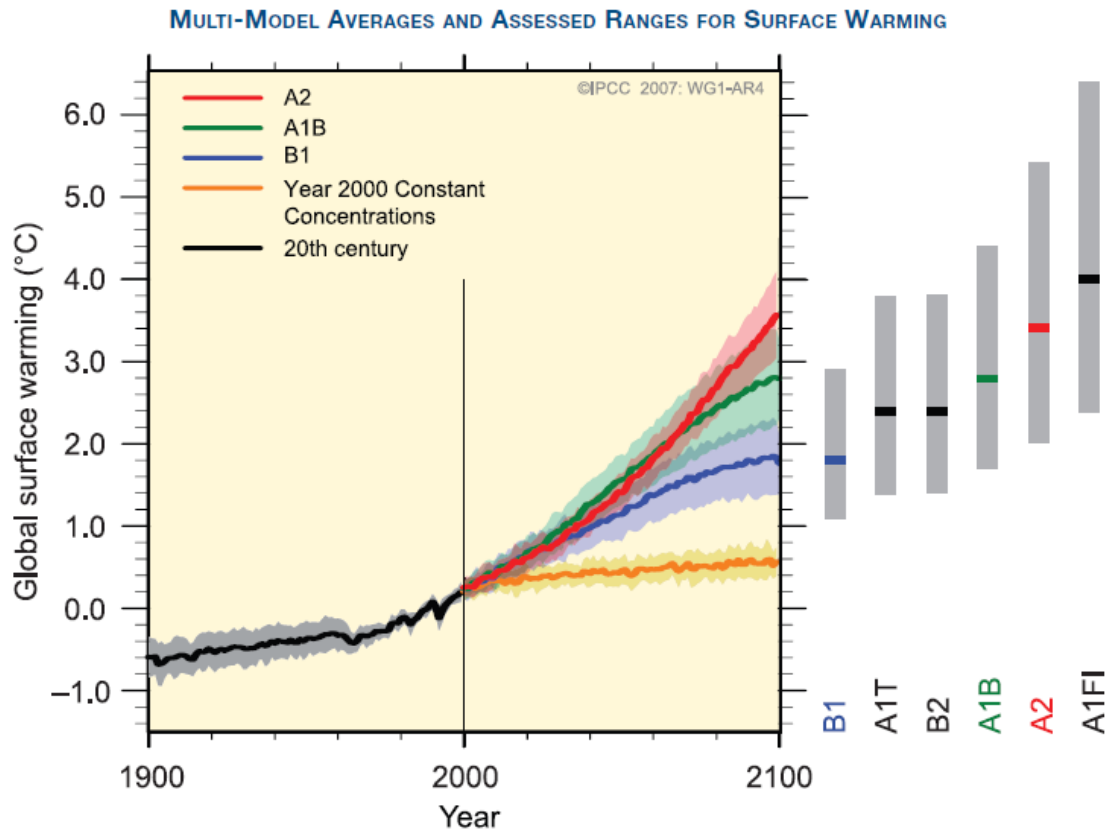


Figure 1.1 Projected global warming of the near surface temperature according to different SRES-scenarios (IPCC, 2007, p.14).

important role at time scales of decades or centuries. Consequently, the climate is able to provide climatic changes on long time scales without any external forcing (Solomon et al., 2007b).

The separation of external influences from internal variability is a difficult task, which has to be performed carefully and is to be based on physical understanding and principles, e.g. concept models or numerical climate models (Hegerl et al., 2007). With the use of climate models it is possible to distinguish between natural and anthropogenic induced climate change (Hegerl et al., 2007); however, this is associated with uncertainty. This uncertainty of the projected climate change is a combination of responses from three main sources (Deser et al., 2012), whose individual attributions are not known. These three contributions are the anthropogenic forcing, the model response and the external/internal natural forcing (Foley, 2010) (a prominent example is the El Niño phenomenon (Collins, 2007)), which can occur on the same spatial and temporal scales (e.g. Hawkins and Sutton, 2009; Solomon et al., 2007b; Tebaldi and Knutti, 2007; Deser et al., 2012). The reasons for these uncertainties are that it is not known how the internal variability will look like in the future (or even in present time) or how the anthropogenic emission will evolve in the next century. This anthropogenic uncertainty concerns mainly the future developments of population density, social and economic sectors, energy or land-use changes and connected greenhouse gas emissions (Deser et al., 2012). Assumptions of some possible future trajectories of such an anthropogenic forcing were summarized in so-called Special Report on Emission scenarios (SRES) by the IPCC Fourth Assessment Report (Solomon et al., 2007b). The impact on the global mean surface temperature can be seen in Fig. 1.1. For instance, the A1B-scenarios describes a

rapid economic growth, a peak of global population growth in the mid-century, a rapid introduction of new technologies and a balanced usage of fossil and non-fossil energy sources. The storyline of A2 and B1 orientate on slower and faster adaptations of the fertile rate and on more efficient technology.

Uncertainties arise from the climate model itself as well (see section 1.2.1 for details) because of the simplifications that have to be met in numerical models (Foley, 2010). Even if models are forced with the same anthropogenic scenario, they produce different results due to different physical and numerical formulations (Deser et al., 2012; Collins, 2007).

As mentioned above, the climate change signal is a composite of different sources, which act partly on the same scales. *Detection* of climate change means to show that an observed change is statistically different from what can be explained by natural variability alone (Hegerl et al., 2001). Thus, climate change detection is a statistical "signal-to-noise" problem (Hegerl et al., 2001). However, a significant change does not imply that the causation is understood. This signal-to-noise problem is approached by applying statistical methods. For this purpose, a wide range of methods have been applied in the literature. These methods include for example, "optimal fingerprints" methods (Hasselmann, 1993) and modifications thereof (e.g. Hegerl et al., 1996; Tett et al., 1999; Stott et al., 2001; Stone et al., 2007), regression methods (Zwiers and Zhang, 2003) or temporal optimal detection (Ribes et al., 2009). At the regional scale, the separation of climate signals from noise becomes more difficult (Ribes et al., 2009) (see a review for instance in Hegerl et al. (2007)).

In this thesis, there is only one possible climate-path simulated: the SRES-scenario A1B. Because of this single realisation, it is not possible to draw conclusions regarding the model or anthropogenic uncertainty. It is hence not the intention to separate different forcing of the climate change but to analyse the impact of the combined forcing on the future climate. This is achieved with classical tests for differences in the mean or distribution, such as a t-test or Wilcoxon-Whitney test.

1.1.2 Climate models

The only tool which is available to simulate physically based potential developments of the future climate are General Circulation Models (GCMs) and Regional climate Models (RCMs). These numerical models simulate the climate of the whole earth and for regions of interest, respectively. GCMs are based on fundamental thermo-dynamical equations and thus provide physical consistent and spatially coherent projections of the present and potential future climate. If a climate model simulates the present climate confidently, the next step is to apply it for the future climate. However, certain assumptions on the forcing factors have to be met which influence the future development. These forcing factors (SRES scenarios) were discussed above and differ mainly with respect to the evolution of greenhouse gas concentrations. GCMs are used to simulate climate on a global scale. Although they are capable to reproduce large-scale circulations and processes, they cannot predict features on the small scale for specific regional areas (Buser et al., 2009). Typical state-of-the-art GCMs have a formal horizontal spacing of about 100 – 200 km but their effective resolution is only about 6–8 grid distances (Rummukainen, 2010). Higher resolutions are in principal possible (Mizuta et al., 2006) but the computational cost makes such models currently not feasible (Rummukainen, 2010). A compromise will be used in the new generation GCMs as for example in the ICON model¹ of the German Weather Service (DWD) and the Max-Planck-Institute for Meteorology. While a coarse resolution is used for most regions of the globe a much finer scale is set up for the focus region of Germany.

For adequate climate change impact studies, projections on the regional scale are needed. To achieve this, the scale-gap from the global model to the regional model has to be bridged. Several downscaling methods have been developed for this purpose, see for instance the reviews in Fowler

¹ <http://www.mpimet.mpg.de/en/science/models/icon.html>

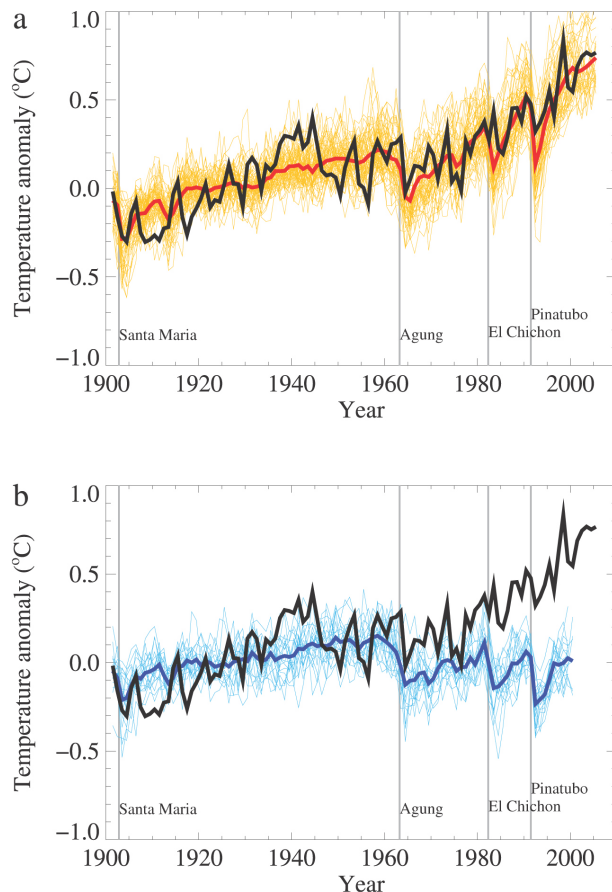


Figure 1.2 Comparison of global mean temperature anomalies of observations (black solid line) and climate models that (a) include natural and anthropogenic forcing (ensemble mean: red solid line, individual models: thin red lines) and (b) only natural forcing (ensemble mean: blue solid line, individual models: thin blue lines). Major volcanic eruptions are marked by grey vertical lines (adapted from Hegerl et al. (2007, p.684, Fig. 9.5) based on Stott et al. (2006, p.3059, Fig. 1)).

et al. (2007), Maraun et al. (2010) and Benestad et al. (2008). Downscaling can be accomplished by either using statistical methods which relate GCM model output (e.g. the mean sea level pressure) to local observations (e.g. precipitation) via empirical functions and weather generators or dynamically by nesting a RCM into a GCM in a so-called nesting chain (Giorgi, 1990). A third approach are GCMs with a variable-resolution grid (Fox-Rabinovitz et al., 2008) as mentioned before. More simple approaches, such as the delta-approach, add only the climate change signal from GCMs to observations (Giorgi and Mearns, 1991) or interpolate the coarse grids of the GCM to finer scales (Maurer et al., 2007). Downscaling is either used to amend observations where the coverage is sparse or to refine the informations from GCMs. Statistical downscaling is rather cheap with respect to computational costs but lacks of physical or spatial consistency. A dynamical downscaling requires much more computing resources but provides a physical consistent and spatially coherent representation of the meso-scale climate (Lafon et al., 2013). Furthermore, the RCM is consistent with the driving GCM which is an important characteristic if the output should provide realistic driving data for impact models (Lafon et al., 2013).

Moreover, RCMs add further details to the GCM simulations and are used for research and application studies (Rummukainen, 2010). RCMs are forced by GCM model output at the lateral boundaries.

Consequently, the quality of the RCM strongly depends on the quality of the driving GCM. Additional difficulties arise because of the scale jump at the boundaries which cause instabilities which have to be reduced by damping functions. Otherwise gravity waves are reflected at the boundaries (Rummukainen, 2010) which influence the complete model domain. The damping is done in a so-called sponge zone at the boundaries (usually 4-10 grid points wide, (Rummukainen, 2010)) where the transition from the large to the regional scale takes place. This sponge zone is thus omitted from the subsequent analysis. A current issue of interest is the impact of the domain size on the simulations in RCMs, for instance the well known 'Big Brother/Little Brother' experiment (Denis et al., 2002; Denis and Laprise, 2003). The setup of this study was to first run a RCM with a high resolution over a large domain (Big Brother), then to apply a low-pass filter to coarsen the model output and to force a second RCM over a varying smaller domain with this filtered data at the boundaries (Little Brother). By comparing the 'truth' (Big Brother) with the the Little Brother runs, conclusions on the impact of the domain size, resolution and location can be drawn (Rummukainen, 2010; Denis and Laprise, 2003). For instance, it was found that a scaling factor of 12 (i.e. a resolution which is 12 times finer) is the upper limit to achieve adequate results with a RCM and 12 h as a minimum for updating the boundary data. However, the scale jump in a nesting chain is usually up to 6-8 but often smaller. If large jumps are necessary a multiple-nesting strategy should be applied (Rummukainen, 2010). In principle, the smaller the RCM-domain the larger the dependence on the driving data. In a too large domain, the RCM can act freely and thus deviate in the synoptic pattern from the parent GCM. Usually RCMs are nested in a one-way coupling into a GCM; thus, there is no feedback to the large-scale system. Currently, there are only a few studies whose ambition was to nest RCMs in a two-way-coupling procedure (Lorenz and Jacob, 2005).

GCMs and RCMs are used to detect and to assess climate change, for instance the human 'fingerprint' on the global climate could be identified and made visible by Stott et al. (2006) (see Fig. 1.2). One of the simplest approaches for climate models is to show that "[...] the recently observed global warming is significantly stronger than model-based estimates of the internal variability of the climate system and thus the warming is detectably stronger than internal variability" (Hegerl and Zwiers, 2011). In summary, climate models offer the possibility to systematically study the impact of different forcing sources and intensities on the past, present and future climate. But it has to be remembered that such models are simplifications of the real world and are afflicted with errors due to these simplifications.

1.2 Bias-correction

1.2.1 *Origin of the bias*

Climate models cannot simulate all physical processes which are intrinsic to nature for several reasons: on the one hand, certain processes are still not fully understood or known and on the other hand, difficulties arise from the representation of physical processes in the differential equations. While the first issue addresses a theoretical problem the latter is rather a technical problem. Therefore climate models are only approximations of the natural environment. Further approximations concern the differential equations which have to be solved by numerical methods and achieve only a certain degree of accuracy. Finally, considering that physical processes act on many scales, e.g. micro-physics act on much smaller temporal and spatial scales than, for example, frontal dynamics and that differential equations are solved on discrete grids, it is obvious that sub-scale processes have to be parameterized (Fowler and Ekström, 2009; Frei et al., 2003; Hagemann et al., 2004; Fowler et al., 2007; Themeßl et al., 2012; Maraun et al., 2010). This is either because their scales are much smaller than the grid size or their complexity would result in unfeasible models. Vice versa the spatial resolution of climate models is insufficient to capture all sub-grid-scale processes (Rauscher et al., 2010; Fowler et al., 2007). Parameterizations are hence

introduced to represent the average effect of the small-scale processes on the grid-scale. However, this averaging introduces errors.

These errors originate already within the GCMs (Sharma et al., 2007; Hansen et al., 2006; Feddersen and Andersen, 2005) and are transferred into the nested regional climate models (RCMs) through the lateral boundaries. But the transfer itself is also a source of errors since the scale mismatch between a GCM and a RCM requires a nesting procedure. How the dampening of the boundary problem is done, i.e. which function type is used, is another issue (Rummukainen, 2010). But errors originate from RCMs as well for the same reasons as in the GCM (Rummukainen, 2010). In summary, model errors can arise from three sources: (i) transfer from GCMs to RCMs (boundary problem), (ii) numerics and parametrizations and (iii) insufficiently resolved surface properties (e.g. orography) (Schmidli et al., 2006; Gutjahr and Heinemann, 2013). Especially precipitation is prone to errors due to its highly non-linear nature and spatial variability. Consequently, the accuracy is highly depending on the model resolution, numerical scheme and parameterization (Fowler et al., 2007; Maraun et al., 2010; Fowler and Ekström, 2009; Bachner et al., 2008). Associated with precipitation errors is the parameterization of moist convection (e.g. Tiedtke (1989)), since such convection parameterizations were developed for coarser spatial resolutions in the tropics (Hohenegger et al., 2008). This does not only result in errors of total precipitation but also in a wrong onset and dynamics of convection and cloud formation (Guichard et al., 2004). The correct representation of precipitation depends on numerous processes such as cloud micro-physics, radiative transfer, atmospheric dynamics and boundary layer interactions, which act over a variety of space and time scales (Hagemann et al., 2011).

A bias is defined as a systematic error which is not depending on time. This is a major assumption for all bias-correction (BC) methods (Piani et al., 2010a), although this hypothesis is not trivial (Trenberth et al., 2003). Typical types of biases are the simulation of too many wet-days (Fowler et al., 2007) with low-intensity rain ('drizzle-error') (Bachner et al., 2008; Boberg et al., 2009), incorrect temperature extremes (Ines and Hansen, 2006) or wrong seasonal amplitudes (Teutschbein and Seibert, 2012; Christensen et al., 2008). Bias-correction aims to remove these errors for mainly two reasons: (i) to avoid a quality reduction of the results that would be inevitable if the analysis is based on uncorrected data and (ii) to prevent errors from being transferred from the climate model to subsequent impact models (e.g. hydrological models). This issue was addressed in many recent studies (Piani et al., 2010a; Hagemann et al., 2011; Haddeland et al., 2012; Wood et al., 2004; Christensen et al., 2008; Maraun et al., 2010; Teutschbein and Seibert, 2010). A bias-correction is thus necessary to prevent this error from being transferred (Muerth et al., 2013; Sharma et al., 2007; Hansen et al., 2006). In particular, it was found that bias-correction improves the results of subsequent hydrological models (Muerth et al., 2013; Casper et al., 2012).

1.2.2 *Bias-correction methods*

There is a large diversity of methods available for correcting model biases; a fact that impedes an adequate choice based on the intended purpose (Gudmundsson et al., 2012). Their development, performance and limitations are in the focus of ongoing research. Recently, a discussion arose whether climate models should be corrected or not. Ehret et al. (2012) state that bias-correction methods are often used in an invalid way. For instance, BC violates conservation principles by altering spatio-temporal field consistencies because feedback mechanisms are disregarded if climate variables are corrected separately. Moreover, they argue that bias-correcting climate model output narrows the uncertainty range of simulations, which is not based on physical reasoning and that the time-invariance of transfer functions is at least uncertain (Ehret et al., 2012). However, it is difficult to model time-dependent corrections since there is no information on the future available that could be included to judge whether transfer functions change and in what direction. Although, some approaches have been applied regarding this

issue (e.g. Li et al., 2010; Buser et al., 2009), it has to be kept in mind that corrections aim to provide the best possible output by combining state-of-the-art models and correction methodologies. A more sophisticated step would be to correct the model physics per se, but this is a task for a much longer time scale and is not helpful for end-users on the short term. Bias-correction is thus a short-cut given the current limitations of climate models. As long as the associated uncertainty of such corrected data sets is communicated to the end-user, a clear description of the applied method is given and the results, based on the uncorrected and corrected data sets, are juxtaposed, the application of bias-correction methods is justifiable and consistent with the demands of Ehret et al. (2012) and Maraun et al. (2010).

A good overview of several BC approaches can be found in Themeßl et al. (2012). In general, it is assumed that the model error is stationary, or time invariant. That is, the empirical relationships of the correction functions do not change with time and are thus valid for future projections (e.g. Piani et al., 2010a). A simple methodology, which originated from the hydrological sciences, is the 'delta change approach' (Hay et al., 2000; Fowler and Kilsby, 2007; Graham et al., 2007; Déqué, 2007). Therein, the climate signal is extracted by subtracting e.g. the annual mean precipitation of the present climate from the future projection, so that the model error has been cancelled. Then this obtained signal is added to observational data to construct a time series of the future climate. However, this approach is only applicable for climate change analysis studies but not for correcting present climate simulations. Furthermore, it only accounts for changes in the mean but not for changes in variability or extremes, which are necessary to consider in climate change studies (Hagemann et al., 2011). It is expected that climate change will affect extremes to a larger extent than the means (Trenberth et al., 2003; Katz and Brown, 1992) and hence the correction of the frequency and the magnitude of extremes is of major importance.

More sophisticated methods have been introduced by e.g. Boé et al. (2007); Leander and Buishand (2007); Lenderink et al. (2007); Piani et al. (2010a); Themeßl et al. (2011); Dobler and Ahrens (2008). Themeßl et al. (2011) classifies the bias-correction methods in indirect and direct methods. Indirect methods include for instance multiple linear regressions (e.g. Hay and Clark, 2003) or re-sampling approaches (analogue method) (von Storch and Navarra, 1999) in which large-scale predictors are used to create local weather statistics. Direct methods have been used much more frequently in recent studies, for instance the local intensity scaling approach (LOCI) (Widmann et al., 2003; Leander and Buishand, 2007; Schmidli et al., 2006), which applies scaling factors to the means at the locations of observations or quantile-matching (QM) (Panofsky and Brier, 1968; Boé et al., 2007; Déqué, 2007; Michelangeli et al., 2009; Sennikovs and Bethers, 2009; Wood et al., 2004; Themeßl et al., 2011; Gutjahr and Heinemann, 2013), which relates modelled and observed climate directly by extending the correction from means to the entire distribution.

Most often only precipitation and 2 m temperature are corrected because they are the key-variables for impact models (Christensen et al., 2008; Piani and Haerter, 2012). Haddeland et al. (2012) extended the bias-correction also for radiation, humidity and near-surface wind, but these corrections had less effect on the climate changes projections of their hydrology model. Furthermore, the variables are usually corrected independently. This may lead to inconsistent fields and is a known issue in impact studies (Hagemann et al., 2011; Chen et al., 2011). Recently, Piani and Haerter (2012) showed that it is possible to correct precipitation and temperature jointly with significant improvements. They applied a bias-correction in a copula approach to correct precipitation and temperature from the output of the regional climate model REMO. This could be even achieved by a rather simple relation between temperature and precipitation, that is, even if the temperature is binned into less than 5 interval-classes, the copula is improved significantly. These results show that only a few temperature classes are needed within the copula to correct the precipitation adequately.

1.2.3 *Quantile-matching*

The bias-correction methods which are used in this thesis are all modifications of the quantile-matching approach (Panofsky and Brier, 1968; Boé et al., 2007; Déqué, 2007) and were separately applied to precipitation and temperature. The QM approach was selected because in former performance tests it was one of the top ranked methods and therefore recommended for statistical bias-correction (Gudmundsson et al., 2012; Themeßl et al., 2011). The reasons are the following: QM (i) has a simple and non-parametric configuration and (ii) corrects all moments simultaneously. In the QM, a transfer function is established between observations and model data that assigns a corrected value to every uncorrected value. This transfer function can be constructed in manifold ways. In general, two groups can be distinguished: parametric and non-parametric (empirical) QM. While in the non-parametric QM no assumption on the underlying probability density function (PDF) have to be met (Gudmundsson et al., 2012), parametric QM either assumes a specific parametric distribution family or models a transfer function as a function of some parameters.

This flexibility allows to correct extremes and low intensities differently, a characteristic that is especially necessary in impact applications such as flood risk or drought studies (Hay and Clark, 2003; Dobler and Ahrens, 2008). This would not be possible in a simple linear correction. Because of these qualities, the QM is supposed to be the most suitable choice for the correction of precipitation and temperature with regard to a subsequent extreme value analysis.

However, the original non-parametric QM approach after Panofsky and Brier (1968), had a drawback in correcting extreme values outside the range of the calibration period. Therefore, it has been extended with a linear correction term by Boé et al. (2007) (hereafter referred to as eQM). Additionally, the problem of over-fitting may arise due to the large degrees of freedom, i.e. the distribution is fitted too closely to specific characteristics of the chosen calibration period, which may tamper the correction in the validation period (in particular the future scenario). However, in a cross-validation framework it could be shown that this is not the case (Gudmundsson et al., 2012; Gutjahr and Heinemann, 2013).

Besides this non-parametric or empirical method parametric QM approaches reduce the degrees of freedom considerably but have to disaggregate the correction problem into two parts if applied to precipitation data (Gudmundsson et al., 2012). Most often, this is a mixture of a Bernoulli and a Gamma distribution (e.g. Piani et al., 2010a): the Bernoulli distribution is used to model the probability of precipitation occurrences (i.e. number of wet days) and the Gamma distribution to model the intensities (e.g. Thom, 1968; Cannon, 2008). This method is referred to as gQM in the following. Within this thesis a new parametric QM approach is introduced that mixes a Bernoulli with a Gamma and a General Pareto distribution (GPD) (Gutjahr and Heinemann, 2013). Because of this extension of a extreme value distribution the tails are expected to be corrected more accurately. The GPD requires three additional parameters and thus has more free parameters compared to gQM. However, both methods reduce the degrees of freedom considerably with respect to the empirical correction. On the other hand, due to this reduction the problem of parameter-estimation arises. As a consequence, it could be argued that sufficiently longer time series may be advantageous for fitting the distributions properly to the data in the parametric QM.

A further issue arises from the role of time-scales in the bias-correction of climate model output. Haerter et al. (2010) argue that correcting a variable with respect to one time-scale, e.g. on a daily basis, can deteriorate the characteristics of the variable on another, e.g. monthly values. They propose to use a so-called cascade bias-correction in which transfer functions are successively applied to different scales (from hourly to monthly).

1.3 Extreme value analysis

1.3.1 *Extreme value analysis under stationarity*

The analysis of extreme events and its associated extreme value theory (EVT) has become a major statistical discipline over the last 50 years (Coles, 2001). This is because extremes have an enormous impact on the natural environments, socio-economical systems and the human health (Ghil et al., 2011; Zhang and Zwiers, 2013). Connected, but not inherently, with extremes or severe events are substantial losses, e.g. human lives, financial losses or loss of species. These severity is measured by the risk, which depends on the concurrence probability of an extreme event, the exposure to the event and the vulnerability or damage (Stephenson, 2008). However, the term "extreme" is not consistently used throughout the climate literature, which is why there is no uniform definition (Stephenson, 2008). The term "extreme" can refer to a climate variable or to an impact and the linkage may be not straightforward (Zhang and Zwiers, 2013). Kantz et al. (2006) gave a definition which Ghil et al. (2011) propose to follow. According to Kantz et al. (2006), "[...] extreme events (i) are rare, (ii) they occur irregularly, (iii) they exhibit an observable that takes extreme values and (iv) are inherent to the system under study, rather than being due to external shocks". In the IPCC Special Report on Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation (IPCC, 2012), this definition was refined for climate extremes, stating that "an extreme (weather or climate) event is generally defined as the occurrence of a value of a weather or climate variable above (or below) a threshold value near the upper (or lower) ends of the range of observed values of the variable" (Seneviratne et al., 2012). The latter definition focuses on extremes as part of the distribution tails. Many examination methods of the EVT are based on this definition.

Extreme events exhibit a variety of different attributes so that it is not possible to describe them with a single quantity (Stephenson, 2008). After Stephenson (2008) "[...] such attributes are: (i) rate of occurrence, (ii) magnitude (intensity), (iii) temporal duration and timing, (iv) spatial scale and (v) multivariate dependencies". Therefore, numerous techniques were developed to describe these attributes. The Joint Commission for Climatology (CCI)/Climate Variability and Predictability Project (CLIVAR)/Joint Technical Commission for Oceanography and Marine Meteorology (JCOMM) Expert Team on Climate Change Detection (ETCCDI) defined a set of descriptive indices, which were based on the European Climate Assessment indices (Klein Tank and Können, 2003) and first introduced by Frich et al. (2002). Some of these extremal indices will be used in this thesis and are explained in detail in section 3.3.1.

EVT is applicable to give a sound probabilistic foundation for describing extreme events in a variety of scientific disciplines (Ghil et al., 2011; Beirlant et al., 2004). The EVT is based on asymptotic arguments for sequences of identical and independent distributed (i.i.d.) random variables (Coles, 2001). The theory then describes a limiting distribution after which suitably rescaled maxima follow asymptotically as the sample size $n \rightarrow \infty$. This limiting distribution is the Generalized Extreme Value (GEV) distribution, which is the generalisation of the three "classical" extreme value distributions: Gumbel, Fréchet and Weibull (Coles, 2001). The selection of extremes or maxima out of a sequence is a crucial part of the EVT and no trivial task. Usually, so-called 'block maxima', i.e. the maximum value of certain time intervals (e.g. years, months), are used for assessing and predicting extreme events (Ghil et al., 2011). This is a statistically wasteful approach, though, since a large amount of information is lost for the analysis. In order to exploit more information, a different approach for selecting maxima was introduced. Within the "peaks-over-threshold (POT)" approach maxima are defined as exceedances over a specific threshold. If scaled appropriately they asymptotically follow the Generalized Pareto distribution (GPD) according to the second EVT theorem or Pickands-Balkema-De Haan theorem (De Haan and Ferreira, 2006). In its basic form, the EVT was developed to model and describe extremes in a stationary

climate (Gumbel, 1958). But recent modifications enables the application under non-stationarity as well (see section 1.3.2).

Extreme value models, such as the POT-model, can be used to calculate return levels, i.e. extremes which are likely to occur once in a specified time interval or return period. These models are valuable for two reasons: (i) they can be used to quantify the intensity of extreme events that may occur on much longer return periods than covered by the observations and (ii) confidence intervals, which quantify the uncertainty of return levels, can be calculated from the estimated model parameters. "This is precisely the task of statistical analysis in climate risk research: to quantify the uncertainties, to give error bars or confidence intervals of our estimates and projections" (Mudelsee, 2011). A pioneering study in which EVT have been applied to climate model output of a GCM in a control and future climate was that of Zwiers and Kharin (1998). After Katz (2010) this paper functioned as a catalyst to accept the application of extremes value statistics to numerical model output.

1.3.2 *Extreme value analysis under non-stationarity*

Under climate change it is expected that extremes change as well (Solomon et al., 2007a). Climate models can be used to simulate the future climate and with the help of EVT changes in extremes can be evaluated (e.g. Wehner, 2010; Kharin and Zwiers, 2005). But before the impact of a future emission scenario on the extremes is investigated, it is good to know how extremes changed in the past. This helps to strengthen the confidence in simulating future changes by climate models (Zhang and Zwiers, 2013). For instance, it was found that the late 20th and early 21st-century in Europe was warmer than any time during the last 500 years and that the summer in 2003 was by far the hottest (Luterbacher et al., 2004). For the future in Europe it is projected that heat-waves will occur more often (from 1 heat-wave within 10 years in 1961-1990 to 1 heat-wave every year in 2071-2100; a heat-wave was defined by at least 6 consecutive days where the maximum temperature exceeds the 90th percentile of the period 1961-1990) and that they will have an increased amplitude (about +4°C) (Fischer and Schär, 2010).

From a statistical point of view, climate change introduces non-stationarity, which causes a violation in the EVT because the distribution of maxima does not remain constant, i.e. the assumption of identical distributed values does not hold anymore. After Solomon et al. (2007b), this is a severe problem if climate change effects on extremes are to be evaluated. Fortunately, there has been research undergone to include non-stationarities (e.g. Coles, 2001). Until quite recently, only a few studies applied the EVT in context of climate change (Katz, 2010). A newer study is for instance the work of Cooley (2009) who investigated the effect of a slowly changing climate on the frequency of extreme events. As a catalyst served the paper of Zwiers and Kharin (1998). The adaptation to non-stationarities is usually achieved by making the parameters of the GEV/GPD time-dependent, i.e. introducing trends (Katz, 2010). However, identification of changes in the extremes depend on the applied analysis technique (Trenberth et al., 2007; Trömel and Schönwiese, 2005).

Why extremes are expected to increase in a warmer climate is schematically illustrated in Fig. 1.3. If the probability distribution function of temperature is shifted in its mean to a warmer climate but remains constant in its shape (Fig. 1.3a), then the frequency of hot extremes is increases. If only the shape, e.g. the variance of daily temperature, is increased (Fig. 1.3b) then both cold and hot extremes increase in frequency. If the mean and the variance increases simultaneously, then even more hot extremes may occur (Fig. 1.3c). The change of cold extremes depends in this case on the magnitude of both terms, which act in opposed directions. For instance, a larger variability may balance the reduction of cold extremes, whereas a decrease in variance may accelerate the reduction.

In case of precipitation it is more complicated. Generally, the formation of precipitation extremes requires processes that cause air to rise on the one hand, e.g. through orographic uplifting, unequal heating of the atmosphere or potential vorticity dynamics and on the other hand the presence of moisture

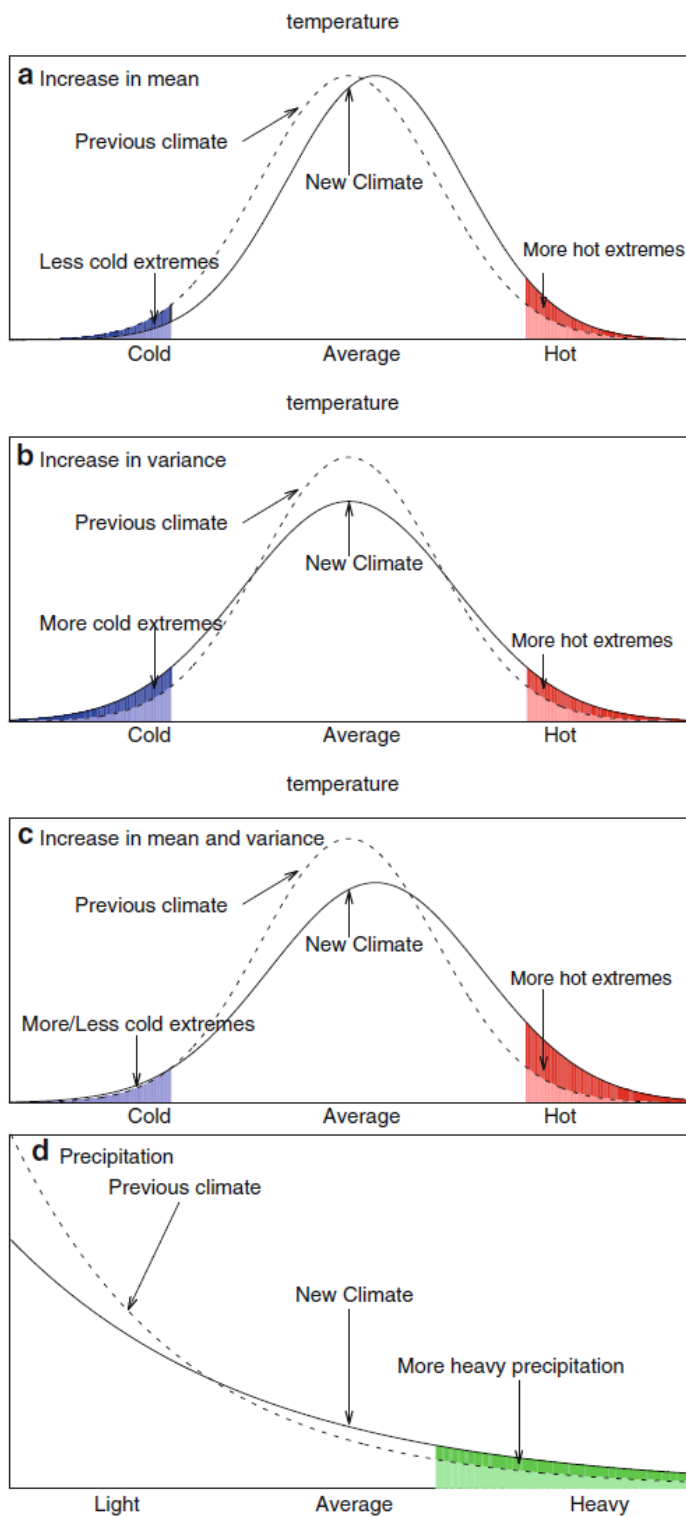


Figure 1.3 Schematic illustration of changes in the probability distribution of temperature and its connected changes in extremes (Zhang and Zwiers, 2013, p.4).

(Trenberth et al., 2003). The potential atmospheric water vapour amount rises due to an increase in temperature according to the Clausius-Clapeyron equation (O’Gorman and Muller, 2010). This increase is about $7\% \text{ K}^{-1}$ (Allan and Soden, 2008; Trenberth et

al., 2003). evaporation (as long as water is present), the potential moisture content of the atmosphere should increase as well. It is, however, unlikely that this increase occurs uniformly, although climate models predict surface temperature warming for most parts of the earth (Trenberth et al., 2003). Furthermore, dynamical processes such as convergence and subsidence influence the local moisture distribution. Low-level moisture convergence is the major process for producing heavy rain events. Under these circumstances, it is possible that the rate of potential moisture increases and even exceeded $7\% \text{ K}^{-1}$ due to latent heat feedbacks (Trenberth et al., 2003).

1.3.3 Multivariate extreme value theory

After Katz (2010), there is a "challenging problem of statistical modelling climate extremes across space", i.e. to evaluate extremes which are spatially dependent, such as in climatic fields of numerical model output. One approach to investigate how extremes vary by location is to fit extreme value distributions to every location separately, i.e. assuming independence between locations (or grid boxes) (e.g. Kharin and Zwiers, 2000; Zwiers and Kharin, 1998; Frei et al., 2006). If the spatial dependency is to be modelled explicitly, methods of the multivariate extreme value analysis have to be applied. However, this is much more complex since the family of multivariate extreme value distributions cannot be summarized parametrically (Cooley, 2009; Ghil et al., 2011), which is why this topic is an active research field (e.g. Resnick, 1987; Heffernan and Tawn, 2004; Beirlant et al., 2004; De Haan and Ferreira, 2006). Spatial data can also be used to reduce the uncertainty of parameter estimation (Cooley, 2009). One possible approach is to combine data from similar locations, for example from neighbouring grid boxes, to extend the time series. This method is called regional frequency analysis (RFA). Good examples are the studies of Fowler et al. (2005) and Ekström et al. (2005), who applied RFA and the GEV on extreme rainfall in the U.K. under climate change.

1.4 Aims and outline of this study

In this study, the impacts of climate change on precipitation and temperature are investigated with the regional climate model COSMO-CLM (CCLM) for the region of Rhineland-Palatinate and Saarland in Germany, Lorraine in France and Luxemburg (SLL&RLP). The climate conditions were simulated as time-slices experiments for the present (1991–2000, C20), the near-future (2041–2050, A1B_1) and the far future (2091–2100, A1B_2). To provide physical consistent and spatially coherent high-resolution fields of atmospheric variables for the SLL&RLP region, the output of the GCM ECHAM5/MPI-OM (Roeckner et al., 2003) was dynamically downscaled in a multiple nesting chain. The first nest of the CCLM was done with a horizontal resolution of 18 km by Hollweg et al. (2008) and was followed by a second nest of the CCLM with a resolution of 4.5 km and ended in a third nest with a resolution of 1.3 km.

The following questions were in the focus of the analyses:

- How does the future climate develop in the SLL&RLP region with respect to mean and extreme values of precipitation and temperature?
- Is there any added value by increasing the horizontal resolution from 4.5 km to 1.3 km?
- Does a bias-correction improve the output of the CCLM?

- What is the impact of the bias-correction on the climate change signals (means and extremes)?

The analyses are performed on the two horizontal scales: 4.5 km and 1.3 km. A resolution of 1.3 km enables an explicit treatment of moist convection processes (or at least parts of it) and is expected to improve the simulation of precipitation. The effect of the horizontal resolution on the regional climate is investigated to rate the added value of such high-resolution climate models, which are computational cost intensive. The CCLM with a resolution of 1.3 km is currently at an experimental stage. Due to such enormous demands of computer sources, there are only a few studies which use climate models on a 1 km scale (e.g. Knote et al., 2010; Kendon et al., 2012) and they were most often performed as time-slice experiments as well. For the region of SLL&RLP there is only one experiment carried out with an older CCLM version by Knote et al. (2010). However, Knote et al. (2010) only simulated the summer months of 10-year periods (1961–1970, 2091–2100), which did not allow the projections of future extremes in other seasons. They found the strongest change in daily minimum temperature extremes for mountainous regions while the strongest change in daily maximum temperature occurs in flat and lowland regions.

An exception is the study of Kendon et al. (2012) who performed simulations at 1.5 km over the United Kingdom for the period 1989–2008. They found that rainfall is simulated much more realistic by the high-resolution run compared to a 12 km run but extremes tend to be too intensive. The bias in the coarser simulation is ascribed to deficiencies in the representation of convection. Furthermore, they state that diurnal cycles are improved.

Within this thesis continuous 10-year periods were simulated with a newer CCLM model version for the first time covering the SLL&RLP region. This study may give new insight and informations on the climate change of mean and extreme values in this area. Extreme indices and POT models are fit to a grid-box individually to keep a manageable amount of free parameters. Furthermore, the variables are assumed to be spatially independent and thus are modelled separately. The non-stationarity is accounted by adjusting the threshold of the POT models in every time period, which is approximately equivalent to making it time dependent.

Furthermore, bias-corrections of the model errors are performed and the influence on the climate change signal in terms of means and extremes are analysed. Since bias-correction is nowadays a standard technique that is applied by researchers but also by end-users, it is essential how such correction procedures influence or modify the climate change signal that is retrieved from GCM/RCM chains. The effect of the bias-correction is not only analysed with respect to the mean values but, even more important for impact models, with respect to extremes as well. Different state-of-the-art bias-correction techniques, among them a newly developed, are applied and compared regarding the performance to reduce model errors on the one hand and the influence on the climate signals on the other.

The thesis is outlined as follows: an overview of the regional climate model COSMO-CLM is given in chapter 2 and the observational data sets are presented. In chapter 3 the statistical methods for climate change detection, bias-correction and the extreme value analysis are described in detail. Chapter 4–7 presents the results of the hindcast simulations, the performance of the bias-corrections, the simulated climate change signals and the results of the extreme value analysis. Chapter 8 concludes with an examination of the outcome of the previous chapters and indicates further steps.

Climate model and data sets

2.1 The regional climate model COSMO-CLM

The regional climate model COSMO-CLM (CCLM) is the climate version of the COSMO¹ ("COntortium of SMall-Scale MOdelling") model (Steppeler et al., 2003) which evolved from the former "Lokal-Modell" (LM) of the German Weather Service (DWD) and is applied and further developed by the Climate Limited-area Modelling Community² (CLM, Rockel et al. (2008)). The version of the here used CCLM is 4.8_clm11. The latest officially released version is 4.8_clm17.

2.1.1 Model structure

The COSMO-CLM is a non-hydrostatic limited-area atmospheric prediction model developed for applications on the meso- β and meso- γ scale (Schättler et al., 2009). The CCLM relies on the primitive thermo-dynamical equations describing compressible flow in a moist atmosphere. It is set up in a rotated geographical coordinate system to allow nearly equal-sized grid boxes in the model domain. The atmospheric model variables are staggered on an Arakawa-C/Lorenz grid with scalars (e.g. temperature, pressure) defined at the centre of a grid box and the velocity components on the box faces (Schättler et al., 2009) (Fig. 2.1). The vertical wind component is calculated at half levels (Fig. 2.2).

The grid boxes of the CCLM are organized in a rotated geographical coordinate system with a terrain-following vertical coordinate. The rotated coordinate system is necessary because of the convergence of the meridians at the poles. The usage of the standard geographical coordinate system would arise two problems: (i) the poles are singularities that violates the numerical schemes and (ii) the horizontal resolution is not constant with latitude but a function of it. By rotating the North Pole in such a way that the new equator runs through the centre of the model domain ensures approximately equidistant grid-boxes. Terrain-following coordinates are required because otherwise the model levels would intersect with the orography.

The default grid-spacing for the meso- β scale is 7 km in the official forecast mode by the DWD (COSMO-EU). For the meso- γ scale a resolution of 2.8 km is used in the domain of Germany (COSMO-DE) (Schättler et al., 2009). At this scale it is expected that deep convection and its feedback on the

¹ <http://www.cosmo-model.org/>

² <http://www.clm-community.eu/>

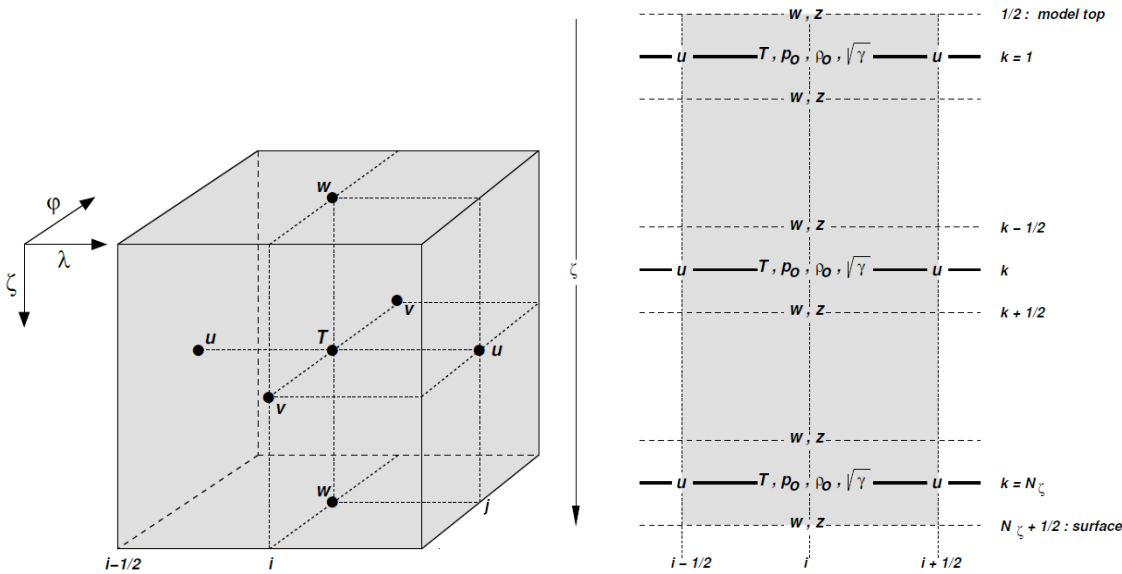


Figure 2.1 Schematic sketch of a grid box with **Figure 2.2** Vertical staggering of grid boxes (Schättler et al., 2009).

larger scales of motion are resolved explicitly and that convection and the orographic impact on the organization of convection is simulated much more realistically and likewise severe weather events. In the climate mode (CLM), the COSMO-CLM has no prescribed horizontal resolution although the resolutions of the numerical weather prediction (NWP) models, COSMO-EU and COSMO-DE, are used as standard resolutions as well. In contrast to the NWP mode no data assimilation takes place and vegetation parameters are updated on a monthly basis, which is ignored for short-time weather forecasts.

The CCLM is coupled to a multi-layer soil module (TERRA-ML, Schrodin and Heise (2001)), which can be used as a stand-alone model itself (e.g. Grasselt et al., 2008). The soil module provides updates of the surface boundary data and models hydrological processes based on Doms et al. (2007) and Ament and Simmer (2006) and the exchange with the atmosphere. Exchange processes with the biosphere are modelled by a Biosphere-Atmosphere scheme (BATS, Dickinson et al. (1993)). The TERRA-ML considers only the dominant vegetation and soil type within a grid cell. The spatial distribution of those parameters were retrieved from the ECOCLIMAP data base (Masson et al., 2003; Champeaux et al., 2005).

2.1.2 Physical parameterizations

Processes in the atmosphere act over a large span of temporal and spatial scales. Since atmospheric models, as the CCLM, are organized on staggered grid boxes and thus limited in spatial and temporal resolutions, certain processes cannot be solved explicitly by the numerical schemes. Examples for such sub-scale processes are turbulence, micro-physics or convection. In order to account for such processes, parameterizations are used which model the aggregated effect of subgrid-scale processes on the grid-scale. For instance, parameterizations do not aim to model individual eddies but their ensemble effect within a grid-box. Convection, however, cannot be defined as a subgrid-scale process as easily. Considering the horizontal scale, individual up-and downdrafts of a convection cell are obviously subgrid-scale processes but in terms of time-scales, the time-step of the model is in the order of the typical-time scale

of penetrating convection. Thus the scales are not separated, which yields a mixture of the scales (Doms et al., 2007).

The following subgrid-scale processes are parameterized in the CCLM:

- Subgrid-scale turbulence
Prognostic turbulent kinetic energy (TKE)-closure (level 2.5 after Mellor and Yamada (1974)), which includes effects of subgrid-scale condensation and thermal circulations.
- Surface layer fluxes
TKE-based scheme including a laminar sub-layer.
- Grid scale clouds and precipitation (micro-physics)
Bulk-formulation of a three-category-ice scheme with prognostic cloud ice, snow and graupel. Sub-grid scale clouds (water and ice) are included.
- Moisture convection
The Tiedtke-scheme (Tiedtke, 1989) is used for the 4.5 km resolution and a shallow convection scheme, a component of the Tiedtke-scheme, for the 1.3 km resolution.
- Radiation
The scheme of Ritter and Geleyn (1992) is used for parameterizing radiative fluxes and heating rates.

Convection-parameterizations

In the COSMO-CLM there are currently two options for the moist convection scheme implemented: The Tiedtke mass-flux scheme (Tiedtke, 1989) and the shallow convection scheme, which is a component of the Tiedtke-scheme. In general, convection has a major impact on the vertical distribution of temperature and moisture but its horizontal scales are too small to be resolved by mesoscale RCMs and thus the overall effect of convection has to be parameterized (Doms et al., 2007). Several processes and feedbacks have to be considered with moist convection. For instance, latent heat release due to phase transitions of cloud droplets or due to evaporation of precipitation. Heat, moisture and momentum are transported vertically in updraft and downdraft regions, whereas lateral exchange with the environment takes place via entrainment and detrainment (Doms et al., 2007).

In the Tiedtke mass-flux scheme it is assumed that large-scale moisture convergence and boundary layer turbulence trigger convection (Stensrud, 2007). The processes mentioned above are calculated in a simple 1-d bulk cloud model. Within the cloud model there is only one up- and one downdraft per grid cell. The budget equations of mass, heat, moisture and momentum are calculated for up- and downdrafts separately and the entrainment and detrainment rates are treated as functions of the grid-scale variables and hence are related to the grid-scale dynamics. Then convective tendencies are calculated from integrating these equations vertically from the cloud base up to the cloud top. The formation of convective precipitation is parameterized as:

$$g_p = K_p(z)q_u^c. \quad (2.1)$$

Thus the convective precipitation amount g_p depends linearly on the updraft cloud water content q_u^c and on a height dependent conversion function K_p (Doms et al., 2007) which is defined as:

$$K_p(z) = \begin{cases} 0, & \text{if } z \leq z_b + \Delta z_c \\ \beta_p, & \text{if } z > z_b + \Delta z_c \end{cases}. \quad (2.2)$$

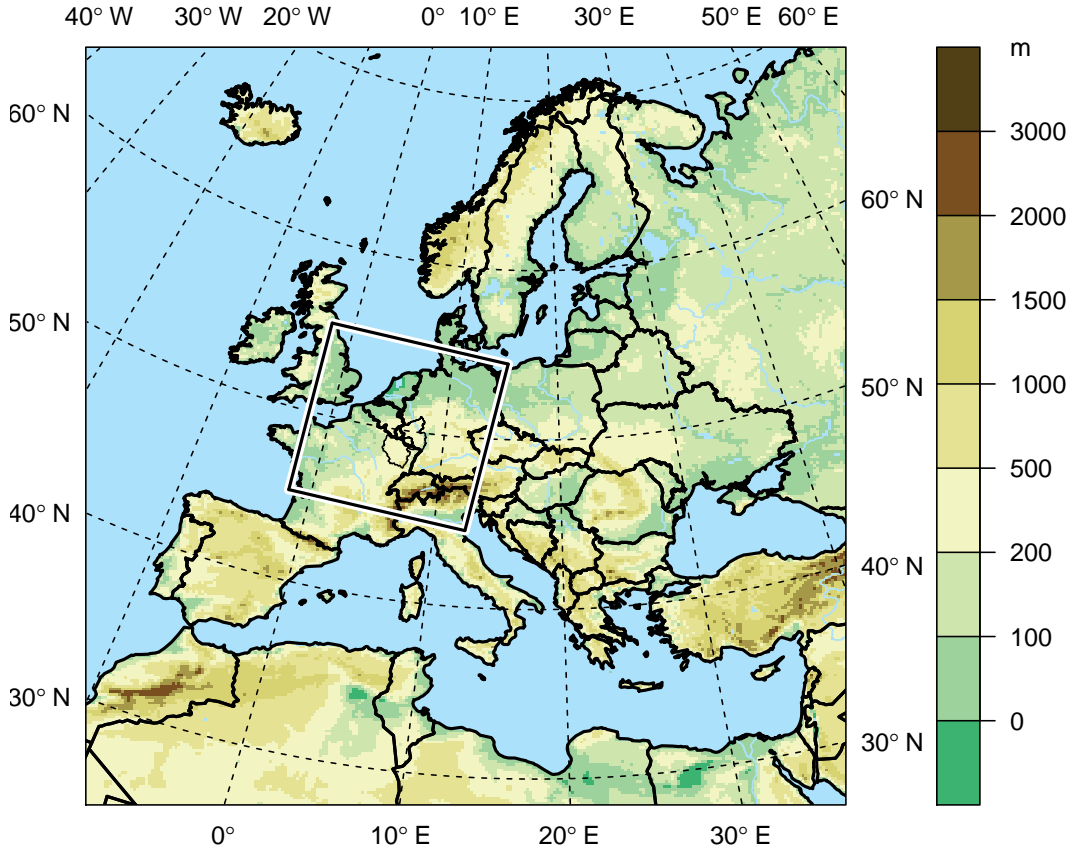


Figure 2.3 Model domain of the CLM18 together with the height of orography. The black box marks the domain of the CCLM4.5 model.

The relation suppresses convective precipitation in a region Δz_c above the cloud base z_b and otherwise the parameter for the linear dependency is set to $\beta_p = 2.0 \cdot 10^{-3} \text{ s}^{-1}$ and $\Delta z_c = 1500 \text{ m}$ over water and $\Delta z_c = 3000 \text{ m}$ over land (Doms et al., 2007). The flux of convective precipitation is finally calculated from:

$$\frac{\partial P}{\partial z} = -\rho(g_p - e_d - e_p), \quad (2.3)$$

with e_d the evaporation of precipitation within the downdraft and e_p the evaporation below the cloud base. While it is assumed that the downdraft is constantly at saturation, rainwater evaporates to maintain the saturation and evaporative cooling is accounted for in the heat equation. The evaporation rate of e_p is calculated from a modified microphysical Kessler-scheme (Kessler, 1969; Doms et al., 2007) that non-linearly relates evaporation to a moisture deficit weighted by the area-fraction of the grid cell which is covered by convective precipitation.

The Tiedtke scheme allows for three different types of convection, which are mutually exclusive

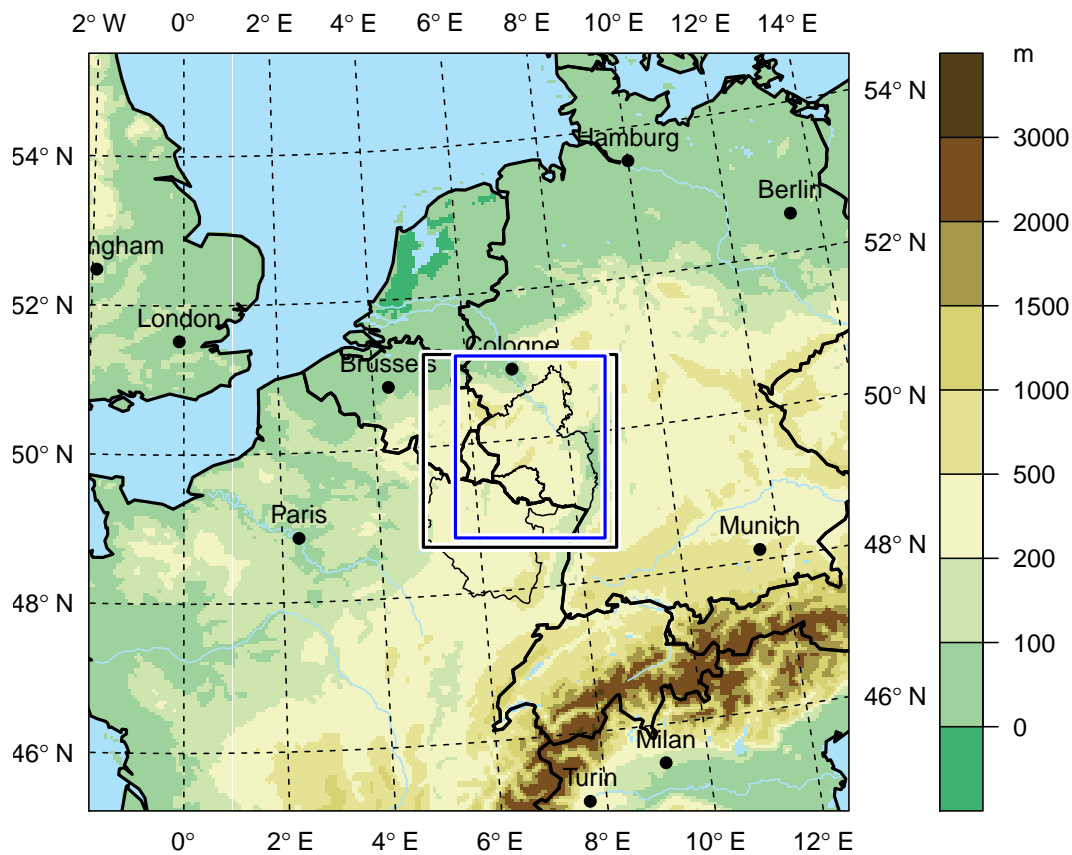


Figure 2.4 Model domain of the CCLM4.5 along with the height of orography. The black box marks the domain of the CCLM1.3 model and the blue box is the investigation area.

and differ in the vertical extend: (i) shallow convection, (ii) mid-level convection and (iii) penetrative or deep-convection convection. While penetrative and shallow convection originate in the surface boundary layer the midlevel convection does not. Penetrative convection is often related to large-scale convergence, whereas shallow convection can also be formed in slightly divergent flow and is often driven by evaporation from the surface. Thus in the penetrative convection closure it is assumed that the moisture convergence, which is related to the updraft mass flux by the vertical turbulent flux of specific humidity, is mainly forced by advection. In the shallow convection the same relation is used but the main source of moisture is supplied from the surface due to evaporation (Doms et al., 2007). The shallow convection scheme does not parameterize precipitation as in the deep convection scheme.

2.1.3 Model configuration

The regional climate simulations in this study were performed with the COSMO-CLM in version 4.8.clm11 coupled to the multi-layer soil model TERRA-ML (Schrodin and Heise, 2001; Heise et al.,

2006; Doms et al., 2007) within the project *Impacts of Global Change on biological resources, legislation and standards*¹ and were funded by the Research Initiative Rhineland-Palatinate. The CCLM simulations were run with a horizontal resolution of 0.04° (≈ 4.5 km, CCLM4.5) and 0.0118° (≈ 1.3 km, CCLM1.3). For the downscaling a nesting chain was used which started with the GCM ECHAM5/MPI-OM (Roeckner et al., 2003). The first nest of the COSMO-CLM was done by Hollweg et al. (2008) in a resolution of 0.165° (≈ 18 km). The output of this model runs were then used to force the CCLM4.5 which in turn forced the CCLM1.3. The model domain of the CLM18 is shown in Fig. 2.3, that of CCLM4.5 in Fig. 2.4 along with the CCLM1.3 domain. The focus region of this study is shown in Fig. 2.5 for the 4.5 km and 1.3 km resolution, respectively. The CCLM4.5 simulations have 254×254 grid boxes and 40 vertical levels, whereas the CCLM1.3 has 220×220 grid boxes and 40 vertical levels. An official and evaluated model configuration neither exists for a horizontal resolution of 4.5 km nor for 1.3 km. Therefore the model configuration was newly adapted for the CCLM4.5 and CCLM1.3 simulations. The configuration of CCLM4.5 is based on a former configuration from the CLM3.2, which was used in the *LandCaRe 2020* project (Berg et al., 2008). That of the CCLM1.3 is partly identical with the configuration used by Knote et al. (2010). However, those configurations are from former CCLM versions and had thus to be adapted to the newer model version. The main difference between both models, besides the resolution, is the treatment of moist convection. While the Tiedtke-scheme is used in CCLM4.5 for parameterizing deep-convection, only the shallow-convection scheme is active in the CCLM1.3. The time-integration of the prognostic equations are solved by a 2 time-level time-split Runge-Kutta numerical scheme of 3rd order. The time splitting is necessary to allow for a much smaller time step for the treatment of acoustic and gravity waves. The time steps in the CCLM4.5 and CCLM1.3 were set to $dt = 45$ s and $dt = 12$ s, respectively. The detailed model configurations can be found in Appendix A.

2.2 Data sets

An overview of all climate simulation data which have been used within this thesis is given in Tab. 2.1 and all observation data sets are listed in Tab. 2.2.

2.2.1 Initial and lateral boundary data

To assess climate change and its impacts on the study region in high spatial resolution, a dynamical downscaling chain (Fig. 2.6) was applied to bridge the gap from a global to a local scale. The starting point of this nesting chain was the General Circulation Model (GCM) ECHAM5/MPI-OM of the Max Planck Institute (MPI) (Roeckner et al. (2003), which was also described in a special section of the *Journal of Climate*, e.g. Jungclaus et al. (2006)) and part of the IPCC AR4 report. As a first downscaling step, the CCLM model with a horizontal resolution of 18 km (CCLM18), 257×271 grid boxes and 32 vertical levels was nested into the GCM by the Model & Data group of the MPI in the frame-work of the consortial runs experiment (Hollweg et al., 2008). These CCLM18 runs were performed for the periods 1955–2000 (C20 scenario) and 2001–2100 (A1B scenario). In addition, a hindcast simulation (hc) run driven by ERA-40 (Uppala et al., 2005) for the period of 1991–2000 (CCLM18hc, (Keuler et al., 2012)) was used as forcing data. The model domain of the CCLM18 and CCLM18hc runs are shown in Fig. 2.3.

2.2.2 CCLM runs

The first set of model runs were hindcasts (hc) for the period 1991–2000 forced by the CCLM18hc run (nested into ERA-40, without any assimilation). This backtesting allows to assess the model performance.

¹ <http://www.uni-trier.de/index.php?id=40193&L=2>

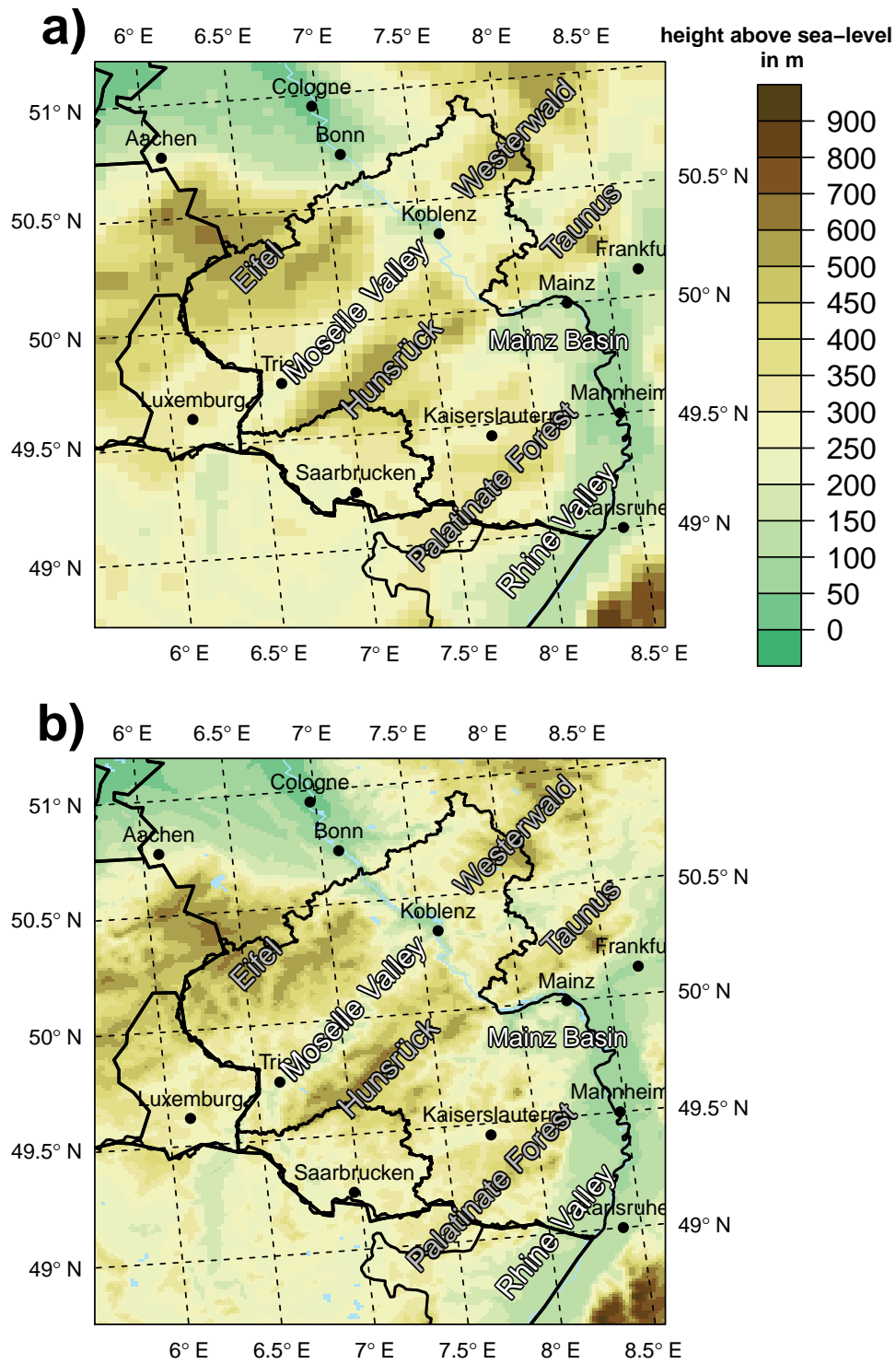


Figure 2.5 Investigation area: Rhineland-Palatinate and the Saar-Lor-Lux region with orography in two resolutions. In a) the domain in 0.04° (4.5 km) resolution from the CCLM4.5 model and in b) in a resolution of 0.018° (1.3 km) from the CCLM1.3 model. Lakes and rivers are mapped in blue.

Table 2.1 Overview of the regional climate model runs with CCLM.

time period	model	resolution	boundary data	source
1991–2000 (hc)	CLM18hc	0.165° (18 km)	ERA-40	Keuler et al. (2012)
	CCLM4.5hc	0.04° (4.5 km)	CCLM18hc	this study
	CCLM1.3hc	0.0118° (1.3 km)	CCLM4.5hc	this study
1991–2000 (C20)	CLM18c20	0.165° (18 km)	ECHAM5/MPI-OM	Hollweg et al. (2008)
	CCLM4.5c20	0.04° (4.5 km)	CCLM18c20	this study
	CCLM1.3c20	0.0118° (1.3 km)	CCLM4.5c20	this study
2041–2050 (A1B_1)	CLM18a1b1	0.165° (18 km)	ECHAM5/MPI-OM	Hollweg et al. (2008)
	CCLM4.5a1b1	0.04° (4.5 km)	CCLM18a1b1	this study
	CCLM1.3a1b1	0.0118° (1.3 km)	CCLM4.5a1b1	this study
2091–2100 (A1B_2)	CLM18a1b2	0.165° (18 km)	ECHAM5/MPI-OM	Hollweg et al. (2008)
	CCLM4.5a1b2	0.04° (4.5 km)	CCLM18a1b2	this study
	CCLM1.3a1b2	0.0118° (1.3 km)	CCLM4.5a1b2	this study

The hindcast runs were performed with the 0.04° (CCLM4.5hc) and 0.0118° (CCLM1.3hc) resolution. In a hindcast run a RCM is forced by reanalysis data at the boundaries, which allows a direct comparison with observational data sets.

The next simulations were performed for the C20 period 1991–2000 (CCLM4.5c20 and CCLM1.3c20) with the CLM18c20 data as forcing. To simulate the climate for the near future 2041–2050 (A1B_1) and for the end of the century 2091–2100 (A1B_2), the CCLM4.5 was forced by CLM18a1b1 and CLM18a1b2 (CCLM4.5a1b1, CCLM4.5a1b2). Afterwards the CCLM1.3 was nested into the 4.5 km simulations (CCLM1.3a1b1, CCLM1.3a1b2). In total, eight simulations were performed, which cover 10-year periods respectively (see Tab. 2.1).

The restriction of simulating 10-year periods was due to the required computing time for such high-resolution experiments. For instance, the simulation of 10 years with a horizontal resolution of 4.5 km, 254×254 grid boxes, 40 vertical levels and a time step of 45 s requires ≈ 35000 CPUh and 10 years with a resolution of 1.3 km, 220×220 grid boxes, 40 vertical levels and a time step of $dt = 12$ s requires ≈ 130000 CPUh. The output of the CCLM simulations were stored with a temporal resolution of one hour.

2.2.3 Observational data

For the comparison of the CCLM simulations with observations, three gridded data sets were available (Tab. 2.2). However, such high-resolution observations with up to 1 km are rare and of limited length. The data set with the coarsest resolution is the free-available ECA&D E-OBS data set (Haylock et al., 2008) from the ENSEMBLES project (van der Linden and Mitchell, 2009). This data set has a horizontal resolution of 0.22° (≈ 22.5 km) and a daily temporal resolution. The E-OBS data set contains the variables: precipitation, 2 m-temperature, and daily minimum/maximum temperature and is only used to briefly evaluate the CCLM18hc forcing data set (see the complete evaluation in Hollweg et al. (2008)).

The next data set is a product of the German Weather Service (DWD) named REGNIE, which is the result of an interpolation of precipitation gauge stations within Germany. There were approximately 4500 observation stations with daily measurements used for the interpolation (Bartels et al., 2006). The REGNIE data set is a well-known in the hydrological and meteorological communities (Rauthe et al., 2013). For example it was used in climate simulation studies (Bellprat et al., 2012; Berg et al., 2012) or in validating the soil-module TERRA-ML of the CCLM (Grasselt et al., 2008).

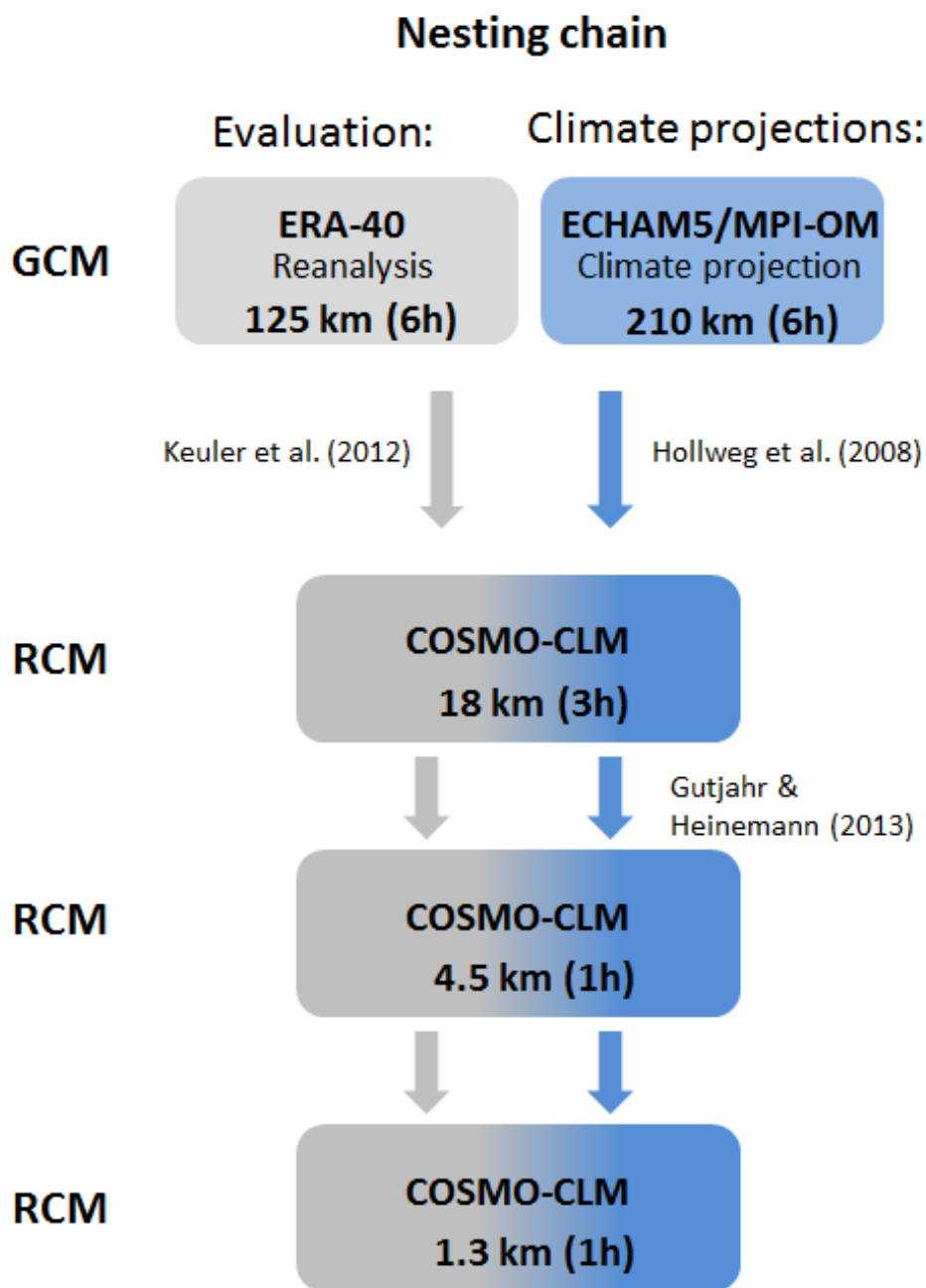


Figure 2.6 Nesting chain of the dynamical downscaling consisting of a reanalysis branch (grey) and a climate projection branch (blue). The first nesting step of the CCLM in 18 km resolution into the General Climate models and ERA-40, respectively, has been done by two other groups. See Keuler et al. (2012) and Hollweg et al. (2008) for more details of the model configurations and evaluations. In parentheses the increment of saving model output is given, which is identical to the increment that is used to update the lateral boundaries for the following nested model.

Table 2.2 Overview of the observational data sets.

time period	data set	resolution	source
1991–2000	ECA&D E-OBS	0.22° (22.5 km)	Haylock et al. (2008)
	REGNIE	0.0084° × 0.0167° (0.9 km × 1.2 km)	DWD
1993–2002	INTERMET	1 km × 1 km	LUWG

The target grid of REGNIE has an original resolution of $0.0083^\circ(\text{W} - \text{E}) \times 0.0167^\circ(\text{N} - \text{S})$ ($\approx 0.9 \text{ km} \times 1.2 \text{ km}$), but to compare it with the CCLM simulations, it was bilinearly interpolated onto the CCLM4.5 and CCLM1.3 grid.

An additional high-resolution product, which was available for the state of Rhineland-Palatinate, is the INTERMET data set (Dobler et al., 2004) of the state office for environment, water and factory inspectorate Rhineland-Palatinate (LUWG). This data set is again a product of interpolated precipitation and 2 m temperature stations (and other variables, as humidity and wind), but uses a different, smoother approach that is based on a combination of kriging and fuzzy logic (Dobler et al., 2004). The original resolution is 1 km x 1 km in a non-rotated coordinate system (Gauß-Krüger, GK2), therefore the INTERMET data set is interpolated on the CCLM grids as well. REGNIE and INTERMET are used for evaluating the CCLM4.5 and CCLM4.5 hindcast and climate simulations. Since INTERMET was only available from 1993 onwards, all comparisons are restricted to the period 1993–2000.

2.2.4 Short evaluation of the forcing data

The variables precipitation, 2 m temperature and surface pressure at mean sea level (PMSL) of the forcing data of CCLM18hc were compared to the E-OBS data set. Fig. 2.7 shows the difference in annual total precipitation of period 1991–2000 of CCLM4.5 compared to E-OBS in the domain of west-central Europe and the CCLM4.5 model domain (black box). Focusing on the CCLM4.5 domain, the CCLM18hc is wet biased in the southern domain, in particular in the Alps where the bias is about $+50 \text{ mm year}^{-1}$. Marginally stronger wet biases occur in the eastern part of Germany, France, Luxemburg and the state of Rhineland-Palatinate with $0 - 20 \text{ mm year}^{-1}$. In the domain of interest (Saar-Lor-Lux and RLP) the bias is only up to $+15 \text{ mm year}^{-1}$. Dry biases occur along the coasts of Germany, the Netherlands and Belgium with about $-0 - 15 \text{ mm year}^{-1}$, along the Rhine Valley and in the Black Forest in Germany and the Vosges in France with about -30 mm year^{-1} . A more detailed evaluation of the CLM18hc can be found in Keuler et al. (2012) and the evaluation of the climate runs were published by Hollweg et al. (2008).

In the C20 run (CLM18c20), the annual precipitation of Germany is overestimated by about 127–178 mm. The wet bias is nearly systematic in all months. Hollweg et al. (2008) state that this bias was transferred from the GCM ECHAM5/MPI-OM into the CLM18c20. Thus, it can be expected that this wet bias is transferred to the CCLM4.5c20 as well. The 2 m temperature is too warm in winter (about $+1.25^\circ\text{C}$ but too cold in summer (about -1.75°C).

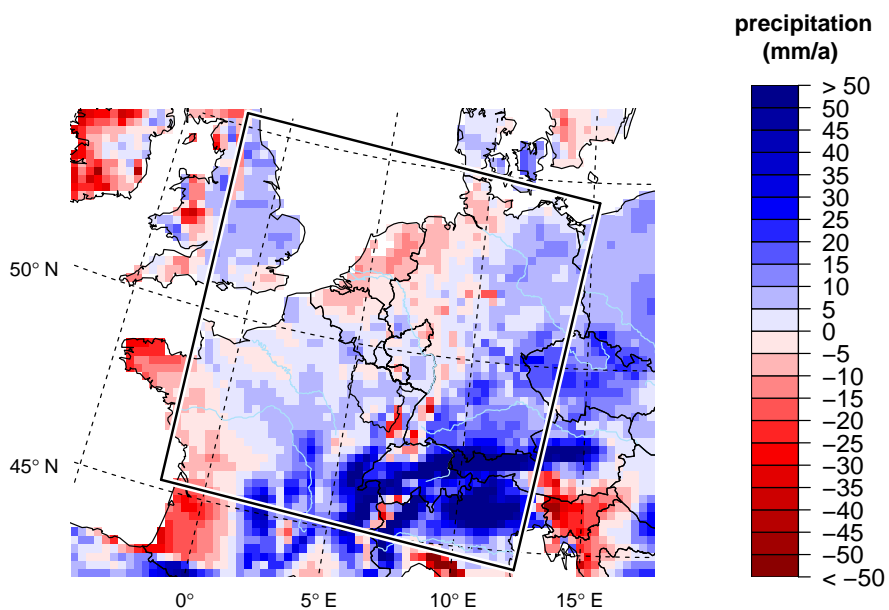


Figure 2.7 Differences of simulated precipitation from CLM18hc minus E-OBS for the time period 1991–2000. The black box marks the CCLM4.5 domain.

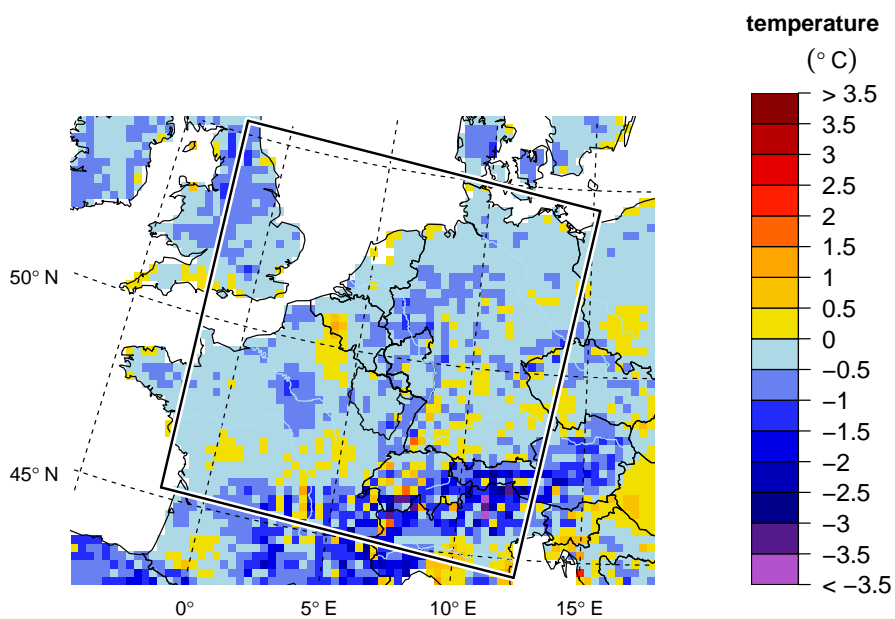


Figure 2.8 Differences of simulated 2 m temperature from CLM18hc minus E-OBS for the time period 1991–2000. The black box marks the CCLM4.5 domain.

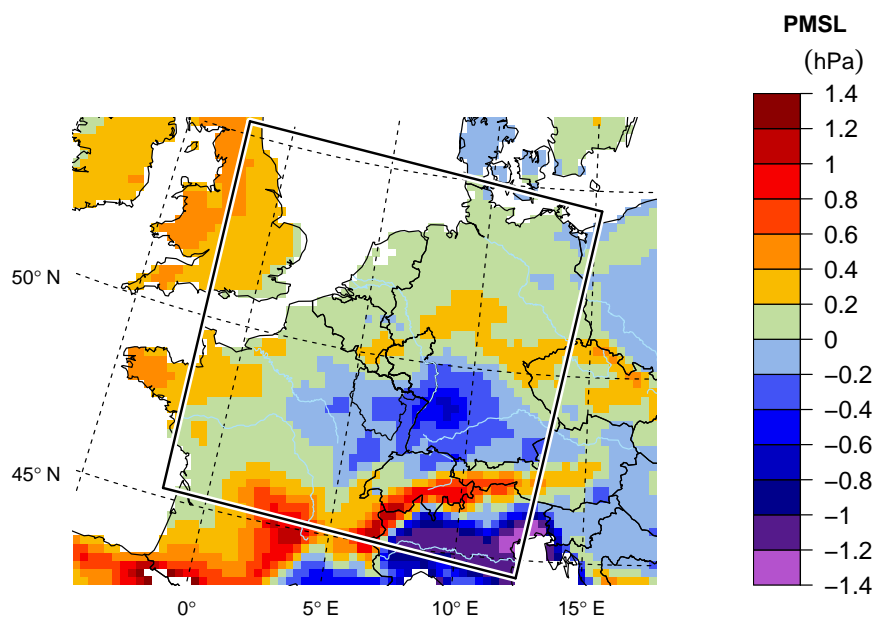


Figure 2.9 Differences of simulated pressure at mean sea level from CLM18hc minus E-OBS for the time period 1991–2000. The black box marks the CCLM4.5 domain.

Statistical methods and bias correction

3.1 Climate change signal detection

3.1.1 *Statistical tests and the field significance*

The detection of climate change is the demonstration that the climate has changed in some statistical sense without ascertaining the reasons (Zhang and Zwiers, 2013). The climate change signal is thus the potential change in the state of climate that is statistically significant, i.e. larger than the natural climate variability. This can be understood as a signal to noise problem. In this thesis, the change in annual and seasonal means are analysed for the variables precipitation and 2 m temperature. Statistical tests are applied to time series of the present and future climate. These tests enfold t-tests, Wilcoxon-Mann-Whitney (U) tests and Kolmogorov-Smirnov (KS) tests (see e.g. Wilks (2011) for detailed descriptions). The t-test is used to determine any changes in the means by assuming that the annual or seasonal means are approximately normal distributed. Although the t-test is rather robust versus non-normality (Rasch and Guiard, 2004) a non-parametric alternative (U-test), which does not require any specifications on the distribution, was applied additionally. The third test, the KS-test is a non-parametric hypothesis test which compares cumulative distribution functions for equality. The t-test and the U-test were applied as two-sided tests, since it is not known a-priori whether there will be more or less rain or how the temperature develops in the future. The KS-test is performed as the two-sample test, comparing two empirical distributions. It should be emphasized that within this thesis only one realisation of the future climate path is simulated. Thus all subsequent results are restricted to one possible climate path. Therefore the model uncertainty and the anthropogenic uncertainty cannot be incorporated into the analyses.

The three statistical tests are implemented into a "false discovery rate" (FDR) procedure to account for *multiplicity* in the framework of field tests (Wilks, 2006). Multiplicity denotes the problem of jointly evaluating tests under spatial dependency, e.g. fields of a meteorological variable. This joint evaluation is also termed "field significance" (Wilks, 2006). To clarify the problem of multiplicity, Ventura et al. (2004) give an example. Consider a level of significance of $\alpha = 5\%$ and at 1,000 locations a test for trend is applied. Then the average number of falsely rejected null hypothesis (H_0) will be $0.05 \times 1,000 = 50$, which is unacceptably high. Transferred to the climate change question, this would mean 50 out of 1,000 grid points are wrongly identified as significant. One of the first who discussed this problem in a climatological context were Livezey and Chen (1983). They developed a method to account for

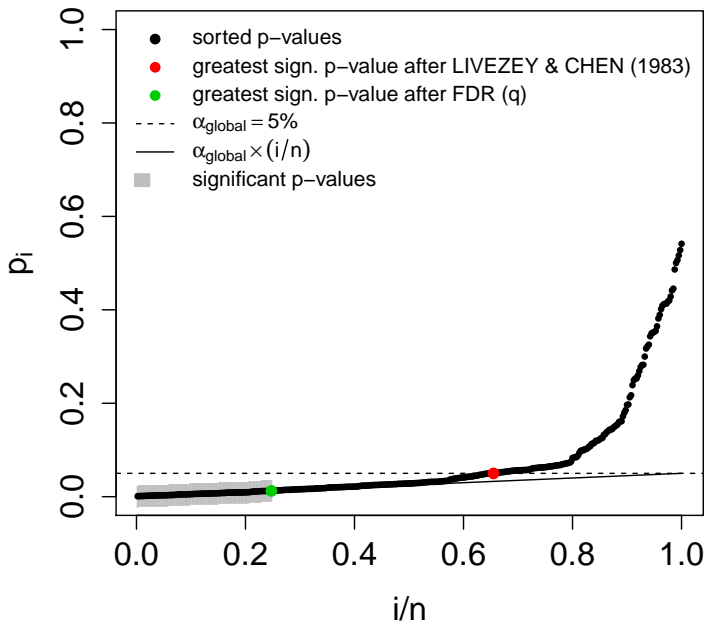


Figure 3.1 Illustration of the FDR-procedure after Benjamini and Hochberg (1995) with $q = \alpha = 5\%$. The *black solid line* represents the sliding scale which restricts the significant p-values (highlighted in grey). The green point marks the largest p-value which satisfies eq. 3.1. All p-values smaller than this green point are significant considering FDR. In contrast, the red point marks the largest p-value which is significant in terms of the classical counting test after Livezey and Chen (1983).

this problem, referred to as the "counting method", which assumes that the p-values follow a binomial distribution. Then it is possible to calculate the probability of observing the actual number of rejected H_0 . If this probability is smaller than α , then the field is significant. This method is not appropriate for the purpose of this study, because it is necessary to decide on every local p-value, not on the field as a whole.

A better method for this purpose is the classical FDR procedure of Benjamini and Hochberg (1995) which rejects H_0 at all locations i for which $p_i \leq p_k$:

$$k = \max_{i=0, \dots, n} \left\{ i : p_i \leq q \frac{i}{n} \right\}. \quad (3.1)$$

The p-values p_i are sorted in ascending order $i = 1, \dots, n$. It can be shown that the false discovery rate q is equal to α (Wilks, 2006). In Fig. 3.1, this procedure is illustrated. The restriction on significant p-values is obvious. The largest significant p-values fulfilling equation 3.1 is marked as a green dot. All p-values smaller than this are significant as well (indicated by the grey shadings).

3.1.2 Power analysis of the t-test

The application of t-tests for detecting significant climate change signals offers furthermore the possibility to calculate the *power* of the tests, i.e. the probability of detecting a signal if there is a true signal. Or expressed in a statistical sense: the power gives the probability of rejecting H_0 if H_0 is actually false, which corresponds to $\text{power} = 1 - \beta$, where β denotes the "type II error", i.e. the incorrect

rejection of a true H_0 . The power analysis is based on Cohen (1988) and exploits the mathematical relationship between the power, the significance criterion (α), the sample size and the so-called effect size (ES) (Cohen, 1992). The effect size of a t-test is just the difference in means divided by the common within-population standard deviation, which is an absolute value in case of a two-sided test. This effect size is named *Cohen's d*. The common within-population standard deviation is used here in its simpler form (Cohen, 1988):

$$\sigma_{xy} = \sqrt{(\sigma_{C20}^2 + \sigma_{A1B}^2)/2}. \quad (3.2)$$

Thus *Cohen's d* can be calculated as:

$$d = \frac{|\bar{x} - \bar{y}|}{\sigma_{xy}}. \quad (3.3)$$

To make the interpretation of the power analysis more transparent, the numerator and denominator of eq. 3.3 are disentangled and set in relation rather than discussing *Cohen's d* directly. The numerator of eq. 3.3 is denoted as D which is just the absolute climate signal whereas eq. 3.2 can be interpreted as the common variance of C20 and A1B or climate noise due to inter-annual or inter-seasonal variability.

In general, the power of a t-test increases with the sample size, with a larger d and with a smaller α criterion.

3.2 Bias-correction

3.2.1 Quantile-matching methods

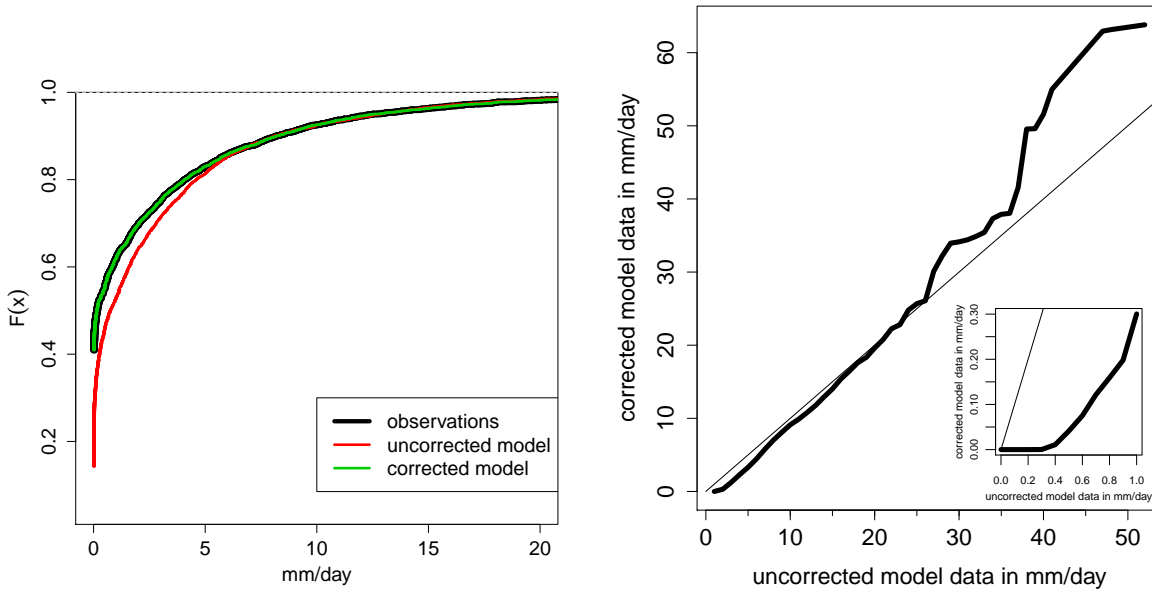
To bias-correct the daily CCLM precipitation fields, four variants of the quantile-matching approach (QM) are applied. The QM was chosen out of a pool of possible BC methods because it is widely adopted (Lafon et al., 2013; Piani and Haerter, 2012; Rojas et al., 2011; Schoetter et al., 2012; van Roosmalen et al., 2011) and performed generally better than other methods as for instance linear correction (Thiemeßl et al., 2011; Gudmundsson et al., 2012; Teutschbein and Seibert, 2012; Berg et al., 2012).

The QM adjusts all moments of the probability distribution function (PDF) of any model variable (Hagemann et al., 2011; Piani et al., 2010a). Therefore, as a first step PDFs are constructed from model and observed data. Next, the PDFs are integrated to obtain cumulative distribution functions (CDFs). Finally, so-called *transfer functions* are created which translate uncorrected model data into corrected model data. After the correction, the CDF of the model should equal the observed one. The four variants differ only with respect to the construction of the PDFs/CDFs, which yields automatically different transfer functions.

The reference method uses empirical distribution functions (Panofsky and Brier, 1968; Boé et al., 2007; Déqué, 2007; Sennikovs and Bethers, 2009; Michelangeli et al., 2009) and is hence referred to as empirical quantile-matching (eQM). This method is expected to produce the best correction but depends on many degrees of freedom and may not be stationary due to possible over-fitting and hence may violate this assumption in future periods. However, for climate change applications it is assumed that the transfer function stays constant with time (Piani et al., 2010a), which is not a trivial assumption (Trenberth et al., 2003).

The QM method is constructed by calculating the empirical PDFs but uses the cumulative distribution functions (CDFs) for the correction:

$$y = F_{\text{obs}}^{-1}(F_{\text{CCLM}}(x)), \quad (3.4)$$



(a) Schematic illustration of the quantile-matching

(b) Schematic illustration of a transfer function

Figure 3.2 Schematic illustration of the quantile-matching result and the corresponding transfer function. In a) The cumulative distributions of the observations (black) and the CCLM model (green) match after the correction whereas the uncorrected model produces too many "drizzle". In b) a typical transfer function is shown (solid line) which adjusts low intensities non-linearly (see the subplot) and gets asymptotically linear for higher intensities, although strong deviations from linearity may occur. The thin line represents the $x = y$ line.

with y the corrected precipitation value, x the original precipitation value and F_{obs}^{-1} the inverse function of the CDF of the observations, which is named quantile function, and accordingly F_{CCLM} is the CDF of the CCLM model. The probability of observing x mm day⁻¹ (or less) in the CCLM is thus transferred to the quantile of the observed CDF, matching exactly this probability. The right hand side in eq. 3.4 is referred to as *transfer function*. This procedure is graphically illustrated in Fig. 3.2a. A typical transfer function is shown in Fig. 3.2b. In the subplot it can be seen that intensities < 0.3 mm day⁻¹ are set to 0 to account for the drizzle-error and in general, that low intensities are corrected in a non-linear way, whereas higher intensities are corrected asymptotically linear. Furthermore, because the correction is based on empirical cumulative distributions, the function is not smooth.

The second QM method replaces the empirical distributions, which depend on df = length of time series, by fitting a parametric model to the "perfect" transfer function of eQM. This method was introduced by Piani et al. (2010b) and is henceforth referred to as Piani transfer function (PTF). The transfer function of PTF is defined as:

$$y = (a + bx) \left(1 - e^{-(x-x_0)/\tau} \right), \quad (3.5)$$

with a and b an additive and multiplicative parameter, $x_0 = -a/b$ is a dry-day correction factor which sets all precipitation values below x_0 to zero and τ denotes the rate at which the asymptote is reached. The PTF reduces the degrees of freedom to $df = 3$ since only three parameters have to be fitted.

Another parametric method, which was applied by Piani et al. (2010a), is called gQM. The gQM

replaces F in eq. 3.4 by a gamma distribution which is fitted to observed and modelled data separately. The gamma distribution depends on only two parameters and is commonly used for representing the PDF of precipitation (Yang et al., 2010; Wilks, 2011). However, this is not valid in every region in Europe as shown by Vlcek and Huth (2009) and should therefore be tested a priori. In this study, it was not tested explicitly but referring to the results of Vlcek and Huth (2009) (Fig. 2c therein), our domain of interest allows the approximation with a gamma distribution. The gamma distribution is defined as:

$$f(x) = \frac{(x/\beta_g)^{\alpha_g-1} \exp(-x/\beta_g)}{\beta_g \Gamma(\alpha_g)} \quad x, \alpha_g, \beta_g > 0, \quad (3.6)$$

with β_g the scale parameter, α_g the shape parameter and Γ the gamma function. The gamma distribution is not defined for $x = 0$ mm/d. Therefore, the correction process has to be a dual step (Piani et al., 2010a). First, the number of dry days is corrected by optimizing a threshold value, i.e. all values smaller than this threshold are set to zero, such that:

$$\min_{x_d \in \mathbb{R}_+} |n_d - m(x_d)|. \quad (3.7)$$

With n_d the number of dry days from the observations, $m(x_d)$ the number of dry days in the model depending on the threshold x_d where all $x < x_d$ are set to zero. After the threshold is found, the two parameters of the gamma distribution from eq. 3.6 are fitted only for wet days with the maximum likelihood estimation (MLE). This is done for observations and model data separately. Afterwards, these fitted PDFs are integrated and the resulting CDFs are used to replace the empirical CDFs in eq. 3.6.

Finally, a new method is applied which replaces the empirical PDFs (or CDFs) in eq. 3.4 by a combination of a gamma and a General Pareto distribution (GPD) and is referred to as GQM. The latter is a heavy-tailed extreme value distribution (Coles, 2001) (see eq. 3.12). The assumption is that the combination of a gamma distribution for the main body and an extreme value distribution for the tails is better suited for correcting precipitation, in particular with regard to extremes.

As a first step, the same threshold x_d for dry day correction from equation 3.7 is applied. Second, for the threshold u the 95th percentile is used as proposed by Yang et al. (2010) who combined two gamma distributions. This implies that values below the 95th percentile are assumed to follow a gamma distribution, whereas values exceeding this threshold are assumed to follow a general Pareto distribution:

$$y = \begin{cases} F_{\text{obs, gamma}}^{-1}(F_{\text{CCLM, gamma}}), & \text{if } x < 95\text{th percentile} \\ F_{\text{obs, GPD}}^{-1}(F_{\text{CCLM, GPD}}), & \text{if } x \geq 95\text{th percentile} \end{cases}. \quad (3.8)$$

There are now four parameters: a scale (β_g) and shape (α_g) parameter for the gamma distribution and a scale (σ) and shape (ξ) parameter for the GPD for observations and the model, respectively. Along with the threshold u , there are 5 parameters to be estimated. It has to be noted that the threshold u is different for the observations and the model. Therefore, it may happen that $x > 95\text{th percentile}$ of the observations in equation 3.8, but at the same time $x < 95\text{th percentile}$ of the model. In such a situation, x is not regarded as an extreme value. In addition, outliers occur in our model data. They were defined as single high extreme values with a probability density of $F_{\text{CCLM, GPD}}(x) > 1 - 10^{-8}$. These outliers have not been corrected, because such events are too far on the tail of a GPD due to the short time periods of 10 years. That is, these extremes would be transformed to unrealistically high values after applying eq. 3.8.

Altogether, there are four bias-correction methods of the QM-family applied to daily precipitation fields. For the temperature correction only the empirical correction (eQM) is used seasonally, but no parametric alternative. The above mentioned methods are applied to every individual grid box of the CCLM model and to the corresponding grid box of the observations yielding a transfer function for every

grid box. The grid boxes are hence treated as independent. Then a performance test is run by constructing the transfer functions from the CCLMhc, REGNIE and INTERMET data in a cross-validation framework. The procedure is described in the next section (section 3.2.2). After the performance test, a relative skill score is calculated in order to provide a ranking of the methods.

Gutjahr and Heinemann (2013) showed that good results are achieved if station data are used to correct model data. However, a simple inverse-distance interpolation may not be sufficient to provide adequate spatial variability because errors are introduced from interpolation, particularly in areas with strong orography. Furthermore, Maraun (2013) argues that if point measurements (station data) are used to correct model data, which are spatial averages, the scale-mismatch may introduce undesired side-effects. Because of this reason, it was decided to use REGNIE and INTERMET, which were produced by more sophisticated interpolation schemes.

3.2.2 Cross-validation

In order to evaluate the performance of the bias-correction methods, a temporal cross-validation (CV) is applied. As a first step the time series of the CCLM, which are 10 years long, are used to calibrate the transfer functions of the above mentioned QM methods. Then, the same 10 years are used for validating the QM methods. The performance is measured by a skill score (MAEx, see section 3.2.3). These corrections are henceforth referred to as *baselines*. In a second step, one identical year is omitted from the observations and the model time series during the calibration period. Then the transfer functions are calibrated with the remaining 9 years. The QM methods are validated by applying the corrections to the omitted year, which was not part of the training data. The reason why this has to be done is that the eQM method uses empirical distributions and would by definition provide the best performance if the calibration and validation period are identical. This would neither be a fair comparison nor would there be any information about the performance of correcting new or unknown data such as future data. If one year is omitted and corrected by a transfer function which was calibrated by the remaining 9 years, it is possible to judge the bias-correction since the truth (observations) is known.

Every year is omitted once from the calibration period, which results in 10 validations per method. The error from the CV is then added to the baseline and compared with the skill score of the uncorrected model. The CV is performed at every grid-box, but the skill scores are spatially averaged afterwards.

The CV is run two times: once for correcting the hindcast runs and once for the C20 runs. Because there is no serial correlation between the C20 runs and the observations, the CV is performed differently. The omitted year does not have to be the same year in the model and the observations, for example in the model the year 1991 is omitted whereas the year 1998 is omitted from the observations. Since only the distributions should match, this approach is reasonable. In this case, there are 100 validations per method.

3.2.3 MAEx

For the comparison of the overall performance of the bias-correction methods, Gudmundsson et al. (2012) introduced a mean absolute error (MAEx) skill score. The MAEx integrates the differences of quantiles within 0.1 wide probability intervals. The index x indicates the upper limit of such an interval (e.g. 0.1 for the 10 % interval). Each interval is further subdivided into 10 steps. Thus, per interval the differences of 10 quantiles are integrated and divided by 10. For example, MAE_{0.4} aggregates the absolute differences of the quantiles between the 30 % and the 40 % percentile divided by 10. The best performance is achieved if MAE_x = 0 and the performance gets worse with increasing MAE_x scores.

The MAE_x is calculated for MAE_{0.1}, MAE_{0.2}, ..., MAE_{1.0}. Hereby, MAE_{0.1} evaluates the differences in the dry part and MAE_{1.0} in the extreme part of the distributions. The mean of the above MAE_x

is called MAE, analogue to the definition of Gudmundsson et al. (2012). In the following, the MAE and all individual MAE_x denote spatial averages.

3.3 Extreme value analysis

3.3.1 *Extreme indices*

Extremes are most often discussed and analyses with respect to daily precipitation and temperature (Sillmann et al., 2013a), for example heavy precipitation events, heat-waves or dry-spells. Extremes are usually defined by either setting a threshold which is exceeded or fallen below or by fitting an extreme value model. Such a threshold may be a fixed value or vary with time, which is often the case in connection with percentile-based thresholds (Sillmann et al., 2013b). The Expert Team on Climate Change Detection and Indices (ETCCDI) has defined a set of 27 extreme indices (Klein Tank and Können, 2003), which are used in multiple applications due to their robustness and straightforward calculation (Sillmann et al., 2013a).

Most of these extreme indices, however, describe moderate extremes with reoccurrence times of one year or shorter (Sillmann et al., 2013a). They can be divided into threshold-, percentile-, duration-based and absolute indices.

Threshold-based indices use a certain threshold and count for example the number of very heavy precipitation days (R20) with more than 20 mm per day. Frost days (FD) and tropical nights (TR) count days on which the minimum temperature falls below 0 °C or above 20 °C. These indices are of interest in climate change impact studies, for example frost days are important in agricultural practices, whereas tropical nights are of importance in context of human health Sillmann et al. (2013a); Tebaldi et al. (2006).

Percentile-based indices are rely on annual or seasonal percentiles, which may be fixed or variable over a time-period. For instance, the R95PTOT is the proportion of the precipitation above the 95 % percentile on the total amount of precipitation. Whereas HWFI measures the longest period of at least 6 consecutive days on which the daily mean temperature is above the 90 % percentile in a sliding 5-day window over the calendar days. In this case, the threshold is time-varying and will be higher in summer than in winter.

Duration indices count the longest period of consecutive days which exceed a defined threshold. From Tab. 3.1 these are the heat-wave duration index (HWDI), the cold-wave duration index (CWDI) and the consecutive dry-day index (CDD). The CDD is the only index that describes the lower region of the precipitation distribution and is often interpreted as an indicator of drought. However, droughts are more complex phenomena which do not solely depend on precipitation, therefore the CDD is more a indicator for meteorological droughts (Sillmann et al., 2013a). For such purposes, other indices should be considered as well (Tebaldi et al., 2006), e.g. the Standardized Precipitation Index (SPI, McKee et al. (1993)) or the Standardized Precipitation-Evaporation Index (SPEI, Vicente-Serrano et al. (2010)).

The reference period for the indices is 1993–2000 for the evaluation of the hindcast runs with observations (section 7.1.1) and 1991–2000 for the analyses of the climate simulations and the climate change signal.

However, in order to investigate extremes with much longer reoccurrence intervals or to estimate extremes with a given re-occurrence probability, it is made use of the Extreme Value Theory (EVT) and the peak-over-threshold model.

Table 3.1 Subset of extreme temperature and precipitation ETCCDI indices.

Index	Index description	Indicator definition	unit
CDD	Consecutive dry days	Maximum number of consecutive days when precipitation < 1 mm/day	days
CDDn	Number of dry periods	Number of periods with > 5 consecutive dry days	number
R20	Number of very heavy precipitation days	Annual count when precipitation ≥ 20 mm	days
RX5DAY	Max 5-day precipitation amount	sum of highest 5-day consecutive precipitation period	mm
RX5DAYn	Number of high precipitation events	number of 5-day consecutive precipitation periods ≥ 50 mm	number
R95PTOT	Precipitation percent due to R95p days	Percentage of precipitation days > 95th percentile on the total	%
FD	Frost days	Annual count when daily minimum $T_{\min} < 0$ °C	number
CFD	Consecutive frost days	Maximum number of consecutive days with $T_{\min} < 0$ °C	days
TR20	Tropical nights	Annual count when daily minimum temperature > 20 °C	number
HWDI	Heat wave duration index	Annual maximum period > 5 consecutive days with $T_{\max} > \bar{T}_{\max} + 5$ °C	days
HWDI _n	Number of heat waves	Annual count of heat waves of ≥ 5 consecutive days	number
HWFI	Warm-spell days index	Count maximum period of (≤ 6 days) when $T_{\text{mean}} > 90$ th percentile of a 5-day sliding window over the calendar days of a reference period	days
HWFI _n	Number of warm-spells	Number of periods with (≤ 6 days) when $T_{\text{mean}} > 90$ th percentile of a 5-day sliding window over the calendar days of a reference period	number
CWDI	Cold wave duration index	Annual maximum period > 5 consecutive days with $T_{\min} > \bar{T}_{\min} - 5$ °C	days
CWDI _n	Number of cold waves	Annual count of cold waves of ≤ 6 consecutive days	number
CWFI	Cold-spell days index	Count maximum period (≤ 6 days) when $T_{\text{mean}} < 10$ th percentile of a 5-day sliding window over the calendar days of a reference period	days
CWFI _n	Number of cold-spells	Number of periods (≤ 6 days) when $T_{\text{mean}} < 10$ th percentile of a 5-day sliding window over the calendar days of a reference period	number
GSL.l	Growing season length	Count number of days between GSL.d and the first occurrence after 1 July of at least 6 consecutive days with $T_{\text{mean}} < 5$ °C	days
GSL.d	Begin of GSL	First occurrence of at least 6 consecutive days with $T_{\text{mean}} > 5$ °C	Julian day

3.3.2 Fitting of a POT model

The Generalized Pareto Distribution (GPD)

To analyse extremes which are much rarer than defined by the ETCCDI indices, statistical theories have to be considered. Extremes can be well described by the Extreme Value Theory (EVT). The first theorem of the EVT, the Fisher-Tippett-Gnedenko theorem (Fisher and Tippett, 1928; Gnedenko, 1943), states that suitable rescaled maxima of an identical distributed and independent (i.i.d.) sample follow asymptotically one of the extreme value distribution families: Gumbel, Fréchet and Weibull, which differ substantially in their tail-behaviour (Coles, 2001). It can be shown that these three distribution families can be combined into the Generalized Extreme Value (GEV) distribution (Ghil et al., 2011). This central theorem can be seen as an analogue to the central limit theorem for mean values (Coles, 2001).

The classical extreme value analysis applies the GEV to 'block maxima' of a time series, i.e. annual maxima. This is however statistically wasteful as argued in section 1.3.1. To account for this problem,

the 'peaks-over-threshold' (POT) method uses exceedances above a certain threshold u to define extreme values. According to the second EVT theorem or Pickands-Balkema-De Haan theorem (Pickands, 1975; De Haan and Ferreira, 2006) the intensities of these extremes can be approximated to follow a the Generalized Pareto distribution (GPD) and their frequencies to be a Poisson point process (Ghil et al., 2011). The GPD is one of three basic extreme value families that are usually applied in extreme value analysis. The other two are the GEV and the Point Process (PP) (see section 1.3.1).

Given a sequence of independent and identical distributed (i.i.d.) random variables X_1, X_2, \dots and their maxima

$$M = \max(X_1, \dots, X_n), \quad (3.9)$$

if rescaled by some suitable scaling parameters $a > 0$ and b , their probabilities converge to a max-stable distribution as the sample size $n \rightarrow \infty$:

$$P((M - b)/a \leq x) \rightarrow G(x), \quad (3.10)$$

where $G(x)$ is the GEV distribution:

$$G(x; \mu, \sigma, \xi) = \exp \left\{ - \left[1 + \xi \left(\frac{x - \mu}{\sigma} \right) \right]^{-1/\xi} \right\}, \quad (3.11)$$

with $\mu, \sigma > 0$ the location and scale parameter and ξ the shape parameter. If a threshold u is now chosen high enough, the exceedance probability of a variable X over a threshold u ($X - u$), with the condition that $X > u$, is the GPD:

$$H(x) = \begin{cases} 1 - \left(1 + \frac{\xi x}{\tilde{\sigma}} \right)^{-1/\xi}, & \text{if } \xi \neq 0 \\ 1 - \exp \left(-\frac{x}{\tilde{\sigma}} \right), & \text{if } \xi = 0 \end{cases}, \quad (3.12)$$

with u the threshold, $\tilde{\sigma} = \sigma + \xi(u - \mu)$ the reparametrized scale parameter and ξ the shape parameter. The parameters of the GPD are uniquely associated with those of the GEV, especially the shape parameter ξ (Coles, 2001), which is identical. This duality means that the shape parameter is dominant for the GPD as well as for the GEV.

The choice of the threshold is the difficult part in the POT analysis and is a trade-off between bias and variance (Ghil et al., 2011). A lower threshold increases the sample size but may violate the asymptotic assumption since the tail does not converge to it (bias). In contrast, if the threshold is increased, the sample size reduces, which in turn increases the variance because fewer data are available for fitting the parameters (Coles, 2001; Ghil et al., 2011). There are graphical procedures such as the mean residual life (MRL) (Coles, 2001) plot or stable regions of the parameters are identified by plotting them as a function of threshold. The threshold is then chosen as the smallest value of such an interval. The selection of a threshold is the major weakness of the POT approach (Pandey et al., 2001; Simiu and Heckert, 1996). How the threshold is chosen in this thesis, is described in the next section.

The requirement of identical distributed or stationary data is usually ensured by detrending a time series. Therefore, all time series were detrended by removing a linear trend before the application of the POT model.

Threshold selection

There are several methods for choosing an appropriate threshold. The estimates of σ and ξ can be plotted as functions of a varying threshold, then if the parameters are nearly constant in an interval. Then the threshold is chosen as the smallest value within this interval to keep it as low as possible (Coles, 2001).

Another exploratory method is the mean residual life plot in which the mean of the exceedances above the threshold u is plotted as a function of the threshold. The change of the means is expected to be linear at suitable threshold levels. The threshold is then chosen as the smallest value of such a linear section. However, both methods can be difficult in praxis and are not suited for large data sets, such as climate fields where an individual visual inspection would be necessary for every grid point. Thus, in this thesis the threshold selection is a compromise between accurateness and practical usability. The threshold is set to the local 95 % percentile of each grid box in case of precipitation (Gutjahr and Heinemann, 2013) and to the 90 % percentile in case of minimum and maximum temperature. This procedure has been already used within the study area by Knote et al. (2010). This threshold selection represents a non-parametric approach and is purely based on statistical considerations. However, for other variables, such as discharge, there may be direct physical thresholds to consider (Lang et al., 1999).

The threshold-selection is repeated for every individual grid-point of a data set or scenario. This means that for example the threshold for the C20- and A1B_2-scenario may be different. In this way non-stationary is introduced. For longer and continuous time series this is usually achieved by letting parameters be time-dependent; for instance, the threshold could have a linear trend.

Declustering

In a POT approach it is likely that an extreme exceeding a threshold is accompanied by a neighbouring threshold exceedance, which leads to a clustering of extremes. This is because of the dependency of extremes and thus violates the i.i.d. assumption. Therefore, a declustering method has to be applied. A simple approach is to use only the largest event within a cluster of a certain interval size that is defined by the so-called 'run length', and omit all others exceedances. This accounts for the fact that extremes are close to independence if they are far enough apart (Leadbetter et al., 1989). If any two extremes are closer than r_l data points, they belong to one cluster and hence only the highest extreme is selected. Or in other words, $\pm r_l$ values around an extreme are masked out. The estimation of r_l itself is again a trade-off between bias and variance. For simplicity, it is assumed that the dependence of extremes in time series can be expressed by the auto-correlation function (ACF) (see e.g. Wilks, 2011). Then, the estimated r_l is just the largest time lag d of the ACF which is significant at the 95 % level of a moving average process MA(d-1) (Knote et al., 2010). The procedure is illustrated in Fig. 3.3. The running length is calculated for every individual year and afterwards the mean of these running lengths is used for the declustering. This annual calculation is necessary because otherwise in a seasonal EVT analysis, e.g. for summer precipitation extremes, extremes in one year would possibly be depending on those of the year before.

Parameter estimation and confidence intervals

The parameters of the GPD are fitted with the maximum likelihood estimation (MLE)¹. Although the L-moments optimization (Hosking, 1990) is more robust to outliers if small sample sizes are used, the MLE optimization provides direct methods for estimating the uncertainty of the parameters. If a set of unknown parameters θ is to be estimated, which are elements of a parameter space Θ , then a likelihood function can be defined which gives the probability of the observed data as a function of θ (Coles, 2001):

$$L(\theta) = \prod_{i=1}^n f(x_i; \theta). \quad (3.13)$$

¹ The fitting is performed with the R-function `gpd.fit()` of the *ismev*-package.

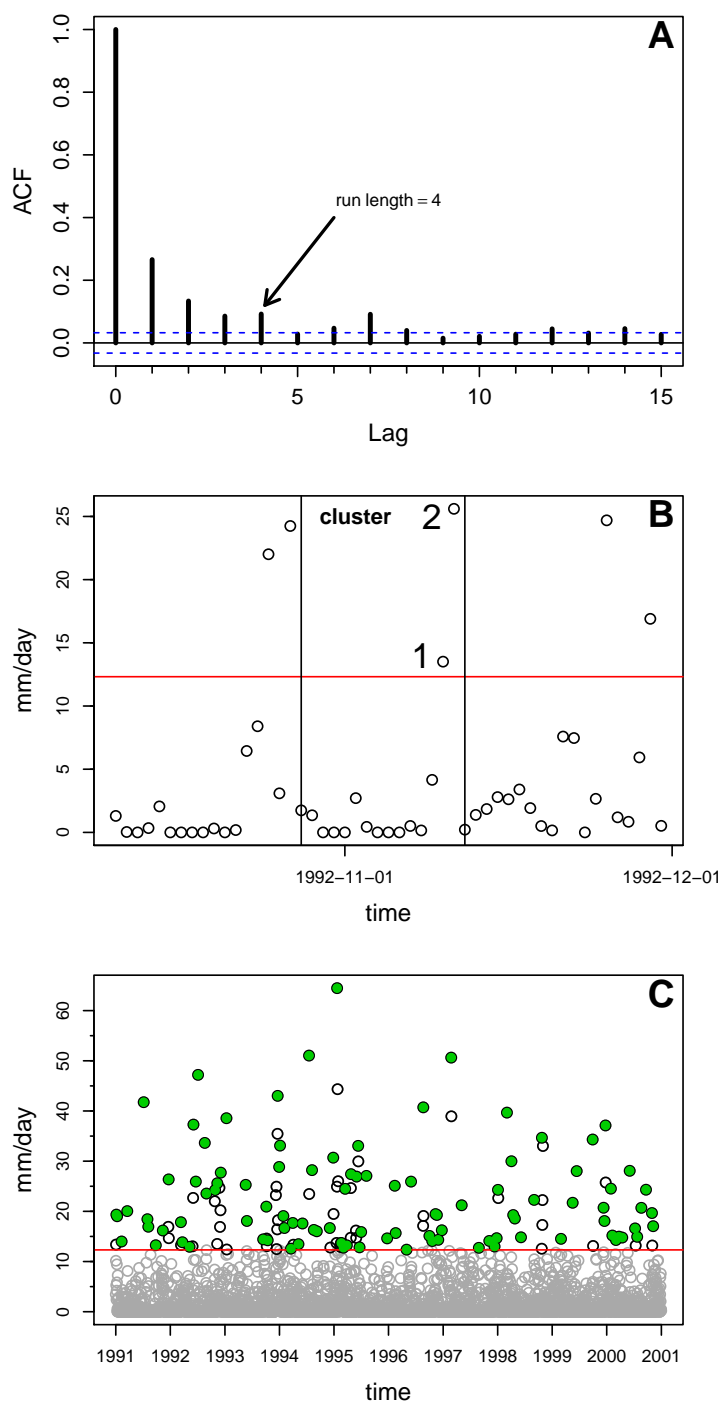


Figure 3.3 Schematic illustration of the running length declustering after Leadbetter et al. (1989). In a) the auto-correlation function is shown along with the run length of the maximum lag which is significant at the 95 % significance interval of a χ^2 -test. In b) an example of an extreme-cluster is shown that contains two extremes labelled as 1 and 2. Because $2 > 1$ only extreme number 2 is retained. In c) the time series after the declustering with a run length=4 is shown.

However, it is more convenient to use the negative logarithm, which yields the negative log-likelihood (nllh) function:

$$\ell(\theta) = -\log L(\theta) = -\prod_{i=1}^n \log f(x_i; \theta). \quad (3.14)$$

For every parameter set the nllh is calculated. Then the maximum likelihood estimator is defined as:

$$\hat{\theta}_{mle} = \operatorname{argmax}_{\theta \in \Theta} \ell(\theta). \quad (3.15)$$

Thus, the most likely estimate of a parameter set is that with the highest nllh because it assigns the highest probability to the observations. Or in other terms, it is searched for those parameters that maximize the nllh function. The MLE is solved by differentiating eq. 3.14 with respect to θ and setting the derivative equal to zero:

$$\frac{d\ell(\theta)}{d\theta} = 0 \quad (3.16)$$

The log-likelihood of the GPD, if $\xi \neq 0$ is thus:

$$\ell(\sigma, \xi) = -k \log \sigma - (1 + 1/\xi) \sum_{i=1}^k \log(1 + \xi y_i / \sigma) \quad (3.17)$$

and in the case where $\xi = 0$:

$$\ell(\sigma) = -k \log \sigma - \sigma^{-1} \sum_{i=1}^k y_i \quad (3.18)$$

A known problem occurs in estimating parameters for the GPD (or GEV) from a small sample size (Coles and Dixon, 1999; Martins and Stedinger, 2000), which is in particular true if in such a case only seasonal data are considered. Although other estimation methods, such as the L-moments (Hosking, 1990), may be better suited for such situations, Martins and Stedinger (2000); Coles and Dixon (1999) show that a censored MLE estimation produces better results without losing the flexibility of MLE. In general such a censoring can be done for the shape and scale parameter, but for this study it is restricted to the shape parameter ξ . As a censor function, or prior function in Bayesian terms, the beta function with the parameters $p = 6$ and $q = 9$ with mean = 0.10 and variance = 0.122² is used. By applying this prior, it is ensured that maximum likelihood estimate of ξ is within the interval $[-0.5, 0.5]$ and is penalized if it deviates from 0.10.

The beta prior (a shifted beta distribution) has the following form (Martins and Stedinger, 2000):

$$P(\xi) = (0.5 - \xi)^{p-1} (0.5 + \xi)^{q-1} / B(p, q), \quad (3.19)$$

with $B(p, q) = \Gamma(p)\Gamma(q)/\Gamma(p+q)$ the beta function (where Γ is the gamma-function). Note that in Martins and Stedinger (2000) the shape parameter is defined as $k = -\xi$ so that the signs in the numerator of equation 3.19 are reversed. The beta prior is shown in Fig. 3.4 for the interval $[-0.5, 0.5]$. Then the likelihood is penalized with:

$$L_{pen}(\theta) = L(\theta) \times P(\xi). \quad (3.20)$$

which introduces an additive penalty term in the equations 3.14-3.18.

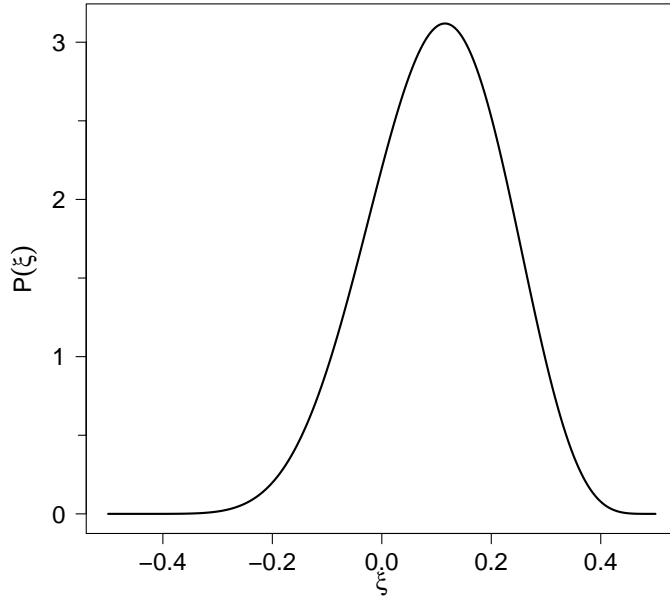


Figure 3.4 Beta-prior for the shape parameter ξ of the GPD after Martins and Stedinger (2000) with $p = 6$, $q = 9$, mean = 0.10, variance = 0.122^2 for the interval $[-0.5, 0.5]$.

Confidence intervals (CI) for the maximum likelihood estimations can be constructed in different ways. The standard method is the so-called *delta-method*. If the true parameters (which are not known) are defined as θ_0 then for large n and under suitable regularity conditions (Coles, 2001) the MLE estimation of θ_0 is assumed to follow a multivariate normal distribution (Beirlant et al., 2004; Venzon and Moolgavkar, 1988), i.e. each individual component of θ_0 follows a normal distribution (Coles, 2001):

$$\hat{\theta}_i \sim N(\theta_i, \tilde{\psi}_{i,i}), \quad (3.21)$$

then the confidence intervals can be calculated from:

$$\hat{\theta}_i \pm z_{\alpha/2} \sqrt{\tilde{\psi}_{i,i}}, \quad (3.22)$$

with θ_i an individual parameter component (e.g. the scale or shape parameter) and $\tilde{\psi}_{i,i}$ the inversion of the observed information matrix, which is an approximation of the expected curvature of the log-likelihood surface and $z_{\alpha/2}$ is the $(1 - \alpha/2)$ quantile of the standard normal distribution (Coles, 2001).

However, if the distribution of the parameters is markedly skewed, then this CI construction is not valid anymore. This is often the case for MLE-estimates from small sample sizes, which causes deviations of the asymptotic properties (Venzon and Moolgavkar, 1988). An alternative and more robust estimation is derived from the profile log-likelihood, which usually gives more accurate CI estimations (Beirlant et al., 2004; Coles, 2001):

$$\ell_p(\theta_i) = \max_{\theta_{-i}} \ell(\theta_i, \theta_{-i}), \quad (3.23)$$

with θ_i an individual parameter component as before and θ_{-i} denotes all other components of θ excluding θ_i . Thus for each value of θ_i , the log-likelihood is maximized with respect to all other components of

θ (Coles, 2001). The profile-likelihood method reduces the log-likelihood maximization to a single parameter of interest by treating all other as nuisance parameters (Venzon and Moolgavkar, 1988). In case of the GPD, the profile log-likelihoods can be calculated for the scale- and the shape-parameter. For instance, if $\theta_i = \sigma$ then $\theta_{-i} = \xi$ and vice versa.

The construction of the CI is then based on the deviance function which is assumed to follow a χ^2 -distribution (Coles, 2001):

$$D_p(\theta_i) = 2(\ell(\hat{\theta}_0) - \ell_p(\theta_i)) \sim \chi^2. \quad (3.24)$$

The deviance is a measure of the distance on the log-likelihood plane between the likelihood from all parameters and the likelihood of a subset of parameters. Since θ -estimations with small deviances correspond to a large likelihood, the confidence region is specified as:

$$C_\alpha = \{\theta_i : D_p(\theta_i) \leq c_\alpha\}, \quad (3.25)$$

where c_α is the $(1 - \alpha)$ quantile of the χ^2 distribution. Thus the confidence region has the prescribed probability of $(1 - \alpha)$ to contain the true parameter θ_0 (Coles, 2001).

Return level estimation

Return levels are defined as the level or value z_m which is exceeded on average once every M -year (i.e. return period). After Coles (2001) the return levels for a POT model are calculated from:

$$z_m = \begin{cases} u + \frac{\sigma}{\xi} \left[(M n_y \zeta_u)^\xi - 1 \right], & \text{if } \xi \neq 0 \\ u + \sigma \log(M n_y \zeta_u), & \text{if } \xi = 0 \end{cases} \quad (3.26)$$

with u, σ, ξ the location, scale and shape parameter, M the M -year return period, n_y the number of observations per year and ζ_u the exceedance probability.

The confidence intervals are constructed by the use of the profile log-likelihood method rather than the delta-method. This is done by transposing eq. 3.26 with respect to σ , substitute this equation into eq. 3.17, and differentiate the obtained one-parameter log-likelihood function with respect to ξ . As a function of z_m this yields the profile-likelihood for the M -return year. Within this thesis the focus was laid on the 10-year and 20-year return period. The reason for this choice is that the sample size, which was used to calculate the return levels, was 10 years long and thus the 10-year return period events were simulated by the regional climate model whereas the 20-year events are entirely based on extrapolations from the fitted POT models. 10-year (20-year) return period events occur with the probability of 10% (5%) per year and are expected on average once every 10 years (20 years). The confidence intervals of the return levels are constructed from the profile log-likelihood method.

Evaluation of the hindcast simulations

4.1 Evaluation of the hindcast runs

In this chapter the hindcast runs CCLM5hc and CCLM1.3hc are evaluated, which allows to assess the model skill of simulating the present climate. The hindcast runs of the CCLM were compared to the observational data sets REGNIE and INTERMET. The time period is the intersection period of the observations and the CCLM simulations: 1993–2000. Unfortunately, the observations do not cover the complete model domain so that the evaluation is restricted to a sub-area, which consists approximately of the state of Rhineland-Palatinate except for the southernmost part. The remainder of the domain has been masked out before the analysis and REGNIE and INTERMET were remapped onto the CCLM grids by a first order conservative interpolation¹ (Jones, 1999). The evaluation was performed for the variables precipitation and 2 m temperature and implied annual means, monthly means, daily variability and diurnal cycles.

Before the output of the CCLM is compared with the observations, the difference in the annual mean precipitation between REGNIE and INTERMET for the period 1993–2000 is shown in (Fig. 4.1). The comparison reveals a distinct pattern: regions with high orography, e.g. the Hunsrück, receive much more precipitation per year in REGNIE than in INTERMET and vice versa river valleys, such as the Moselle Valley, receive less than in INTERMET. In the first case, annual mean precipitation values in mountainous areas are more than 250 mm higher than in INTERMET, whereas the values about 100 mm lower along rivers; for some grid boxes the difference is even more extreme. One reason for these deviations is the weight that was given to the orography in the interpolation. In REGNIE the dependence of precipitation on orography is considerably stronger than in INTERMET. Thus the field is much smoother in INTERMET compared to REGNIE. Another reason for the positive anomalies in high-orographic regions is the correction for undercatchment in REGNIE, which mainly affects such regions. Overall, the annual mean precipitation is about 120 mm higher in REGNIE than in INTERMET.

4.1.1 Annual means

The first comparison was made between the annual means of precipitation and 2 m temperature of the CCLM simulations and the REGNIE/INTERMET data set for the period 1993–2000 (Fig. 4.2).

¹ For the interpolation the CDO routine *remapcon* was used.

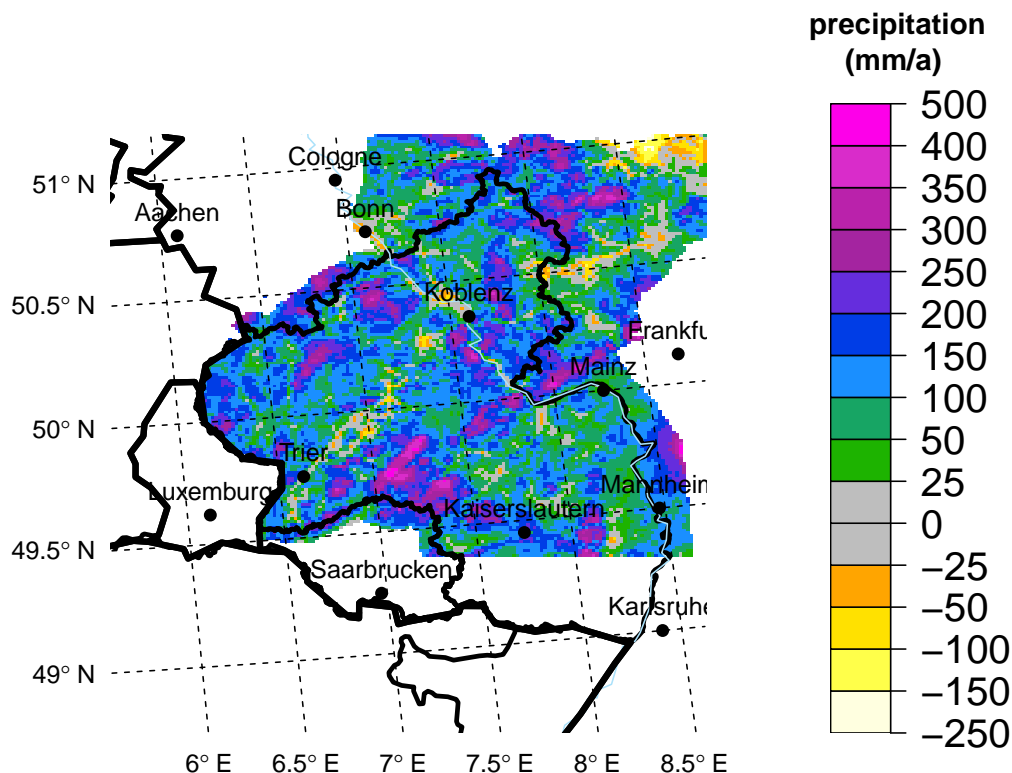


Figure 4.1 Differences of the annual mean precipitation (1993–2000) of REGNIE - INTERMET at a horizontal resolution of 1.3 km.

Precipitation

The spatial distribution of the annual mean precipitation of the CCLM simulations is shown in Fig. 4.2a-b. In general, the annual precipitation is correlated with the orography: higher sums occur in elevated areas whereas lower values are found in flatter terrain and river valleys. The maximum of almost $1500 \text{ mm year}^{-1}$ is simulated by the CCLM4.5hc in the Westerwald. The lowest annual precipitation, about 580 mm year^{-1} , is simulated in the flat Rhine Valley close to the city of Mannheim. The CCLM4.5hc simulates 860 mm year^{-1} in the domain-average whereas the CCLM1.3hc is generally drier with 781 mm year^{-1} . This becomes in particular apparent in the Westerwald and the Rhine Valley. The maximal value is about $1290 \text{ mm year}^{-1}$ and the minimum about 490 mm year^{-1} (see Tab. 4.1).

Both models are dry-biased compared to REGNIE (Fig. 4.2c-d). The bias is about $-56.89 \text{ mm year}^{-1}$ (-6.2%) in the CCLM4.5hc and $-136.25 \text{ mm year}^{-1}$ (-14.8%) in the CCLM1.3hc. This dry-bias is most pronounced in the Eifel, Hunsrück and Palatinate Forest. At the northern boundary of the CCLM1.3hc domain the error is mainly caused by boundary effects, which slightly extend into the investigation area. In comparison with INTERMET the bias is more variable. Now, CCLM4.5hc is wet-biased with about $+63.96 \text{ mm year}^{-1}$ ($+8\%$), mainly along the great river valleys of Rhine and Moselle, but the dry-bias in the CCLM1.3hc reduces to $-15.40 \text{ mm year}^{-1}$ (-1.9%).

The deviations can be ascribed to the different interpolation methods which have been used in

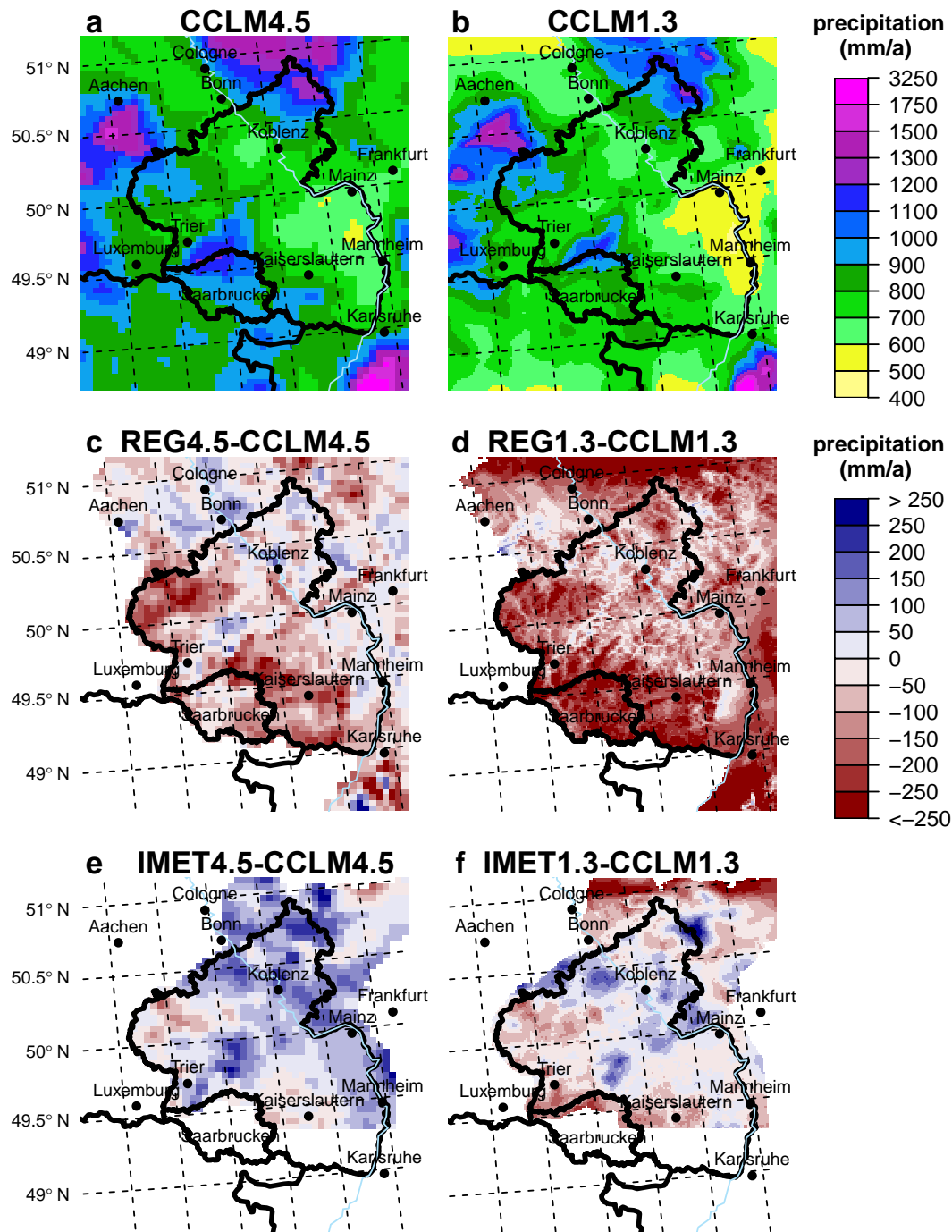


Figure 4.2 Comparison of the annual mean precipitation of CCLM4.5hc and CCLM1.3hc versus REGNIE and INTERMET for the period 1993–2000. Top row: annual mean precipitation sums of a) CCLM4.5hc and b) CCLM1.3hc, centre row: differences c) CCLM4.5hc - REGNIE4.5, d) CCLM1.3hc - REGNIE1.3, bottom row: differences e) CCLM4.5hc - INTERMET4.5, f) CCLM1.3hc - INTERMET1.3.

Table 4.1 Comparison of domain-averaged annual precipitation and 2 m temperature. The values are all calculated for the region which is shown in Fig. 4.1.

data set	precipitation				temperature		
	mean mm/year	bias mm/year	sd mm/year	max/min mm/year	mean °C	bias °C	sd °C
CCLM4.5hc	860	63.96 (-56.89) ¹	121	1490/579	8.80	-0.29	0.68
CCLM1.3hc	781	-15.40 (-136.25) ¹	119	1288/489	8.92	-0.16	0.67
INTERMET	796	-	131	1315/495	9.10	-	0.79
REGNIE	917	120.86	133	1580/570	-	-	-

¹ In brackets the bias compared to REGNIE is shown.

Table 4.2 Bias of seasonal mean precipitation and 2 m temperature from the CCLMhc in comparison with REGNIE (REG) and INTERMET (INT). The bias is shown for the 4.5 km and 1.3 km resolutions.

season	precipitation								temperature	
	REG4.5		REG1.3		INT4.5		INT1.3		INT4.5	INT1.3
	mm	%	mm	%	mm	%	mm	%	°C	°C
DJF	-8	-3.12	-12	-4.8	+29	+13	+25	+11.2	-1.00	-1.08
MAM	+13	+6.3	-8	-3.7	+41	+23.4	+20	+11.5	-0.34	-0.15
JJA	-15	-7.0	-62	-29	+14	+7.5	-34	-18.3	+0.27	+0.61
SON	-46	-18.9	-59	-24	-18	-8.2	-31	-14	-0.12	-0.06

REGNIE and INTERMET, i.e. the algorithm and the number of stations. Furthermore, a Richter-correction (Richter, 1995) for undercatchment due to wind was applied in case of REGNIE and hence systematically higher precipitation values result. The main difference between both interpolation methods seems to be the weighting which was given to the orography. It is much more pronounced in REGNIE than in INTERMET in which precipitation fields are interpolated much smoother. In summary, the CCLM captures the spatial structure of the annual precipitation realistically and achieves annual precipitation values in the observed range.

2 m temperature

The evaluation of the annual mean 2 m temperature is shown in Fig. 4.3. The 2 m temperature fields of the model runs were compared to the INTERMET data set for the period 1993–2000. In general, the warmest temperatures (up to 12 °C) occur in the Rhine Valley and the temperatures are slightly colder along other river valleys, e.g. the Moselle Valley or along the river Nahe. The coldest annual mean temperatures occur in the Westerwald (about 6.4 °C) but other local minima were simulated in the Hunsrück and Eifel areas. The domain-averages of the annual mean temperatures are 8.8 °C with regard to CCLM4.5hc and 8.92 °C in case of CCLM1.3. Both CCLM runs produce a slight cold-bias of about -0.29 °C and -0.16 °C in the CCLM4.5hc and CCLM1.3hc, respectively.

The reduction of the cold-bias is owed to the better resolution of the orography, since temperature is

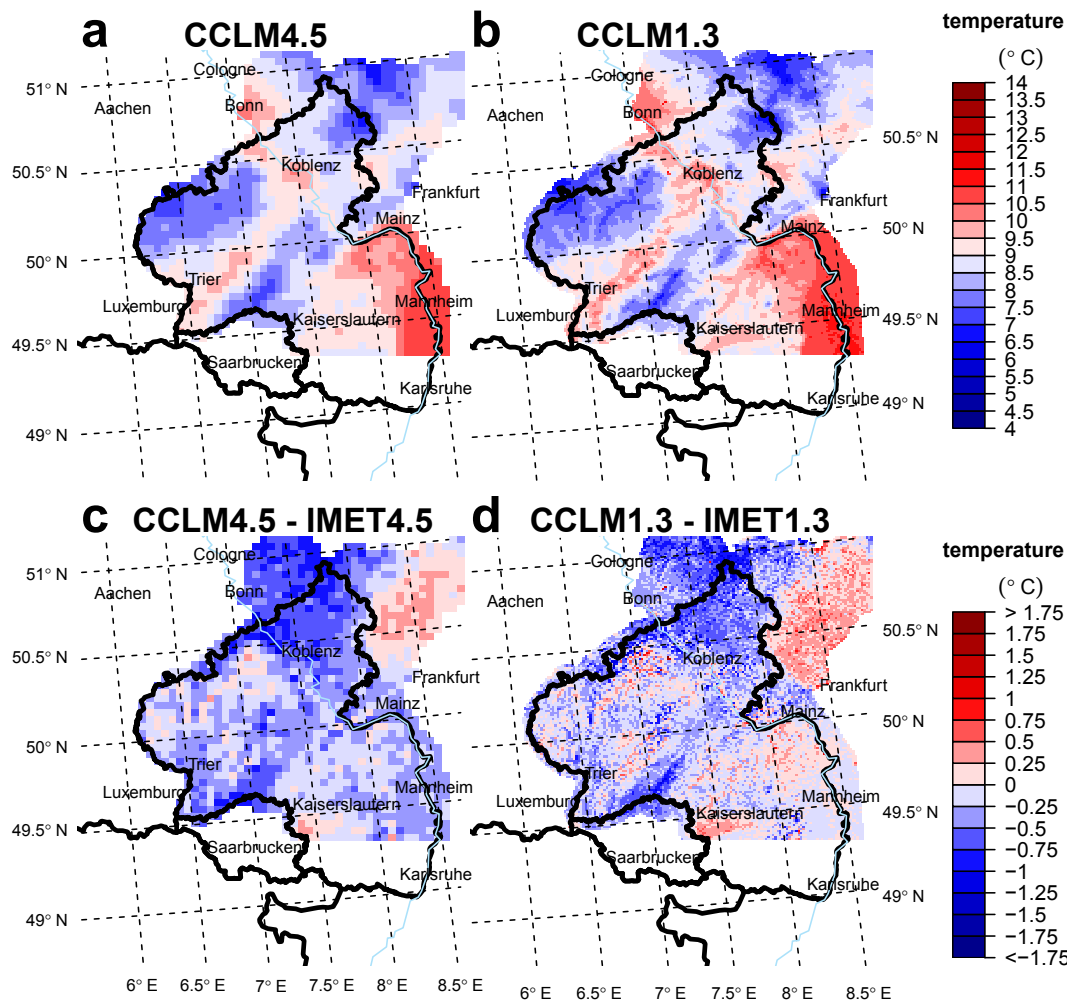


Figure 4.3 Comparison of the annual mean 2 m temperature of CCLM4.5hc and CCLM1.3hc versus INTERMET for the period 1993–2000. Top row: annual mean 2 m temperature of a) CCLM4.5hc and b) CCLM1.3hc, bottom row: differences e) CCLM4.5hc - INTERMET4.5, f) CCLM1.3hc - INTERMET1.3.

height-dependent. Overall, the bias is however small. By comparing Fig. 4.3a and Fig. 4.3b much more small-scale features are present in the CCLM1.3hc temperature field. For instance, secondary valleys of the Nahe or Moselle river catchment or individual mountain ridges in the Hunsrück become visible. However, the cold-bias mainly occurs in mountainous areas what can be seen from Fig. 4.3c-d. Finally, the standard deviation in the CCLM runs are virtually equal, but slightly ($0.1\text{ }^{\circ}\text{C}$) too low compared to INTERMET. It has to be kept in mind that mountainous areas are only sparsely covered with synoptic stations and interpolated fields are likely to be most erroneous in such areas. In summary, both CCLM runs are in good agreement with the observations.

4.1.2 Monthly means

In a next step monthly means of precipitation and temperature were compared with REGNIE and INTERMET. Therefore, the hourly output of the CCLM was aggregated to monthly sums and means,

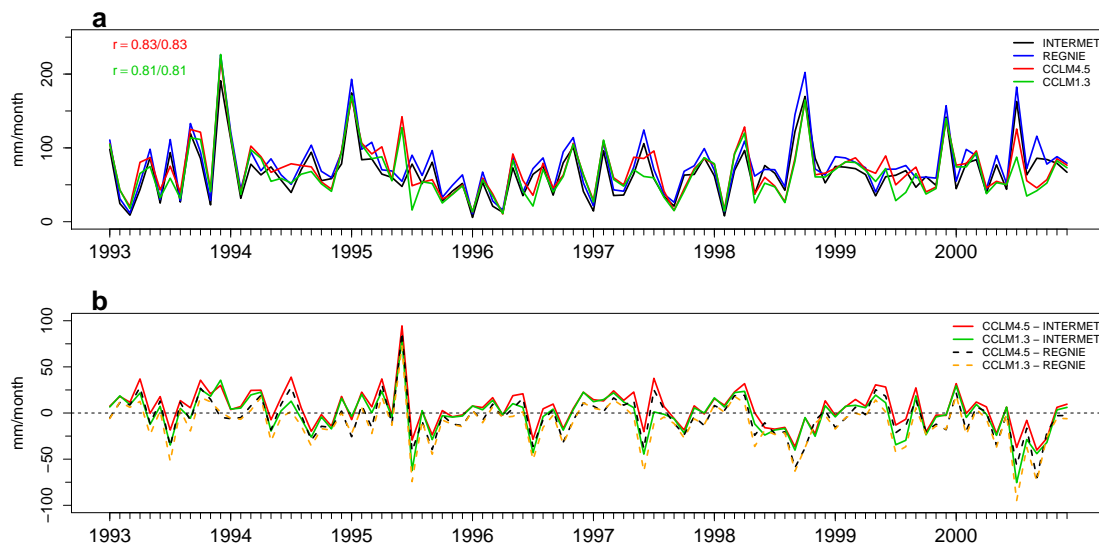


Figure 4.4 Monthly time series of precipitation for the period 1993–2000 in a) and in b) the differences of CCLM minus INTERMET and REGNIE, respectively.

respectively.

Precipitation

The time series of monthly precipitation for the period 1993–2000 is shown in Fig. 4.5a. Focusing on the comparison of REGNIE (blue line) and INTERMET (black line) first, the above mentioned systematic positive offset due to the Richter-correction becomes visible. The CCLM runs are in good agreement with the observations; the Pearson's correlation coefficients are $r = 0.83$ for the CCLM4.5hc and $r = 0.81$ for the CCLM1.3hc, respectively. However, in absolute terms, the CCLM1.3hc yields systematically lower monthly sums than the CCLM4.5hc; in particular in summer, e.g. 1997. Fig. 4.5b shows the differences of CCLM minus REGNIE and CCLM minus INTERMET. The differences vary between $\pm 50 \text{ mm month}^{-1}$ for most of the time except in the summer of 1995 and 2000. In the summer of 1995 there is a distinctive peak in both CCLM runs (about $+100 \text{ mm month}^{-1}$), which did not occur in the observations. The reason for this deviation cannot be deduced from this plot, but it is possible that the CCLM fails to reproduce the synoptic situation in this period. Another spike, though of opposite direction, occurred in the summer 2000 where the CCLM runs simulate about $-75 \text{ mm month}^{-1}$ too less precipitation.

The comparison of the climatological monthly means reveals more clearly the deviations which were already found in the previous figures. Fig. 4.5a shows the absolute values of the climatology whereas Fig. 4.5b shows the deviations from the observations. With regard to INTERMET, both models overestimate the precipitation in winter and in early spring but underestimate it in late summer and autumn. The bias of the CCLM4.5hc varies in winter between $10 - 20 \text{ mm month}^{-1}$ and in summer between -8 to $-12 \text{ mm month}^{-1}$ while the higher resolution model increases the bias in summer, in particular in July ($-32 \text{ mm month}^{-1}$), but reduces the bias slightly in winter to $10 - 15 \text{ mm month}^{-1}$. Compared to REGNIE the bias in December, January and February reverses to a slight dry-bias, whereas the summer dry-bias increases even more. The maximum reaches $-45 \text{ mm month}^{-1}$ in July in the CCLM1.3hc, which accounts for about 50% of the total sum.

In summary, the models overestimate precipitation in spring and underestimate it in summer and

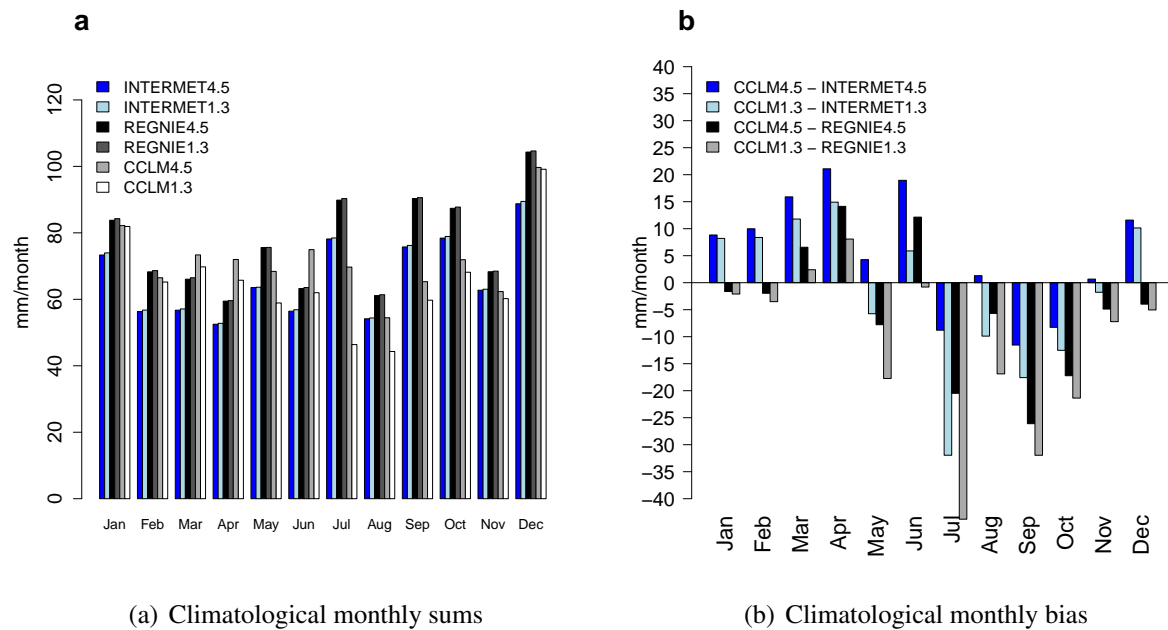


Figure 4.5 Comparison of climatological monthly precipitation sums in a) and in b) the differences of CCLM minus INTERMET and REGNIE, respectively.

autumn. In winter, a slight underestimation was found from the comparison with REGNIE but a slight overestimation in the comparison with INTERMET.

2 m temperature

The monthly mean time series of the CCLM runs and INTERMET are shown in Fig. 4.6a. The temporal correlation with INTERMET is $r = 0.99$ for both models. However, in absolute terms there are some deviations with respect to INTERMET. The models tend to simulate too cold temperatures in winter and slightly too warm temperatures in summer. This can be seen in Fig. 4.6b for individual months and in Fig. 4.7 for the climatological means. From Fig. 4.7b the cold bias in winter of the CCLM4.5hc is about -1.3°C in January and reduces to -0.3°C in April. The bias of CCLM1.3hc depicts a similar development but does nearly vanish in April. From May–September, the CCLM is generally warm biased with about $+1^{\circ}\text{C}$ (except in June where the CCLM4.5hc is slightly too cold). The warm bias of CCLM1.3hc is about $+0.3^{\circ}\text{C}$ higher than in the CCLM4.5hc. From October to December the bias gets negative again with values between -0.5°C and -1.2°C .

4.1.3 Taylor diagrams

For assessing the performance at the daily time scale, the precipitation and 2 m temperature have been aggregated from hourly to seasonal daily fields and averaged for every calendar day. Finally, the fields were spatially averaged. The performance is summarized in a Taylor diagram (Taylor, 2001) which combines Pearson's correlation coefficient, the standard deviation (σ) and the root-mean-squared error (RMSE). For a better comparison, the standard deviations have been normalized with the observed standard deviation. In Fig. 4.8a the precipitation is compared with REGNIE. Without seasonal splitting a correlation of $r = 0.65$ is achieved for both models with only a marginal better σ with regard to

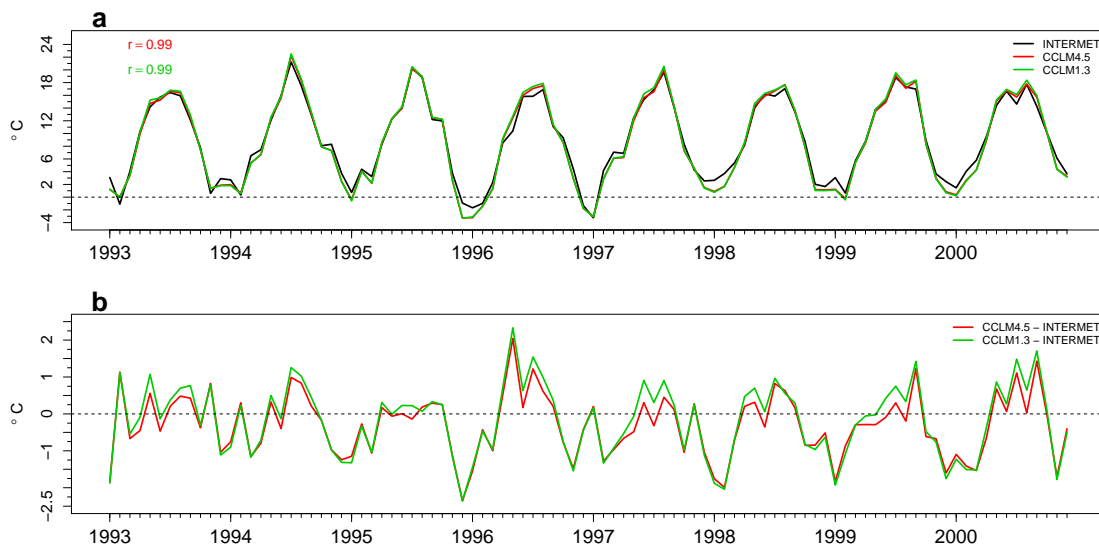
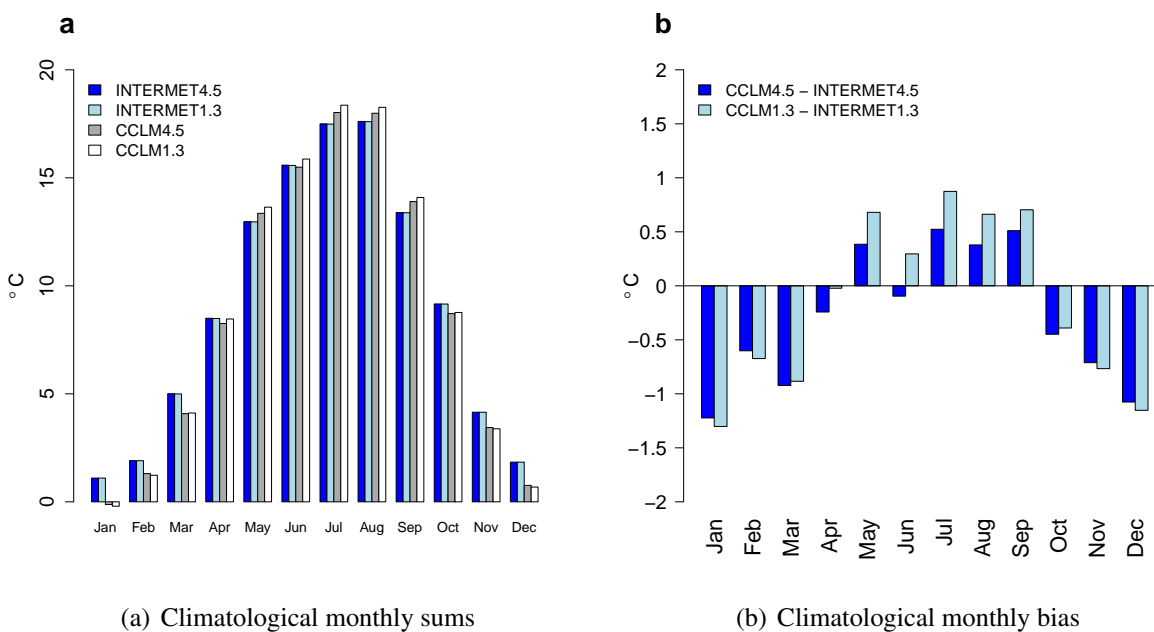


Figure 4.6 Monthly time series of 2 m temperature for the period 1993–2000 in a) and in b) the differences of CCLM minus INTERMET.



(a) Climatological monthly sums

(b) Climatological monthly bias

Figure 4.7 Comparison of climatological monthly means of 2 m temperature in a) and in b) the differences of CCLM minus INTERMET.

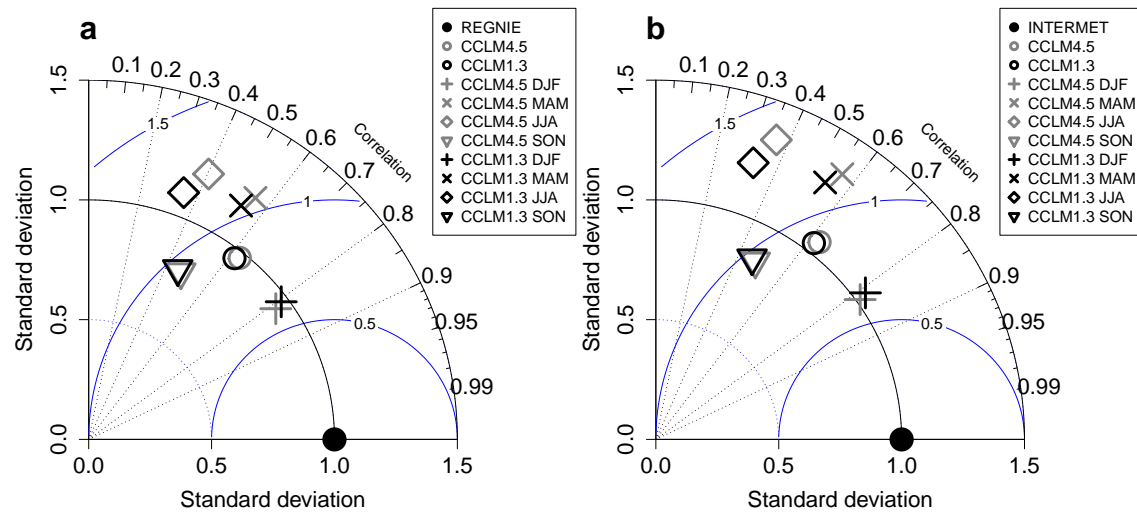


Figure 4.8 Taylor diagrams of seasonal monthly precipitation for the period 1993–2000. In a) CCLM compared to REGNIE and in b) CCLM compared to INTERMET.

CCLM1.3hc. Both CCLM perform best in winter ($r = 0.8$, σ close to 1 and $RMSE \approx 0.6$) and worst in summer ($r = 0.3 - 0.4$, $\sigma = 1.2 - 1.3$ and $RMSE = 1.1 - 1.3$). The correlation in summer is slightly higher with regard to CCLM1.3hc but σ is even higher. In spring and autumn, the performance is only moderate. This result could be expected since the synoptical forcing is much stronger in winter than in summer, where precipitation is mainly formed by local convection. The timing and occurrence of convection depend much more on local conditions and thus have a larger variability in space and time.

Similar results were obtained from the comparison with INTERMET (Fig. 4.8b) except that the deviation of σ is larger. The reason is that the INTERMET precipitation fields are much smoother than the REGNIE fields and hence exhibit lesser spatial variability.

The comparison of the daily 2 m temperature yields better results (Fig. 4.9). In general, the correlation is ≥ 0.9 , the $RMSE \leq 0.5$ and the $\sigma = 1 - 1.2$. There is hardly any difference between the models, only in autumn (SON) and spring (MAM) is the standard deviation closer to 1 with regard to CCLM1.3hc. In winter, the correlation is somewhat lower ($r = 0.9$) but σ is matched precisely.

4.1.4 Standard deviation

In this section the spatial pattern of the standard deviation of precipitation and 2 m temperature is investigated. The analysis is restricted to daily standard deviation since the results for hourly σ are comparable (not shown). Furthermore, in case of precipitation only the comparison with REGNIE is shown because the results of the comparison with INTERMET add no new information except the above mentioned smoothness of the precipitation field. By increasing the spatial resolution, it is expected that small-scale features (e.g. valley circulations, orographic induced convection) are better simulated. This should reduce averaging effect and cause higher variability. In a statistical sense, the variance should increase while the mean remains constant.

In Fig. 4.10, the spatial fields of standard deviation of daily precipitation are shown. The spatial pattern is clearly connected to the elevation height. The CCLMs (Fig. 4.10a-b) simulate lower variability in flat areas, e.g. in the Rhine Valley ($\sigma \leq 4 \text{ mm day}^{-1}$) and higher variability in mountainous areas, for instance in the Westerwald, Eifel and the Hunsrück where $\sigma > 7 \text{ mm day}^{-1}$. This pattern is also

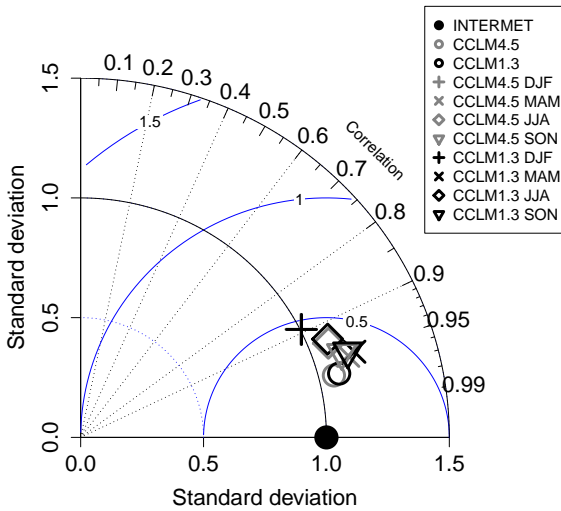


Figure 4.9 Taylor diagram of seasonal monthly 2 m temperature for the period 1993–2000.

present in the REGNIE data set Fig. 4.10c-d. The comparisons to REGNIE Fig. 4.10e-f yield a very good agreement with deviations of only $\pm 0.5 \text{ mm day}^{-1}$ for most part of the domain. One exception is the considerably higher standard deviation ($\sigma = +3 - 4 \text{ mm day}^{-1}$) near the city of Koblenz in both CCLM simulations. This deviation is owed to a very high intensive precipitation event which was simulated at 1995-08-07, but did not occur in the observations. This is the same anomaly which was noticed in the monthly mean time series Fig. 4.4.

The fields of the standard deviation of 2 m temperature show larger deviations. In Fig. 4.11a-b the spatial pattern is again associated with the orography. The largest variability is simulated in the Rhine Valley ($\sigma = 8 - 9^\circ\text{C}$) and the lowest variability in higher areas ($\sigma \approx 7^\circ\text{C}$). Comparing the σ -field of CCLM1.3hc with that of CCLM4.5hc, the largest differences occur in the Rhine Valley. Firstly, the area of high variability is broader in the CCLM1.3hc and secondly, small localized regions with the highest variability ($\sigma \approx 9^\circ\text{C}$) occur close to the river Rhine, which are absent in the CCLM4.5hc. In addition, the region with lower variability is more structured, i.e. the mountain areas are separated by valleys.

The spatial field from INTERMET (Fig. 4.11c-d), has a much reduced orography-dependent gradient compared to CCLM. Most of the domain has a standard variation of $\sigma = 7 - 7.5^\circ\text{C}$, only the western heights vary less than 7°C . This becomes especially visible in the difference plots (Fig. 4.11e-f). The largest deviation is about 2°C in the Rhine Valley with respect to CCLM1.3hc. This indicates that the interpolation method may produce too less variability in the temperature fields. However, this has to be confirmed with synoptic station data. In summary, the CCLM simulations provide a realistic pattern of variability, although they deviate in the Rhine Valley from the observations. The standard deviation of the CCLM1.3hc depicts more small-scale features than the CCLM4.5hc, which is clearly due to a better representation of the orography and thus adds information to the model output.

4.1.5 Histograms and diurnal cycle

Histograms

To further analyse the daily and hourly characteristics of precipitation and 2 m temperature, daily and hourly histograms were calculated. In winter (Fig. 4.12a), the frequencies of daily precipitation are in

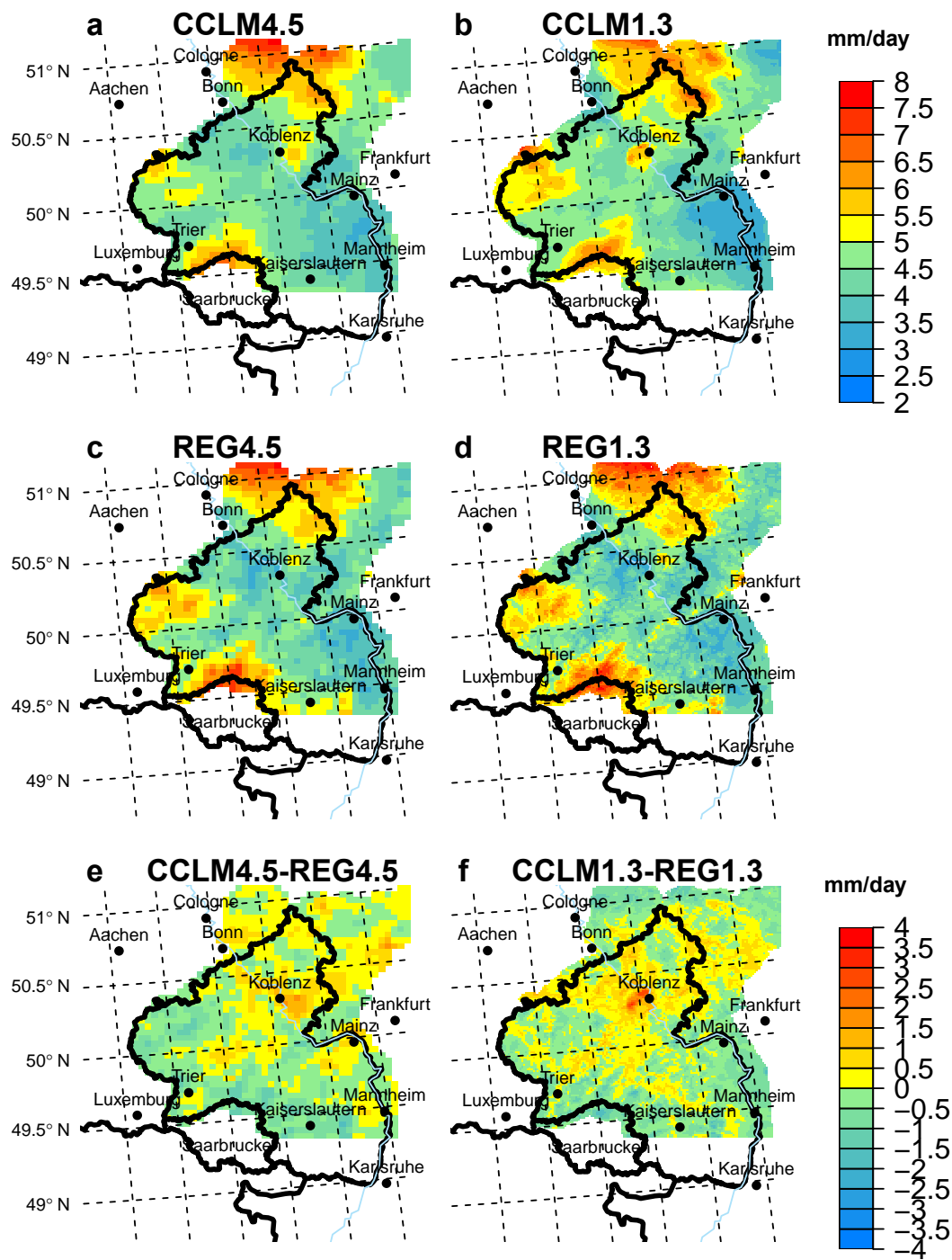


Figure 4.10 Standard deviations of daily precipitation from CCLM and REGNIE for the period 1993–2000. a) CCLM4.5hc, b) CCLM1.3hc, c) REGNIE4.5, d) REGNIE1.3, e) CCLM4.5hc - REGNIE4.5 and f) CCLM1.3hc - REGNIE1.3.

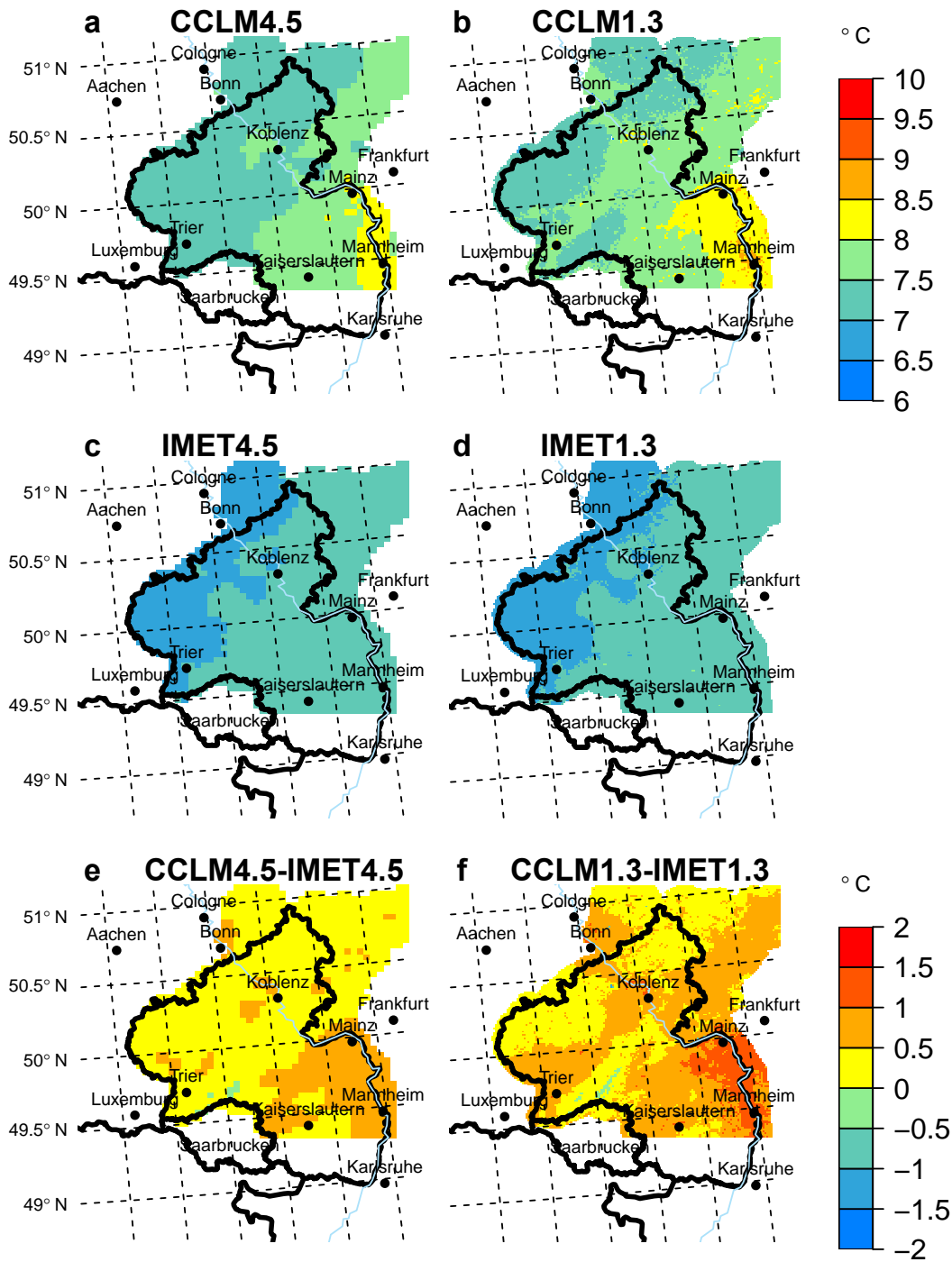


Figure 4.11 Standard deviations of daily temperature from CCLM and INTERMET for the period 1993–2000. a) CCLM4.5hc, b) CCLM1.3hc, c) INTERMET4.5, d) INTERMET1.3, e) CCLM4.5hc - INTERMET4.5 and f) CCLM1.3hc - INTERMET1.3.

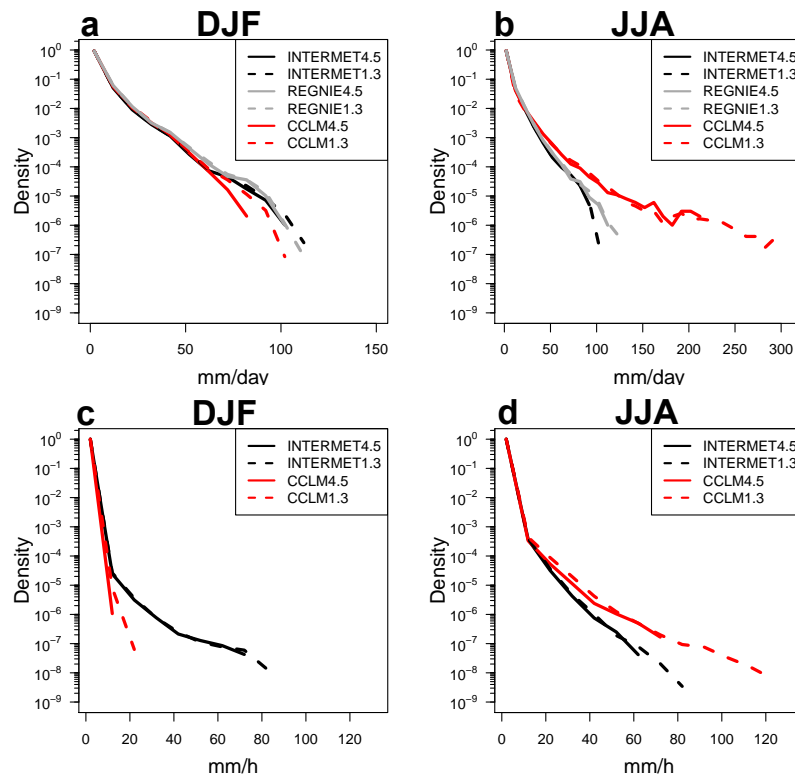


Figure 4.12 Histograms of precipitation: a) daily, winter (DJF), b) daily, summer (JJA), c) hourly, (DJF) and d) hourly, (JJA).

good agreement with the observations. The extremes ($> 60 \text{ mm day}^{-1}$) are strongly underestimated by the CCLM4.5hc model; the maximal value is only about 80 mm day^{-1} , whereas in the observations about 120 mm day^{-1} occurred (note that there are discrepancies between REGNIE and INTERMET due to the interpolation scheme and the number of considered stations). The CCLM1.3hc is able to reproduce those extremes, although the frequencies of extremes are slightly underestimated. In summer (Fig. 4.12a), the error is reversed. Precipitation values of about $> 50 \text{ mm day}^{-1}$ are overestimated in both models and in particular values $> 100 \text{ mm day}^{-1}$. The maximal observed daily precipitation sums are $100 - 120 \text{ mm day}^{-1}$ whereas CCLM4.5hc produces a maximum of about 210 mm day^{-1} and CCLM1.3hc even 300 mm day^{-1} .

On the hourly timescale, there is an underestimation of winter intensities as well (Fig. 4.12c). The CCLM4.5hc simulates only values of maximal 15 mm hour^{-1} and CCLM1.3hc of about 20 mm hour^{-1} , which is clearly less than the observed 80 mm hour^{-1} from INTERMET. In summer, the CCLM overestimates precipitation values $> 20 \text{ mm hour}^{-1}$, the maxima are about 75 mm hour^{-1} in case of CCLM4.5hc, 120 mm hour^{-1} in case of CCLM1.3hc and between $80 - 100 \text{ mm hour}^{-1}$ in INTERMET.

In order to analyse the frequency of low intensities, a quantile-quantile (QQ) plot is used (Fig. 4.14). To emphasise the general deviations, domain-averages were used for the QQ-plot. Again, higher intensities are underestimated (Fig. 4.14a), but low intensities are in good agreement with REGNIE. If compared to INTERMET (Fig. 4.14c) however, CCLM slightly overestimates low intensities and even stronger the high intensities. Larger deviations were found for summer low intensities (Fig. 4.14b,d). While the CCLM4.5hc only slightly underestimates the low intensities up to 10 mm day^{-1} compared to

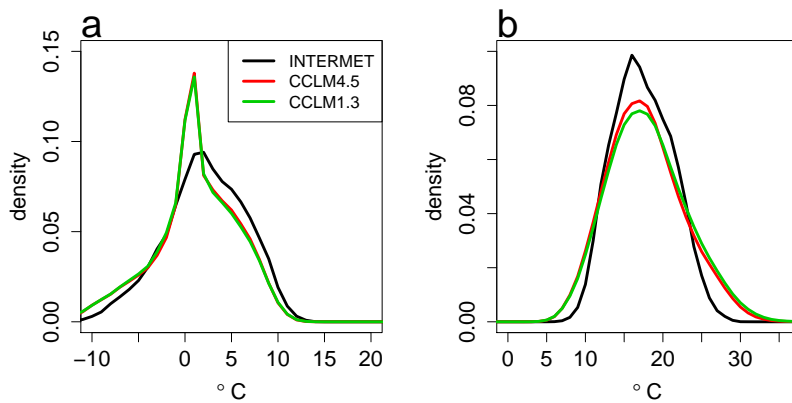


Figure 4.13 Hourly PDFs of 2 m temperature (1°C wide intervals): a) for winter (DJF) and b) for summer (JJA).

REGNIE (Fig. 4.14b) and almost fits perfectly compared to INTERMET (Fig. 4.14d), the CCLM1.3hc strongly underestimates them in both cases. Higher intensities are overestimated from both models, though.

According to the histograms and the QQ-plot it can be concluded that CCLM underestimates wintertime precipitation intensities, which are mainly linked to large-scale or synoptical forcing, whereas in summer, high intensities are strongly overestimated. Furthermore, the summer dry bias originates from an underestimation of low intensities ($< 10 \text{ mm day}^{-1}$).

Fig. 4.13a shows the PDF of hourly 2 m temperature in winter. The shape of the PDF is captured by the CCLM although there are some differences. The CCLM produces temperatures in the interval $0-1^\circ\text{C}$ approximately 5% more often than INTERMET. The reason for this overestimation in frequency was not investigated further, but it gives reason to hypothesise that an error in the surface module TERRA-ML causes temperatures of 0°C too often due to melting of snow and ice. In addition, the 2 m temperature is not a direct variable in the CCLM output, but is extrapolated. This high frequency of 0°C was also found in the forcing data (CLM18) by Hollweg et al. (2008) and according to their results it is most pronounced in the region of Germany. Further deviations in both simulations were found in the interval of -10°C to -5°C , where cold temperatures occur too frequently. Temperatures between $2-12^\circ\text{C}$, however, are less frequent than observed. These deviations produce in total a cold-bias in both CCLMs. In summer (Fig. 4.13b), the PDF of the 2 m temperature of CCLM is too widespread because of a higher variance and consequently temperatures at the tails of the PDF occur too frequent whereas values close to the mean are underestimated.

Diurnal cycles

The mean diurnal cycles of domain-averaged precipitation in winter (DJF) and summer (JJA) are shown in Fig. 4.15. In winter (Fig. 4.15a), the CCLM overestimates the hourly intensities with respect to INTERMET, particularly in the night-time and late evening. The simulations differ mainly between 9-16 UTC where the CCLM1.3hc simulates to less precipitation, which is probably due to the shallow-convection scheme. In summer, the CCLM4.5hc does match the absolute intensities, but the onset of convection is too early (early midday). The Tiedtke-scheme produces convective precipitation mainly between 7 and 11 UTC while in the observations the convection is most pronounced in the late afternoon at 17 UTC. However, if the hourly values are accumulated, this temporal bias vanishes. Therefore, the

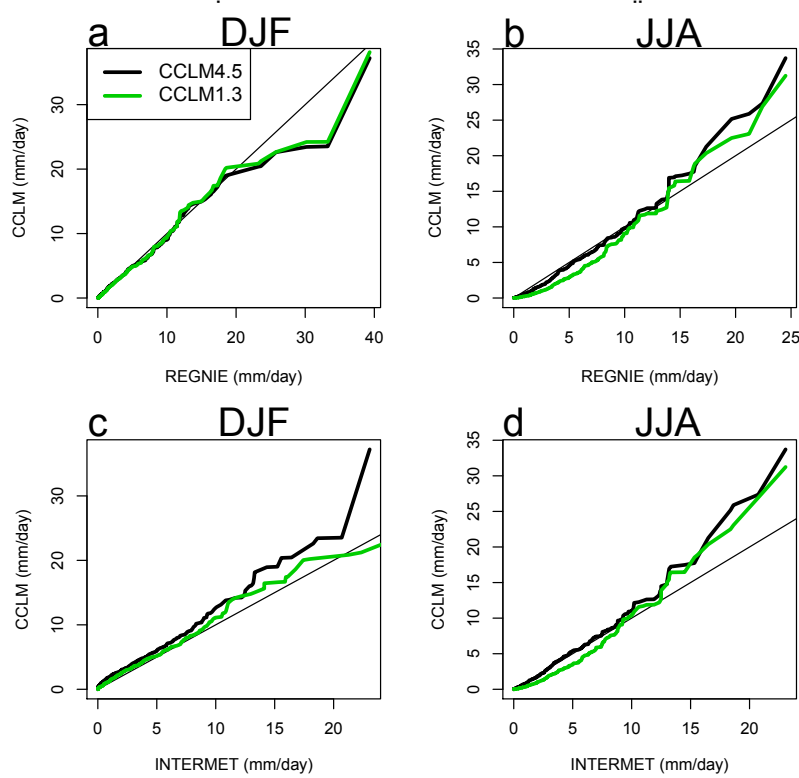


Figure 4.14 Quantile-Quantile plots of domain-averaged daily precipitation. In a) the CCLM runs are compared to REGNIE for winter (DJF), in b) for summer (JJA). In c) the CCLM runs are compared to INTERMET for winter and d) summer, respectively.

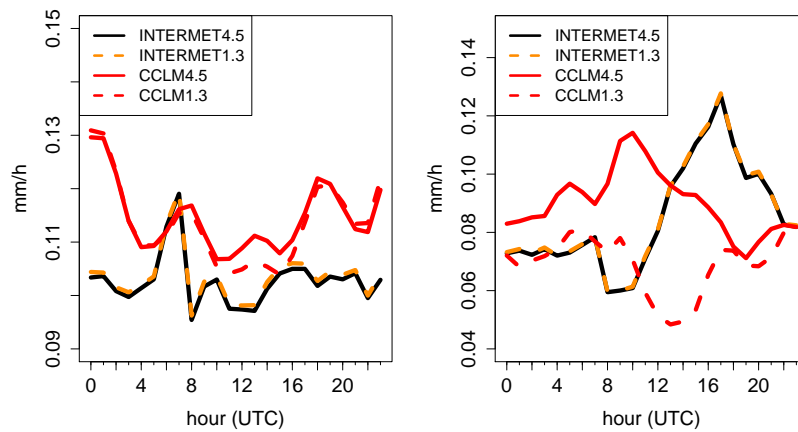


Figure 4.15 Mean domain-averaged diurnal cycle of precipitation for (left) winter (DJF) and (right) summer (JJA) in the period 1993–2000.

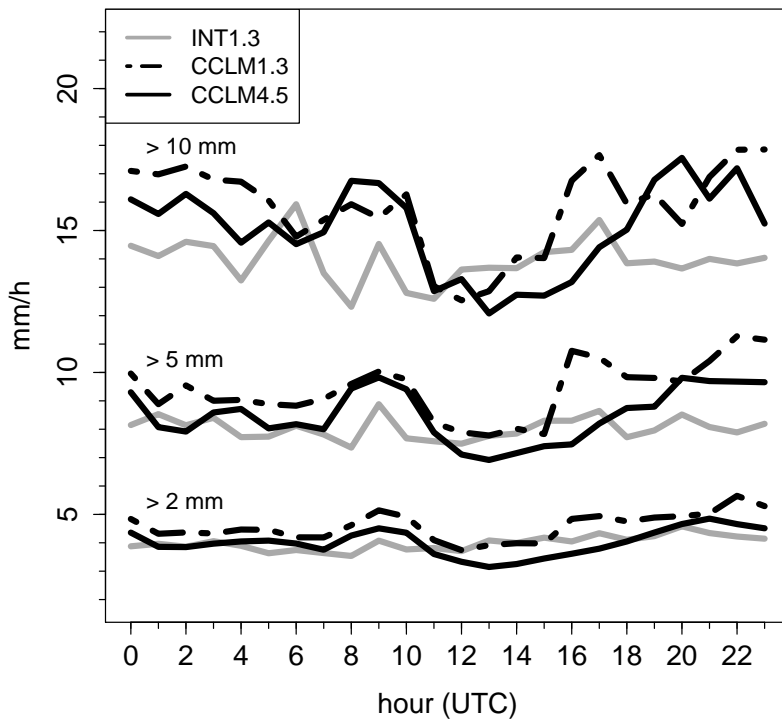


Figure 4.16 As in Fig. 4.15 (right) but for precipitation values exceeding a threshold of 2, 5 or 10 mm hour⁻¹. As reference only the diurnal cycles of INTERMET1.3 are plotted.

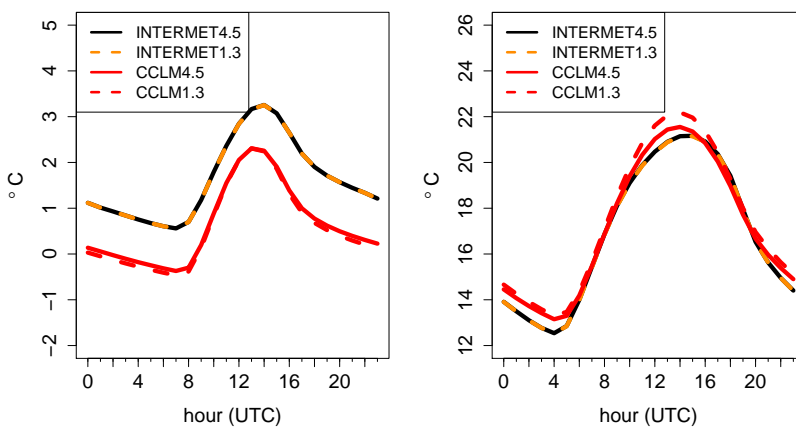


Figure 4.17 As Fig. 4.15 but for 2 m temperature.

bias of CCLM4.5hc is small in the summer. Distinctly different is the magnitude and temporal sequence with regard to CCLM1.3hc. The diurnal cycle matches the observations between 0–8 UTC but at midday there is a drop in intensity with a minimum at 13 UTC. In the late afternoon, the intensity rises but does not reach the observed peak. However, if the diurnal cycle is calculated for precipitation exceeding a threshold of 2, 5 or 10 mm hour⁻¹ (Fig. 4.16), then the convective precipitation is considerably increased compared to CCLM4.5hc. It seems, that there is a problem in producing convective precipitation of low intensities. To assess the reason for this error a systematic sensitivity study of the model would be necessary. This was however out of scope due to the limitation of computing time. This error explains the large dry-bias of the CCLM1.3hc in summer. Combined with the information from Fig. 4.14d, the error mainly affects the formation of low-intensity precipitation (0 – 10 mm day⁻¹), whereas more extreme days are not or only minor affected.

The diurnal cycles of 2 m temperature are shown in Fig. 4.17. In winter, there is a systematic cold-bias of approximately the same magnitude in both models. This bias is about -1 °C at a maximum. But the temporal pattern is well captured by both CCLM. In summer, there is a slight warm-bias in the night (nocturnal summer boundary layer) and at midday (10 – 16 UTC). The warm-bias at midday is slightly larger in the CCLM1.3hc. This is caused by the change from the Tiedtke to the shallow-convection scheme. The shallow convection produces less clouds at midday, which leads to warmer soil surface temperatures (not shown). By this feedback, convection is enhanced but too less precipitation reaches the surface (not shown).

4.2 Summary and discussion

In this chapter the COSMO-CLM hindcast simulations were compared to the gridded observational data sets REGNIE and INTERMET. The CCLM was run at two horizontal resolutions: 4.5 km and 1.3 km, for the time period 1993–2000. The evaluation was performed for the variables precipitation and 2 m temperature. Both models are able to reproduce the spatial distribution of the observed annual precipitation field although a dry-bias of about -6.2% and -14.8% was found in comparison with REGNIE for the 4.5 km and 1.3 km run, respectively. However, if compared with INTERMET this bias reduces to -1.9% in the CCLM1.3 or gets even positive ($+8\%$) in the CCLM4.5. Possible reasons might be the different interpolation approach and that REGNIE has been corrected for undercatchment by the Richter-correction. Thus it is of essential importance which data set is used for the evaluation since the bias depends on the quality of the observations. But the availability of two data sets with a high-resolution of about 1 km is rarely given for the same area. Furthermore, the INTERMET data set has a temporal resolution of 1 hour that allows for sub-daily spatial comparisons, e.g of diurnal cycles, which would be restricted to station locations otherwise.

The comparison on a monthly basis revealed a time-varying bias of the absolute values but the temporal correlations are $r = 0.81 - 0.83$. The largest correlation coefficient was found for winter and the lowest for the summer season. In the winter season, the CCLM simulations are wet-biased, whereas in summer there is a severe dry-bias, in particular in July with about -50% in the CCLM1.3. This bias is caused by too less precipitation in the afternoon and evening, which indicates that the convective part of precipitation is possibly too low. From the quantile and histogram plots it became evident that in particular intensities < 10 mm day⁻¹ occurred less frequently, whereas more extreme intensities did occur more often than in the observations. This underestimation of convective precipitation maybe related to the shallow convection parameterization and the soil-moisture feedback. Although many preconditions for enhanced convective precipitation were given in the CCLM1.3 (not shown) there is too less rain precipitating under such circumstances. For example, the soil moisture is lower in the CCLM1.3 than in the CCLM4.5 which should favour the production of thermals and the triggering of deep convection

(Hohenegger et al., 2008). According to this, the production of turbulent kinetic energy is enhanced in the CCLM1.3 (Fig. B.1 in the appendix) but more detailed investigations of the boundary layer processes are necessary to solve this problem in the future. However, for precipitation $> 2 \text{ mm hour}^{-1}$ a clear enhancement of the convective precipitation was found. This leads to the conclusion that only convective precipitation events with weak intensities are underestimated whereas events with higher intensities are enhanced in the 1.3 km simulation compared to the coarser model. However, these intensities are higher than in the observations so that convective events with large amounts of precipitation are overestimated by the CCLM1.3. Another possible explanation for this mismatch could be the density of the meteorological network. If the distance between two stations is smaller than the horizontal extension of a thunder cell or the area of rain, then no precipitation would be registered at those stations.

In winter, it was found that extremes are underestimated with respect to REGNIE and slightly overestimated with regard to INTERMET. However, at the hourly timescale the extremes are clearly overestimated in summer and underestimated in winter. Finally, the standard deviation of daily precipitation is well captured by the CCLM. Deviations vary between $\pm 0.5 \text{ mm day}^{-1}$. Larger differences were only caused by a single event near Koblenz which did not occur in the observations.

The annual and seasonal temperature is well simulated by the CCLM with correlation coefficients of $r = 0.99$. However, a seasonally varying bias was found as well. In winter, the daily mean temperature is about 1°C too cold in the spatial average and in summer a warm-bias of $+0.27^\circ\text{C}$ was found. The daily standard deviations were considerably higher simulated in the Rhine-Valley compared to INTERMET, which may be due to the smooth interpolation of the temperature field. The structure of the diurnal cycles in summer and winter were captured, although there is a systematic cold-bias in winter and a warm-bias in the nocturnal and midday summer temperatures. The latter is caused by reduced cloud formation in the CCLM1.3 during midday whereby more shortwave radiation reaches the surface.

An anomalous high frequency of 0°C was detected in winter, which was also found by Hollweg et al. (2008) in the forcing data. According to their argumentation, this is a systematic error that occurs when snow and ice are melting and due to the computation of the 2 m temperature. In summer, the warm extremes occur more frequently than in INTERMET but the core of the distribution is underestimated. As a consequence, the PDF is too broad which confirms the findings of a larger daily variability.

Bias correction of the hindcast and climate simulations

In this chapter the performance of the bias-correction (BC) is investigated. The BC was applied to the CCLM4.5hc/CCLM1.3hc and to the CCLM4.5/CCLM1.3 precipitation and temperature fields of the C20 scenario. However, the performance analysis was restricted to the simulations with a resolution of 4.5 km.

Four variants of the quantile-matching were applied to the daily precipitation fields of the CCLM and their performance was compared and quantified by the skill score MAE_x (see section 3.2.3) in a cross-validation. The skill score measures the differences of quantiles in 10 % wide interval classes. This and the additional computation of higher moments of the probability distribution functions allow a more detailed comparison than a restriction to the means and standard deviations. The applied methods were the empirical quantile-matching (eQM) (Boé et al., 2007; Déqué, 2007) and the parametric alternatives: PTF (Piani et al., 2010b), gQM (Piani et al., 2010a) and GQM (Gutjahr and Heinemann, 2013). The temperature fields were corrected with the eQM approach but the results were not tested within a cross-validation.

Finally, the corrected and uncorrected fields of precipitation and temperature were compared with observations.

5.1 Cross-validation

The performance of the BC methods was analysed with the help of the skill score MAE_x , which was introduced by Gudmundsson et al. (2012) (see section 3.2.3). For this analysis it is necessary to divide the data set into a calibration and a validation period. This was done via a cross-validation (CV) by omitting one year from the calibration period, calculate the transfer functions and then apply them to the omitted year. The MAE_x is consequently only calculated from this one year long validation period. In the performance test of the hindcast simulation the identical year was omitted from the model and from the observations. In contrast, the year which was omitted in the C20 simulation and in the observations is not necessarily the same. In the first case, the observations and the model are pairwise correlated and a mismatch of the years would introduce unnecessary noise. In the second case, there is no time-correlation and thus the years do not have to be omitted jointly. However, the error from the CV in the second case is expected to be higher than in the first case.

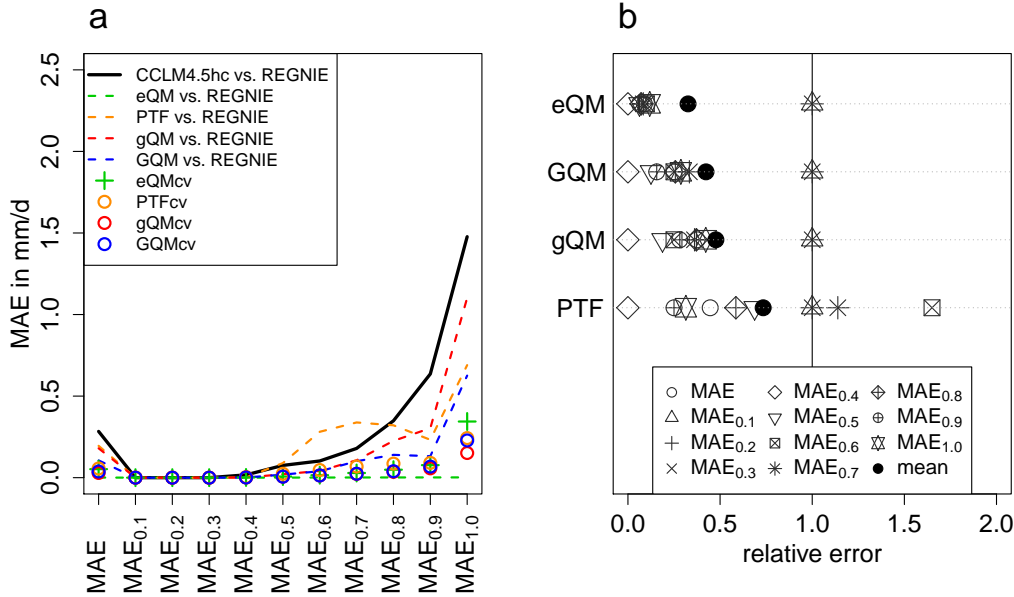


Figure 5.1 MAE_x scores from the temporal cross-validation of CCLM4.5hc versus REGNIE4.5 for the period 1991–2000. a) The MAE_x skill scores for 10 % wide percentile intervals (within an interval 10 substeps were used to calculate MAE_x). The means of these 10 MAE_x are plotted in a). In b) the relative skill scores are shown, i.e. the baseline (coloured dashed lines) plus the uncertainty from the CV (coloured points) divided by the uncorrected reference scores (solid black line). Note that a relative error of 1.0 means no changes.

The MAE_x were calculated at every grid box and have been spatially averaged afterwards. Fig. 5.1a shows the performance results for hindcast simulation. The skill scores of the uncorrected model are zero or close to zero for the intervals $MAE_{0.1-0.4}$ and increasing thereafter by approximately an e-function to a maximum of 1.5 mm day^{-1} at $MAE_{1.0}$. The mean error (MAE) is about 0.3 mm day^{-1} .

If the complete period, i.e. 10 years, is used for the calibration the performance is obviously the best possible for all methods but the empirical correction (eQM) outperforms all other methods and yields MAE_x values equalling zero over all intervals. This is no surprise however, because by definition the eQM should yield the best result given the large degrees of freedom, which is equal to the length of the time series. Thus if all data of the time series are used for calibration, a perfect fit is achieved. The first of the parametric correction methods, PTF, which fits an approximation to the empirical transfer function, shows a different behaviour for intervals larger than $MAE_{0.4}$. The error rises again from $MAE_{0.5}$ onwards and even overshoots the error from the uncorrected model. Not before $MAE_{0.9}$ does the error fall below the reference error but rises to approximately 0.9 mm day^{-1} at $MAE_{1.0}$. Therefore, in the mean (MAE) there is only a minor improvement.

A better performance for intervals $< MAE_{0.8}$ was found for gQM and GQM (Fig. 5.1). Until $MAE_{0.7}$, the corrections perform similar and the MAE_x are close to zero. At higher percentiles the performance bifurcates. The error of gQM rises and reaches approximately 1.0 mm day^{-1} at $MAE_{1.0}$, which is the worst performance of all methods at this interval. The MAE_x of GQM remains relatively low and rises only at $MAE_{1.0}$ to an error of 0.6 mm day^{-1} . The effect of the additional extreme value distribution for the tail of the distribution becomes visible here and clearly demonstrates the value of such a mixing approach despite the increase of free parameters.

The points in Fig. 5.1a represent the errors from the cross-validation, i.e. the additional error if new data is to be corrected. In general, this error is relatively low but larger at the extreme part because of

the rareness of extremes in the calibration period. Thus, transfer functions are in particular sensitive to extremes and to a change in frequency and intensity between the calibration and validation period. At $MAE_{1.0}$ the cross-validation error from the eQM is slightly larger compared to the parametric methods. This may indicate a slight overfitting in the empirical correction. In Fig. 5.1b the relative error is shown. Both errors from Fig. 5.1a were added and divided by the reference error. Thus, if this relative error is < 1 an improvement was achieved, if the relative error = 1 there is no improvement and > 1 means a worsening. In addition, the relative errors of all MAE_x and of MAE have been averaged, which was used for the ranking. The relative errors of eQM are altogether below 1, except for $MAE_{0.1-0.3}$ where there is no improvement since the error is already zero so that the relative error is set to 1. The relative error results in the following ranking: the best performance was achieved by the eQM, followed by GQM, gQM and the worst performance was achieved by the PTF.

The same analysis was carried out for the CCLM4.5c20 precipitation fields (Fig. 5.2). In general, the MAE_x is higher compared to the analysis in Fig. 5.2a, in particular between $MAE_{0.3}$ and $MAE_{0.9}$. The relative performance of the baselines remain the same, except that gQM fails to correct the extreme part and produces even a larger bias. PTF reduces the error now in all intervals but still shows the local maxima in the intervals $MAE_{0.4-0.9}$. The errors from the CV increase remarkably but this could be expected because more combinations were implied in the CV. Particularly, the errors at $MAE_{1.0}$ worsens to 2.0 mm day^{-1} for PTF, 2.0 mm day^{-1} for eQM and GQM and 2.2 mm day^{-1} for gQM. This confirms that the extreme part of a distribution is most sensitive to the choice of the BC method, whereas the mean values are more robust to the choice.

The relative errors Fig. 5.2b yield the same ranking as before, although there are now distinct differences in the extreme part. First, the relative error of $MAE_{1.0}$ is larger than before and second, only eQM and GQM show an improvement in the tail-correction (relative error < 1) whereas gQM and PTF degrade the errors in the tails.

In summary, the eQM and GQM method did perform best, in particular for extreme values. Although eQM showed the largest CV-error at $MAE_{1.0}$ in the CV of the C20 simulation, the error is of the same order as from the parametric methods. The worst performance at the tail was found for gQM. Thus there seems to be no problem with over-fitting in the eQM. For further applications in chapter 6 and 7, the methods eQM, GQM and gQM are applied. PTF is dropped because it showed the worst skill over all quantiles and gQM remains because the added value of GQM should be examined.

5.2 Correction of precipitation and temperature fields

After the final set of BC methods has been chosen, the corrections were applied to the precipitation and temperature fields of the C20 simulations (4.5 km and 1.3 km resolution) and the results were validated with observations.

5.2.1 Precipitation

The first comparison was made between the annual mean precipitation of REGNIE and the CCLMc20 simulations. In Fig. 5.3a the annual mean precipitation of the uncorrected CCLM4.5c20 simulation is shown and in Fig. 5.3b from REGNIE4.5. The spatial structure of the precipitation pattern is reproduced by the CCLM which is mainly characterized by higher annual means in orographic regions and lower values in flatter terrain such as river valleys. However, the absolute values are overestimated by the CCLM4.5c20 and yield a wet bias over the complete domain (Fig. 5.3c). The wet bias is about 21 % in the spatial average but individual deviations range from -7% to 59% (see Tab. 5.1 column 1-3). The regions with the strongest bias are the Rhine Valley (Mannheim-Mainz-Koblenz-Bonn) and the

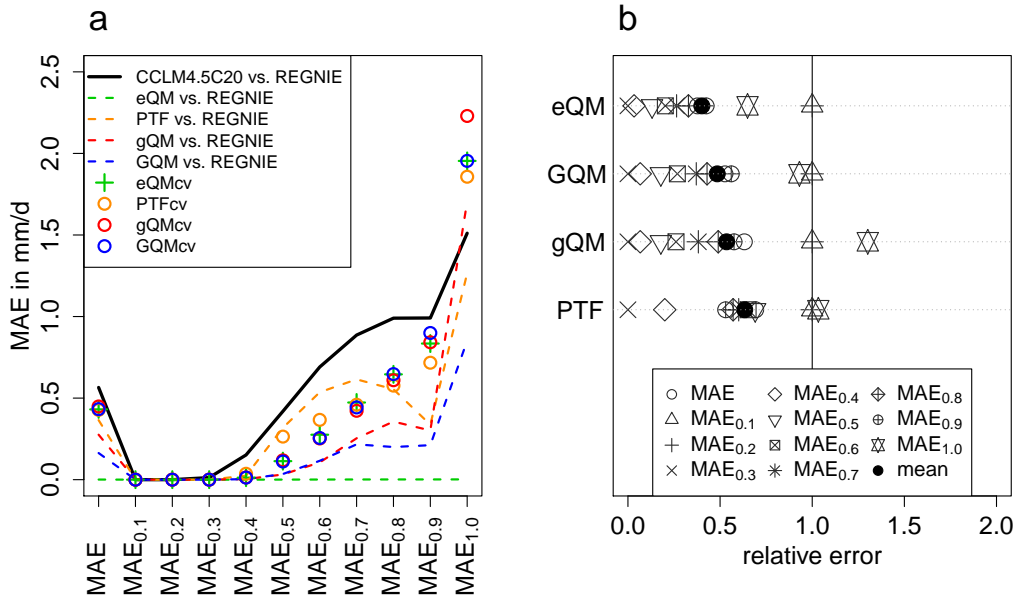


Figure 5.2 MAE_x scores from the temporal cross-validation of CCLM4.5c20 versus REGNIE4.5 for the period 1991–2000. a) The MAE_x skill scores for 10 % wide percentile intervals (within an interval 10 substeps were used to calculate MAE_x). The means of these 10 MAE_x are plotted in a). In b) the relative skill scores are shown, i.e. the baseline (coloured dashed lines) plus the uncertainty from the CV (coloured points) divided by the uncorrected reference scores (solid black line). Note that a relative error of 1.0 means no changes.

mountainous areas of the Westerwald (north of Koblenz) and the Hunsrück (east of Trier). Here the bias is between 20 – 50 %. Individual spots at Koblenz and Bonn reach up to 60 %. This wet-bias was mainly transferred from the forcing model (CLM18) to the CCLM4.5 (not shown, see Hollweg et al. (2008)).

After applying the BC-methods in the baseline configuration the wet-bias reduces considerably. The remaining bias after applying eQM is shown in Fig. 5.3d. The spatial average bias is now -0.08% and ranges between -0.28% and 0.04% . Thus the bias is negligibly small. In contrast, after applying gQM (Fig. 5.3e) a bias of the order of $+2\%$ ranging from -0.4% up to 5.8% remains. After applying GQM (Fig. 5.3f) the bias is considerably reduced and the remaining bias is about $+0.4\%$ in the domain average and ranges from -1.1% to 2.9% . However, at some individual grid boxes outliers occurred with a remaining bias of up to $\pm 10\%$. They concentrate in the south of the validation area. As mentioned by Gutjahr and Heinemann (2013) this area maybe sensitive to the assumption of a combination of a gamma and a GPD distribution. Compared to the original bias these deviations are small, though. This spatial analysis of the annual precipitation fields yield the same ranking as the cross-validation experiments in the section before.

In Fig. 5.4, the same plots are shown with regard to the correction of CCLM1.3c20. The annual precipitation pattern is well captured by the model compared to the observations (Fig. 5.4a-b). With a resolution of 1.3 km the CCLM simulates overall less precipitation than in the 4.5 km resolution. However, according to the analysis in chapter 4 this systematical reduction in precipitation is caused by too less precipitation in summer. The wet-bias decreases as a consequence (Fig. 5.4c) and is about $+10.45\%$ (see Tab. 5.1) in the spatial average. Again, the precipitation is mainly overestimated along river valleys. Underestimated are regions in the mountainous parts of the Palatinate Forest, Hunsrück and Eifel with up to -20% . Along the northern boundary of the domain a dry band of the same order is visible, which was also noticed in Fig. 4.2. This bias is due to boundary effects which slightly extends

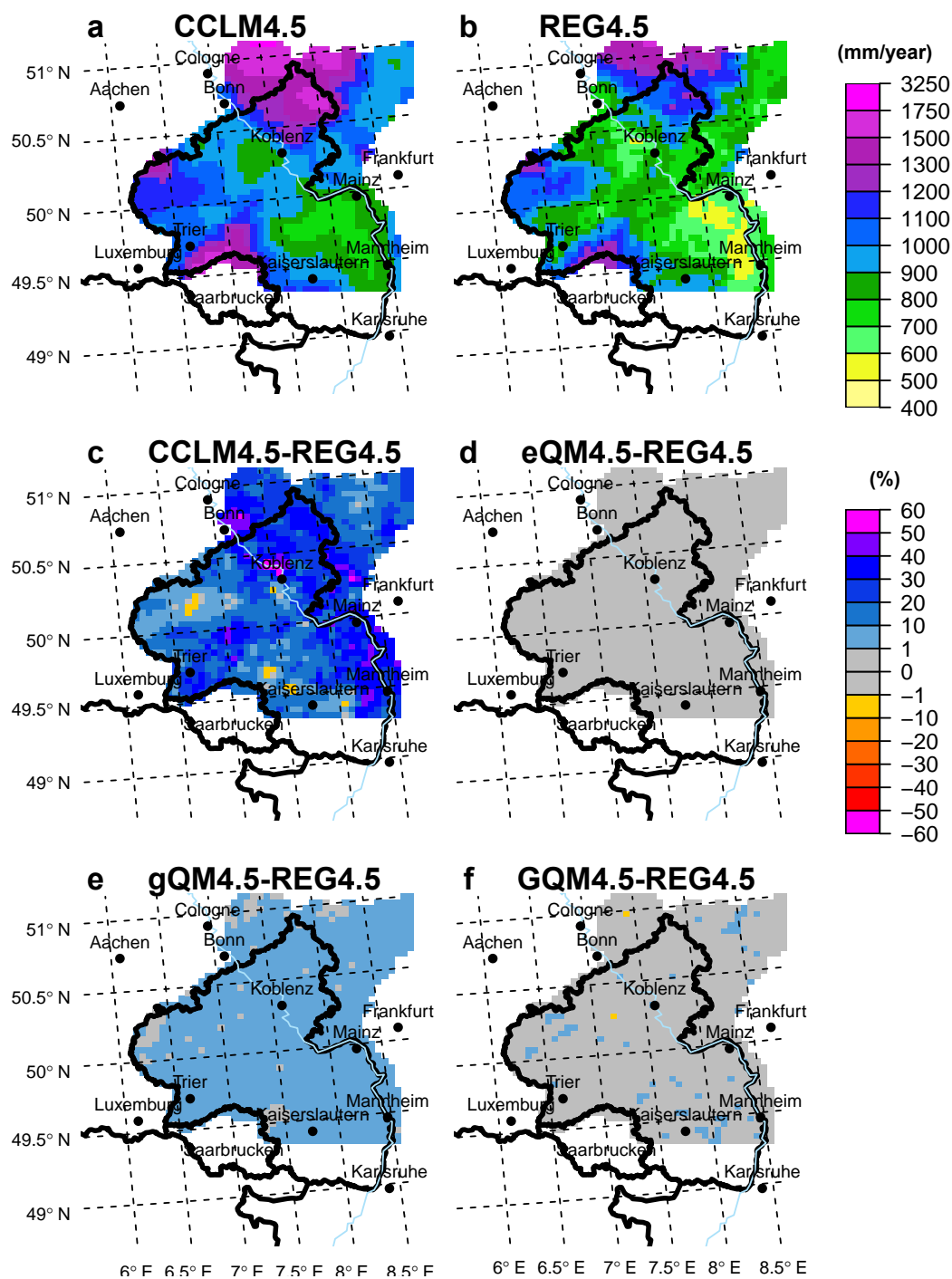


Figure 5.3 Precipitation fields of CCLM4.5c20 before and after bias-correction compared to REG4.5. Annual mean precipitation of a) CCLM4.5c20 and b) REG4.5 and the differences CCLM4.5 minus REG4.5 in c) uncorrected, d) after eQM, e) after gQM and f) after GQM.

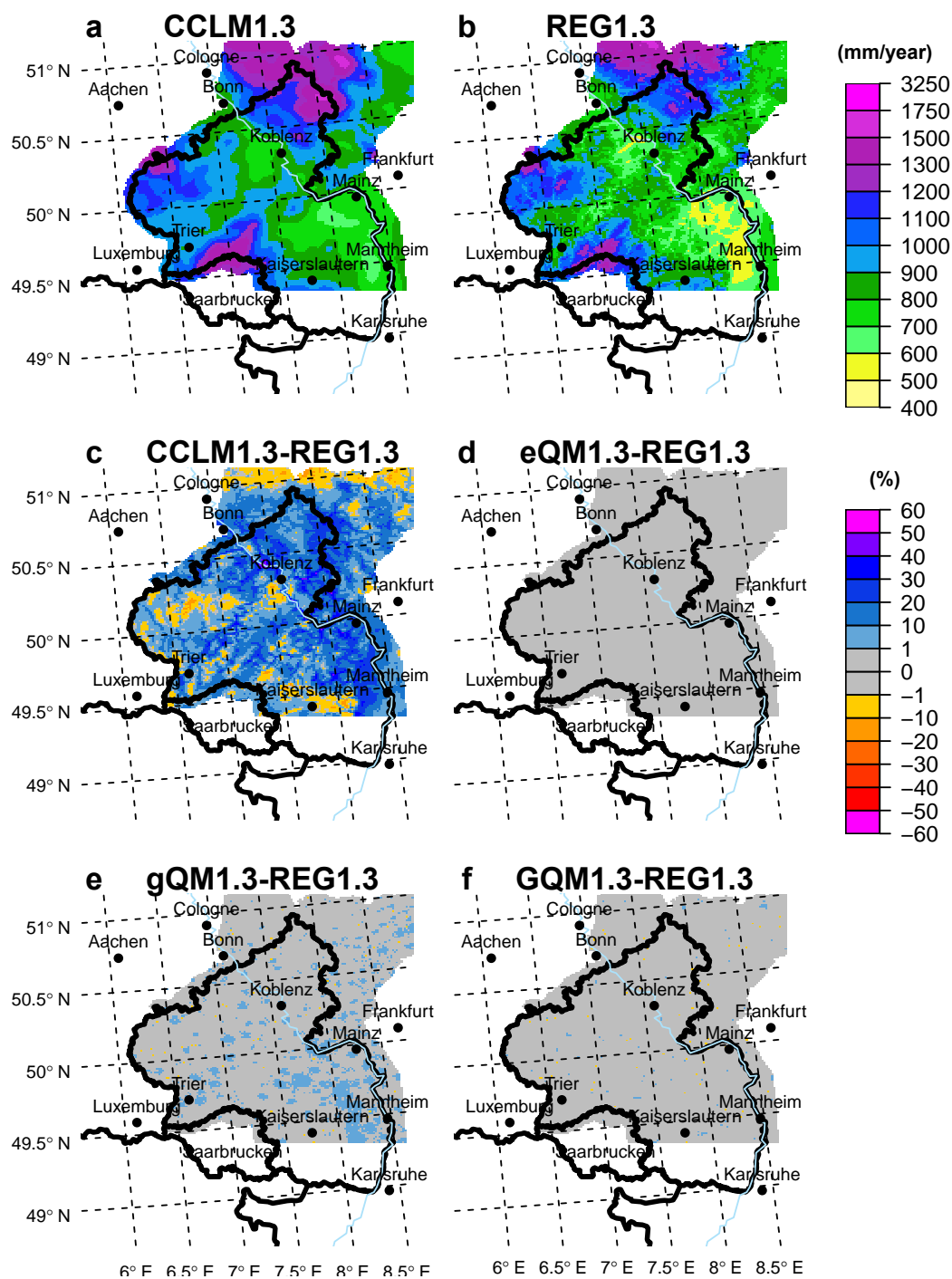


Figure 5.4 Precipitation fields of CCLM1.3c20 before and after bias-correction compared to REG1.3. Annual mean precipitation of a) CCLM1.3c20 and b) REG1.3 and the differences CCLM1.3 minus REG1.3 in c) uncorrected, d) after eQM, e) after gQM and f) after GQM.

Table 5.1 Distribution moments of daily precipitation from CCLM4.5c20 before and after bias-correction in comparison with REGNIE4.5.

	mean bias (%)	min. bias (%)	max. bias (%)	μ (mm day ⁻¹)	σ (mm day ⁻¹)	skewness (mm day ⁻¹)	kurtosis (mm day ⁻¹)
4.5 km							
REGNIE				2.45	4.80	3.83	26.42
CCLM	21.19	-6.62	58.85	2.94	5.05	3.79	31.86
eQM	-0.08	-0.28	0.04	2.44	4.80	3.83	26.47
gQM	1.98	-0.41	5.77	2.49	5.53	4.99	46.57
GQM	0.40	-1.07	2.88	2.45	4.92	3.82	28.79
1.3 km							
REGNIE				2.46	4.85	3.86	26.98
CCLM	10.45	-17.20	52.48	2.69	4.96	3.74	27.97
eQM	0.02	-0.16	0.03	2.46	4.85	3.86	26.97
gQM	0.60	-1.60	4.35	2.47	5.04	4.23	34.07
GQM	0.13	-1.46	3.09	2.46	4.88	3.87	28.04

into the analysis domain. The bias-correction with eQM effectively removes the bias (0.02 %)(Fig. 5.4d). Although the performance of gQM is improved compared to the correction of the 4.5 km, it is still not satisfying. The mean bias is 0.60 % in the domain-average, which is considerably higher compared to eQM. From (Fig. 5.4e) patchy spots with deviations $> 1\%$ are visible. They follow no identifiable pattern but seem to be more frequent in the south of RLP. Finally, GQM performs similarly well as in the correction of the 4.5 km model. The mean bias is 0.13 % and the range of deviation of individual grid boxes is slightly less than in the gQM correction, but generally of the same order (Tab. 5.1).

Next, the correction of the whole distribution of daily precipitation was analysed. Therefore the first four moments of the distributions were calculated as proposed by Lafon et al. (2013) and in addition, quantile-quantile (QQ) plots were used to analyse the complete distributions. In Tab. 5.1 the columns 4-7 corresponds to the mean, standard deviation, skewness and kurtosis of the daily precipitation distributions. In both resolutions, the uncorrected CCLM overestimate the mean daily precipitation: 2.94 mm day⁻¹ (2.69 mm day⁻¹) compared to 2.45 mm day⁻¹ (2.46 mm day⁻¹). After the correction, the mean is adequately adjusted by all methods. The second moment, the standard deviation (σ), is about 4.80 mm day⁻¹ in REGNIE whereas both CCLM simulate slightly to high σ , about 5 mm day⁻¹ in the area-mean. The eQM and GQM methods adjusted the σ accordingly but gQM even increased the σ to values > 5 mm day⁻¹. This new bias became also visible in the two remaining higher moments. The skewness and the kurtosis were both too high after the gQM correction in both resolutions. This indicates that the first two moments are corrected sufficiently by all methods but considerable differences occur in higher moments.

The complete distributions were calculated at every grid box and domain-averaged afterwards and compared in terms of quantile plots. In Fig. 5.5 the results are shown as QQ-plots. In the CCLM4.5c20 simulation (Fig. 5.5a), the uncorrected quantiles deviate mainly in the lower tail, between 0 – 10 mm day⁻¹, and in the upper tail > 40 mm day⁻¹. Although the discrepancies are largest in the upper tail, the lower intensities occur much more frequent and thus contribute larger to the wet-bias. The eQM correction adjusts the whole distribution perfectly and the quantiles are located along the $x = y$ line. Despite the successful correction of the low intensities by the gQM method, intensities > 10 mm day⁻¹ are strongly overestimated. This error even grows with increasing intensities. From this

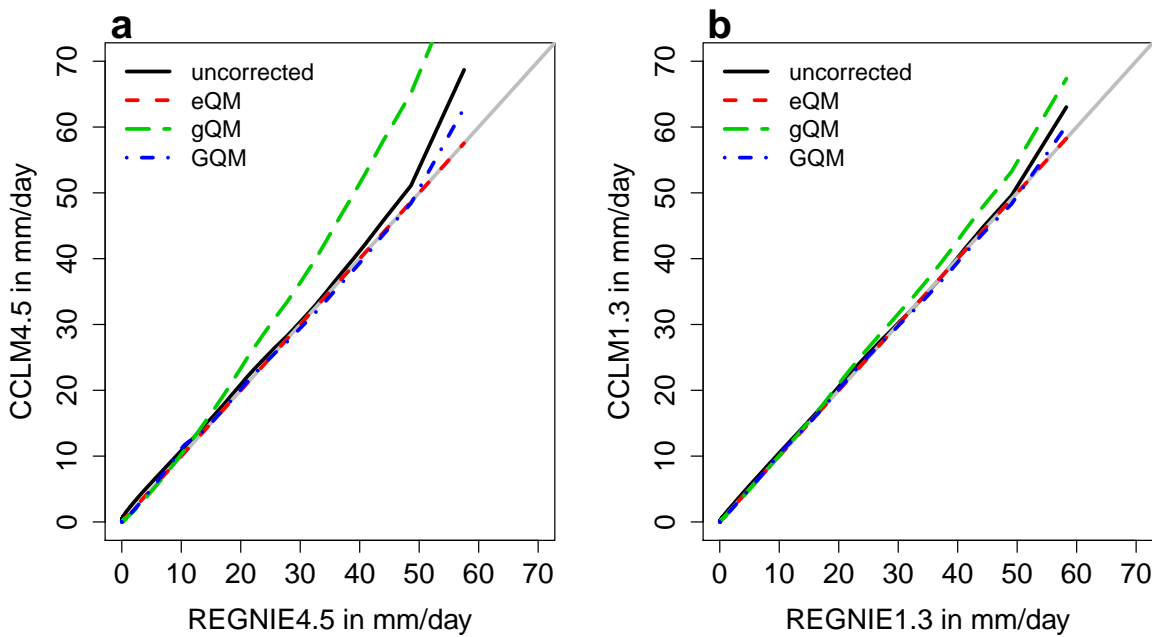


Figure 5.5 QQ-plots of domain-averaged precipitation from a) CCLM4.5c20 and b) CCLM1.3c20 before and after bias-correction compared to REGNIE4.5/REGNIE1.3. The solid grey line marks the $x = y$ line.

graph it is obvious why higher moments are not in agreement with the observations. Far better performs the GQM method. The quantiles after the GQM correction follow closely the $x = y$ line as well. The only noteworthy deviation occurs at the most extreme intensities which are slightly overestimated. A similar structure of the quantiles from CCLM1.3c20 can be seen in Fig. 5.5b, although the overestimation of the quantiles by gQM is less but still worse than in uncorrected model.

Finally, the seasonal performance was analysed with respect to CCLM4.5c20. The data sets were split into the meteorological seasons: winter (DJF), spring (MAM), summer (JJA) and autumn (SON). Then the seasonal quantiles were compared to REGNIE before and after the bias-correction. The results are shown in Fig. 5.6. The most striking deviations between the uncorrected precipitation and the observations occur in the extreme tail in winter (Fig. 5.6a) and summer (Fig. 5.6c). In summer, the extreme values are overestimated and in winter underestimated. In spring (Fig. 5.6b) and autumn (Fig. 5.6d) the deviations are less. After the correction with eQM, which was calibrated from the complete 10-year period, i.e. only one transfer function was applied to all seasons, the deviations reduce in all seasons, except in winter where there is nearly no correction. The performance of gQM (Fig. 5.6e-h) is similar to what was shown in Fig. 5.5. The mid to upper tail is largely overestimated in all seasons. Lastly, the precipitation quantiles after GQM are similar to those of after the eQM correction. Only the values in the upper tail in summer are slightly higher. Likewise, in autumn there is nearly no correction visible. As a consequence, the extremes of the winter season are still too low compared to the observations.

5.2.2 2 m temperature

The mean annual temperature simulated by the CCLM4.5c20 run (Fig. 5.7a) is about 8.08°C in the spatial mean and about 9.20°C in INTERMET4.5 (Fig. 5.7b), thus the CCLM is 1°C too cold in average.

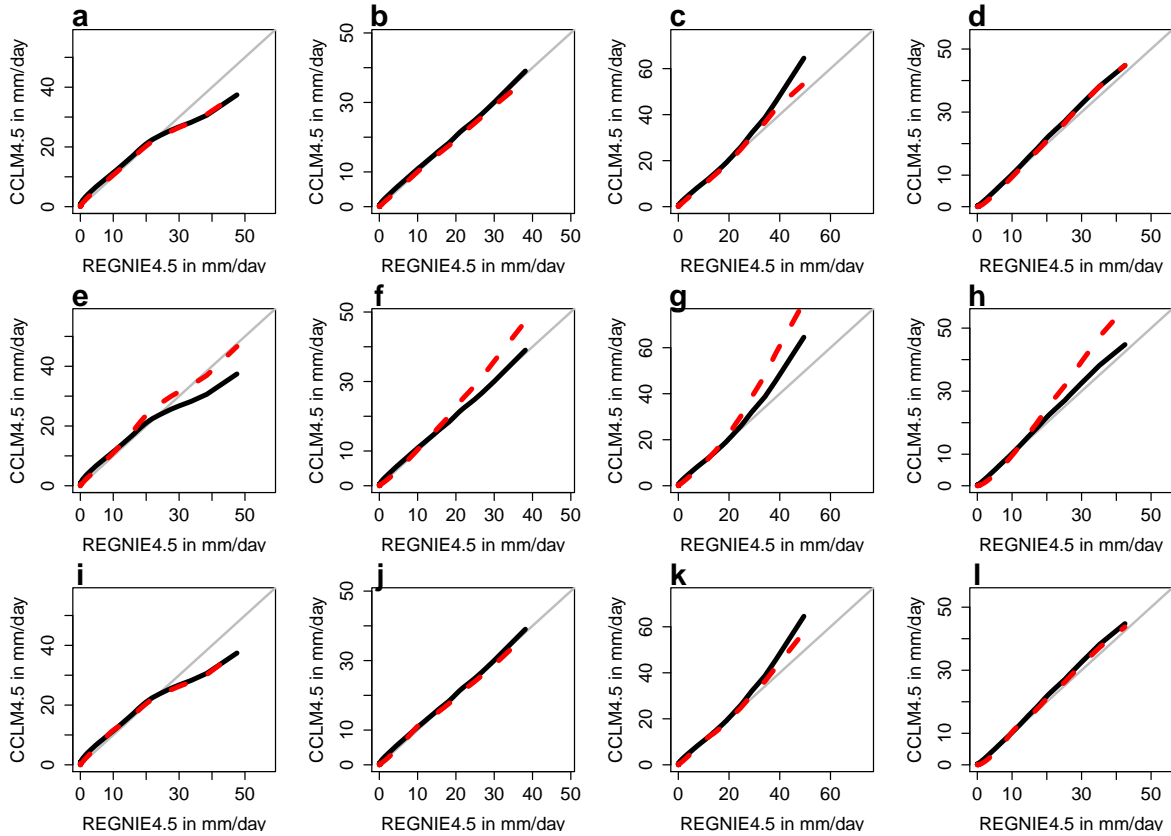


Figure 5.6 Seasonal QQ-plots of daily precipitation from CCLM4.5c20 before (black solid lines) and after the bias-corrections (red dashed lines) in comparison with REGNIE4.5. In a-d) the eQM correction for DJF, MAM, JJA and SON, in e-h) for gQM and in i-l) for GQM.

The warmest region in the model and in the observations is the Rhine Valley with annual mean values of 10.1°C and 11.1°C , respectively. The coldest minimum temperatures were in average simulated in the Westerwald with 5°C but in INTERMET they occurred in the Eifel with 6.8°C . Aside from this deviation the general pattern of the spatial temperature distribution is well reproduced by the CCLM and is clearly depending on the orography.

The differences CCLM minus INTERMET are shown in Fig. 5.7c. The annual mean temperature is simulated too cold over the complete domain as mentioned before but in particular in the Westerwald and in the Hunsrück. In these regions the cold-bias is even -2°C . However, it should be noted that especially in orographic regions the observational network is not as dense as in flatter areas. Therefore, at least parts of the cold-bias may be caused by missing observations in these areas. The applied interpolation algorithm may hence produce too warm temperatures depending on the weight which was given to the orography during the interpolation. However, the INTERMET data set is the only one with such a high resolution that covers large parts of the focus domain.

The correction of the temperature was done with the seasonal eQM method on daily temperatures and the differences are shown in Fig. 5.7d. The cold-bias of the annual mean is completely removed in the whole domain. The seasonal analysis of the quantiles is again analysed in terms of QQ-plots (Fig. 5.8). In addition to the seasonal correction, a eQM correction with only one transfer function for all seasons was applied for comparison. In winter (Fig. 5.8a) and autumn (Fig. 5.8d) the quantiles of the uncorrected model temperature fit already well with the observed quantiles. Only in autumn the

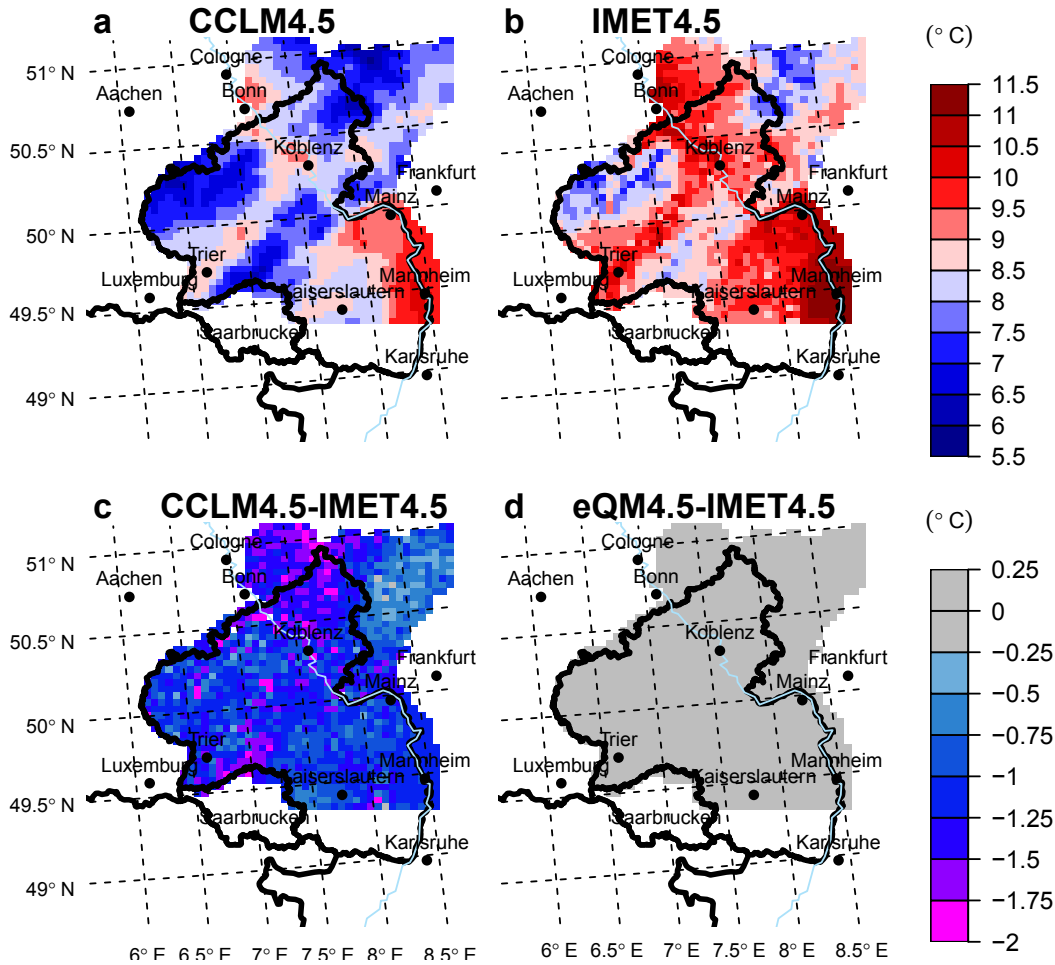


Figure 5.7 Temperature fields of CCLM4.5c20 before and after bias-correction compared to INTERMET4.5. Annual mean precipitation of a) CCLM4.5c20 and b) INTERMET4.5 and the differences c) CCLM4.5c20 minus INTERMET4.5 and d) eQM4.5 minus INTERMET4.5.

coldest temperatures are slightly overestimated, which is worsened by a non-seasonal correction. More pronounced deviations appear in spring and summer (Fig. 5.8b-c). The lower tail in spring is about -5°C too cold in the CCLM. The non-seasonal eQM is not able to correct this error, whereas a seasonal bias-correction removes this deviation. In summer, there are larger discrepancies in the core of the distribution and a small warm-bias in the hot extremes. The standard eQM introduces a warm-bias in the cold extremes, but is circumvented by the seasonal eQM. In the upper tail both methods yield a perfect correction. The choice whether to use seasonal or non-seasonal eQM does not have an impact on the further analysis, but the seasonal correction is used for the analysis in chapter 6 and chapter 7.

5.3 Summary and discussion

The uncorrected model output of the CCLM for the 20th century is wet and cold-biased compared to the observations. The average wet-bias of 21.19% in annual precipitation is mainly caused by the forcing data. The CLM18c20 simulated about $+200\text{ mm year}^{-1}$ too much precipitation in central Europe (Hollweg et al., 2008). This corresponds to a bias of about $+21.8 - 25.1\%$ depending on

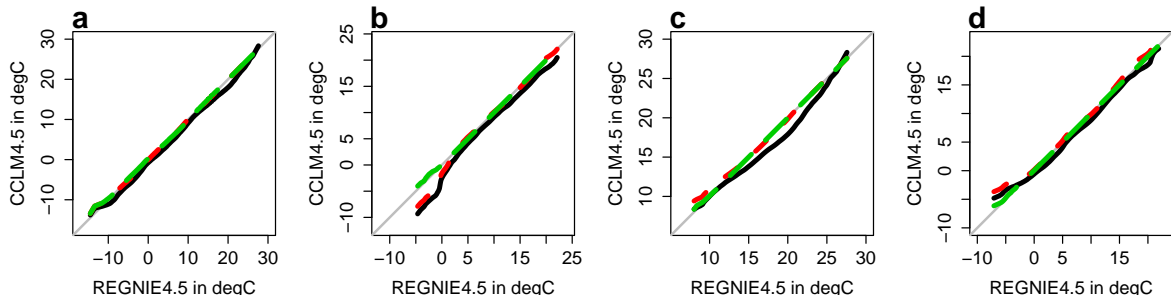


Figure 5.8 Seasonal QQ-plots of daily 2 m temperature from CCLM4.5c20 before (black solid lines) and after the bias-correction with eQM (red dashed lines) and seasonal eQM (green dashed lines) in comparison with INTERMET4.5 for a)-d) DJF, MAM, JJA and SON.

the observational data set. Thus, the biases of the here performed simulations are close to the bias of the forcing. The bias of the annual mean 2 m temperature is about -0.29 K in the CCLM4.5c20 and -0.16 K in the CCLM1.3c20. In the forcing model this error is between -0.08 K and 0.18 K over Germany. Thus the higher resolution run is slightly colder. Furthermore, the bias is larger in the seasonal means: about -1 K in winter and $+0.27$ in summer. This seasonal pattern is opposite to what was found in the forcing. There, the temperatures were too warm in winter but too cold in summer (Hollweg et al., 2008).

The analyses of the precipitation correction revealed the following performance ranking: the empirical method, eQM, did outperform all other methods. The best parametric method was the GQM method, which adjusted the upper tail considerably and reduced the bias. Compared to eQM, there remained a slight overestimation of the extremes but on the other hand, much less parameters have to be fitted in the GQM. The ranking of the methods is equal to the findings in Gutjahr and Heinemann (2013), who corrected the same precipitation fields by using the three nearest synoptical stations to a grid box and interpolated the corrected values by an inverse-distance weighting. If however, a gridded data set is used to calibrate the transfer function. Considerably less severe outliers occurred in the GQM correction. As discussed by Maraun (2013) it is not feasible to apply a quantile-matching for bias-correction (or downscaling) if the observations have a much higher resolution than the climate model. In such a configuration the quantile-matching does not only correct the bias but simultaneously attempts to gap the scale mismatch. Thus it is attempted to explain local variability by grid box variability. Interpolation schemes can reduce this scale mismatch if grid boxes, which contain a gauging station, are not exactly the support values.

A problem of overestimation did not arise, except that the cross-validation error from eQM in the correction of the hindcast runs was slightly higher than from the parametric methods. But this was only the case in the correction of the hindcast runs and not in the correction of the C20 run. The empirical quantile-matching is found to be the optimal bias-correction method. This finding is similar to the rankings of Themeßl et al. (2012), Gudmundsson et al. (2012) and Gutjahr and Heinemann (2013). Lafon et al. (2013) found that eQM outperforms for instance gQM as well, but the sensitivity to the choice of the time period, which is used for the calibration, was higher than in the parametric methods. They concluded that gQM is the best compromise but with the restriction that the gQM correction is only valid if the precipitation follows a gamma distribution. In this study, however, there was no indication that eQM was more sensitive than the parametric methods. But the analysis of sensitivity was limited to the first moment. Higher sensitivity may arise for higher moments as reported by Lafon et al. (2013). The correction with gQM showed a good performance for lower intensities and in adjusting the daily mean but strongly overestimated moderate to severe extremes. The reason for this overestimation is

that the gamma distribution declines too rapidly in the tail and hence too low probabilities are assigned to extremes. If such a low probability is now transferred to a quantile in the observed PDF, then small deviations in probability lead to large differences in the corresponding quantile. The GQM relaxes this error because the GPD is a heavy-tail distribution, i.e. the tail behaviour does not decline as fast as for a gamma distribution. The extension for adjusting extremes in the GQM method clearly circumvents this wrong "correction" of the upper tail. Although three additional parameters have to be fitted within the GQM, the increase in performance makes it feasible. However, the treatment of outliers in the gQM and GQM method is handled with an empirical approach by omitting all values with a probability of $F_{\text{CCLM,GPD}}(x) > 1 - 10^8$ from the correction. This problem may reduce if gQM and GQM are applied to sufficiently longer time series but remains an open topic for further research.

These results show that it is not sufficient to solely rank bias-correction methods with respect to the mean. Higher moments and the behaviour at the tails should be included into the ranking because extremes are of essential importance in impact studies, for instance in hydrological applications.

The correction of the 2 m temperature was achieved with a seasonal empirical correction method (eQM). It was shown that this method removes any bias of the annual mean if applied to daily temperature values. A seasonal analysis revealed that the largest bias occur in the lower tail in spring and in the centre part of the distribution in summer. In winter and autumn the deviations are only minor. In comparison with the standard eQM it was found that especially in the lower tails of the distributions the bias is not removed but rather increased. The seasonal eQM is chosen for further analysis because of its overall better performance, although the bias in the summer hot and winter cold extremes, which are in the focus in chapter 7, are removed by both methods.

Climate change signal analysis

6.1 Climate change signals from uncorrected model data

The climate change signals (CCS) are calculated on an annual and a seasonal basis for the variables precipitation and 2 m temperature. A CCS is classified as significant if the p-value of an applied statistical test is lower than or equal to the significance level, i.e.: $p \leq \alpha$. In this study the significance level was set to: $\alpha = 5\%$. The tests were performed locally at every grid box and have been further restricted by the False-Discovery-Rate (FDR) field significance after Benjamini and Hochberg (1995). The FDR is usually used as a global statistical test to determine whether a complete field is significant. Thereby, it controls the false discovery rate, which was set to $\alpha_g = 10\%$. The false discovery rate is the rate at which the null hypothesis is rejected per chance. Here the FDR is only used to control the type-I error. The results are shown with the outcome of the t-test only, but the Wilcoxon- and KS-test yielded similar results.

Firstly, the signals of the uncorrected CCLM output are presented for the complete model domain of Rhineland-Palatinate and the Saar-Lor-Lux region. Afterwards, the impact of the bias-correction on the CCS are demonstrated in the BC-target region (see chapter 5).

6.1.1 Annual signals

The annual mean precipitation of the regional climate models CCLM4.5c20 and CCLM1.3c20 are shown in Fig. 6.1a-b. More than 1200 mm is simulated in regions with high orography and about 600 – 900 mm in flatter areas. There are distinct differences visible between the 4.5 km and 1.3 km resolution. The higher resolution resolves orography features in more detail, which affects the precipitation fields as well. Furthermore, the annual precipitation is less in the CCLM1.3c20 for most of the domain (see chapter 4). This is mainly caused by the summer dry-bias. The CCS vary about $\pm 10\%$ in the future periods 2041–2050 and 2091–2100 but none of the signals are significant at the 95% level ($\alpha_g = 10\%$). However, if the FDR procedure is not applied then small areas show significant signals.

The annual mean temperatures are shown in Fig. 6.2. In the simulations, the warmest region is the Rhine Valley with up to 11 °C and the coldest regions are the Black Forest, Eifel, Westerwald and Hunsrück with less than 6 °C. The climate change signals of the the near (2041–2050) and far (2091–2100) future are homogeneously in the complete domain with values between +1.5 – 2 °C and +3.5 – 4 °C. The warming signals were rated as significant according to the statistical tests.

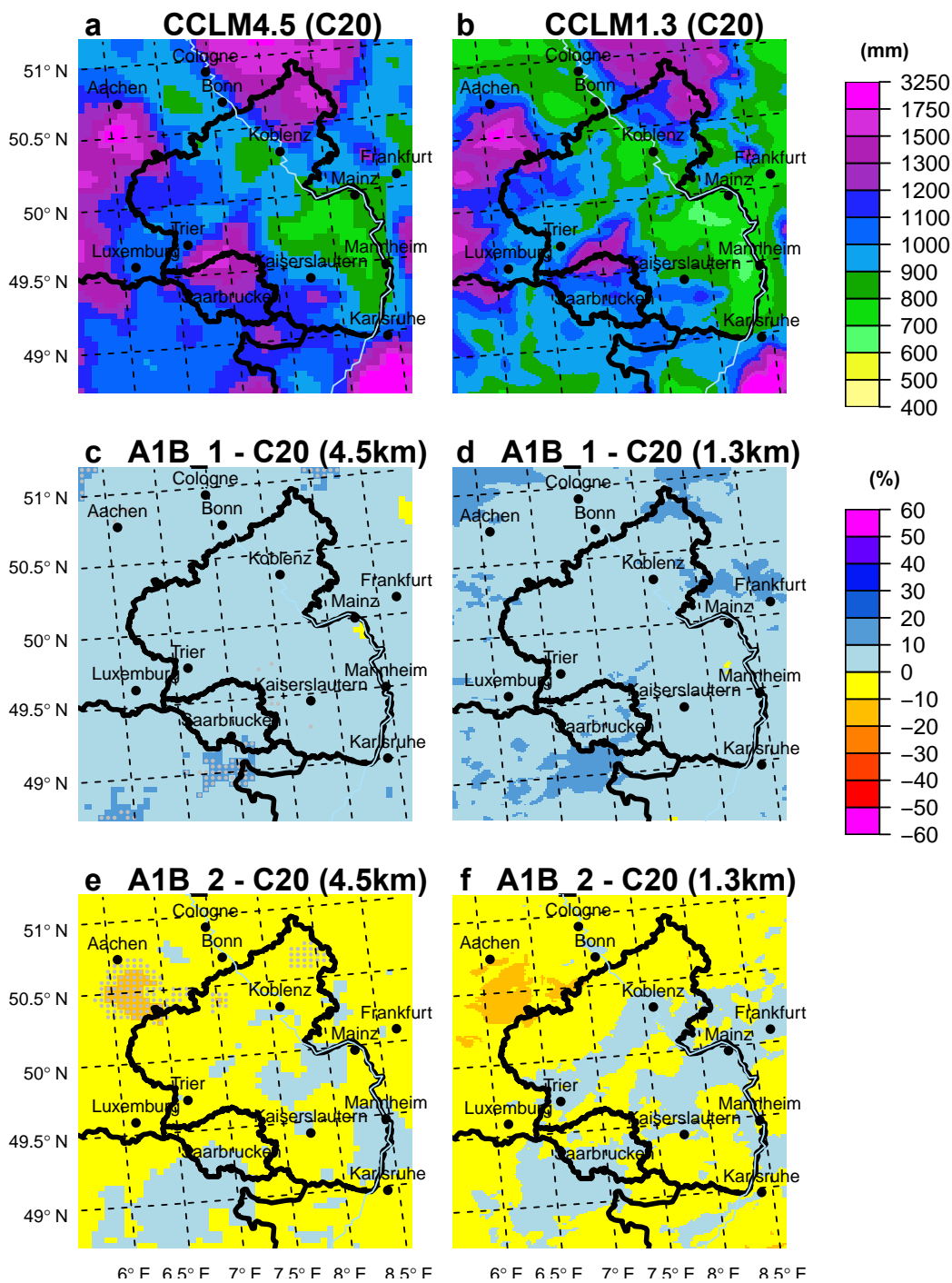


Figure 6.1 Annual mean precipitation of the period 1991–2000 (C20): a) CCLM4.5c20, b) CCLM1.3c20, and the changes until 2041–2050 (A1B_1): c) CCLM4.5a1b1 - CCLM4.5c20, d) CCLM1.3a1b1 - CCLM1.3c20 and until 2091–2100 (A1B_2) e) CCLM4.5a1b2 - CCLM4.5c20 and f) CCLM1.3a1b2 - CCLM1.3c20. Grey dots denote significant grid boxes of two-sided t-tests ($\alpha = 5\%$) if no FDR is applied.

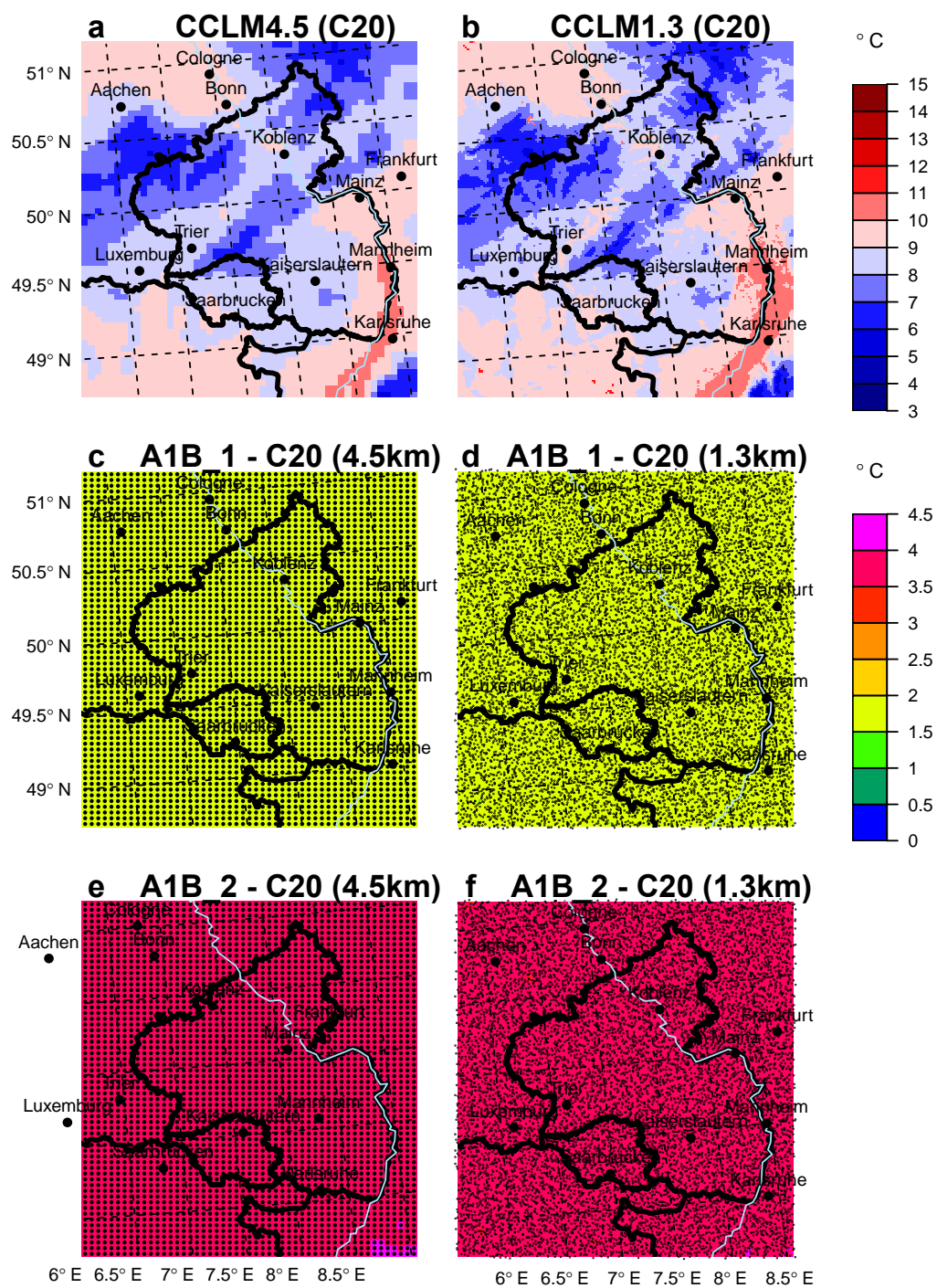


Figure 6.2 As in Fig. 6.1 but for the annual mean temperature in 2 m height. Black dots denote significant grid boxes of two-sided t-tests ($\alpha = 5\%$, $\alpha_g = 10\%$). Note that only a subset of the significant grid boxes is shown for the CCLM1.3 (b,d,f) to maintain the visibility.

Table 6.1 Seasonal climate change signal of domain-averaged precipitation (mm season^{-1}) and the daily standard deviation (mm day^{-1}) for the 21st century (C20) and the differences to future projections.

model	MAM		JJA		SON		DJF		Annual	
	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
1991–2000										
CCLM4.5	247	4.5	263	5.6	260	5.3	351	5.5	1121	7.0
CCLM1.3	221	3.3	198	5.0	243	5.3	339	5.5	1001	7.1
$\Delta 2050$										
CCLM4.5	6 (+2 %)	0.1	-25 (-10 %)	-0.2	81 (+31 %)	0.8	7 (+2 %)	0.4	70 (+6 %)	0.3
CCLM1.3	9 (+4 %)	0.1	-20 (-10 %)	-0.1	78 (+32 %)	0.8	8 (+2 %)	0.4	74 (+7 %)	0.4
$\Delta 2100$										
CCLM4.5	31 (+13 %)	0.6	-67 (-25 %)	0.4	2 (+1 %)	0.5	-2 (-1 %)	0.4	-35 (-3 %)	0.5
CCLM1.3	26 (+12 %)	0.7	-55 (-28 %)	0.2	2 (+1 %)	0.4	-1 (± 0 %)	0.4	-28 (-3 %)	0.4

6.1.2 Seasonal signals

The seasonal climate changes signals were calculated by first aggregating the daily precipitation (temperature) fields to seasonal sums (means). The seasons were defined as: December-January-February (DJF), March-May-April (MAM), June-July-August (JJA) and September-October-November (SON). This resulted in 10 sums (means) for every season for every climate scenario. Statistical tests were then applied to the seasonal C20 and A1B_1/A1B_2-periods as described in section 3.1.1.

Precipitation

The winter (DJF) precipitation of the C20-period and the CCS with respect to the near future (A1B_1) and the far future (A1B_2) are shown in Fig. 6.3. The majority of the winter precipitation is simulated by both models (Fig. 6.3a-b) and differences due to the increased resolution of the orography. For instance, single mountain peaks as the *Hohes Venn* in the Eifel (50.5°N , 6.2°E) become visible in the CCLM1.3c20 simulation or the lowermost area in the Rhine Valley, which receives less rain than in the CCLM4.5c20.

Until 2041–2050, both models simulate a slight decrease in winter precipitation of up to -20% (Fig. 6.3c-d). Stronger reduction occurs in the CCLM4.5 only at the eastern fringe of the domain and in the north-west of Frankfurt whereas the CCLM1.3 simulates stronger reduction in Luxemburg, France and in the western part of RLP. Positive signals of about $+0 - 10\%$ are simulated for parts of central and southern RLP, Saarland, France, north of Cologne and in the Mainz Basin between the cities of Mainz and Frankfurt. However, none of the signals are significant at the significance level $\alpha = 5\%$ ($\alpha_g = 10\%$). Until 2091–2100, both models simulate signals of $\pm 10\%$. Thus the magnitude of the winter CCS is much less than the interannual variability. This becomes evident from the power analysis of the t-tests (Tab. 6.2 and Tab. 6.3). The CCS of the area-averaged winter precipitation until 2041–2050 of the CCLM4.5 is 7.57 mm , whereas the common variance (which is a measure of interannual variability of C20 and A1B_1/A1B_2) is one magnitude larger 105.29 mm . Hence the power of the t-test is very low and just the prescribed significance level of 5% . That is, even if the signals are true, the t-tests will not rate them as significant. Or in other words, if signals are found to be significant, it is likely that they are significant only by chance. The power analysis of the CCLM1.3 signals yield similar values.

In summer there is much less precipitation simulated than in winter (Fig. 6.4a-b). The CCLM4.5c20 produces average summer precipitation values $< 450\text{ mm}$, whereas the CCLM1.3 is drier and simulates values $< 350\text{ mm}$. This is due to the dry-bias of the CCLM1.3, which was discussed in chapter 4. Until

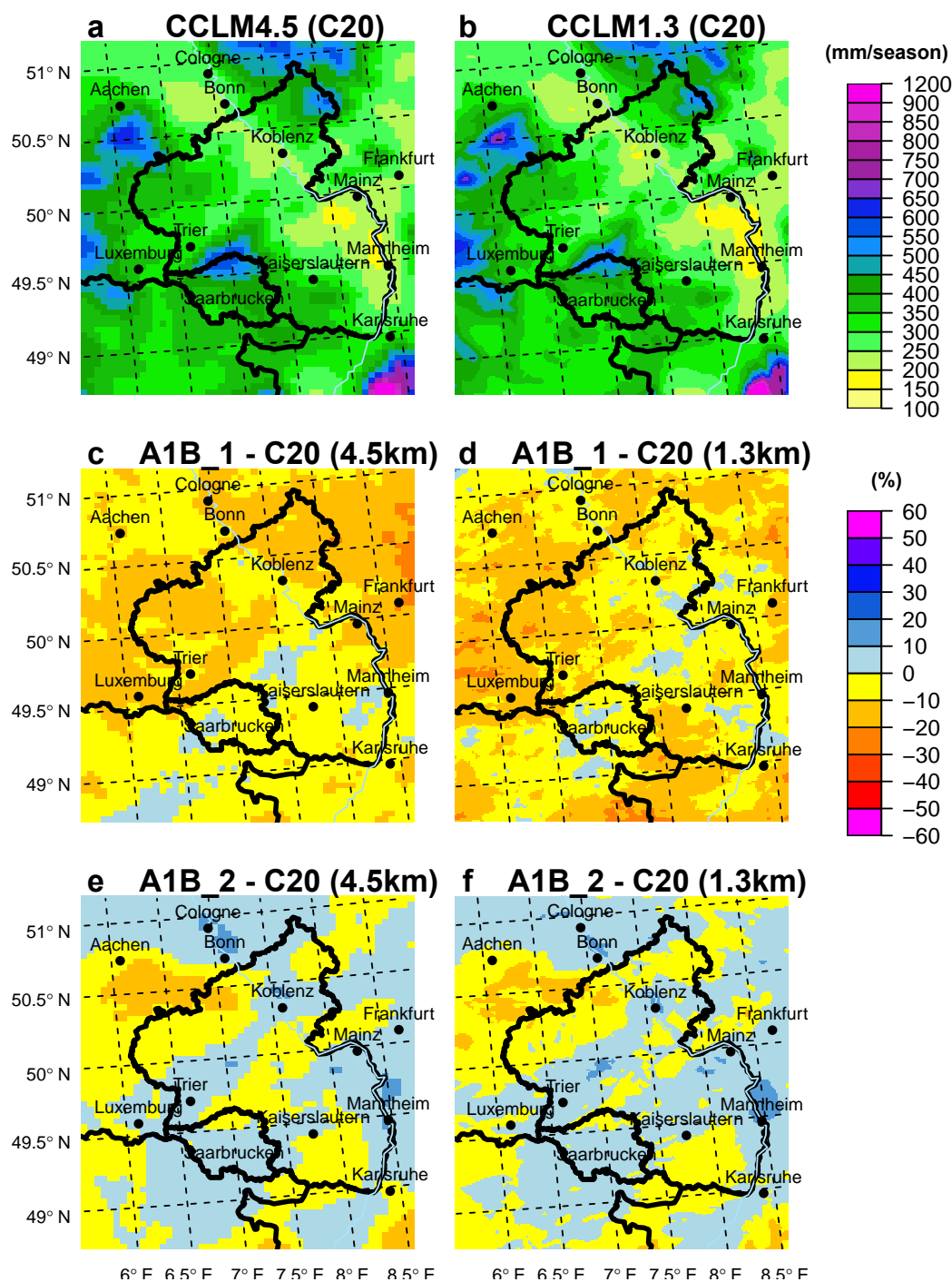


Figure 6.3 Climate change signals of winter (DJF) precipitation: a) DJF mean precipitation from CCLM4.5c20, b) as in a) but from CCLM1.3c20, c) CCLM4.5a1b1 - CCLM4.5c20, d) CCLM1.3a1b1 - CCLM1.3c20, e) CCLM4.5a1b2 - CCLM4.5c20 and f) CCLM1.3a1b2 - CCLM1.3c20. The global significance level of the FDR was set to $\alpha_g = 10\%$.

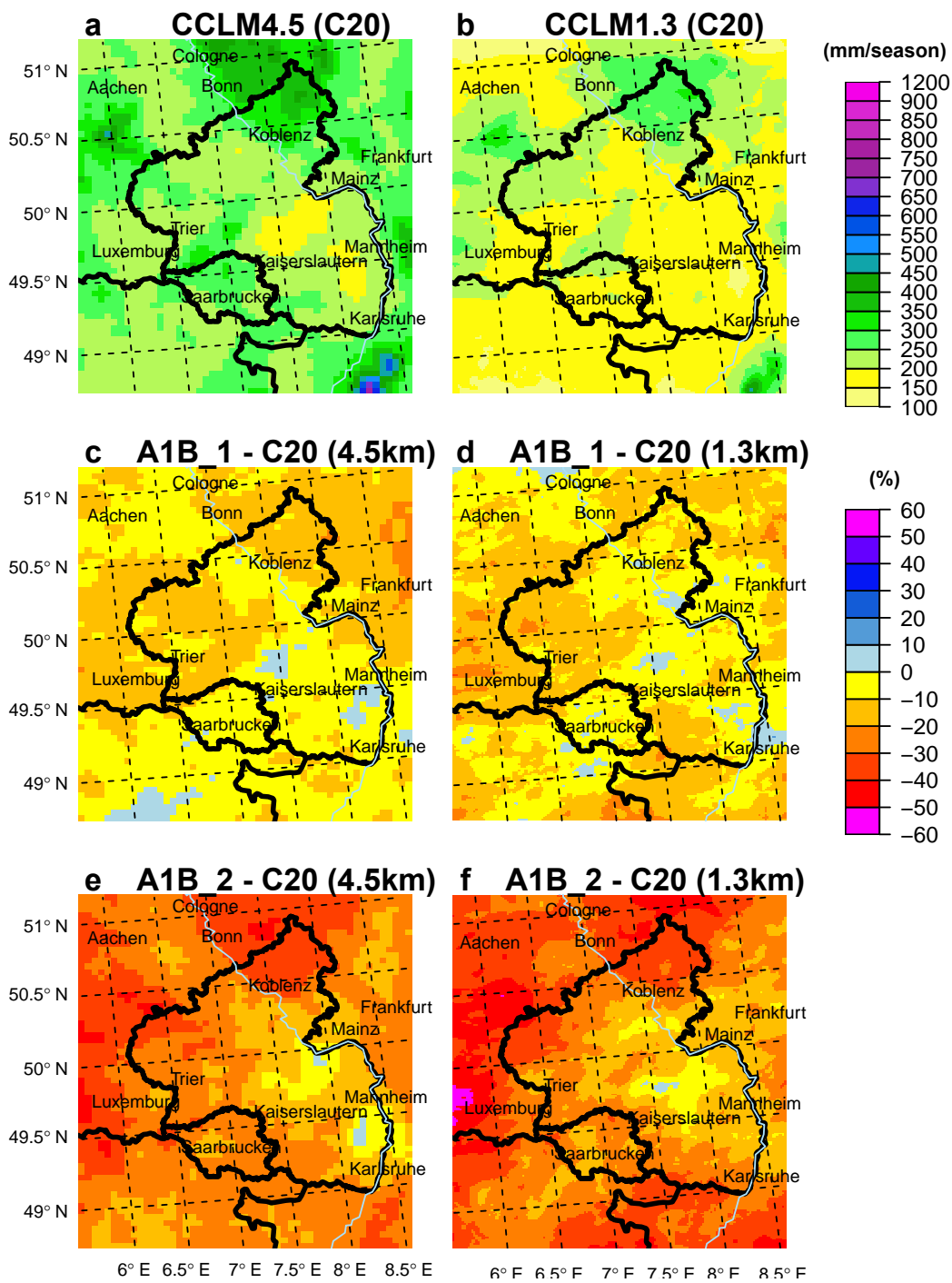


Figure 6.4 Climate change signals summer (JJA) precipitation: a) JJA mean precipitation from CCLM4.5c20, b) as in a) but from CCLM1.3c20, c) CCLM4.5a1b1 - CCLM4.5c20, d) CCLM1.3a1b1 - CCLM1.3c20, e) CCLM4.5a1b2 - CCLM4.5c20 and f) CCLM1.3a1b2 - CCLM1.3c20. The global significance level of the FDR was set to $\alpha_g = 10\%$.

Table 6.2 Power analysis of the t-tests of the seasonal CCLM4.5 precipitation climate change signals.

	power	n(p=0.8) ¹	σ_{xy} ²	D
DJF				
2050	0.05	3040.71	105.29	7.57
2100	0.05	39744.73	79.90	1.59
MAM				
2050	0.06	1019.27	50.81	6.31
2100	0.19	63.03	62.01	31.19
JJA				
2050	0.12	118.96	69.74	25.44
2100	0.57	16.31	66.04	66.90
SON				
2050	0.76	10.97	64.45	81.06
2100	0.05	22880.15	77.78	2.04

¹ Number of years needed to achieve a power of 80 %.

² σ_{xy} denotes the common variance.

Table 6.3 Power analysis of the t-tests of the seasonal CCLM1.3 precipitation climate change signals.

	power	n(p=0.8) ¹	σ_{xy} ²	D
DJF				
2050	0.05	2591.01	104.08	8.10
2100	0.05	44051.87	79.19	1.49
MAM				
2050	0.07	499.10	47.88	8.50
2100	0.16	75.77	56.65	25.95
JJA				
2050	0.12	126.36	57.93	20.50
2100	0.52	18.34	57.77	54.98
SON				
2050	0.77	10.77	61.54	78.20
2100	0.05	18412.41	75.08	2.19

2041–2050, both models simulate a slight decrease of summer precipitation of about -0 – 2%. Although the signal-to-noise ratio is better than in winter (Tab. 6.1, Tab. 6.2), none of the signals is significant. Until 2091–2100, a stronger reduction in summer precipitation is simulated by both models. In particular in the north and north-western of the domain (RLP). The reductions even reach -50% at some spots. Thus the CCS are not affected by the dry-bias in the CCLM1.3. The largest power was found for Δ 2100 in JJA (52 – 57%) and Δ 2050 in SON (76 – 77%). In these two cases the signal is of the order of the climate noise or larger. From the power analyses it can be deduced that a simulation of approximately 20 years would considerably improve the power of detecting summer until the end of the century. The power would rise to 80%, but on the other hand even 30 years would not improve the power for the other seasons. Note, however, that these conclusions are only valid if the climate noise or interannual variability remains constant.

The strongest signal was found for autumn (SON) until 2041-2050. CCLM4.5 as well as CCLM1.3 simulate an increase of about +80 mm (+30%). This increase of precipitation was also found by Hollweg et al. (2008) for the CLM18 simulations, which were used as driving data for CCLM4.5 and CCLM1.3. In absolute terms, Hollweg et al. (2008) found an increase of about 42 – 82 mm in Germany by comparing the periods 2031–2050 versus 1981–2000.

Besides the change in means, the change of the standard deviation (σ) of daily precipitation and daily 2 m temperature was analysed. In Tab. 6.1 the σ of the C20 periods and the change signal are shown. The σ of daily precipitation in C20 is approximately 5 – 5.6 mm day⁻¹ in all seasons except in spring (MAM, 3.3 – 4.5 mm day⁻¹). In spring and summer, the CCLM1.3 simulates lower σ compared to CCLM4.5. Until 2041-2050 there is an remarkable increase in autumn (SON) with +0.8 mm day⁻¹ and a lesser increase in winter (DJF) of +0.4 mm day⁻¹. The strong positive signal in SON indicates that not only the mean increases but also the day to day variability, whereas the mean remains nearly constant in winter but with larger variations on a daily basis. For instance, it can be suspected that precipitation is distributed on fewer days resulting in days with more extreme precipitation on the one hand but more

Table 6.4 Seasonal climate change signal of domain-averaged 2 m temperature ($^{\circ}\text{C}$) and the daily standard deviation (in brackets) for the 21st century (C20) and the differences to future projections.

model	MAM		JJA		SON		DJF		Annual	
	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
1991–2000										
CCLM4.5	8.2	5.1	16.1	3.5	8.5	5.0	0.95	4.1	8.44	7.0
CCLM1.3	8.2	5.1	16.3	3.5	8.5	5.0	0.95	4.2	8.51	7.1
$\Delta 2050$										
CCLM4.5	+1.58	-0.5	+1.45	+0.3	+1.74	-0.2	+1.92	-0.5	+1.68	-0.3
CCLM1.3	+1.55	-0.5	+1.37	+0.2	+1.77	-0.1	+1.93	-0.6	+1.64	-0.4
$\Delta 2100$										
CCLM4.5	+2.24	-0.5	+4.39	+0.9	+4.18	+0.7	+4.11	-0.7	+3.73	+0.2
CCLM1.3	+2.22	-0.6	+4.13	+0.7	+4.13	+0.7	+4.16	-0.8	+3.66	+0.1

dry days on the other hand. A slight negative signal of -0.1 mm day^{-1} to -0.2 mm day^{-1} is only simulated for summer.

Until 2091-2100, this pattern shifts to a positive signal in all seasons. The signal is about $+0.4 - 0.7 \text{ mm day}^{-1}$ except in JJA where the CCLM1.3 simulates only $+0.2 \text{ mm day}^{-1}$. Although the summer total precipitation decreases, the probability of extremes may rise due to an increase in variance.

2 m temperature

The interannual mean of the 2 m temperature in winter (DJF) along with the simulated climate change signals are shown in Fig. 6.4. The spatial averages and the CCS of the daily standard deviation are given in Tab. 6.4. The spatial distribution of the temperature is characterized by warmer temperatures in the river valleys and flat areas and colder mean temperatures in mountainous areas. The maximum mean temperatures in summer occur in the Rhine Valley, in the Cologne Basin (north of Aachen) and in France (Fig. 6.4a-b). The coldest temperatures occur in the Black Forest and the Westerwald. The CCS until 2041-2050 is nearly homogeneously with $+1.7^{\circ}\text{C}$ in the complete domain except in the northeastern part and in the Black Forest, where the signal is slightly higher with $+2.5^{\circ}\text{C}$. Until 2091-2100 the CCS increase to $+3.7^{\circ}\text{C}$ in the spatial mean, with signals up to $+5^{\circ}\text{C}$ in the Black Forest and in the Westerwald and slightly smaller CCS over France and Luxemburg. The stippled grid boxes indicate significance at the 95 % significance level (and FDR 90 % level). The CCS are significant in the complete domain and the associated power of the t-tests (see Tab. 6.5 and Tab. 6.6) in winter is 93 % and 100 % for CCLM4.5 and CCLM1.3, respectively. This means that the CCS are detected with a certainty of nearly 100 %.

These high power of the t-tests can be explained by comparing the climate signal to the climate noise. The common seasonal variance in winter until 2041-2050 (2091-2100) is about 1.17°C (1.17°C) and the signal is about 1.92°C (4.11°C). Thus the signal is approximately two times (four times) the climate noise and is hence detected with high power. Although the sample size of 10 years is already small, a power of 80 % is even achieved with 7 and 3 years, respectively.

The spatial distribution of summer temperatures is shown in Fig. 6.6. In the C20 simulations, the warmest summer temperatures are about $16 - 20^{\circ}\text{C}$ in the Rhine Valley, Moselle Valley, Cologne Basin and France. The maximum average is between Mannheim and Karlsruhe (Fig. 6.6a-b) with nearly 20°C . Cooler temperatures are associated with higher orography. The CCS until 2041-2050 is about $1.4 - 1.5^{\circ}\text{C}$ in the spatial average with maxima of up to 2.5°C in France, Luxemburg and along the Rhine Valley (Fig. 6.6c). In the CCLM1.3 (Fig. 6.6d) the CCS is slightly reduced and less pronounced

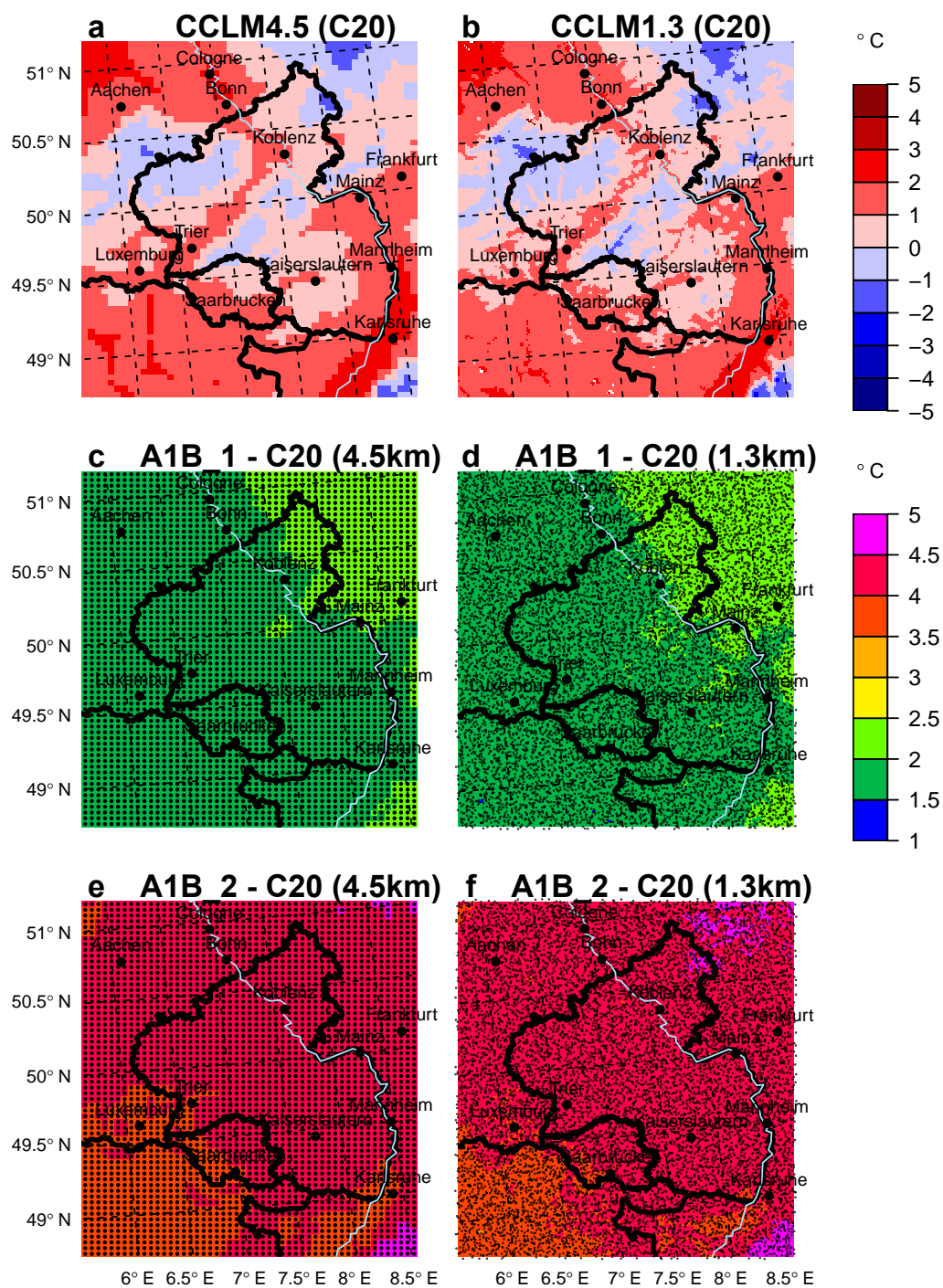


Figure 6.5 Climate change signals of winter (DJF) seasonal mean 2 m temperature: a) DJF mean 2 m temperature from CCLM4.5c20, b) as in a) but from CCLM1.3c20, c) CCLM4.5a1b1 - CCLM4.5c20, d) CCLM1.3a1b1 - CCLM1.3c20, e) CCLM4.5a1b2 - CCLM4.5c20 and f) CCLM1.3a1b2 - CCLM1.3c20. The global significance level of the FDR was set to $\alpha_g = 10\%$.

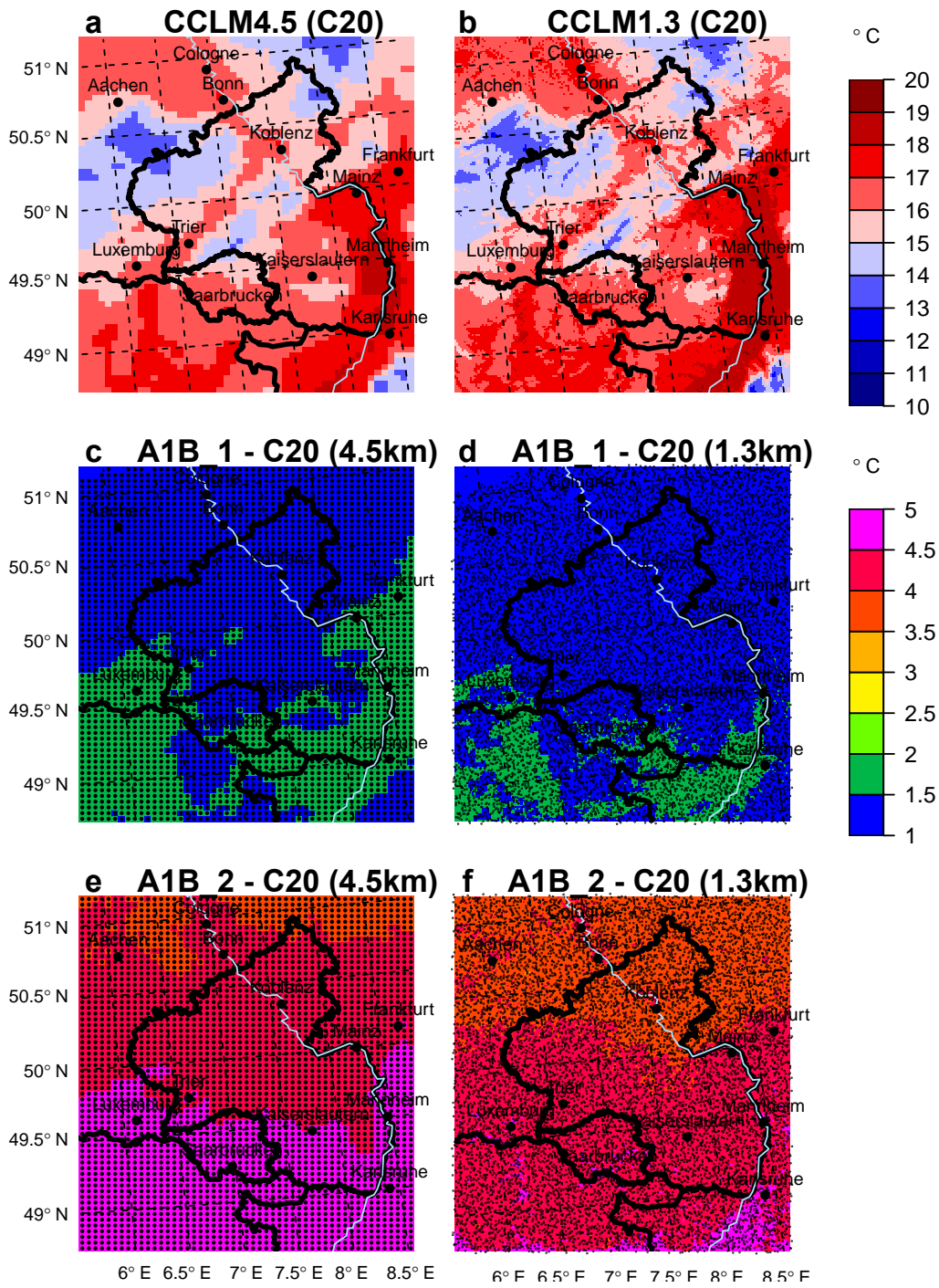


Figure 6.6 Climate change signals of summer (JJA) seasonal mean 2 m temperature: a) JJA mean 2 m temperature from CCLM4.5c20, b) as in a) but from CCLM1.3c20, c) CCLM4.5a1b1 - CCLM4.5c20, d) CCLM1.3a1b1 - CCLM1.3c20, e) CCLM4.5a1b2 - CCLM4.5c20 and f) CCLM1.3a1b2 - CCLM1.3c20. The global significance level of the FDR was set to $\alpha_g = 10\%$.

Table 6.5 Power analysis of the t-tests of the seasonal CCLM4.5 2 m temperature climate change signals.

	power	n(p=0.8) ¹	σ_{xy}^2	D
DJF				
2050	0.93	6.96	1.17	1.92
2100	1.0	2.29	0.95	4.11
MAM				
2050	0.6	15.39	1.51	1.58
2100	0.9	7.69	1.45	2.24
JJA				
2050	0.46	21.46	1.66	1.45
2100	1.0	2.95	1.40	4.39
SON				
2050	0.97	6.0	0.98	1.75
2100	1.0	2.61	1.16	4.18

Table 6.6 Power analysis of the t-tests of the seasonal CCLM1.3 2 m temperature climate change signals.

	power	n(p=0.8) ¹	σ_{xy}^2	D
DJF				
2050	0.93	6.97	1.18	1.93
2100	1.0	2.29	0.96	4.16
MAM				
2050	0.6	15.53	1.49	1.55
2100	0.91	7.65	1.44	2.22
JJA				
2050	0.46	21.57	1.57	1.37
2100	1.0	2.98	1.34	4.13
SON				
2050	0.97	6.03	0.99	1.77
2100	1.0	2.66	1.18	4.13

¹ Number of years needed to achieve a power of 80 %.

² σ_{xy} denotes the common variance.

in the Rhine Valley between Karlsruhe and Frankfurt. At the end of the century, the signals rise to 4.13 – 4.39 °C in average. From Fig. 6.6e-f a more or less zonal structuring of the CCS becomes visible with signals of about +5 °C in the south of the domain and about 3 – 3.5 °C in the north. Along river valleys, the CCS are larger compared to higher elevated surroundings. Again, the CCLM1.3 simulates slightly less intense CCS.

High power is achieved in every season except for the summer signal until 2091–2100. Here the power is only 46 % because the signal of 1.45 °C is smaller than $\sigma_{xy}^2 = 1.66$ °C. Although, the power of detecting temperature signals is quite high with the performed 10 year time slice experiments, an increase of the simulation period to 30 years might increase the power in spring and summer until 2041–2050 considerably.

The simulations indicate an intense shifting of the PDF of the 2 m temperature, which may lead to more intense extremes on its own, but from Tab. 6.4 it becomes obvious that especially in summer and autumn the σ of daily temperature rise from 3.5 °C by about 0.7 – 0.9 °C, an increase of about +20 – 25 %. Thus, in addition to the shift to higher temperatures, the spread of the PDF is increased and according to the scheme in Fig. 1.3 more hot extremes are to be expected. In winter however, the σ decrease about –0.7 °C to –0.8 °C so that there is again a twofold feedback that decreases the winter cold extremes.

6.2 Climate change signals from bias-corrected model data

6.2.1 Effect on precipitation signals

To assess the impact of the bias-correction methods on the climate change signal, the analyses from above were repeated but with the bias-corrected precipitation and temperature. Therefore, the percentage changes in seasonal and annual mean precipitation sums and daily standard deviation were recalculated.

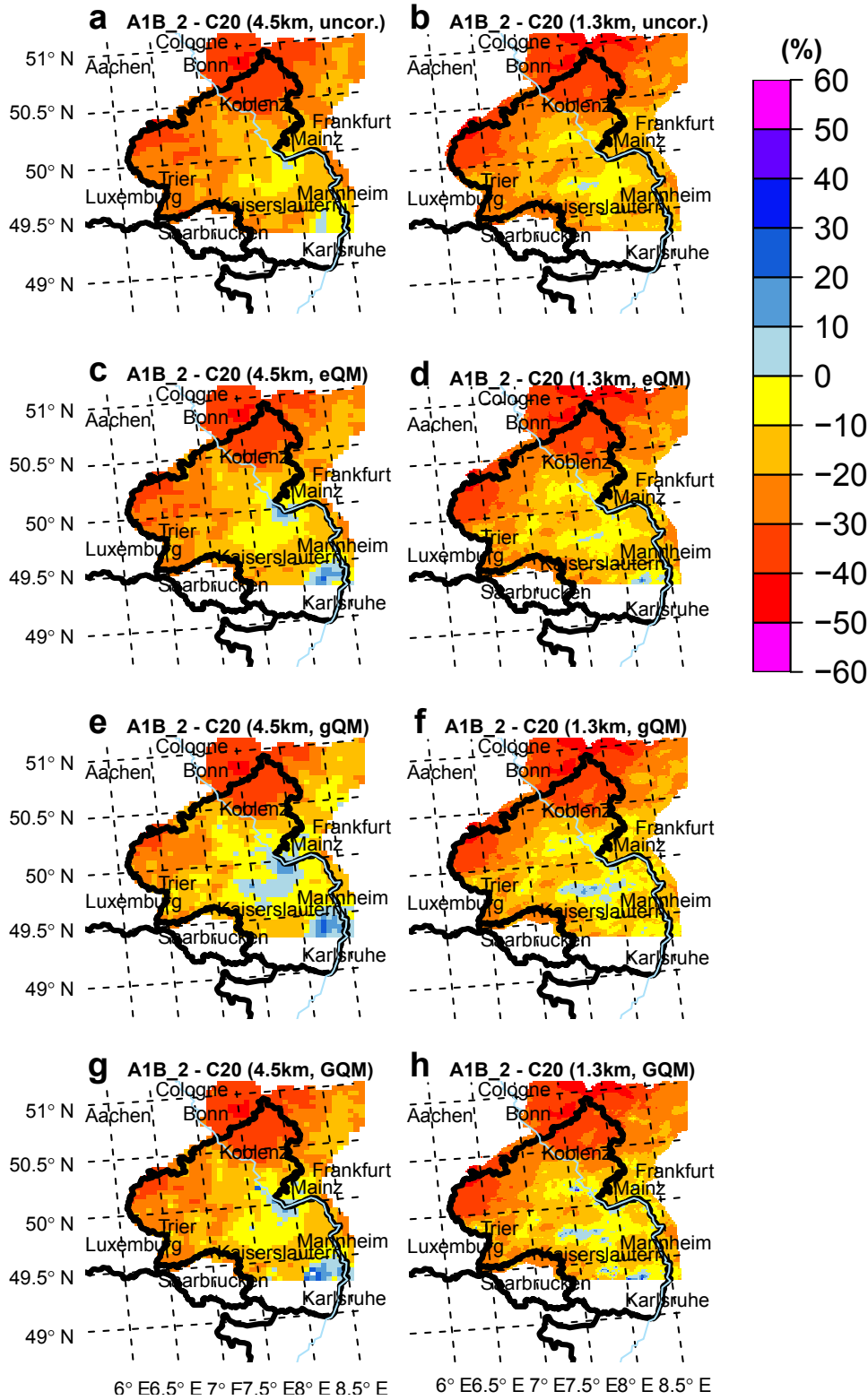


Figure 6.7 Climate change signals of seasonal precipitation in summer (JJA) until 2091–2100 of uncorrected CCLM data (uncor.) and after bias-correction with eQM, gQM and GQM for 4.5 km (a,c,e,g) and 1.3 km resolution (b,d,f,h). Significant grid-boxes (t-test, $\alpha = 5\%$) are marked with black dots. Note that FDR was not applied.

Table 6.7 Seasonal climate change signal (in %) of domain-averaged precipitation (mm season^{-1}) and the daily standard deviation (mm day^{-1}) after bias-correction.

model	MAM		JJA		SON		DJF		Annual	
	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
$\Delta 2050$										
CCLM4.5	+2	+2	-11	-5	+29	+12	+2	+8	+5	+5
eQM4.5	+3	+5	-9	± 0	+32	+12	+3	+8	+7	+7
gQM4.5	+4	+4	-9	-3	+31	+10	+4	+11	+7	+6
GQM4.5	+4	+5	-9	± 0	+32	+14	+4	+12	+7	+8
$\Delta 2100$										
CCLM1.3	+4	+5	-9	-4	+30	+13	+2	+8	+7	+6
eQM1.3	+4	+5	-9	-2	+30	+16	+2	+8	+8	+7
gQM1.3	+4	+5	-9	-4	+32	+15	+3	+9	+8	+7
GQM1.3	+4	+5	-9	+2	+32	+16	+3	+12	+8	+10
$\Delta 2100$										
CCLM4.5	+12	+9	-23	+7	± 0	+8	+1	+8	-3	+8
eQM4.5	+13	+12	-20	+10	+3	+6	+3	+8	± 0	+9
gQM4.5	+15	+16	-16	+11	+4	+7	+5	+7	+2	+10
GQM4.5	+13	+12	-19	+15	+4	+10	+4	+8	± 0	+12
CCLM1.3	+11	+17	-23	+8	+1	+8	+1	+8	-2	+9
eQM1.3	+11	+15	-23	+6	+1	+8	+1	+10	-1	+9
gQM1.3	+12	+19	-20	+10	+2	+8	+3	+8	± 0	+10
GQM1.3	+12	+17	-21	+14	+2	+10	+2	+8	-1	+11

The CCS were again tested with a two sided t-test with $\alpha = 0.05$ and the power analysis was performed on the seasonal data as before. This analysis was however restricted to the BC target region which was defined in chapter 5. Outside this region, no high-resolution observational data were available. It is a comparable area as in the study of (Gutjahr and Heinemann, 2013) except that the southern part of RLP is excluded and a part north of RLP was included.

The complete results of the sensitivity study of the average seasonal and annual precipitation sums are given in Tab. 6.7. In addition, the CCS of the uncorrected data are shown for the this area. In general, the impacts of the bias-correction on the seasonal precipitation sums (μ) are minor. But there is a pattern visible: if a seasonal signal is positive (negative) in the uncorrected model, then the bias-correction adjusts the signal to a more positive (less negative) signal. Thus positive signals are strengthened and negative signals are weakened. For instance, the signal of precipitation in autumn (SON) until 2041-2050 from CCLM4.5 is +29 % in the uncorrected model. This positive signal increases to +31 % after gQM and +32 % after GQM. In contrast, the signal of summer precipitation (JJA) until 2100 from CCLM4.5 is -23 % but only -16 % to -19 %. In the latter case, the difference is about 7 %. However, the correction with eQM have the weakest impact on the CCS so that the signals remain constant or at least close to the uncorrected signals. But overall, there is no change in sign after the bias-correction. The positive annual CCS until 2041-2050 are slightly increased by 1 – 2 % by all BC methods for both models and the negative annual CCS until 2091-2100 either reduce to ± 0 % or even change to a slight positive value (see gQM).

Modifications of the same order were found for the change in daily standard deviation. The largest impact in both models was found for the summer CCS until 2091-2100 where the positive change in σ is increased after the correction with GQM. The CCS changes from the uncorrected signal from 7 % to 15 % and from 6 % to 14 %, respectively.

Since the seasonal CCS of SON until 2041-2050 and JJA until 2091-2100 were the only significant

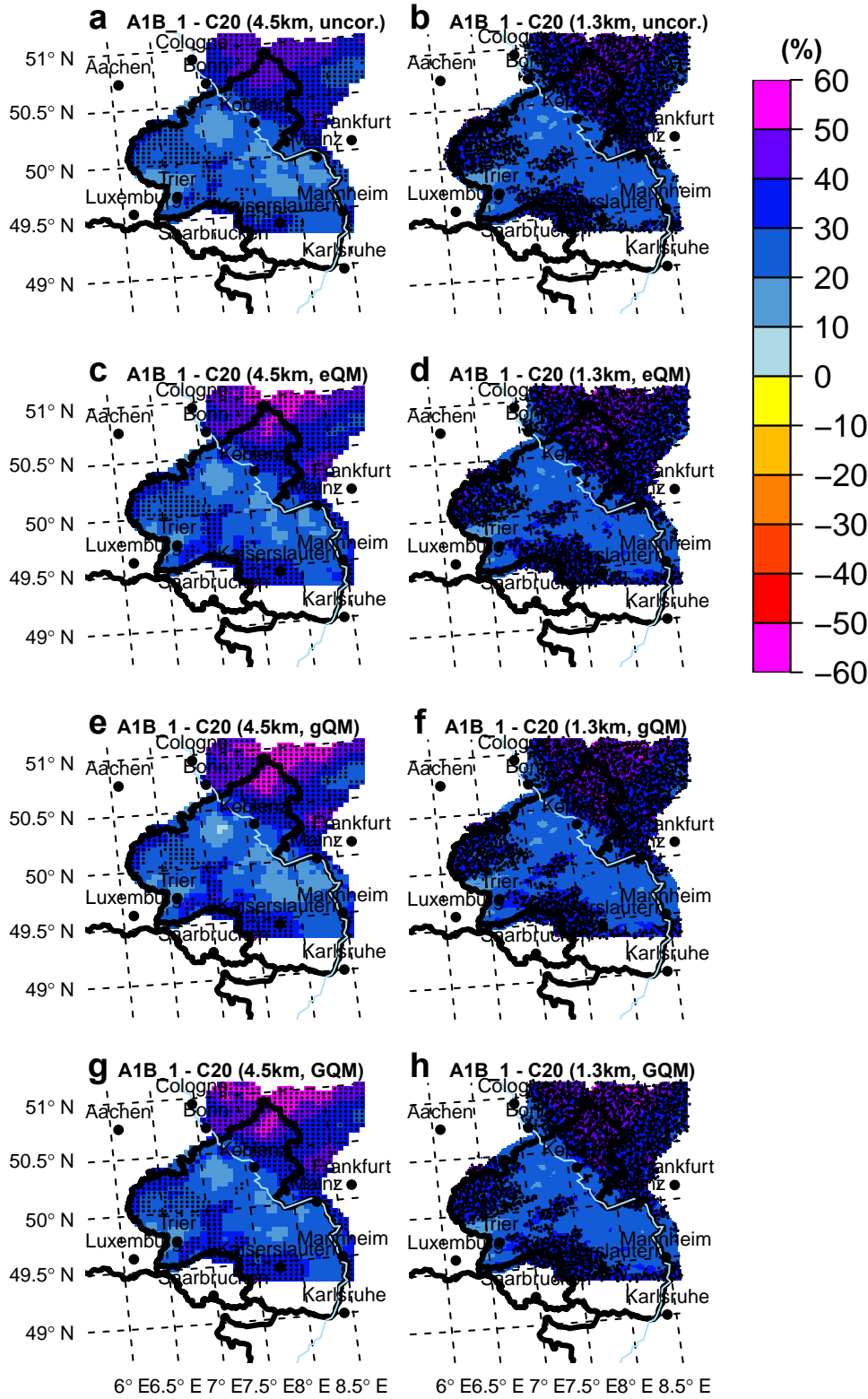


Figure 6.8 Climate change signals of seasonal precipitation in autumn (SON) until 2041–2050 of uncorrected CCLM data (uncor.) and after bias-correction with eQM, gQM and GQM for 4.5 km (a,c,e,g) and 1.3 km resolution (b,d,f,h). Significant grid-boxes (t-test, $\alpha = 5\%$) are marked with black dots. Note that FDR was not applied.

Table 6.8 Effect of the bias-correction on the number of significant grid boxes. The impact is shown for the CCS of SON precipitation until 2050 and of JJA precipitation until 2100. Shown are the percentages of grid points where a two-sided t-test was significant on the 95 % significance level.

CCS of	CCLM4.5	eQM4.5	gQM4.5	GQM4.5	CCLM1.3	eQM1.3	gQM1.3	GQM1.3
SON 2050	54.8 %	55.5 %	47.7 %	57.5 %	58.9 %	59.2 %	55.8 %	58.6 %
JJA 2100	34.5 %	27.8 %	20.6 %	27.2 %	29.1 %	27.6 %	24.0 %	27.3 %

signals, at least in a larger part of the domain, the spatial distribution of the signals was analysed for these two seasons. The first comparison was made for the summer precipitation and its CCS until 2091-2100 (Fig. 6.7). To illustrate a potential impact of the BC on the signals, the FDR restriction was not applied. Thus, in the uncorrected signals of CCLM4.5 (Fig. 6.7a) and of CCLM1.3 (Fig. 6.7b) more significant grid boxes occur than as in Fig. 6.4e-f.

The strong decrease in summer precipitation is mainly restricted to the mountain ranges of the Westerwald in the north of the domain and the Eifel in the west. The strongest CCS is up to -50% in the CCLM4.5 and the CCLM1.3, respectively.

Although the general pattern of the signal is similar between both resolutions, there are some differences in the centre and in the south of RLP. Locations with positive signals of up to 10% are slightly shifted or even missing in the CCLM1.3. However, those signals are not significant and are hence treated as noise. Furthermore, significant grid boxes in the Eifel are more concentrated along the border compared to CCLM4.5 and there are less significant grid boxes in the sector 6.5°E to 7°E and 50°N to 50.05°N . Note however that for illustrative purposes some significant grid boxes have been omitted from the plot.

After applying the eQM correction the main structure of the signal remained but as assessed before, the negative signals are dampened and the positive signals are amplified. For instance, the signal increases to $+30\%$ in the south, between Mannheim and Karlsruhe (Fig. 6.7c). Overall, the number of significant grid boxes reduces for both resolutions. In particular in the region between the Eifel and the Westerwald and in the eastern part of the Westerwald. In the CCLM4.5 34.5% of the grid boxes were significant at the 95% significance level (two-sided t-test). This percentage drops after the bias-corrections (Tab. 6.8). The largest impact showed the gQM method; only 20.6% and hence -13.8% less grid boxes remain significant, whereas the reduction is only about -7% after eQM and GQM. This is also visible in the CCS of the CCLM1.3 although less pronounced. The CCS of the SON precipitation until 2041-2050 are less affected. Here the percentages are approximately constant after the eQM and GQM correction but reduce after the gQM correction by about -3% .

The effect on the bias-correction on the probability to detect a climate change signal, the power of the t-test, is summarized in Tab. 6.9 and Tab. 6.10. The power was calculated at every individual grid box and spatially averaged afterwards. The most striking discrepancies concerns the signals of JJA until 2091-2100. The prior power of 0.58% calculated from the uncorrected CCLM4.5 model data drops considerably after the bias-corrections. The worst case is again the gQM correction after which the power only reaches 26% . Distinctively lower is the drop after applying eQM (43%) and GQM (0.39%). The reason for this declining power is a reduction in the CCS (note that the absolute CCS is shown as D in Tab. 6.9 and Tab. 6.10) while the variances fluctuate around the uncorrected value. Thus the ratio of signal-to-noise is alternated. For instance, the signal after gQM is only about 39 mm , whereas originally it was 62 mm but the common variance remains approximately constant with about 63 mm .

Table 6.9 Power analysis of the t-tests of the seasonal CCLM4.5 precipitation climate change signals before and after BC.

	power	n(p=0.8) ¹	σ_{xy}^2	D
$\Delta 2050$				
DJF				
CCLM	0.05	4040.85	96.73	6.03
eQM	0.06	1583.95	88.29	8.79
gQM	0.06	897.15	93.54	12.38
GQM	0.06	1307.54	90.46	9.92
MAM				
CCLM	0.06	1217.72	48.17	5.47
eQM	0.06	596.70	43.78	7.11
gQM	0.07	484.37	45.78	8.25
GQM	0.07	526.79	45.09	7.79
JJA				
CCLM	0.15	85.90	67.17	28.88
eQM	0.12	118.79	58.57	21.38
gQM	0.11	130.75	61.78	21.49
GQM	0.12	114.44	58.69	21.83
SON				
CCLM	0.71	12.16	62.93	74.75
eQM	0.73	11.67	56.83	69.07
gQM	0.70	12.53	61.07	71.39
GQM	0.74	11.45	57.85	71.04
$\Delta 2100$				
DJF				
CCLM	0.05	30966.95	74.61	1.68
eQM	0.06	1260.66	67.63	7.55
gQM	0.06	591.29	71.82	11.71
GQM	0.06	1094.65	69.50	8.33
MAM				
CCLM	0.16	74.60	57.72	26.65
eQM	0.17	70.77	52.38	24.84
gQM	0.18	66.18	56.17	27.56
GQM	0.17	72.25	53.23	24.98
JJA				
CCLM	0.58	16.19	61.92	62.98
eQM	0.43	22.94	55.24	46.72
gQM	0.26	42.05	63.35	39.17
GQM	0.39	26.01	57.34	45.42
SON				
CCLM	0.05	748909.60	80.75	0.37
eQM	0.05	2005.69	74.14	6.56
gQM	0.06	1117.42	80.57	9.55
GQM	0.06	1342.11	75.57	8.18

Table 6.10 Power analysis of the t-tests of the seasonal CCLM1.3 precipitation climate change signals before and after BC.

	power	n(p=0.8) ¹	σ_{xy}^2	D
$\Delta 2050$				
DJF				
CCLM	0.05	3568.08	97.67	6.48
eQM	0.05	2705.95	93.25	7.10
gQM	0.05	2083.65	94.48	8.20
GQM	0.05	2052.77	93.66	8.19
MAM				
CCLM	0.07	478.85	45.92	8.32
eQM	0.07	416.45	43.34	8.42
gQM	0.07	428.05	43.81	8.40
GQM	0.07	415.56	43.70	8.50
JJA				
CCLM	0.11	136.45	59.52	20.26
eQM	0.10	165.04	55.82	17.27
gQM	0.10	154.15	56.22	18.00
GQM	0.10	166.08	55.55	17.13
SON				
CCLM	0.74	11.43	61.16	75.18
eQM	0.76	10.92	57.80	72.91
gQM	0.75	11.26	59.02	73.17
GQM	0.76	10.89	58.05	73.33
$\Delta 2100$				
DJF				
CCLM	0.05	25222.67	75.28	1.88
eQM	0.05	3170.06	71.74	5.05
gQM	0.05	2524.52	72.81	5.74
GQM	0.05	3002.36	72.28	5.23
MAM				
CCLM	0.16	80.15	55.01	24.49
eQM	0.16	81.11	52.18	23.10
gQM	0.16	77.53	53.39	24.18
GQM	0.16	81.17	52.52	23.24
JJA				
CCLM	0.43	23.04	59.42	50.14
eQM	0.38	26.71	55.89	43.66
gQM	0.32	32.07	58.40	41.50
GQM	0.35	29.27	56.89	42.39
SON				
CCLM	0.05	35977.33	79.10	1.65
eQM	0.05	4598.49	75.31	4.40
gQM	0.05	3614.74	77.17	5.09
GQM	0.05	3626.62	75.84	4.99

¹ Number of years needed to achieve a power of 80 %.² σ_{xy} denotes the common variance.

Table 6.11 Seasonal climate change signal of domain-averaged temperature ($^{\circ}\text{C}$) and the daily standard deviation ($^{\circ}\text{C}$) after bias-correction.

model	MAM		JJA		SON		DJF		Annual	
	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
$\Delta 2050$										
CCLM4.5	1.56	-0.5	1.53	+0.3	1.43	-0.1	1.98	-0.6	+1.69	-0.3
eQM4.5	1.48	-0.5	1.53	+0.2	1.54	-0.1	1.98	-0.5	+1.70	-0.3
$\Delta 2100$										
CCLM1.3	0.79	-0.6	1.38	+0.3	1.32	-0.2	1.99	-0.5	+1.66	-0.4
eQM1.3	0.79	-0.6	1.38	+0.1	1.32	-0.1	1.99	-0.5	+1.63	-0.4
$\Delta 2100$										
CCLM4.5	2.24	-0.5	4.27	+0.8	4.10	+0.7	4.19	-0.7	+3.70	+0.1
eQM4.5	2.17	-0.5	4.25	+0.1	4.97	+0.8	4.21	-0.7	+3.65	± 0
CCLM1.3	2.22	-0.7	4.0	+0.7	4.01	+0.6	4.24	-0.7	+3.62	± 0
eQM1.3	2.22	-0.7	4.0	± 0	4.01	+0.7	4.24	-0.7	+3.51	-0.1

6.2.2 Effect on temperature signals

To analyse the impact of the seasonal eQM correction on the temperature signals and the power of detection the same procedure as for precipitation was repeated. In Tab. 6.11 the summary of the impact on the mean seasonal signal and on the daily standard deviation is given. Overall, the bias-correction modifies the CCS much less than it was the case for precipitation. The spatial distribution of the signal did not change after the BC and all grid boxes remained significant for the mid and end of the century. Therefore maps of the CCS are not shown. The CCS of the seasonal means are robust and nearly not affected from the bias-correction with an exception for the SON signal until 2091-2100 from eQM4.5. The original signal of $+4.10^{\circ}\text{C}$ is amplified to $+4.97^{\circ}\text{C}$. Another distinct impact is the reduction of the CCS of the daily standard deviation in summer. Here the eQM4.5 reduces the large increase of σ from $+0.8^{\circ}\text{C}$ to 0.1°C , the same holds true for eQM1.3. In the other seasons, the change in σ is not affected. Similarly, the power of the t-tests was not affected as well.

6.3 Summary and discussion

The changes in the mean state and in variability, expressed in terms of daily standard deviation, simulated by the regional climate model CCLM were analysed with statistical techniques for a subregion of the Saar-Lor-Lux and RLP domain. The strongest signal of seasonal change of precipitation were detected in autumn until 2041-2050 with $+32\%$ and in summer until 2091-2100 with -28% . The increase in autumn was also detected in the forcing data (CLM18a1b1) and is due to a positive anomaly in this period (Hollweg et al., 2008), which is illustrated in Fig. C.1. The detection of these two signals showed the highest power of the t-tests as well. The autumn signal was detected with a power of $76 - 77\%$ and the summer signal with a power of $52 - 57\%$. Thereby, it was assumed that the seasonal means within the above defined time-slices are distributed normally. Additional Wilcoxon-Mann-Whitney tests (U-tests) revealed the same significant CCS (not shown). Thus the signal detection is robust to the assumption of the population distribution. There were no significant signals detected for the spring and winter season and the changes in annual mean precipitation were not significant as well. The increase in horizontal resolution from 4.5 km to 1.3 km only showed a minor impact on the simulated CCS. Although the CCLM1.3 simulates to less precipitation in summer, which was revealed in chapter 4, the dry-bias did

not affect the relative change in summer.

Besides the changes in seasonal means, the change in daily standard deviation was investigated. Until 2041-2050 there is a large increase in σ for the autumn season of about $+0.8 \text{ mm day}^{-1}$ which has to be seen in connection with the above mentioned positive anomaly. In the other seasons, the signals are not as strong. For example in winter there is an increase of $+0.4 \text{ mm day}^{-1}$ while in spring and summer there is nearly no change. Until 2091-2100 however, an increase of σ of about $+0.2 - 0.7 \text{ mm day}^{-1}$ is simulated for all seasons by both models.

The 2 m temperature rises by about $+1.7^\circ\text{C}$ until 2041-2050 in all seasons and the signal is nearly homogeneous over the domain. Until 2091-2100 the CCS increases to $+3.7^\circ\text{C}$. In the south of the domain and along the Rhine Valley, the signal is even $+5^\circ\text{C}$. Except for MAM2050 and JJA2050, the power of the t-tests is close to 100%. The inflated power in these two seasons is because of an higher inter-annual variability compared to SON and DJF.

To analyse the impact of bias-correction methods on the CCS, three BC methods have been applied to precipitation and one to 2 m temperature and the detection was repeated. The results yield that the bias-correction affects the precipitation signals stronger than the temperature signals. For most seasons the impact on precipitation seasonal means is minor but a systematical effect became visible. The BC amplifies positive signals and dampens negative signals, thus all signals are altered into more positive values. This was also found by Gutjahr and Heinemann (2013) for the same domain but with a different observational data set that was used to calibrate the transfer functions. The strongest dampening is caused by gQM for the summer signals until 2091-2100. The parametric corrections gQM and QQM yield an increase in σ in summer but this effect is absent or of minor importance in the other seasons. It was also found that the number of significant grid boxes were reduced from all BC methods. The largest reductions of about -14% were found for gQM. In contrast, the temperature signals were nearly not affected by the seasonal eQM correction. The only impact was found for the CCS of standard deviation for $\Delta 2100$. Here the CCS is reduced from 0.8°C to 0.1°C and even to $\pm 0^\circ\text{C}$ with respect to CCLM1.3. The reason of this alteration is due to the non-linear correction of the BC methods (Hagemann et al., 2011). This means that lower precipitation values were further reduced while higher values may be increased. This unequal correction alters the PDF and thus the mean and σ . To analyse the impact of bias-correction methods on the CCS, three BC methods have been applied to precipitation and one to 2 m temperature and the detection was repeated. The results yield that the bias-correction affects the precipitation signals stronger than the temperature signals. For most seasons the impact on precipitation seasonal means is minor but a systematical effect became visible. The BC amplifies positive signals and dampens negative signals, thus all signals are altered into more positive signals. This was also found by (Gutjahr and Heinemann, 2013) for the same domain but with a different observational data set which was used to calibrate the transfer functions. The gQM dampens the summer signals until 2100 the most. The parametric corrections gQM and QQM yield an increase in σ in summer but this effect is absent or of minor importance in the other seasons. It was also found that the number of significant grid boxes were reduced from all BC methods. The largest declines were found for gQM of about -14% . In contrast, the temperature signals are nearly not affected by the seasonal eQM correction. The only impact was found for the CCS of standard deviation for $\Delta 2100$. Here the CCS is reduced from 0.8°C to 0.1°C and even to $\pm 0^\circ\text{C}$ with respect to CCLM1.3. The reason of this alteration of the climate change signals is due to the nonlinear correction of the BC methods (Hagemann et al., 2011). This means that lower precipitation values were further reduced while higher values may be increased. This unequal correction alters the PDF and thus the mean and σ .

From this results it can be concluded that there is no significant change in annual precipitation mean but the precipitation is distributed on fewer days which leads to an increase in daily variability. Hence the hypothesis of more frequent and intense extremes in the future arises. The temperature

means shift to warmer values in all seasons and in addition the daily variability increases in summer and reduces in winter. This leads to the expectations that temperature extremes which are defined as threshold-exceedances, such as heat-waves or warm-spells, increase in frequency due to the shift in mean temperature or the first moment of the PDF (e.g. Ballester et al., 2010). Whereas changes in intensities (e.g. changes in uppermost percentile or return periods/levels) are highly sensitive to changes in higher-order statistical moments such as daily temperature variability (Fischer et al., 2012) (refer to Fig. 1.3). Thus, it is likely that summer hot extremes and winter cold extremes warm stronger than the mean. With regard to precipitation, there is no a priori hypothesis for summer extremes from a statistical point of view because the changes in mean and variability are simulated contrary. From a physical point of view an increase in summer extremes is expected because of the Clausius-Clapeyron equation. Thus the atmospheric water vapour content increases with increasing temperature and hence there is more potential moisture in the atmosphere that may produce more extreme precipitation.

Although the temperature signals were found to be robust, the signals of precipitation are much more uncertain. It is stressed here that the analysis was performed as time-slice experiments and thus the base period and the scenario periods do not cover 30 years, which are proposed by the World Meteorological Organization (WMO, 2011). To increase the robustness of the above mentioned conclusions, longer time periods should be considered once the computational costs on such horizontal scales become more feasible. In the next chapter the above mentioned hypotheses on extremes are investigated.

Extreme value analysis of precipitation and temperature fields

The severe impact which extremes may have on social, ecological or natural systems was emphasized in the Special Report on Extreme Events (SREX) of the Intergovernmental Panel on Climate Change (IPCC, 2012). Because of these tremendous consequences, the study of extremes under climate change is an ongoing field of research. The primary tool for such studies are climate models because they provide physically coherent fields of e.g. temperature and precipitation and are based on physical laws. Climate models are subject to constant improvements, for instance in terms of numerical schemes or the horizontal resolution, which requires repetitions of extreme value analysis in order to gain new or improved insight. For a consistent analysis of extreme climate events the Expert Team on Climate Change Detection and Indices (ETCCDI) defined a set of indices based on the work of Frich et al. (2002) that can be applied all over the globe. This allows the comparison of different study regions or data sets in a consistent way (Zhang et al., 2011). To improve their global applicability, the definitions of these indices have been subject to discussions and were finally revised (e.g. Zhang et al., 2005; Tebaldi et al., 2006).

Extremes can be characterized either by percentile-, threshold-, or duration-based indices, such as the ETCCDI, or by analysing statistically the tail-behaviour of the probability distribution function of a variable (Sillmann and Roeckner, 2008). The focus in this study is on precipitation and temperature extremes and therefore a subset of the ETCCDI indices was chosen (hereafter simply indices) (see Tab. 3.1). These indices define rather moderate extremes with short re-occurrence times of about 1 year or even less (Sillmann et al., 2013b) and form a balance between data availability and robustness of change (Zhang et al., 2011). The second part of this study is the statistical analysis of the tails by 'peak-over-threshold' (POT) models. With such an analysis much longer reoccurrence periods can be assessed. For this purpose, the 10- and 20-year return levels of precipitation, minimum and maximum 2 m temperature are estimated from the POT models. It is then analysed whether precipitation extremes increase in a warmer climate as expected from the Clausius-Clapeyron equation and how temperature extremes will change in summer and winter.

The climate simulations were performed with the COSMO-CLM (CCLM) as 10-year long time-slice experiments, which were defined as: 1991–2000 (C20), 2041–2050 (A1B_1) and 2091–2100 (A1B_2). The significance of the change signals is tested with Wilcoxon-Whitney tests at the 90 %, 95 % and 99 % level. The extreme indices and POT models (see section 3.3.2) were applied to the uncorrected model

Table 7.1 Domain-averages and spatial correlations of extreme indices (1993–2000) based on seasonal daily precipitation and temperature. Counting indices are displayed as total counts in 1993–2000. See the description of the indices in Tab. 3.1.

Index	CCLM4.5hc	CCLM1.3hc	INTERMET4.5	INTERMET1.3	REGNIE4.5	REGNIE1.3
JJA						
CDD	22 (-0.13/-0.07)	26 (0.09/0.02)	22	23	20	20
CDDn	31 (0.12/0.22)	36 (0.11/0.20)	32	32	30	30
R20	12 (0.44/0.40)	11 (0.33/0.37)	10	10	11	12
RX5DAY	61 (0.11/0.10)	62 (-0.05/-0.09)	45	45	46	47
RX5DAYn	0.89 (0.02/-0.05)	0.94 (-0.09/-0.06)	0.28	0.30	0.37	0.4
R95PTOT	45.33 (0.49/0.47)	56.01 (0.41/0.43)	43.21	43.49	40.67	40.86
DJF						
CDD	26 (-0.06/-0.02)	25 (-0.04/-0.04)	28	28	23	23
CDDn	19 (0.58/0.60)	21 (0.55/0.56)	23	23	20	20
R20	11 (0.9/0.91)	12 (0.86/0.88)	10	10	12	13
RX5DAY	42 (0.69/0.70)	43 (0.58/0.62)	44	44	47	47
RX5DAYn	0.28 (0.69/0.68)	0.34 (0.55/0.60)	0.29	0.33	0.47	0.52
R95PTOT	33.21 (0.74/0.75)	34.53 (0.72/0.74)	36.48	36.54	33.73	33.78
FD	427 (0.91)	427 (0.88)	380	386	-	-
CFD	34 (0.21)	35 (0.24)	31	31	-	-
TR20	13 (0.19)	16 (0.14)	1	1	-	-
HWFI	5 (0.31)	8 (0.17)	1	1	-	-
HWFI _n	0.72 (0.22)	0.93 (0.09)	0.19	0.18	-	-
HWDI	25 (-0.14)	18 (-0.02)	29	29	-	-
HWDI _n	3 (0.05)	2 (0.09)	4	4	-	-
CWFI	1.53 (NA)	1.24 (0.01)	0.00	0.33	-	-
CWFI _n	0.22 (NA)	0.18 (0.01)	0.00	0.05	-	-
CWDI	55 (0.44)	53 (0.37)	38	37	-	-
CWDI _n	6 (0.25)	6 (0.23)	3	3	-	-
GSL.l	231 (0.87)	231 (0.89)	264	261	-	-
GSL.d	88 (0.86)	88 (0.87)	62	63	-	-

¹ In brackets: comparison with (INTERMET / REGNIE).

output of the CCLM in 4.5 km and 1.3 km resolution. In case of precipitation, the analyses have been repeated for the bias-corrected fields to investigate the impact on extreme statistics.

7.1 Extreme indices

The indices of Tab. 3.1 were firstly calculated from the hindcast simulations (CCLM4.5hc, CCLM1.3hc) to analyse the performance of the CCLM model to simulate extremes in the past. Then the indices have been calculated from the climate simulations (C20, A1B_1 and A1B_2) and finally from the bias-corrected fields to investigate the impact on extreme indices. However, the effect could only be analysed for precipitation indices because neither maximum nor minimum temperature have been corrected and hence all indices which are based on these variables have been omitted. The precipitation and temperature indices were calculated for the summer (JJA) and winter (DJF) seasons only, because it is expected that the characteristics of extremes differs mainly between these two seasons. Furthermore, the analysis was based on grid boxes with a land-fraction of ≥ 0.5 .

Some thresholds are based on the minimum (T_{\min}) and maximum daily temperature (T_{\max}), which was not available from some data sets. For those data sets, T_{\min} and T_{\max} were calculated from hourly time series. Although the variation of temperature within one hour is negligible, T_{\min} (T_{\max}) might be

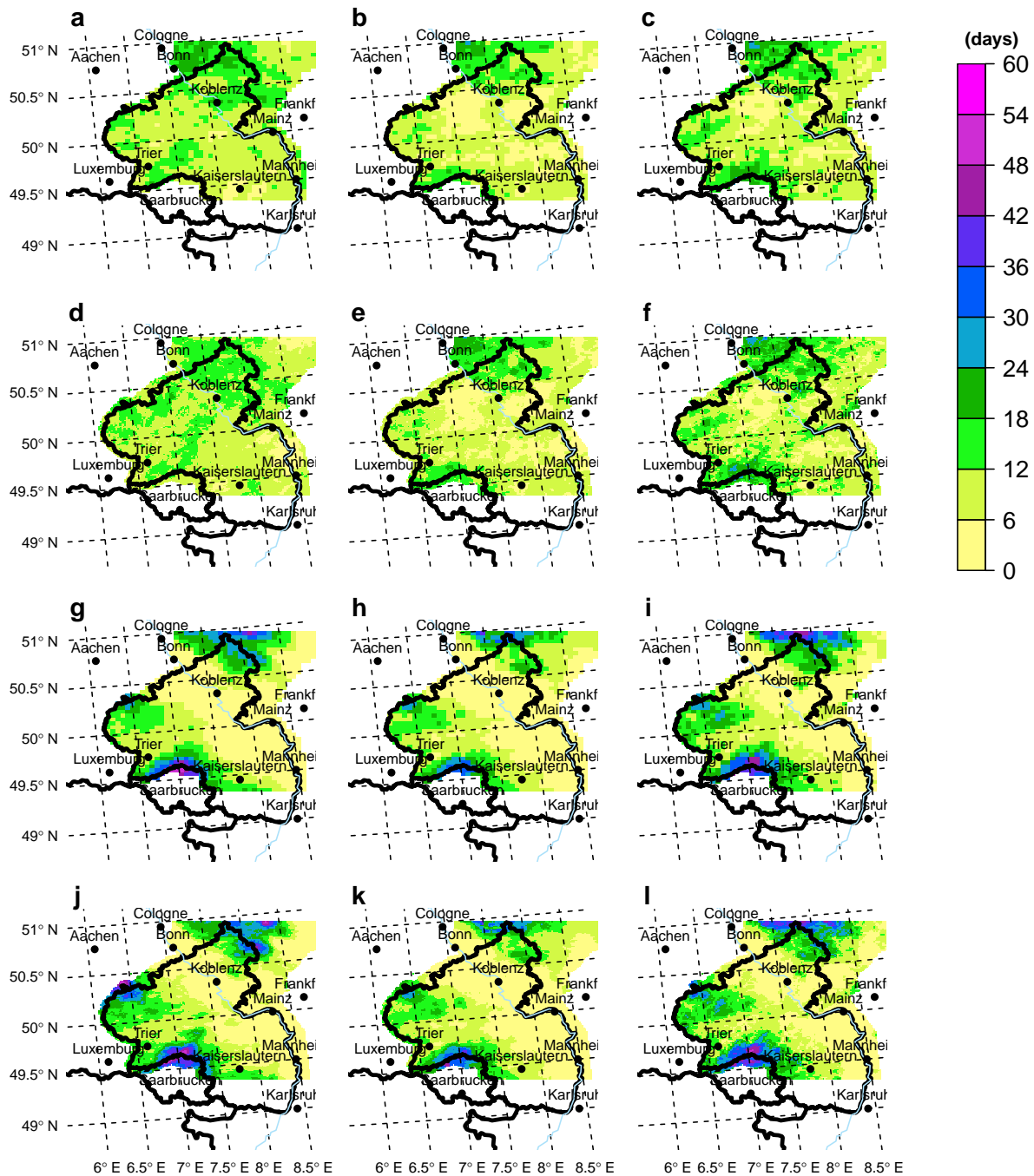


Figure 7.1 Spatial distribution of very heavy precipitation days (> 20 mm, R20) in summer (JJA) for a) CCLM4.5hc, b) INTERMET4.5, c) REGNIE4.5, d) CCLM1.3hc, e) INTERMET1.3, f) REGNIE1.3 and in winter (DJF) g) CCLM4.5hc, h) INTERMET4.5, i) REGNIE4.5, j) CCLM1.3hc, k) INTERMET1.3, l) REGNIE1.3.

overestimated (underestimated). Affected data sets were: INTERMET and the climate simulations of CCLM1.3.

7.1.1 Evaluation of the hindcast

The extreme indices were calculated from the hindcast simulations for the period 1993–2000. As observations, REGNIE was used for calculating the precipitation indices, whereas INTERMET was used for both precipitation and temperature indices. In order to allow a comparison, the observed fields have been remapped onto the CCLM grids with 4.5 km and 1.3 km resolution by a first order conservative remapping procedure (Jones, 1999)¹. Furthermore, the comparison could only be made for a subset of the domain because of the limited spatial coverage of INTERMET and REGNIE. The subset was defined as the intersection area of the observational data sets and covers approximately the state of Rhineland-Palatinate except for the southern part.

To rate the performance of the models, the domain-averages and the spatial variability of the indices were compared with observations. In addition, the spatial distributions of the indices have been compared by the spatial correlation coefficients after Pearson.

Precipitation indices

A summary of the resulting indices in terms of domain averages and spatial correlations are given in Tab. 7.1a. In general, the indices from the simulations are close to the values from the observations, but for some indices larger differences were found.

About 30–32 dry periods (CDDn) were observed within eight summers (1993–2000) and a similar number was simulated by the models; with 36 only slightly more in CCLM1.3hc. Likewise, the maximum length of a dry period (CDD) in summer is 20 – 26 days in the simulations and matches well with the observations, although it is slightly too long in CCLM1.3hc.

The spatial average number of very heavy precipitation days (R20), i.e. $> 20 \text{ mm day}^{-1}$, is well captured by the models and about 10 – 12 days within 8 years; less days in flat terrain but up to 30 days in orographic regions such as the Hunsrück, Westerwald or Eifel (e.g. Fig. 7.1a,d). This pattern is realistic and matches the observations (Fig. 7.1b,c,e,f), although there are some discrepancies between the observational data sets: REGNIE is more heterogeneous, for instance in the centre and southern part of RLP and produces more R20 days in mountainous areas than INTERMET. Concerning the different spatial resolutions of the hindcasts, CCLM1.3hc overestimates the R20 index in the Eifel and the Westerwald in comparison with the observational products.

The spatial mean of the highest precipitation sum of five consecutive days (RX5DAY) is about 45 – 47 mm in the observations but is overestimated by both CCLM with 61 – 62 mm. This is due to a thunderstorm that was simulated near Koblenz by the CCLM simulations but did not occur in the observations. The thunderstorm produced about 200 mm within 5 days. Likewise, precipitation periods with a minimum of 5 consecutive days and at least 50 mm (RX5DAYn) were simulated too frequently by both models and the percentage of extremes ($> 95 \%$ percentile) on the total precipitation (R95PTOT) is overestimated by about +10 % by CCLM1.3hc (Fig. 7.2d). This was already indicated in Fig. 4.12 in section 4.1.5. In average, 41 – 44 % of the total summer precipitation was caused by extremes with regard to the observations and with 45 – 56 % slightly more in the CCLM simulations. At the eastern fringe of the Palatinate Forest and in the Mainz Basin the percentage is even 60 %, particularly in INTERMET1.3 (Fig. 7.2e). The CCLM4.5hc captures these observed proportion but overestimates it in the centre of RLP and in the Eifel region. The spatial correlations of R20 and R95PTOT are about $r = 0.4 - 0.5$ but

¹ implemented in the Climate Data Operators (CDO, <https://code.zmaw.de/projects/cdo>)

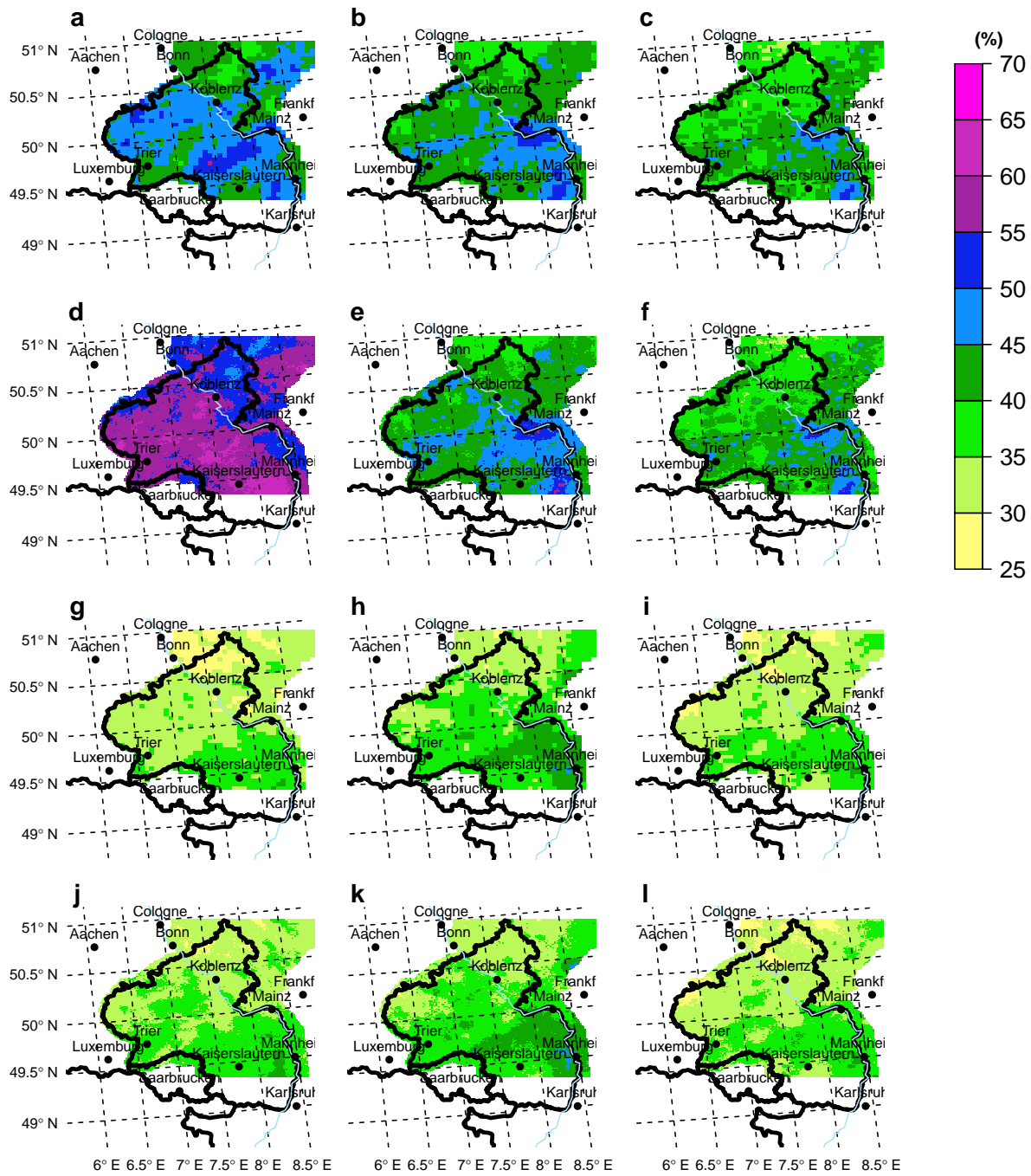


Figure 7.2 Spatial distribution of maximum 5-day precipitation amount (RX5DAY) in summer (JJA) for a) CCLM4.5hc, b) INTERMET4.5, c) REGNIE4.5, d) CCLM1.3hc, e) INTERMET1.3, f) REGNIE1.3 and in winter (DJF) g) CCLM4.5hc, h) INTERMET4.5, i) REGNIE4.5, j) CCLM1.3hc, k) INTERMET1.3, l) REGNIE1.3. As in Fig. 7.1 but for RX5DAY.

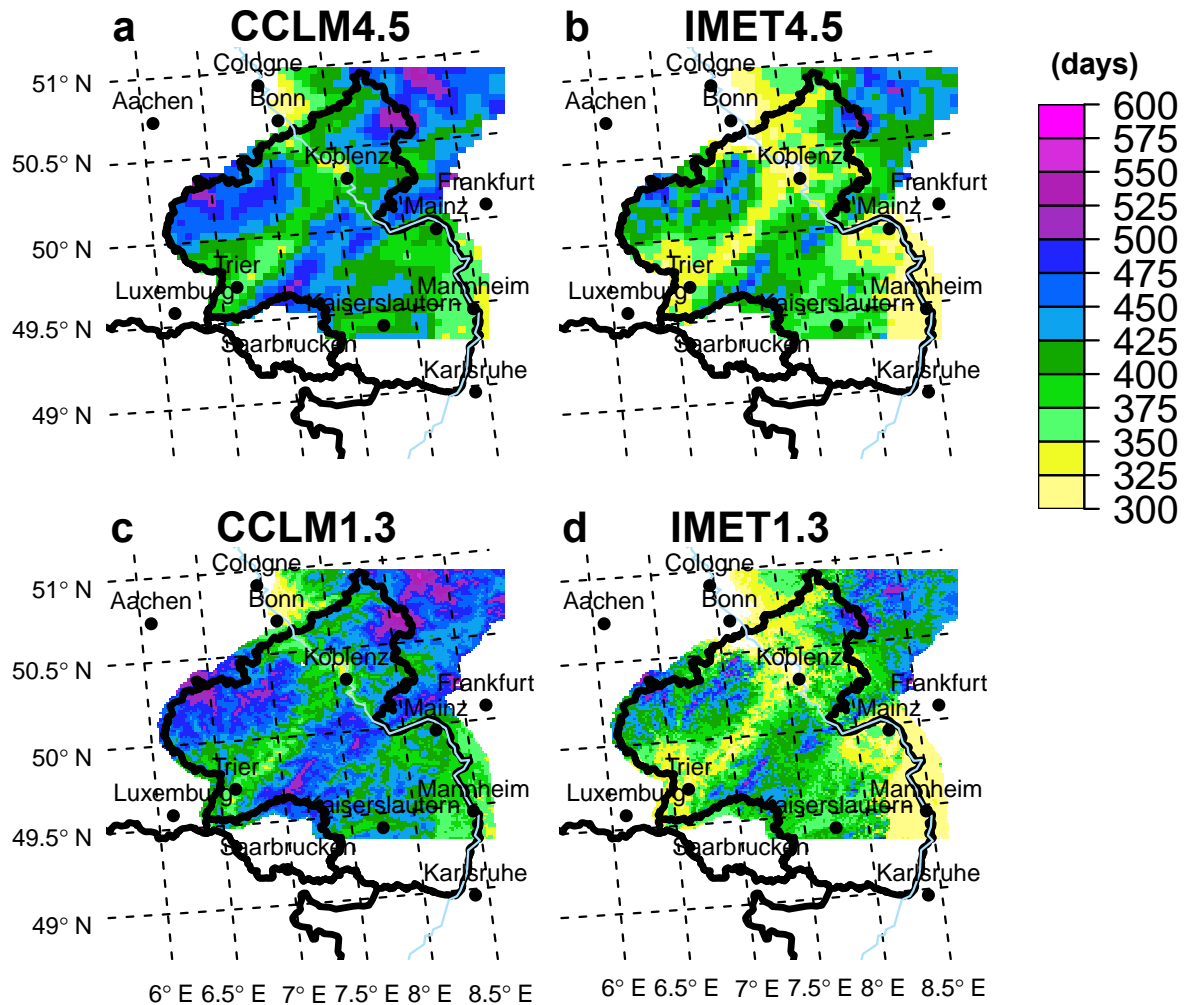


Figure 7.3 Total number of frost days (FD) in winter (DJF) in a) CCLM4.5hc, b) INTERMET4.5, c) CCLM1.3hc and d) INTERMET1.3.

fairly low for summer dry periods (CDD/CDDn) and for most of the other precipitation indices. Overall, the correlations are somewhat lower for CCLM1.3hc.

In winter, the CCLM simulations perform better and there are no larger deviations in the indices. In contrast to the summer season, the number of dry periods (CDDn) reduce to 19-23 but the maximum length of dry periods (CDD) persists as long as in summer. Similarly, R20 occurs equally often as in summer, about 10 – 12 days, but the spatial pattern does change. There is a pronounced negative gradient in R20 days from orographic to flat regions, which is well captured by both CCLMs. In particular in the Hunsrück and the Westerwald very heavy precipitation events occurred on up to 60 days (Fig. 7.1g-l). In INTERMET (Fig. 7.1h,k), this number is considerably lower than in REGNIE (Fig. 7.1i,l). Thus the CCLM simulations match better to REGNIE. RX5DAY reduce in the CCLMs to 42 – 43 mm day⁻¹ and fits well with the observations. Furthermore, high precipitation events (RX5DAYn) decrease in winter. The proportion of extreme precipitation on the precipitation sums reduces by approximately 7% in the observations, 12% in CCLM4.5hc and 22% in the CCLM1.3hc so that there is no overestimation of the

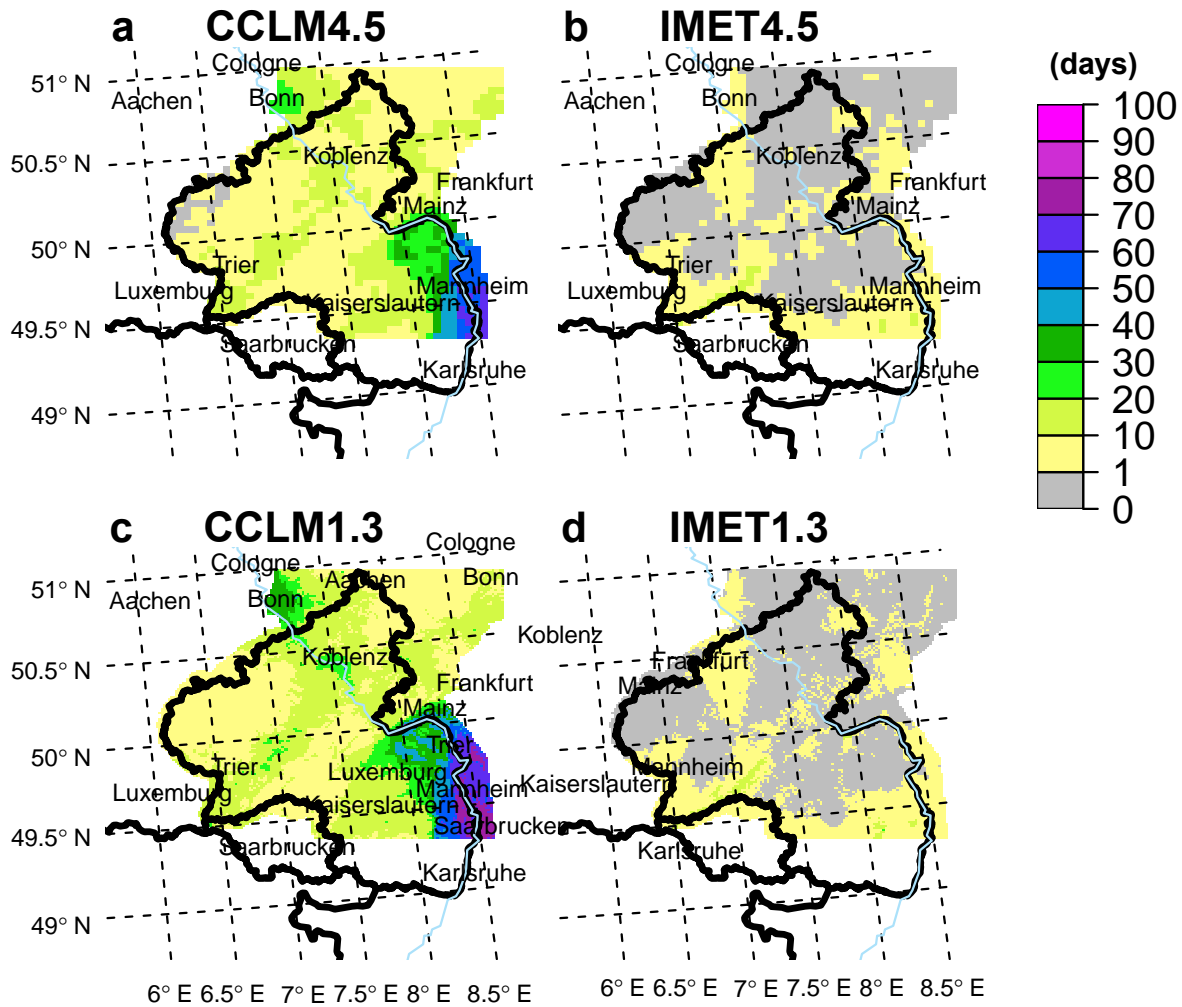


Figure 7.4 Tropical nights (TR20) in summer (JJA) in a) CCLM4.5hc, b) INTERMET4.5, c) CCLM1.3hc and d) INTERMET1.3.

precipitation extreme indices in winter. From Fig. 7.2g-l this reduction is apparent, although there is a gradient from south-east to north-west. In the Rhine Valley the R95PTOT index achieves the largest values in winter with > 40 % in INTERMET and > 35 % in REGNIE. In contrast, the area east of Bonn reaches only 30 – 35 %. These patterns indicate that the CCLM simulations fit better to REGNIE as well. Overall, the spatial correlations rise to $r = 0.55 - 0.91$, which could be expected since winter precipitation is usually better simulated because of the predominantly dependence on the large-scale synoptical situations. Local modifications or forcing play only a minor role. Solely the spatial correlation of CDD is not improved in winter.

Temperature indices

The temperature indices could only be compared to INTERMET, since REGNIE is a precipitation data set. In Tab. 7.1 the temperature indices are listed as well. Within the period of 1993–2000 there were

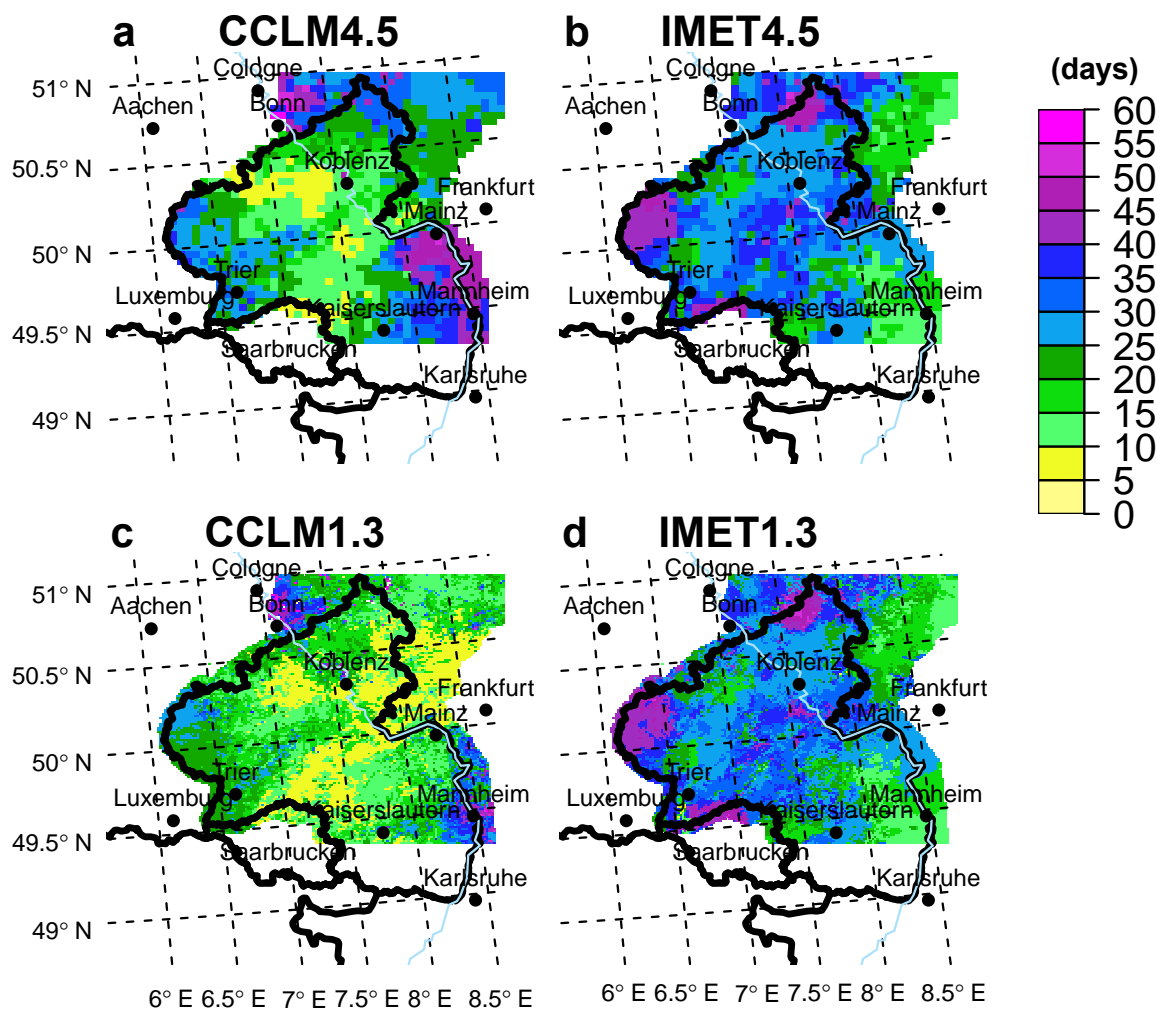


Figure 7.5 Heat-wave duration index (HWDI) in summer (JJA) in a) CCLM4.5hc, b) INTERMET4.5, c) CCLM1.3hc and d) INTERMET1.3.

about 50 more frost days (FD) simulated in the spatial mean by both CCLM than observed, which translates to about one additional week of frost days per winter. In Fig. 7.3 the spatial distribution of FD is shown. Generally, the CCLM models (Fig. 7.3a,c) simulates more FD in the whole domain compared to INTERMET (Fig. 7.3b,d). The pattern from the 1.3 km run is much more heterogeneous compared to that from the 4.5 km run, so that individual mountain peaks and ridges emerge, which are also visible in the observed FD field. Thus there is a clear added value from the 1.3 km run. The comparison of the hindcast runs with observations in chapter 4 revealed a cold-bias of approximately -1.5°C for wintertime. This explains why more frost days occurred in the simulations. On the other hand, the spatial density of temperature stations is lower than for precipitation and is generally sparse in mountainous regions. Therefore it is possible that the observed temperature fields are too smooth in such areas and the result highly depends on the interpolation algorithm. However, periods of consecutive frost days (CFD) are only slightly overestimated and occur most frequently in the mountains (not shown). Similarly cold waves (CWDIn) were simulated more often, about 6 within 8 years in the domain average in contrast to

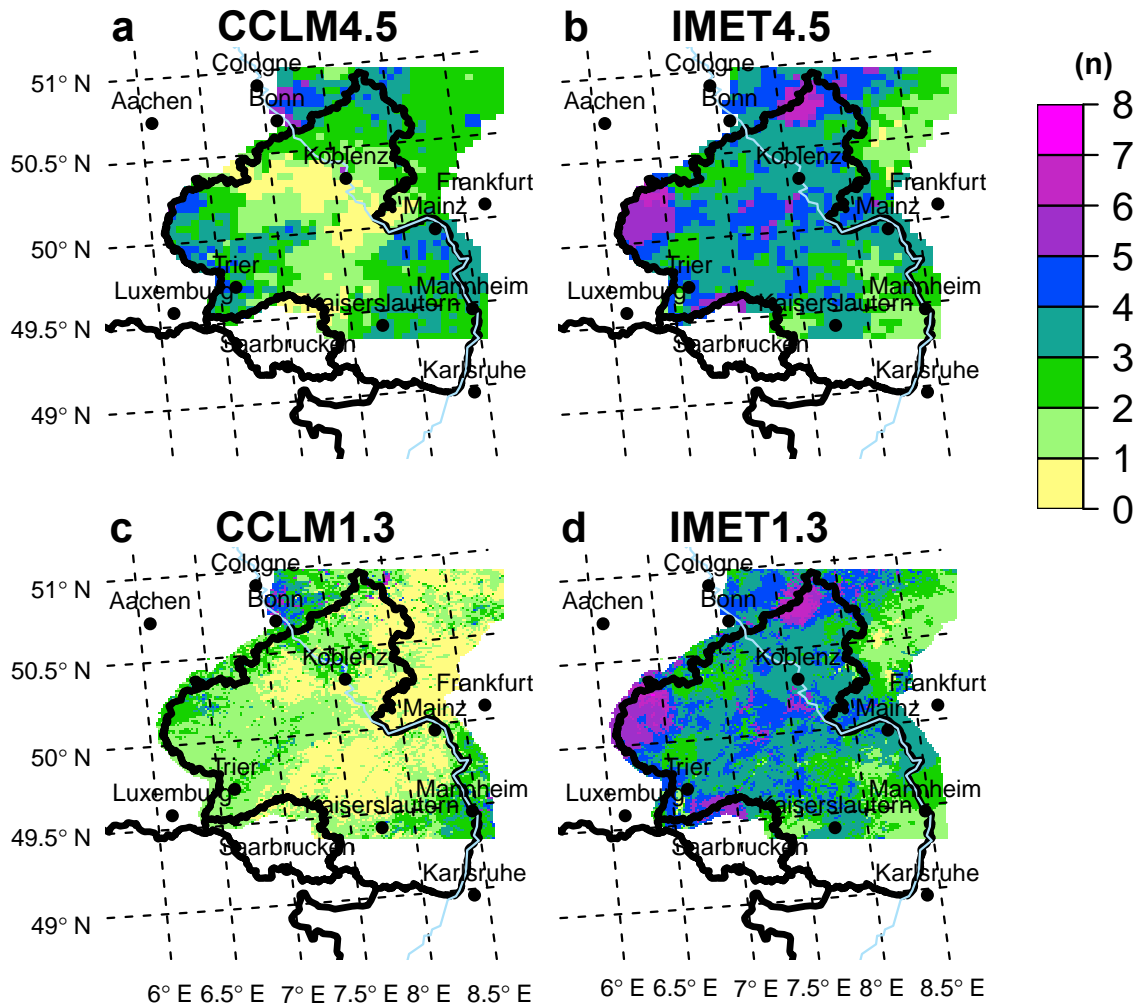


Figure 7.6 Number of heat-waves (HWDIn) in summer (JJA) in a) CCLM4.5hc, b) INTERMET4.5, c) CCLM1.3hc and d) INTERMET1.3.

only 3 observed cold-waves. The longest persisted 55 days, about 17 days longer than observed.

The discrepancies of the hot indices were found to be considerably larger. For example, within the period 1993–2000 there were 13–16 tropical nights (TR20) simulated in the spatial mean but only 1 was observed. In Fig. 7.4 this discrepancy becomes evident. The models simulate up to 100 tropical nights within 8 years in the Rhine Valley, more than 20 in the Mainz and Cologne Basin and in flatter areas of the Palatinate Forest and less than 20 elsewhere in RLP. Only along the Moselle river and near Koblenz does the number of TR20 rise again. Thus, temperature extremes clearly depend on the orography, which is plausible since the warmest temperatures do occur in the Rhine Valley. In a direct comparison the CCLM1.3c20 simulates about 20–30 more tropical nights in the Rhine Valley than the CCLM4.5c20 due to the increased horizontal resolution. Considerably different is the spatial distribution and amount of TR20 in INTERMET, which is only about 1–10. Slightly more tropical nights are visible in the Hunsrück (Fig. 7.4b,d), which does not seem feasible. Consequently, the pattern from INTERMET is not realistic and should be treated with caution, whereas the patterns from CCLM models seem plausible. Even

Table 7.2 Domain-averages of extreme indices calculated from the present (1991–2000, C20) and future climate simulations (2041–2050 (A1B_1), 2091–2100 (A1B_2)). See Tab. ?? for the description of the indices. The indices are based on seasonal daily precipitation and temperature and are given as average per year except for maximum indices such as CDD and RX5DAY. Changes which are significant with respect to the Wilcoxon-Mann-Whitney test are marked with symbols depending on the significance level α : +=0.1, *=0.05 and bold=0.01. The tests were performed with seasonal area-mean values.

Index ¹	CCLM4.5c20	CCLM1.3c20	CCLM4.5a1b1	CCLM1.3a1b1	CCLM4.5a1b2	CCLM1.3a1b2
JJA						
CDD	22	24	25	28	28 ⁺	37 ⁺
CDDn	3	4	3	4	4	5
R20	2	1	2	1	2	2
RX5DAY	61	55	56	56	70	66
RX5DAYn	0.10	0.07	0.09	0.08	0.19	0.14
R95PTOT	35.39	42.55	39.39	44.22	48.57	53.81 ⁺
DJF						
CDD	15	16	27	29	18	19
CDDn	2	2	1	2	2	2
R20	2	2	2	3	2	2
RX5DAY	41	42	43	45	42	43
RX5DAYn	0.03	0.04	0.04	0.05	0.04	0.04
R95PTOT	24.96	26.12	29.35 ⁺	30.37 ⁺	31.77	32.58
FD	46	46	32 [*]	31 [*]	14	14
CFD	30	28	38	21	9	9
TR20	1	1	3	3	13	14
HWFI	8	8	26	26	60	54
HWFI _n	0.24	0.22	0.82	0.77	2.95	2.86
HWDI	13	12	27	25	46	44
HWDI _n	0.46	0.45	0.83	0.76	2.54	2.39
CWFI	7	6	0	0	5	5
CWFI _n	0.11	0.11	0.01	0.01	0.08	0.07
CWDI	16	16	18	10	6	6
CWDI _n	0.47	0.47	0.35	0.17	0.09	0.09
GSL.l	236	236	274 [*]	273 [*]	320	320
GSL.d	77	77	57	58	27	27

¹ The indices CDD, RX5DAY, HWFI, HWDI, CWFI, CWDI denote spatial means of maxima within 10 years, whereas the counting indices CDDn, R20, RX5DAYn, HWFI_n, HWDI_n, CWFI_n, CWDI_n, FD and TR20 are expressed as spatial averages per summer/winter.

more unrealistic patterns are retrieved from the observations for the maximum duration and number of heat-waves (HWDI, Fig. 7.5 and HWDI_n, Fig. 7.6). In the CCLM, the spatial distribution of HWDI and HWDI_n is similar to that of TR20. In INTERMET however, the longest and most frequent heat-waves occurred in the mountain areas of Hunsrück, Eifel and Westerwald and the lowest duration and frequency in the Rhine Valley, which is not plausible. The inter-comparison of the CCLMs exhibits less frequent and less persistent heat-waves, in particular in the Rhine Valley where hot-spots appear as clusters but not as a smooth, coherent area as in the CCLM4.5c20. Finally, the onset of the growing season is simulated to begin on the 88. Julian day (29th of March), about one month later than calculated from the observations and lasts for 231 days, 30 days less than observed.

The spatial correlation is high for the indices FD, GSL.l, GSL.d but weak to moderate for the other indices due to the discrepancies which were discussed above.

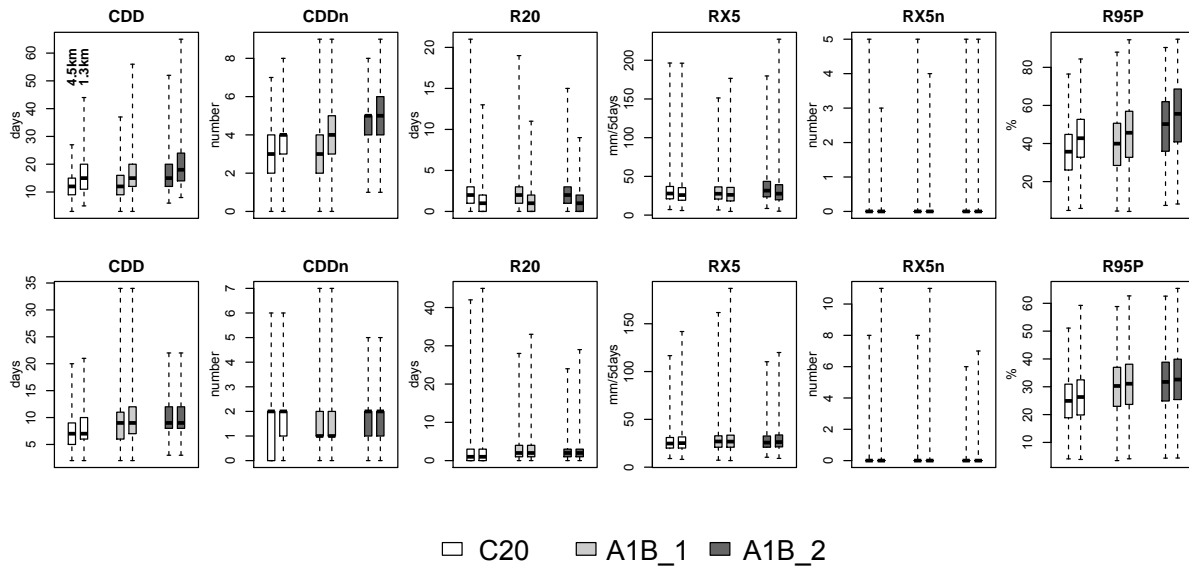


Figure 7.7 Box-and-whisker plots of the ETCCDI precipitation indices for the present climate 1991–2000 (C20) the near future 2041–2050 (A1B_1) and the end of the century 2091–2100 (A1B_2). Top row: summer (JJA) and bottom row: winter (DJF). The bold black horizontal lines mark the medians, the thin horizontal lines span the inter-quartile ranges (25 % to 95 % percentiles) and the vertical dashed lines represents extremes. Each boxplot consists of 10 values per pixel (4.5 km=51 · 62 · 10 = 31620; 1.3 km=172 · 208 · 10 = 357760).

7.1.2 Extreme indices from the climate simulations

In the next step the extreme indices have been calculated from the climate simulations of the present climate (1991–2000, C20) and the future projections for 2041–2050 (A1B_1) and 2091–2100 (A1B_2). The climate simulations were performed in two resolutions, 4.5 km and 1.3 km, which allows to investigate the impact of the resolution on the extreme indices. The significance of the change signals was tested by Wilcoxon-Mann-Whitney tests (Wilks, 2011) but the limited sample size of 10 years and the rareness of extremes may however introduce noise to the tests. Thus they are only used as indicators whether a change signal is large compared to the inter-seasonal variability.

Precipitation indices

A complete overview of the extreme indices is given in Tab. 7.2. The values represent domain-averaged values that have been used to perform the Wilcoxon-Mann-Whitney tests. In addition, the variabilities are shown as box-and-whisker plots in Fig. 7.7.

The longest cumulative dry day (CDD) period, a measure of potential meteorological drought, lasts in summer about 22 days in the spatial average with respect to CCLM4.5c20. Until 2041–2050 there is only a slight increase but until 2091–2100 the maximum CDD rises strongly to 28 days and 37 days in the 4.5 km and 1.3 km model, respectively. The average CDD index per summer is significant at the significance level $\alpha = 0.1$ as well. In addition, the variability increases and the most extreme CDD reach up to 60 days (Fig. 7.7). The CDD index of CCLM1.3c20 is systematically higher than for the 4.5 km model resolution but the relative changes are similar. A comparable behaviour was found for the number of drought events (CDDn). But although the change until 2091–2100 is highly significant, the absolute change amounts to only 1 additional dry period per summer. In the present climate there are about 3-4

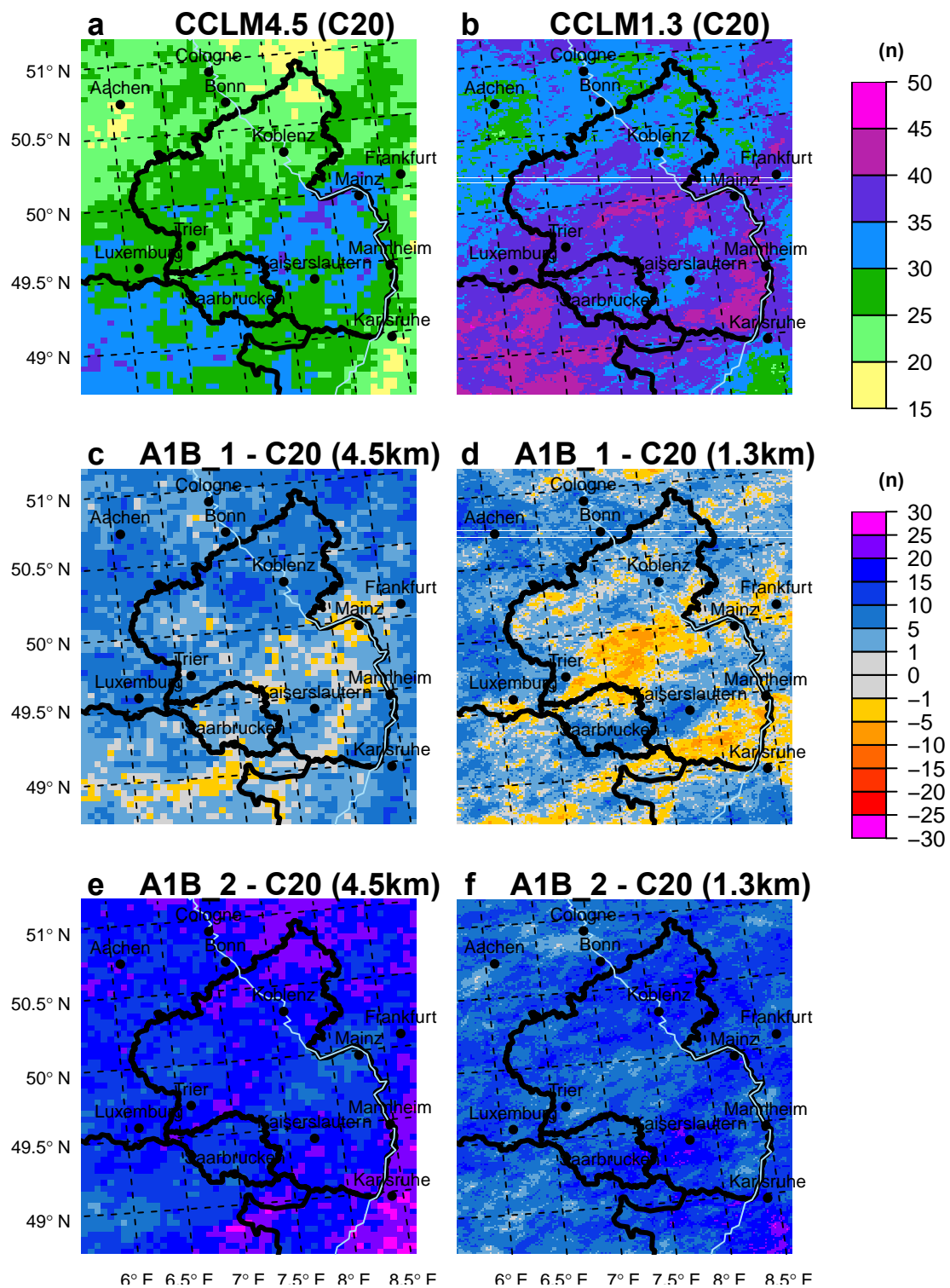


Figure 7.8 Total number of consecutive dry day (CDDn) events in summer (JJA) for present climate 1991–2000 in a) CCLM4.5c20, b) CCLM1.3c20 and the changes until 2041–2050 c) CCLM4.5a1b1 - CCLM4.5c20, d) CCLM1.3a1b1 - CCLM1.3c20 and until 2091–2100 e) CCLM4.5a1b2 - CCLM4.5c20, f) CCLM1.3a1b2 - CCLM1.3c20.

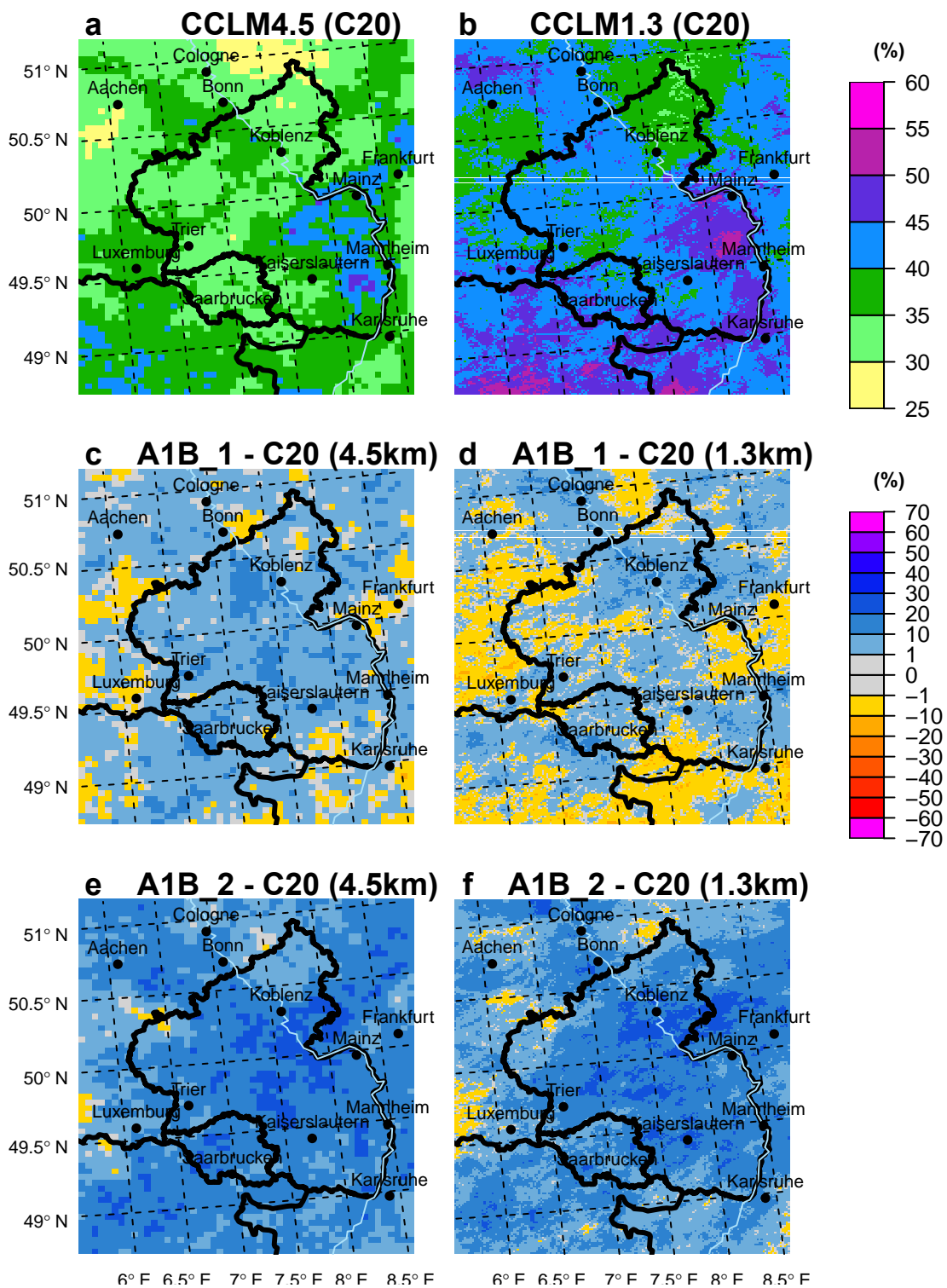


Figure 7.9 Percentage of precipitation days $> 95\%$ on the total summer precipitation (R95PTOT) for present climate 1991–2000 in a) CCLM4.5c20, b) CCLM1.3c20 and the changes until 2041–2050 c) CCLM4.5a1b1 - CCLM4.5c20, d) CCLM1.3a1b1 - CCLM1.3c20 and until 2091–2100 e) CCLM4.5a1b2 - CCLM4.5c20, f) CCLM1.3a1b2 - CCLM1.3c20.

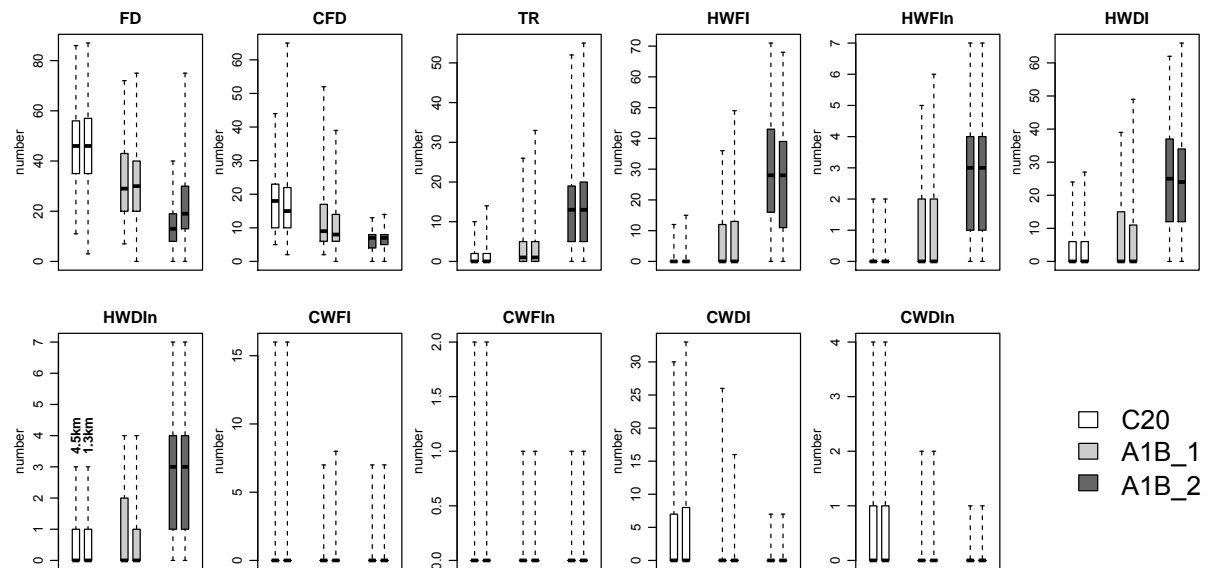


Figure 7.10 Box-and-whisker plots of the ETCCDI temperature indices for the present climate 1991–2000 (C20) the near future 2041–2050 (A1B_1) and the end of the century 2091–2100 (A1B_2). Top row: summer (JJA) and bottom row: winter (DJF). The bold black horizontal lines mark the medians, the thin horizontal lines span the inter-quartile ranges (25 % to 95 % percentiles) and the vertical dashed lines represents extremes. Each boxplot consists of 10 values per pixel ($4.5 \text{ km}=51 \cdot 62 \cdot 10 = 31620$; $1.3 \text{ km}=172 \cdot 208 \cdot 10 = 357760$).

events per summer and until the end of the century this increases to 4-5 events. The systematic higher CDD and CDDn indices are due to the dry-bias in the CCLM1.3c20, which affects predominately the low intensities and thus the days with precipitation values close to the threshold of 1 mm. The total numbers of CDDn in 10 years are shown in Fig. 7.8a-b. The most frequent CDD events were simulated by the CCLM4.5c20 to occur in France, along the Nahe river (roughly in a band from Kaiserslautern to Mainz), which enters the river Rhine in the Mainz Basin and a second maximum extends along the Rhine Valley from west of Karlsruhe northwards into the Mainz Basin. This pattern is realistic and the frequencies match the observations (Fig. D.1b-c) (note that the values in this figure are based on 8 years only). The change signal in the near future is only about ± 1 event per summer (Fig. 7.8c-d) but significant in the spatial mean. Stronger increases of up to +3 events were simulated, e.g. in the Black Forest, by the CCLM4.5a1b2 until 2091-2100 while the signal is lower for CCLM1.3. In the spatial mean, CDDn rises in the whole domain significantly. The maximum change of +30 days is simulated by the CCLM4.5 in the Black Forest and the Rhine Valley. The overall increase is somewhat lower in the CCLM1.3 except for very restricted areas.

The average number of summer days with R20 events is about 2 in all scenarios. Thus there is no significant change visible in the future scenarios. More events were simulated for mountainous areas, e.g. > 40 days in 10 years for the *Hohes Venn* in the Eifel or for the Black Forest. However, it is likely that the events in the Black Forest were overestimated (not shown). The boxplots in Fig. 7.7 show systematically less days in the CCLM1.3. The domain-average of the highest precipitation sum over 5 consecutive days (RX5DAY) is about 55-61 mm in the present climate. Although this index increases by about 10 mm in the far future, it is not significant. Similarly, the minor increase of the number of events with > 50 mm per 5 days (RX5DAYn) is not significant as well. Considerably larger values for RX5DAY (> 200 mm) were simulated in connection with single events, though (see Fig. D.2).

The proportion of extreme events, i.e. the percentage of daily precipitation sums > 95 % percentile

on the total (R95PTOT) is about 35 – 43 % in the spatial mean but takes considerable higher values in the Rhine Valley, France and southern RLP (Fig. 7.9c-d). Here the percentages are > 40 % in the CCLM4.5 but > 50 % in the CCLM1.3. These regions roughly coincide with those of the highest CDDn (compare Fig. 7.9a-b with Fig. 7.8a-b). At the end of the century, the percentages rise significantly, about +12 % in the spatial mean. Regions with the maximum consistency in increase in both future periods are the Rhine Valley between Mainz, Mannheim, Karlsruhe and regions in northern France. The inter-annual variability is increasing as well (Fig. 7.7) towards the end of the century. Overall the proportion of extremes in the CCLM1.3 model is about 5 % higher than in the coarser model. This indicates that more extremes are produced due to the higher resolution and the changes in the convection-parametrization.

In winter, less persistent and less frequent dry periods were simulated. The simulations do not indicate that this will change in the future projections. Likewise, R20, RX5DAY and RX5DAYn remain constant. The only significant signal was found for R95PTOT for which the percentages of 25 – 26 % change about +4 % until 2041-2050 and +7 % until 2091-2100.

Temperature indices

The temperature indices show more distinctive change signals than the precipitation indices. Overall, cold indices reduce in the future projections and hot indices increase, but on different rates. Consider for example the domain-averaged number of frost days (FD) per winter in Tab. 7.2. The 2 m minimum temperature in the simulation of the present climate (C20) falls below 0 °C in winter at 46 days on average. In 2041–2050 (A1B_1) the FD reduces significantly to 32 days per winter and until 2091–2100, the FD reduces even further to only 14 days. This corresponds to a linear reduction of FD of about –3 days per 10 years. However, this trend is uncertain because the simulations are not continuous between 2000 and 2100. The spatial distribution of the FD and the climate change signals are shown in Fig. 7.11. Apparently the FD index depends on the orography. Up to 75 days year⁻¹ were simulated in the mountains but only slightly more than 30 days year⁻¹ in the river valleys. By comparing Fig. 7.11a-b the added value of the 1.3 km resolution is obvious. The higher resolved orography causes a much more heterogeneous spatial distribution of frost days.

The climate change signal until 2041–2050 (Fig. 7.11c-d) is between -10 to -20 days, which confirms the linear estimate (–3 days in about 10 years). The signal in the CCLM1.3 model is again more heterogeneous, although the pattern does not differ very much, and the reduction is slightly stronger in the Westerwald. This indicates that the FD will reduce faster in areas with high orography than in lower areas and this is even more evident from the CCS for the period 2091–2100 (Fig. 7.11e-f). The mountain regions experience 35 less frost days than in the present climate. According to the 1.3 km simulation, the frost days reduce particularly at ridges and peaks. Thus the higher resolutions adds new information to the results.

The CCS of consecutive frost days (CFD) reduces significantly as well and the spatial variability of CFD decreases simultaneously in the A1B_2 period (Fig. 7.10). This reduces the vertical gradients. According to these results, not only do the absolute numbers of frost days decrease but in addition frost periods will be less persistent. Likewise, the number of cold-waves (CWDIn) reduces from about one every 2 years to one every 10 years (Fig. 7.13) and the duration (CWDI) decreases considerably as well. However, the statistics are very uncertain because nearly no cold-wave did occur even in the C20 simulation.

In contrast, all hot indices increase significantly at a more non-linear rate. In the spatial average, tropical nights (TR20) were simulated to occur about once every summer in the present climate, thrice until the mid of the century but 13 tropical nights per summer were simulated for the end of the century. The spatial variability does increase as well, which can be deduced from the spread of the boxplots in Fig. 7.10. The reason for this becomes apparent from Fig. 7.12. The most frequent TR20 were simulated

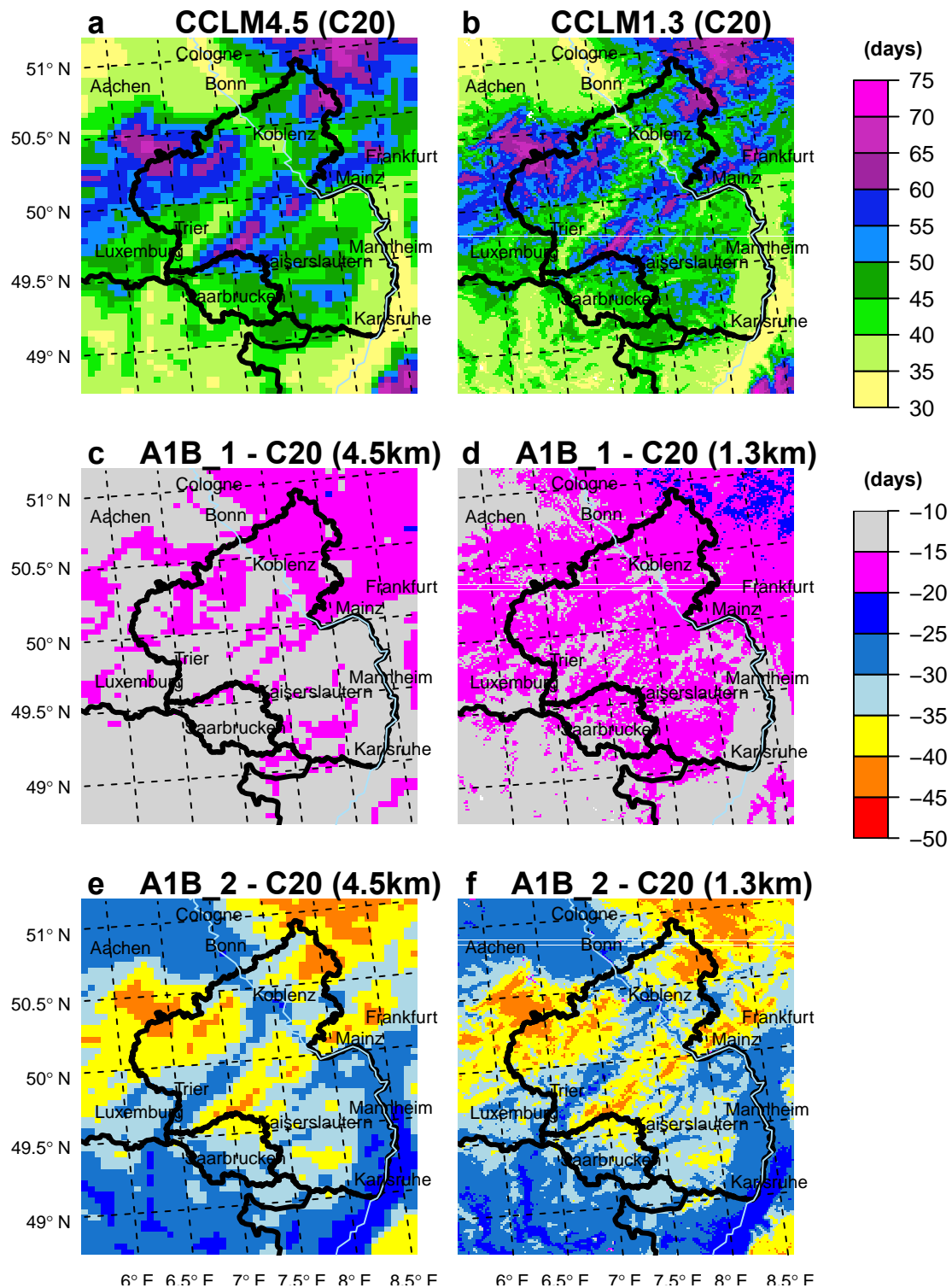


Figure 7.11 Average number of frost days (FD) in winter (DJF) for present climate 1991–2000 in a) CCLM4.5c20, b) CCLM1.3c20 and the changes until 2041–2050 c) CCLM4.5a1b1 - CCLM4.5c20, d) CCLM1.3a1b1 - CCLM1.3c20 and until 2091–2100 e) CCLM4.5a1b2 - CCLM4.5c20, f) CCLM1.3a1b2 - CCLM1.3c20.

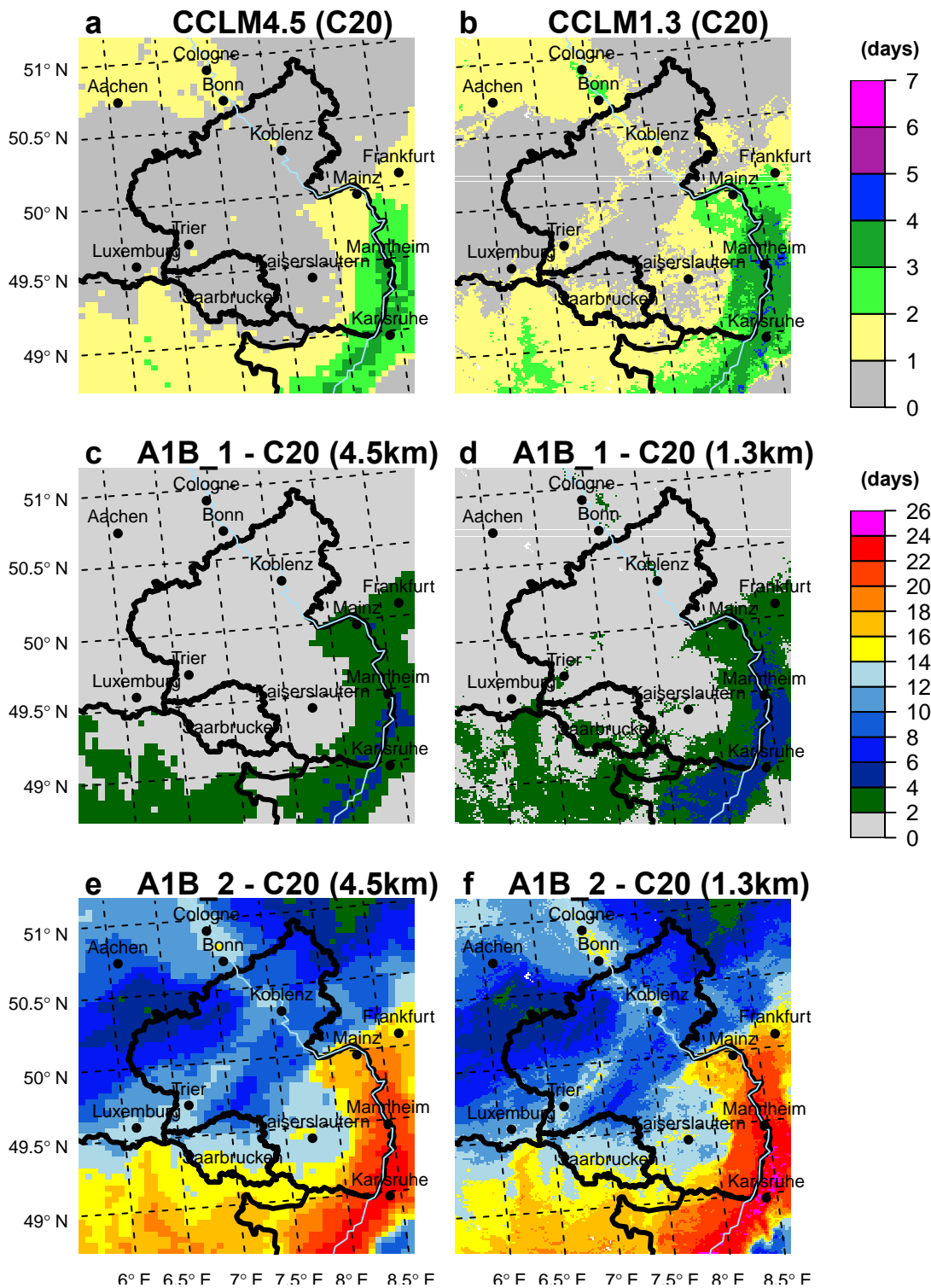


Figure 7.12 Average number of tropical nights (TR20) in summer (JJA) for present climate 1991–2000 in a) CCLM4.5c20, b) CCLM1.3c20 and the changes until 2041–2050 c) CCLM4.5a1b1 - CCLM4.5c20, d) CCLM1.3a1b1 - CCLM1.3c20 and until 2091–2100 e) CCLM4.5a1b2 - CCLM4.5c20, f) CCLM1.3a1b2 - CCLM1.3c20.

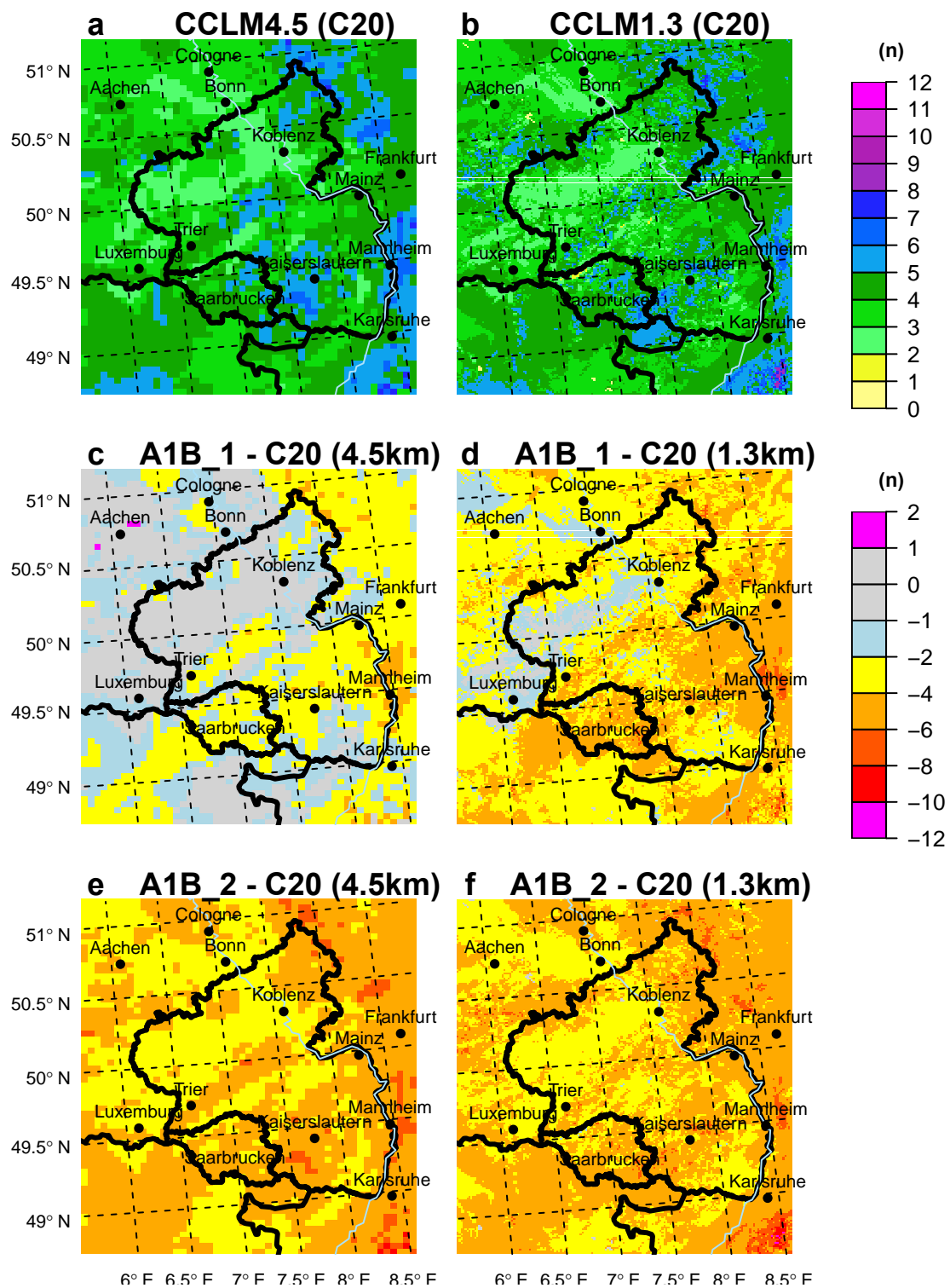


Figure 7.13 Total number of cold-waves (CWDIn) in winter (DJF) for present climate 1991–2000 in a) CCLM4.5c20, b) CCLM1.3c20 and the changes until 2041–2050 c) CCLM4.5a1b1 - CCLM4.5c20, d) CCLM1.3a1b1 - CCLM1.3c20 and until 2091–2100 e) CCLM4.5a1b2 - CCLM4.5c20, f) CCLM1.3a1b2 - CCLM1.3c20.

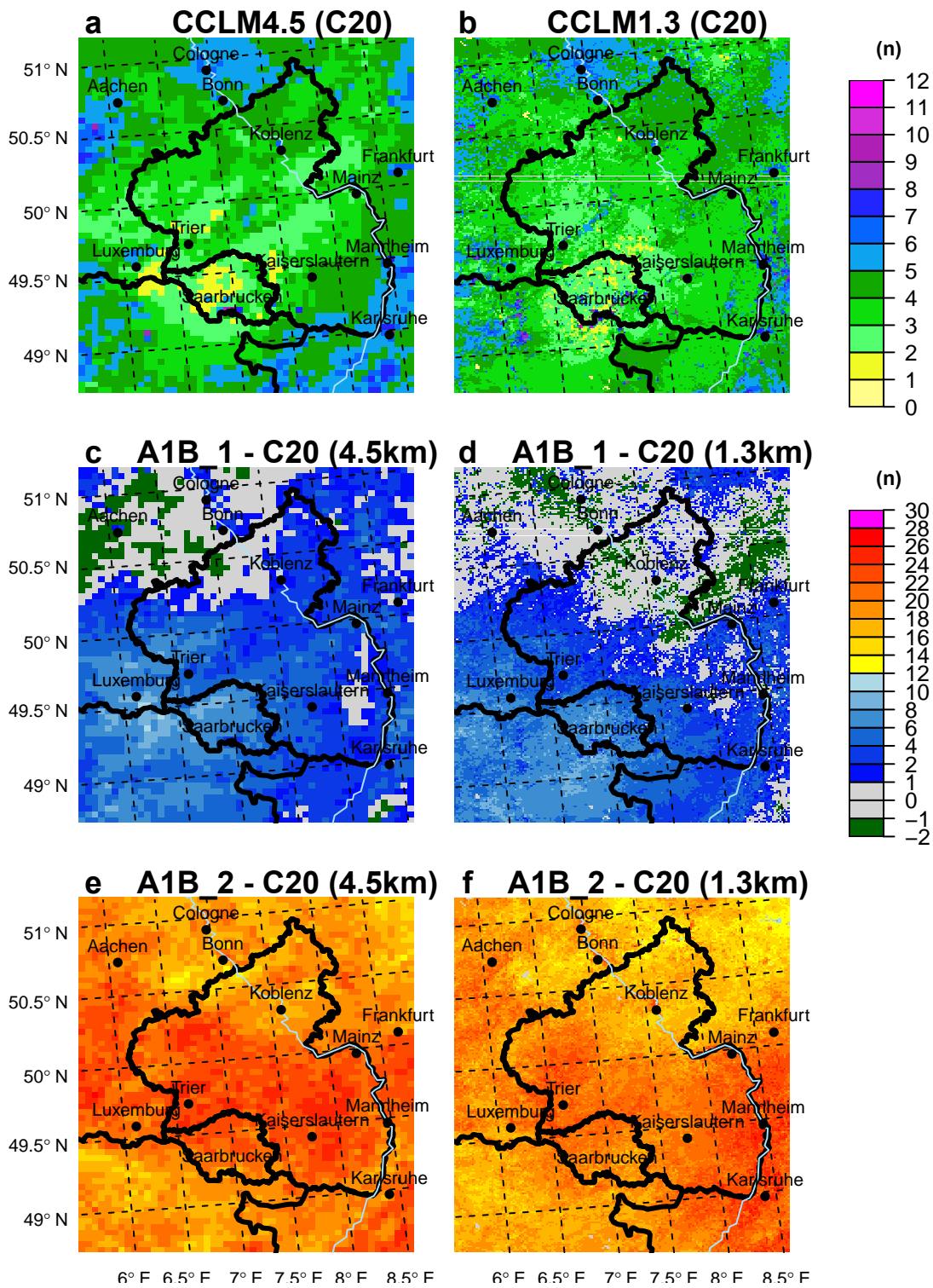


Figure 7.14 Total number of heat-waves (HWDIn) in summer (JJA) for present climate 1991–2000 in a) CCLM4.5c20, b) CCLM1.3c20 and the changes until 2041–2050 c) CCLM4.5a1b1 - CCLM4.5c20, d) CCLM1.3a1b1 - CCLM1.3c20 and until 2091–2100 e) CCLM4.5a1b2 - CCLM4.5c20, f) CCLM1.3a1b2 - CCLM1.3c20.

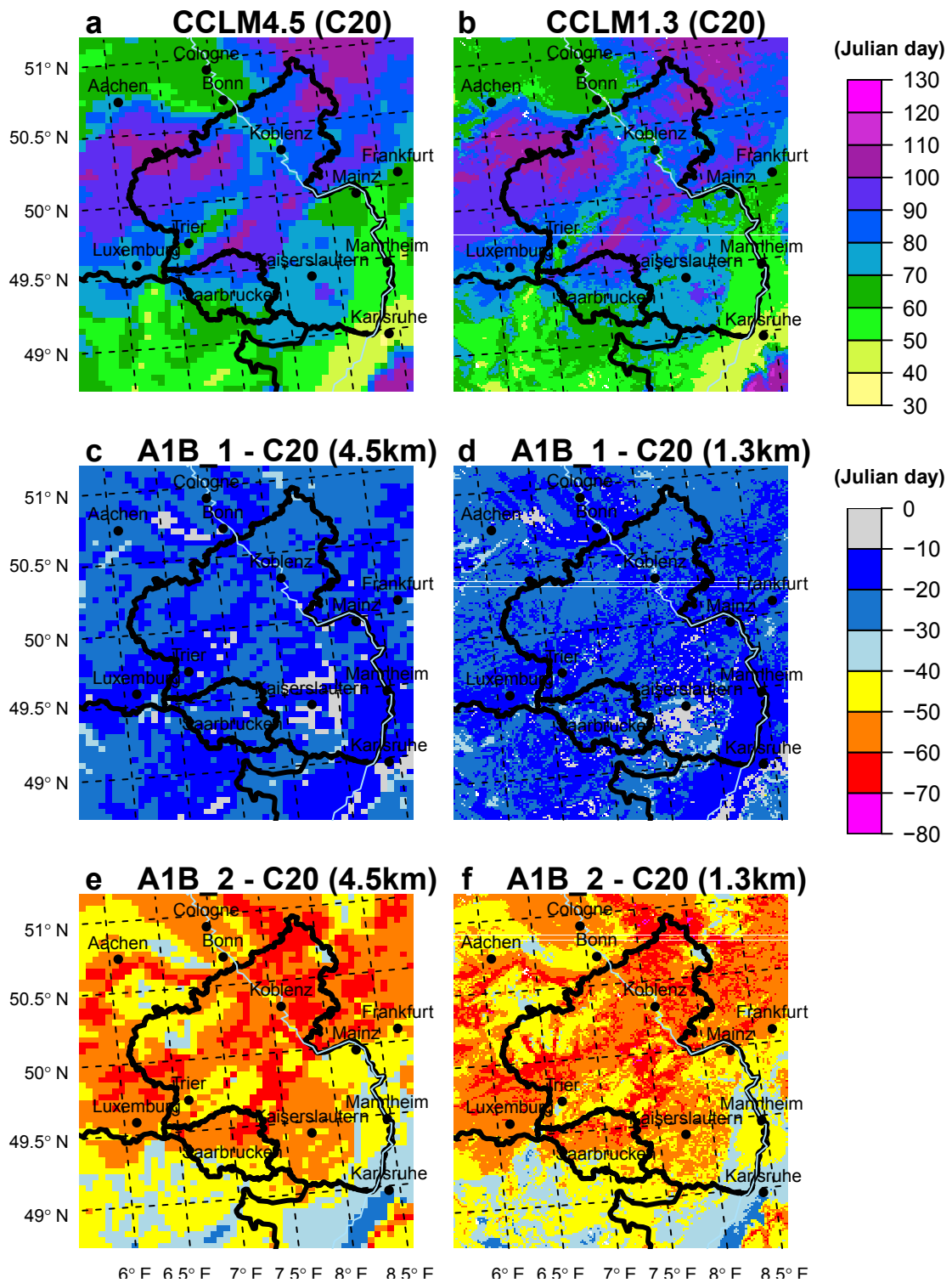


Figure 7.15 Onset of growing season (GSL.d) for present climate 1991–2000 in a) CCLM4.5c20, b) CCLM1.3c20 and the changes until 2041–2050 c) CCLM4.5a1b1 - CCLM4.5c20, d) CCLM1.3a1b1 - CCLM1.3c20 and until 2091–2100 e) CCLM4.5a1b2 - CCLM4.5c20, f) CCLM1.3a1b2 - CCLM1.3c20.

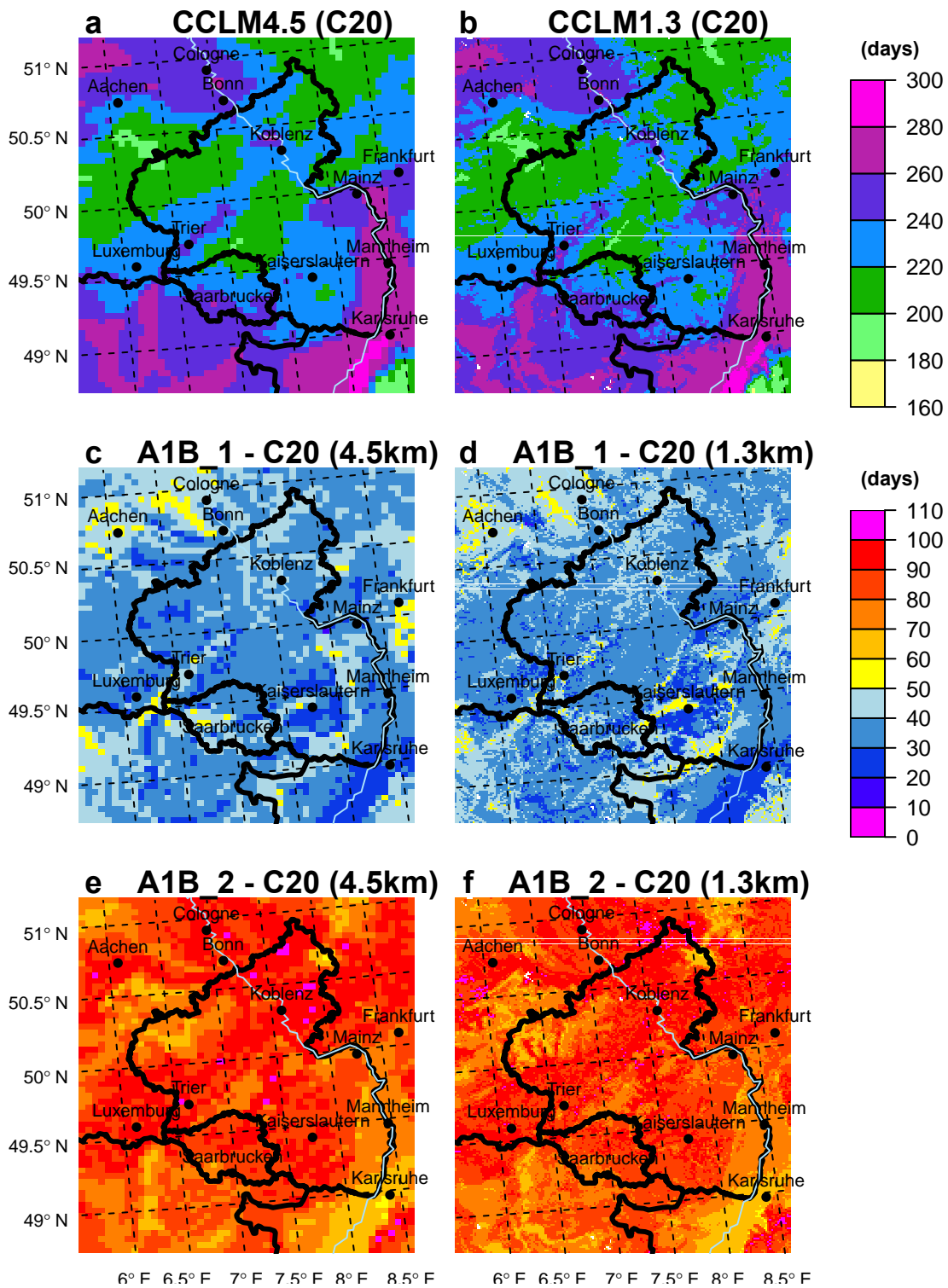


Figure 7.16 Length of the growing season (GSL) for present climate 1991–2000 in a) CCLM4.5c20, b) CCLM1.3c20 and the changes until 2041–2050 c) CCLM4.5a1b1 - CCLM4.5c20, d) CCLM1.3a1b1 - CCLM1.3c20 and until 2091–2100 e) CCLM4.5a1b2 - CCLM4.5c20, f) CCLM1.3a1b2 - CCLM1.3c20.

in the Rhine Valley; up to 4 per summer close to the river in the CCLM4.5c20 and none in the higher areas of RLP and Luxemburg while in the CCLM1.3c20 there are spots with 5 tropical nights per summer and at least 1 along river valleys, such as the Moselle in RLP. Until 2041–2050, the frequency rises to 2–6 additional tropical nights per summer; in particular in the Rhine Valley and France where already the most TR20 were simulated. This signal increases then considerably stronger in the period 2091–2100. The strongest increase is again in the Rhine Valley with up to +26 tropical nights, for instance close to Karlsruhe and further south. Even in the mountains occur some TR20, but the increase in the river valleys is much larger. This leads to an increase of the spatial variability as well. Although the CCLM4.5 simulates a similar pattern, individual features such as the spots with extreme signals are not present in the coarser model. This difference can be up to 10 tropical nights less than in the CCLM4.5 for the same location. Similarly the number of warm-spells (HWFIn) and heat-waves (HWDIn) increase in a non-linear way (Tab. 7.2 and Fig. 7.10).

Heat-waves were simulated preferentially in areas with the most frequent tropical nights (Fig. 7.15a-b). These are the lowlands in the Rhine Valley and the Cologne Basin but also France, Luxemburg and Belgium. In these regions were about 5-7 heat-waves simulated within 10 summers by the CCLM4.5c20. In the CCLM1.3c20 the heat-waves are much more localized and occur even more frequent in some places, in particular between Mannheim and Karlsruhe and in France. Effects of the higher resolution are again visible, for example in the Moselle Valley (from Trier to the northeast) there were only 2 heat-waves simulated in the CCLM4.5c20 but in the CCLM1.3c20 the number rises to 7.

Although the HWDIn decreases for some regions in the period 2041–2050, the index increases by up to 10 additional heat-waves for most of the region (Fig. 7.15c-d), which roughly corresponds to 1 additional heat-wave per 5 years. Until 2091–2100 this number rises non-linearly. In general, there is no strong dependency on the orography in the CCLM4.5 (Fig. 7.15e) but in the CCLM1.3 the correlation is slightly stronger, for instance the centre of the maximum increase is mainly located in the Rhine-Valley. Furthermore, the duration of warm-spells (HWFI) and of heat-waves (HWDI) increase non-linear as well (Fig. 7.10). For instance, the duration of heat-waves rises from 13 days in the spatial mean of the present climate to 46 days at the end of the century. Likewise, the HWFI index increases from 8 to 60 days.

Cold-wave durations (CWDI) decrease from 16 to 6 days in the future projections and their frequency (CWDIn) reduces significantly from about 1 every two winters to 1 every ten winters. No significant reduction was found for cold-spells (CWFIn), though. In general, cold-spells or waves were less frequently simulated in the CCLM than heat-waves or warm-spells.

Finally, the growing season length (GSL.l) and its onset (GSL.d) exhibit both a significant change. In the present climate, the growing season begins on the 18th of March (77 Julian day) and lasts for about 236 days until the 9th of November. In the near future projection, the GSL begins earlier at the 26th of February and lasts until the 27th of November and in 2091–2100 the onset is even at 27th of January and ends at 13th of December. Thus the growing season lasts almost the whole year. However, the GSL index is very simplistic and only based on temperature with no inclusion of plant-physiological characteristics so that the index rather indicates a potential than an actual growing season.

7.1.3 Impact of the bias-correction

The impact of the bias-correction on the precipitation indices is analysed in this section. For this purpose, all indices have been recalculated for the bias-corrected model data for the same region as in section 7.1.1. The correction of the daily precipitation fields was performed with the methods eQM, gQM and GQM (see chapter 5) and the REGNIE data set. At every grid box, a transfer function was calibrated with 10-year long time series of CCLM and REGNIE and then applied to the same time period.

First, the indices of the C20 runs were compared to those of INTERMET and REGNIE in section 7.1.1. Because of the different time lengths of the observations and the CCLM simulations, the indices

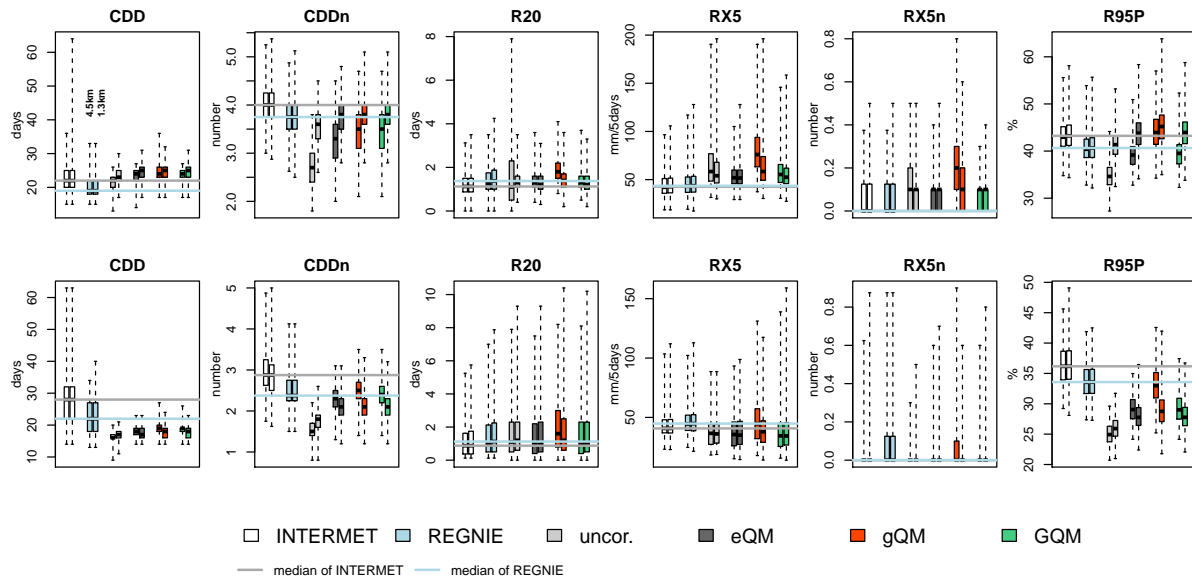


Figure 7.17 Box-and-whisker plots of the ETCCDI precipitation indices for the present climate (1991–2000, C20) calculated from INTERMET, REGNIE, uncorrected CCLM4.5c20/CCLM1.3c20 and from the bias-corrected model data after applying eQM, gQM and GQM. Top row: summer (JJA) and bottom row: winter (DJF). The bold black horizontal lines mark the medians, the thin horizontal lines span the inter-quartile ranges (25 % to 95 % percentiles) and the vertical dashed lines represents extremes.

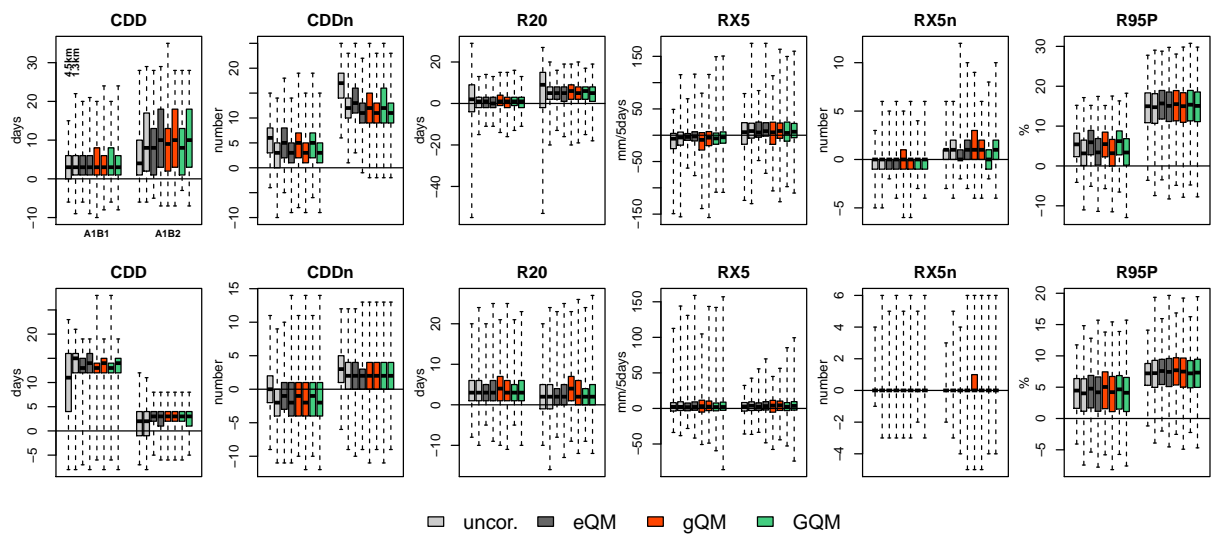


Figure 7.18 Box-and-whisker plots of the climate change signals of the ETCCDI precipitation indices. The signals are shown as differences: 2041–2050 (A1B_1) - 1991–2000 (C20) and 2091–2050 (A1B_2) - 1991–2000 (C20) which is indicated in the first plot in the top row. Top row: summer (JJA) and bottom row: winter (DJF). The bold black horizontal lines mark the median, the thin horizontal lines span the inter-quartile range (25 % to 95 % percentiles) and the vertical dashed lines represent extremes.

have been normalized at every grid box in the summer/winter season, respectively. The results are shown as boxplots in Fig. 7.17.

The dry indices CDD and CDDn increase after the bias-correction. The average CDD per summer is about 22 in INTERMET and 19 in REGNIE. The dry periods from the model were already longer than in REGNIE before the bias-correction so that they increase even further due to the dry day correction. The CDD in CCLM1.3c20 is slightly higher as in the 4.5 km model because of the reduced drizzle error. After the BC the difference of the CDD between the simulations is reduced. In contrast, the number of consecutive dry days (CDDn) were underestimated by both CCLM, in particular by the CCLM4.5c20. After the correction, the CDDn from the CCLM1.3c20 fits well to the observations but it is still too low in the CCLM4.5c20, although the frequency of CDDn increases. The different BC methods do not cause large deviations for CDD and CDDn but eQM produced slightly lower indices compared to the parametric methods. In winter, the impact of the BC methods is similar to that in summer, except that CDD is too low in CCLM1.3c20 as well. Thus the length of dry periods in winter is still too short after the corrections.

The number of very heavy precipitation events (R20) is especially altered in the CCLM4.5c20. The uncorrected model simulated much more variability in the domain than observed and a peak of 8 R20 events per summer for some grid boxes. In contrast, the CCLM1.3c20 simulated R20 values close to the observations and a comparable spatial variability. The bias-corrections with eQM and GQM especially improve CCLM4.5 but gQM introduces too many days with > 20 mm. This error is because of the wrong adjustment of the tail, which was already found in chapter 5. Likewise, the R20 index in winter matches the observations and the BC did not affect the frequency except for gQM.

An improvement was found for the RX5 index in summer. The overestimation of the uncorrected model was reduced by eQM and GQM, whereas gQM worsened the bias and even increased the frequency RX5n. In winter, the BC had nearly no effect on the indices.

Finally, the proportion of the extreme precipitation on the total sum (R95PTOT) was about 5 % too low in the CCLM4.5c20, whereas the percentage in the CCLM1.3c20 did match the observations. After the bias-correction, the percentages were systematically shifted to higher values for both models. Because of this positive effect on the R95PTOT index from CCLM4.5c20, the values are now closer to REGNIE. On the other hand, those of the CCLM1.3c20 are slightly too high compared to REGNIE after the correction but fit well to INTERMET. In winter, the percentage is largely underestimated by the CCLM by about -8 %. Again the BC increases the R95PTOT index but the values are still too low. Only gQM alters the R95PTOT from the 4.5 km model so that they are in the range of REGNIE. However, this is again due to the above mentioned positive bias of gQM and thus a correction for the wrong reason.

The next step was to investigate the impact on the climate change signals. Because the absolute values were changed by the bias-correction, the question arises whether relative changes were affected as well. In Fig. 7.18 the impacts on the CCS are shown as spatial boxplots. Overall the effect on the CCS is rather low, although stronger effects were found for some indices of the CCLM4.5. For instance, the increase of CDD in summer until 2091–2100 (A1B_2) of the uncorrected CCLM4.5 is considerably enhanced (median and variability) by the BC, whereas the signal of CCLM1.3 is nearly unaffected. Or in contrast, the variability of the CDD signal in winter until 2041–2050 (A1B_1) is distinctively reduced. Similar effects were found for CDDn and R20. Larger deviations between the methods occur for RX5n in summer. Here GQM reduces the signal until the period 2091–2100 to 0 in the median and in winter the correction with gQM produces a slight positive deviation. However, RX5 events occurred very rarely in all time periods and the signal is not significant at the 95 % level.

In summary, a large impact of the bias-correction on the absolute values of extreme indices was found. Dry indices were increased by the corrections, the intensity of consecutive precipitation extremes were reduced in summer but the proportion of extremes on the total amount of precipitation was increased

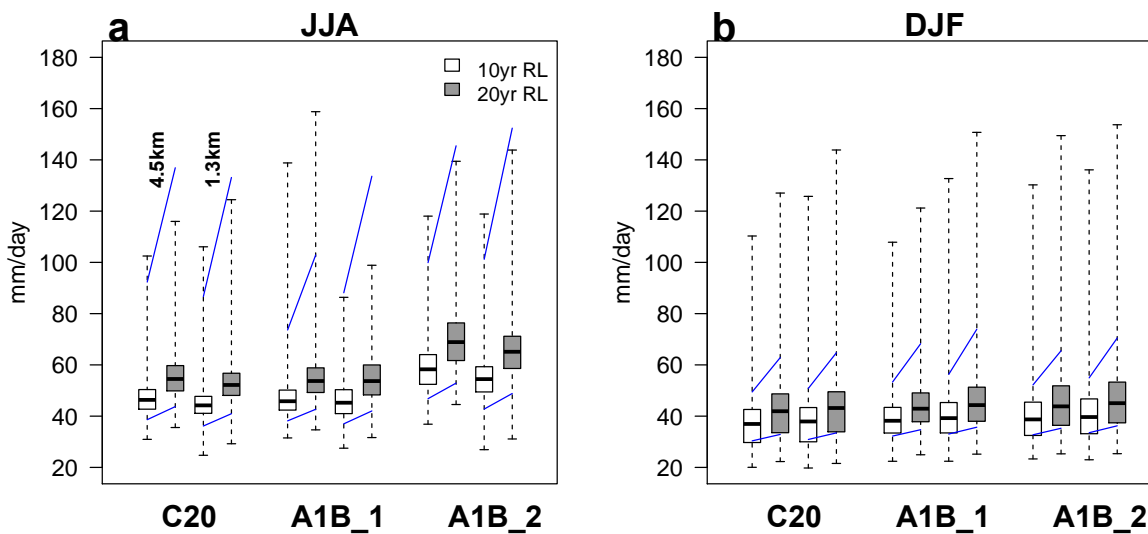


Figure 7.19 Spatial box-and-whisker plots of daily precipitation return levels (RL) for the 10- and 20-year return periods in a) summer (JJA) and b) winter (DJF). The RL were estimated from POT-models of the CCLM4.5 and CCLM1.3 climate simulations and are shown for the present climate (1991–2000, C20), the near-future (2041–2050, A1B_1) and for the end of the century (2091–2100, A1B_2). The blue lines mark the mean of the spatial 95 % confidence interval from the profile-loglikelihood method.

in summer and winter. The relative values remained nearly constant and thus the effect on the climate change signals is only minor.

7.2 Return levels from the climate simulations

The results of the POT models, which have been fitted to daily precipitation, minimum and maximum of 2 m temperature, are analysed in the following. The time periods of the simulations were 1991–2000 (C20), 2041–2050 (A1B_1) and 2091–2100 (A1B_2).

7.2.1 Precipitation

The estimates of the 10- and 20-year precipitation return levels (RL) are shown in Fig. 7.19 as boxplots for the winter and summer season. In summer, the inter-quartile range (25 – 75 % percentile) of the 10-year return levels is between 40 and 50 mm day⁻¹ for both models, although marginally smaller values were estimated for CCLM1.3c20. The most extreme values range between 100 – 130 mm day⁻¹ in the Black Forest (Fig. 7.21a-b). The 20-year return levels rely completely on extrapolation of the POT models and are about 10 mm higher in average. The 95 % confidence intervals from the profile log-likelihood method were calculated at every grid box and spatially averaged (blue lines in Fig. 7.20). They are markedly asymmetric, which confirms the use of the profile log-likelihood method, and increase for the 20-year return levels. In particular, the upper confidence level increases much more steeply than the lower so that the uncertainty widens with longer return periods. It is noteworthy to mention that the

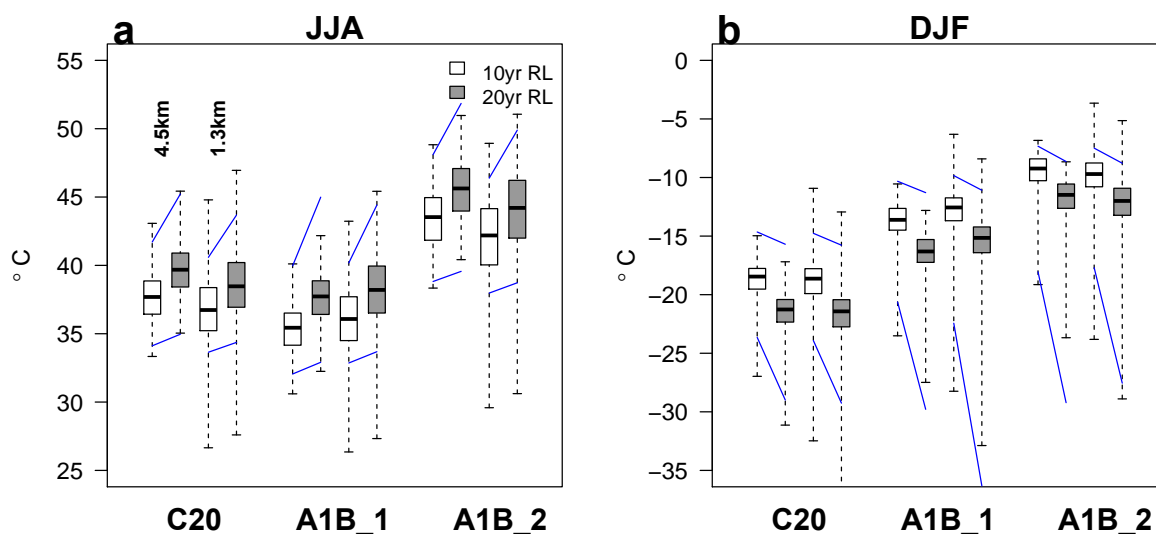


Figure 7.20 Spatial box-and-whisker plots of the 10- and 20-year return levels of a) 2 m maximum temperature in summer (JJA) and b) 2 m minimum temperature in winter (DJF). The RL were estimated from POT-models of the CCLM4.5 and CCLM1.3 climate simulations and are shown for the present climate (1991–2000, C20), the near-future (2041–2050, A1B_1) and for the end of the century (2091–2100, A1B_2). The blue lines mark the mean of the spatial 95% confidence interval from the profile-loglikelihood method.

estimates from the CCLM1.3c20 are slightly lower than from the CCLM4.5c20, although the deviations are small.

In the near future (2041–2050) there are no distinctive changes in the return levels, except that the confidence interval narrows in the CCLM4.5 estimates. Nevertheless, individual grid boxes yield 10-year return levels of up to 160 mm. Fig. 7.21c-d shows the spatial variability with changes of about ± 10 mm for most of the grid boxes. Exceptions are the region between the cities Mannheim, Mainz and Frankfurt with a decrease of about -30 mm (Fig. 7.21c) or the Cologne Basin with an increase of $+30$ mm. This means that the absolute values of the RL change locally but their variability within the domain remains constant, as seen from the box plots.

Pronounced increases were retrieved for the projections of the far future (2091–2100). The median of the 10-year estimates increases to 60 mm day^{-1} and the inter-quartile range spans from 50 to 70 mm day^{-1} . Similarly, the 20-year return levels increase by about $+15$ mm. This indicates intensified summer precipitation extremes in this domain (Fig. 7.21e-f). The change signal is positive in most of the domain with local maxima of more than $+30$ mm. However, there are some differences between the 4.5 km and the 1.3 km model. For instance, the signal in the region between Aachen, Cologne and Bonn is negative in the CCLM1.3 with up to -30 mm while in the 4.5 km model this region receives a positive change.

The winter return levels are about 10 mm lower than in summer and the spread of the confidence intervals is much smaller. Thus the estimation uncertainty of winter extremes is reduced, even though there is a non-linear dependence on the return period. It is, however, likely that the RL were underestimated because of the bias in winter extremes, which was also found in section 4.1.5. No changes of the return levels were found for the future climate and the spatial variability remained nearly constant.

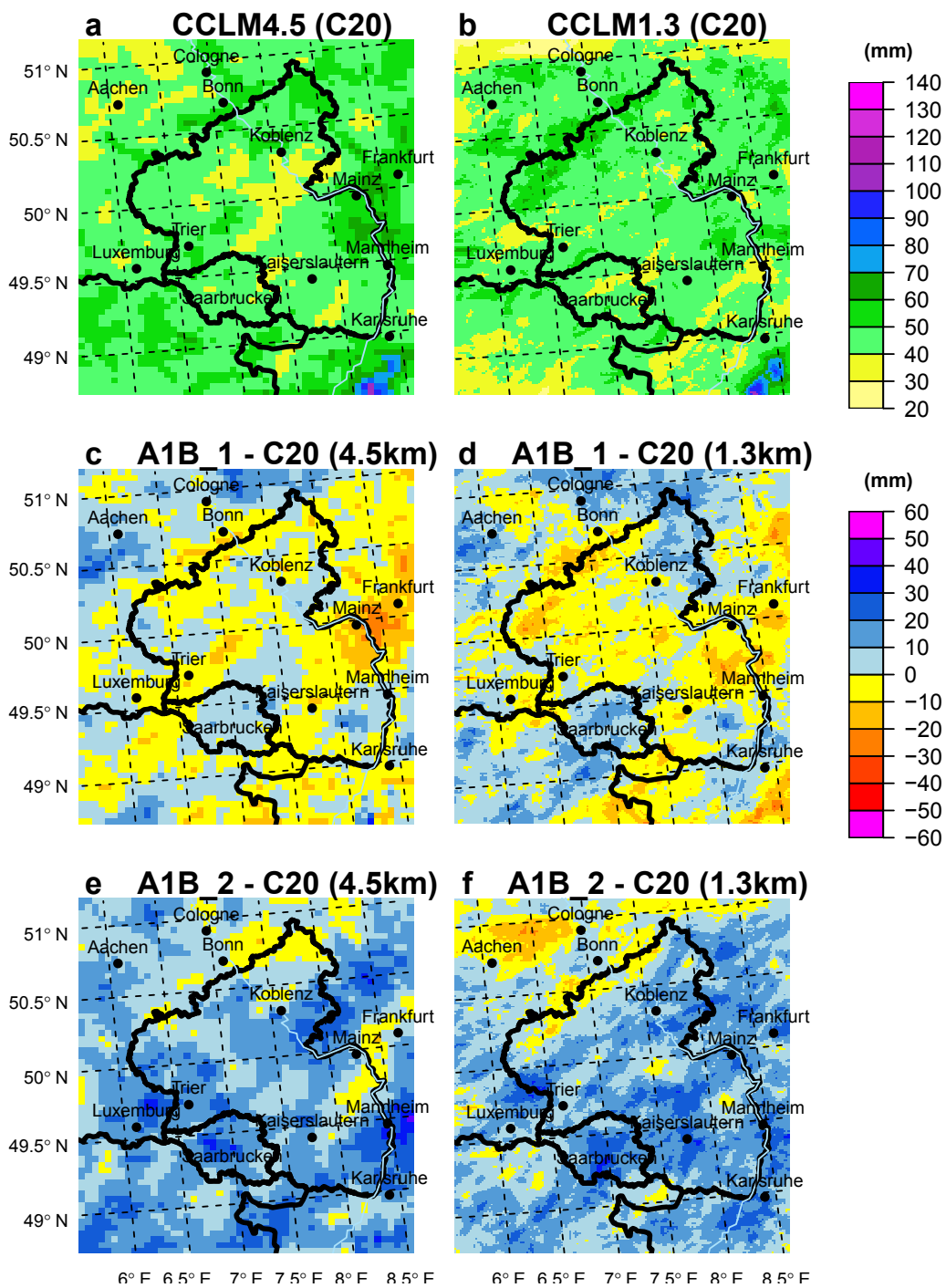


Figure 7.21 10-year return level estimates of daily precipitation in summer (JJA) from the POT models in C20 for a) CCLM4.5 and b) CCLM1.3. The differences until 2041–2050 are shown for the c) 4.5 km and d) 1.3 km model and until 2091–2100 in e) and f), respectively.

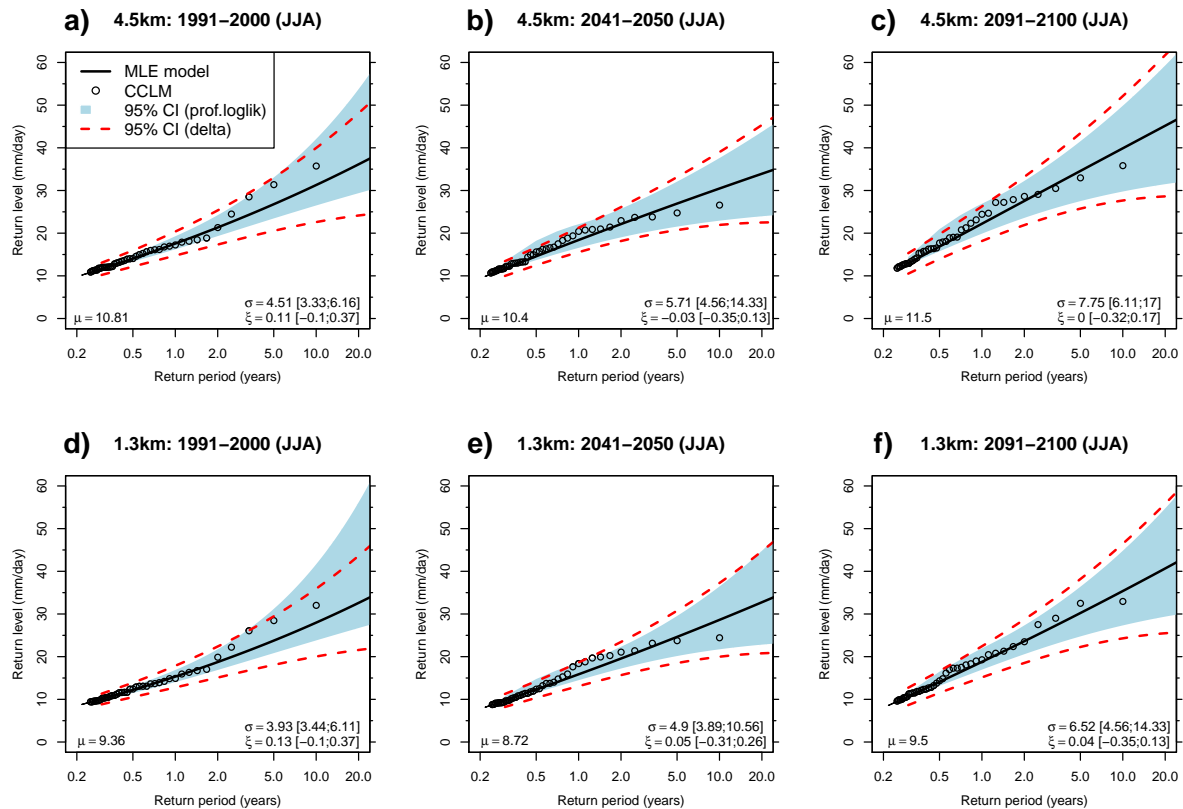


Figure 7.22 Return level plots of daily precipitation in summer (JJA). Top row: CCLM4.5 for the period a) 1991–2000 (C20), b) 2041–2050 (A1B_1) and c) 2091–2100 (A1B_2). Bottom row: CCLM1.3 for d) C20, e) A1B_1 and f) A1B_2. The simulated extremes and their assigned annuality are plotted as points, the estimated return levels from the maximum-likelihood fit of the POT models (MLE model) are shown as solid black lines. In addition, the 95 % confidence intervals (CI) from the profile log-likelihood (blue shading) and from the delta method (dashed red lines) are shown. The thresholds (μ) and the maximum-likelihood estimates of the scale (σ) and shape (ξ) parameters are given along with the 95 % confidence intervals from the profile log-likelihood method.

The main pattern of the RL is similar in both models, but some deviations were caused by single events that may have too much weight in the fitting of the POT models, because they rely on 10-year time periods. As a consequence, if a severe event takes a different trajectory in the 1.3 km model then it is likely that local differences in the return value estimations are caused.

Complete return level curves for domain-averaged daily precipitation in summer are shown in Fig. 7.22. The empirical RL of the CCLM are plotted as point symbols along with their annuality. Because the simulated time series cover 10 years, the largest simulated extreme value is assigned with a return period of 10 years. The estimations from the POT models are shown as black solid lines. The shape parameter (ξ) is positive in the present climate (1991–2000) (Fig. 7.22a,d), which causes the quantile curve to bend upwards. However, the RL estimates of return periods > 2 years are lower in the POT models than from the simulations. In the projection of 2041–2050 (Fig. 7.22b,e), there is nearly no change in the quantile curve, but the uncertainty of the CI reduces. However, the empirical RL with return periods ≤ 2 years increase in intensity but reduce at longer return periods. Thus there is nearly no effect on the POT model estimates. In the period 2091–2100 this is different because now the extremes

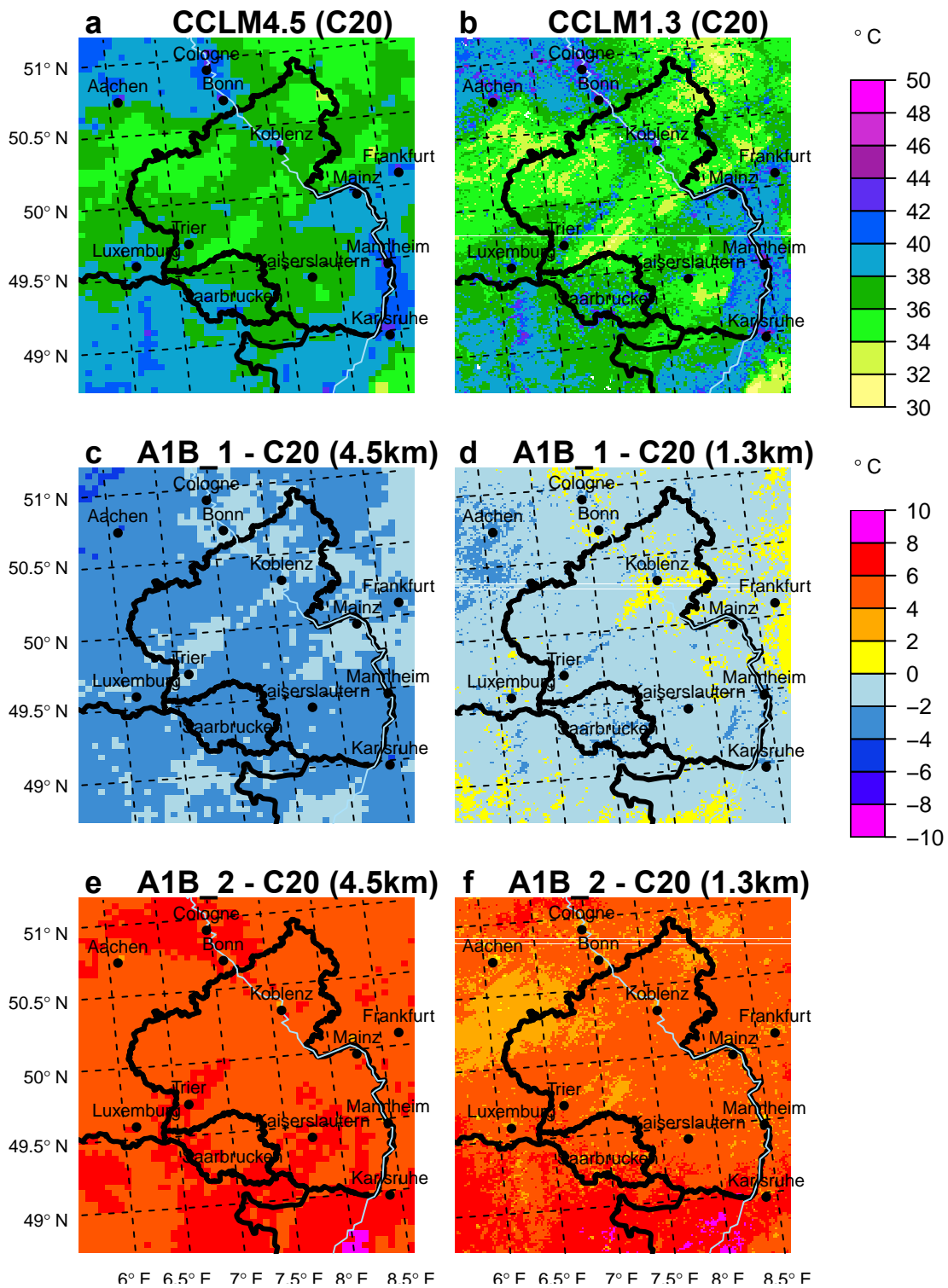


Figure 7.23 10-year return level estimates of maximum 2 m temperature in summer (JJA) from the POT models in C20 for a) CCLM4.5 and b) CCLM1.3. The differences until 2041–2050 are shown for the c) 4.5 km and d) 1.3 km model and until 2091–2100 in e) and f), respectively.

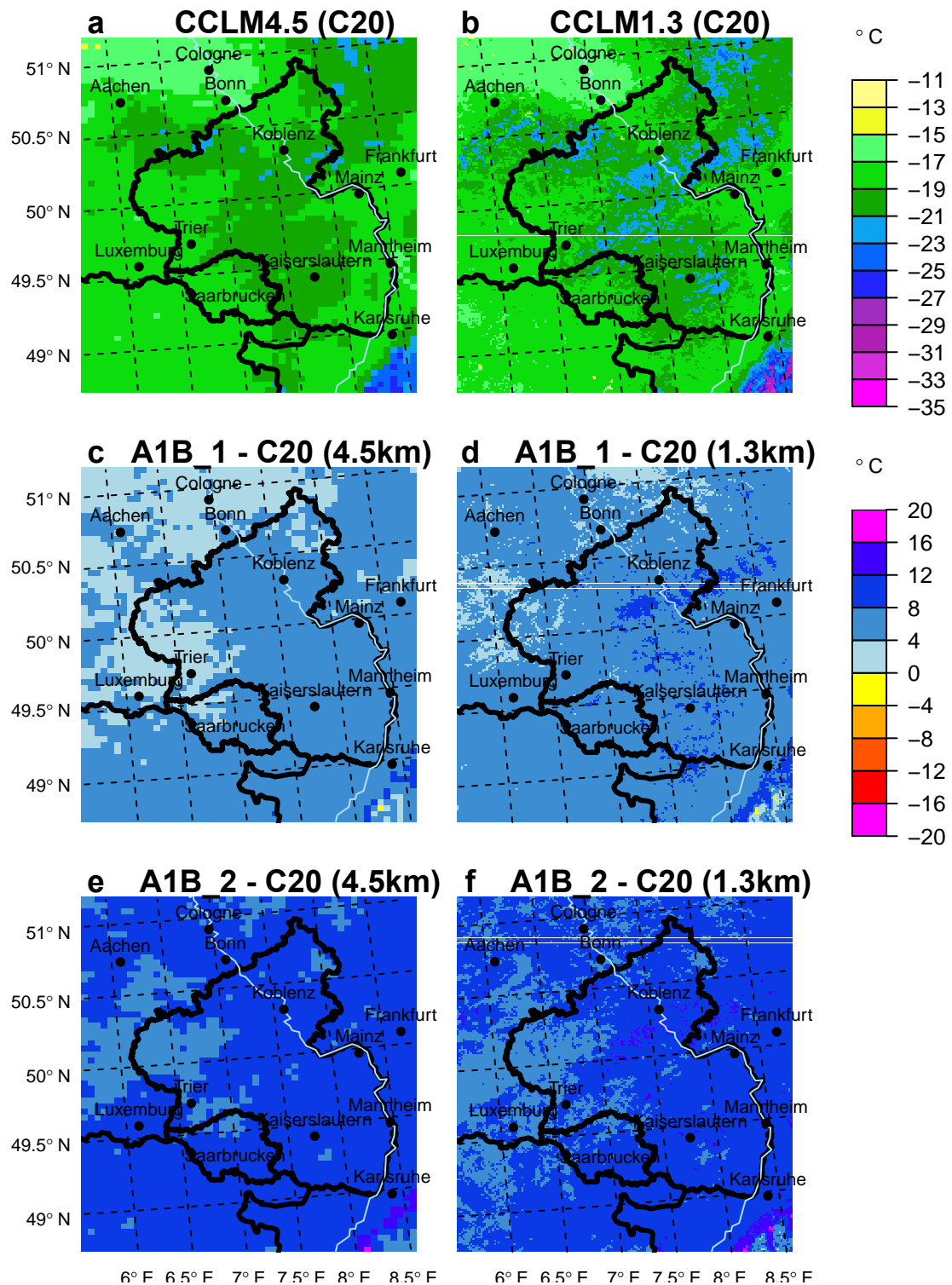


Figure 7.24 10-year return level estimates of minimum 2m temperature in winter (DJF) from the POT models in C20 for a) CCLM4.5 and b) CCLM1.3. The differences until 2041–2050 are shown for the c) 4.5 km and d) 1.3 km model and until 2091–2100 in e) and f), respectively.

with return periods > 2 years have the same intensity as in C20 and the intensities still increase on shorter return periods, which causes the POT model estimates to increase overall. More intense return levels for periods ≤ 2 years in the future also implies that RL with these reoccurrence intervals in the present climate are assigned with shorter return periods in the future projection. As a consequence these extremes will occur more often. For instance, the 2-year return level is about 20 mm day^{-1} in C20 but this RL corresponds to a return period of only 0.75 years in 2091–2100.

This indicates why the R20 index in the section before did not increase significantly but the proportion of extreme values on the total precipitation did. It is the proportion of extremes with return periods ≤ 2 years which causes the increase in the R95PTOT index. This increase is larger in the CCLM4.5 and corresponds to the finding that the R95PTOT index increases stronger the CCLM4.5 as well.

The 95 % confidence interval (CI) from the profile log-likelihood method is markedly asymmetric and increases non-linear with the return period as mentioned before. Thus the more rare an event, the larger becomes the interval which contains the true value to 95 %. For a comparison the CI from the delta-method is plotted as well. Large differences occur for the return levels in the C20 period. Here the CI from the delta method contains lower values than in the CI from the profile log-likelihood. The differences are, however, smaller in the future projections and the CI from the profile log-likelihood is in all cases more narrow, although the uncertainty of the σ parameter increases.

Large uncertainties were also found for the parameter estimation. For instance, in the 4.5 km (C20) POT-model, the maximum-likelihood estimate of the scale parameter is $\sigma = 4.51 [3.33; 6.16]$ and of the shape parameter $\xi = 0.11 [-0.1; 0.37]$. By comparing the confidence intervals, the uncertainty of the shape parameter is much larger than for the scale parameter. Interestingly, the CIs of both parameters increase in the future scenarios.

7.2.2 Maximum and minimum temperature

The estimated 10- and 20-year return levels of daily 2 m maximum temperature (T_{\max}) in summer are shown in Fig. 7.20a and for 2 m minimum temperature (T_{\min}) in winter in Fig. 7.20b. In the C20 period the inter-quartile range of the 10-year return levels covers a range between $36.5 - 39^\circ\text{C}$ in the CCLM4.5 and between $35 - 38^\circ\text{C}$ in the CCLM1.3 with medians of 37.5°C and 36.5°C , respectively. The 20-year return levels are approximately 2°C higher in the CCLM4.5 and 1.5°C higher in the CCLM1.3. Although the estimates for the A1B_1 period are $1 - 2.5^\circ\text{C}$ lower than in C20 they increase about 6°C in A1B_2. Similarly, the spatial variability increases, in particular in the 1.3 km simulation. This indicates that the warming affects regions with already high return levels stronger than regions with lower return levels.

This can be seen in Fig. 7.23. In the present climate (Fig. 7.23a-b) the highest 10-year return levels were estimated in the Rhine Valley, in the Cologne and Mainz Basin and in northern France. The maxima reach about 44°C in these areas, in particular close to large cities as for instance Karlsruhe, Mannheim or Cologne. This is even more pronounced in the field of the CCLM1.3 (Fig. 7.23b). There is a strongly increased spatial heterogeneity that clearly depends on the orography and adds new and more detailed information to the output of the analysis. A good example is the Moselle Valley between Trier and Koblenz. In the CCLM4.5 higher return values, about $36 - 38^\circ\text{C}$, are indicated but there is no differentiation between dale and hill topography. In contrast, the higher resolution of the CCLM1.3 captures the height differences with considerably higher return levels of $40 - 42^\circ\text{C}$ at the valley bottom and $36 - 38^\circ\text{C}$ at the slopes of the flanking hills. On the other hand, less high return levels were estimated for mountain regions such as the Hunsrück, Eifel or Westerwald. The signal for the period 2041–2050 is almost homogeneously negative in the domain with -2°C to -4°C and 0°C to -2°C for CCLM4.5 and CCLM1.3, respectively. But in the period 2091–2100 the return levels increase by about $+4 - 10^\circ\text{C}$ and $+2 - 10^\circ\text{C}$. The strongest warming according to this analysis occurs in areas

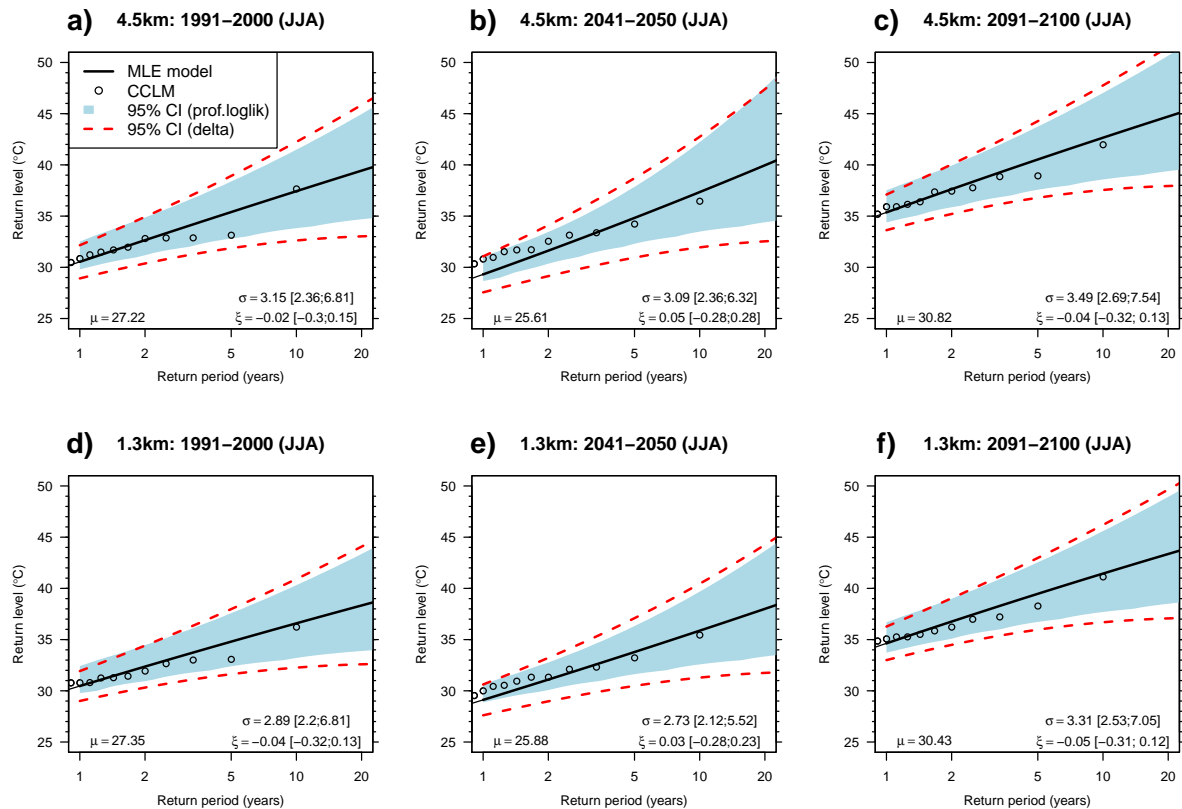


Figure 7.25 Return level plots of 2 m maximum temperature in summer (JJA) for return periods ≥ 1 years. Top row: CCLM4.5 for the period a) 1991–2000 (C20), b) 2041–2050 (A1B_1) and c) 2091–2100 (A1B_2). Bottom row: CCLM1.3 for d) C20, e) A1B_1 and f) A1B_2. The simulated extremes and their assigned annuality are plotted as points, the estimated return levels from the maximum-likelihood fit of the POT models (MLE model) are shown as solid black lines. In addition, the 95 % confidence intervals (CI) from the profile log-likelihood (blue shading) and from the delta method (dashed red lines) are shown. The thresholds (μ) and the maximum-likelihood estimates of the scale (σ) and shape (ξ) parameters are given along with the 95 % confidence intervals from the profile log-likelihood method. Note that a negative shape parameter (ξ) causes the quantile curve (MLE model) to bend upward to more negative temperatures.

with already high return levels, for instance in the south along the river Rhine, northern France or in the Cologne Basin. Less intensive warmings occur in mountain regions with regard to the CCLM1.3, which is not differentiated in the CCLM4.5. Thus again, the results from the high-resolution simulation produces a more orography-dependent signal for temperature extremes.

The maximum-likelihood fits of the POT models to spatial means are shown in Fig. 7.25 along with the CI from the profile log-likelihood and the delta-method. In general, the shape parameter ξ is negative and the CI is asymmetric so that a negative shape is more likely than a positive one, except in the A1B_1 period. A negative ξ implies an asymptotic behaviour of the tail, which is realistic for temperatures. The threshold is about 27 °C in the C20 period, 25 °C in A1B_1 and increases to 31 °C in A1B_2. The fits of both CCLM models are nearly identical except that the return levels in the CCLM1.3 are about 1 °C lower. There is nearly no change in return levels visible until the mid of the century but a strong warming of 5 – 6 °C is found for the A1B_2 estimates, which affects all return periods. The spread of the CI for the return levels and for the parameter estimates remains nearly constant in both future scenarios,

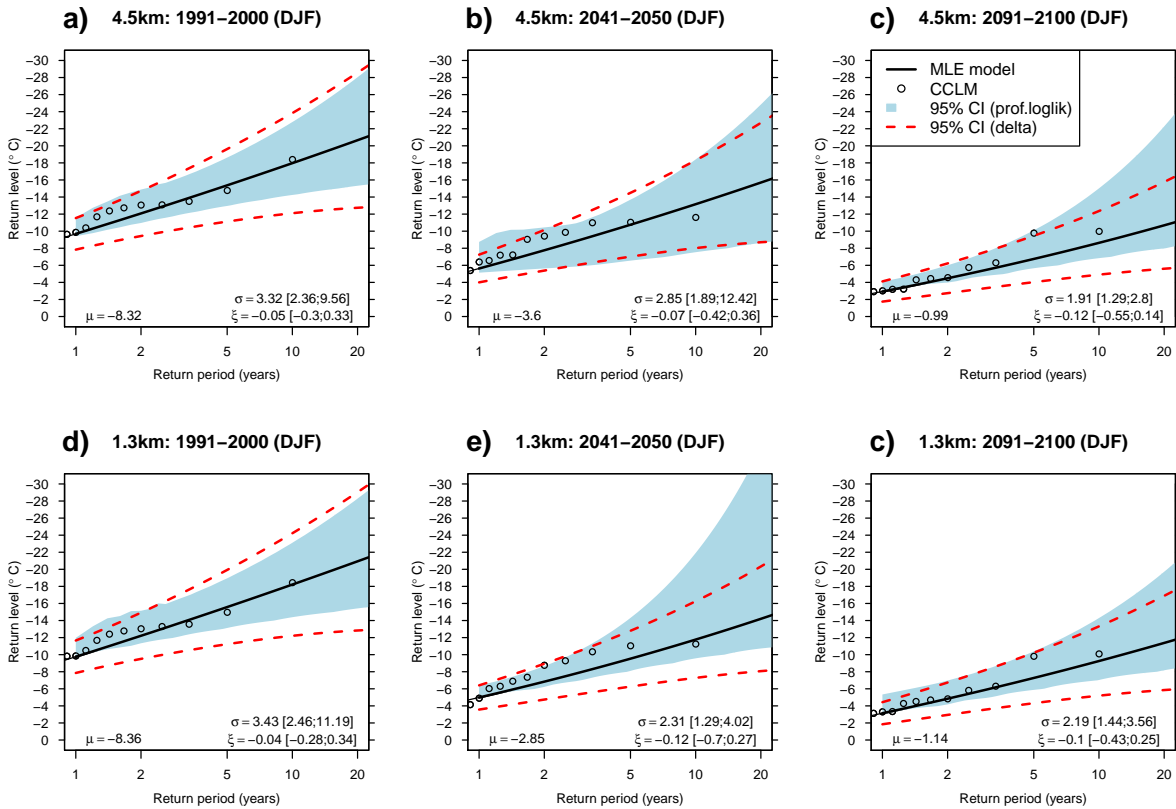


Figure 7.26 Return level plots of 2 m minimum temperature in winter (DJF) for return periods ≥ 1 years. Top row: CCLM4.5 for the period a) 1991–2000 (C20), b) 2041–2050 (A1B_1) and c) 2091–2100 (A1B_2). Bottom row: CCLM1.3 for d) C20, e) A1B_1 and f) A1B_2. The simulated extremes and their assigned annuality are plotted as points, the estimated return levels from the maximum-likelihood fit of the POT models (MLE model) are shown as solid black lines. In addition, the 95 % confidence intervals (CI) from the profile log-likelihood (blue shading) and from the delta method (dashed red lines) are shown. The thresholds (μ) and the maximum-likelihood estimates of the scale (σ) and shape (ξ) parameters are given along with the 95 % confidence intervals from the profile log-likelihood method. Note that a negative shape parameter (ξ) causes the quantile curve (MLE model) to bend upward to more negative temperatures.

although the absolute values differ slightly.

The return levels of the 2 m minimum temperature increase as well in the future scenarios (Fig. 7.20b). The medians of the CCLM4.5 10-year estimates are about -19°C in the present climate, -14°C in A1B_1 and -10°C in A1B_2. The 20-year return levels are 2°C colder but rise similarly. Furthermore, the confidence intervals of the estimates increase distinctively in the A1B_1 and A1B_2 period so that it is likely that the return levels might be even colder than the results from the MLE estimation. The spatial distribution of the 10-year intensities are shown in Fig. 7.24a-b and the changes in the future projections in Fig. 7.24c,d,e,f. The pattern resembles a negative image of the RL field from the maximum temperature. The coldest return levels are estimated in regions with high orography, such as the Black-Forest with $-25 - 27^\circ\text{C}$ and up to -25°C in the CCLM4.5c20 and CCLM1.3c20, respectively. Again, the better resolved orography refines the RL fields. Until 2041–2100 there is a warming signal of $+4 - 8^\circ\text{C}$ for most of the region, the only negative exception are visible for grid boxes in the Black-Forest. Stronger warmings result for instance at the north-western slopes of the Black-Forest. At the end of the

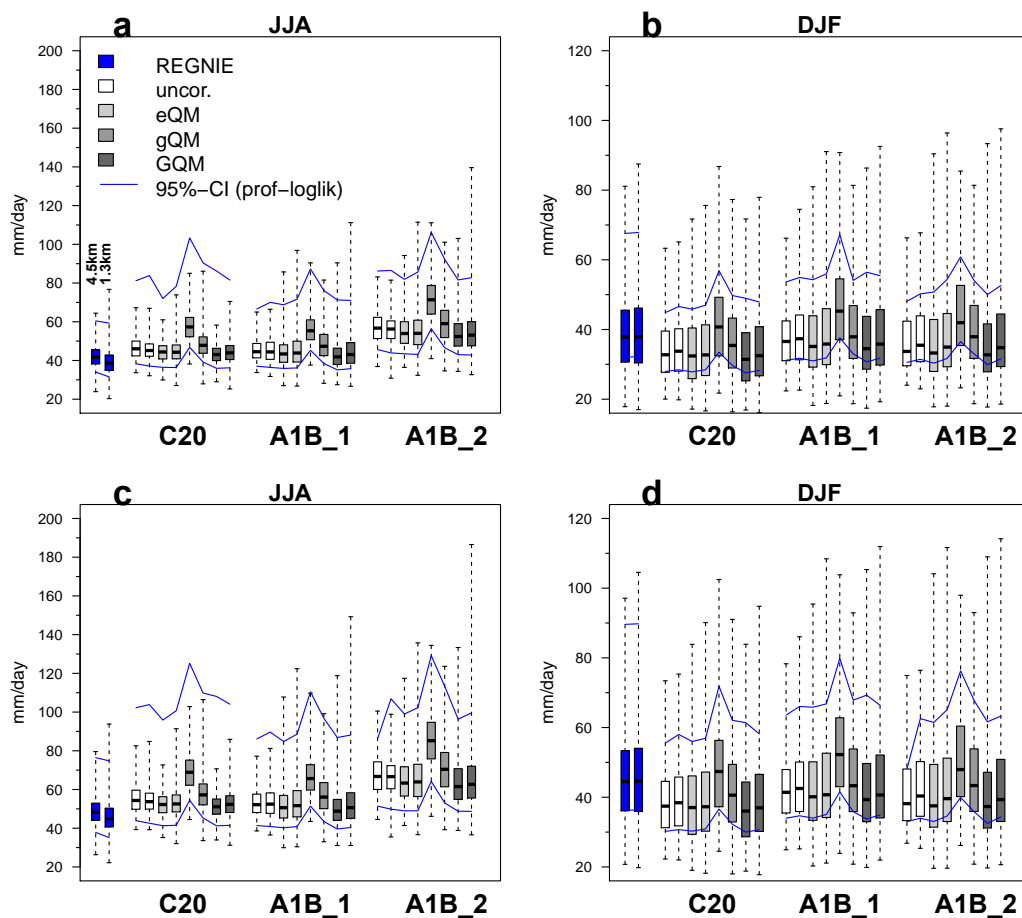


Figure 7.27 Box-and-whisker plots of the 10- and 20-year return levels of daily precipitation from the POT models before and after the bias-corrections. The 10-year return levels are shown in a) for summer (JJA) and b) winter (DJF) and the 20-year return levels are shown in c) and d) for summer and winter, respectively. The blue lines denote the spatial mean of the 95 % confidence interval derived from the profile log-likelihood method.

century, the warming is about $+8 - 12^{\circ}\text{C}$ with more positive values in the same region as mentioned before.

The quantile curves of the fitted POT models show similar findings for periods < 10 years as well. All shape parameters are negative so that the GPD has an upper bound and the increase of the CI in the A1B_1 period is mainly caused by a larger uncertainty of the estimation of ξ .

7.2.3 Impact of the bias-correction on the precipitation return levels

To analyse the impact of the bias-correction on the return level estimation, the POT models have been fitted to the corrected precipitation fields of both CCLM models. Since only the mean temperature has been corrected and not the maxima and minima, the effect on temperature extremes has been omitted from the analysis. The area for which the comparison was made is equal to that in chapter 5.

The spatial distributions of the 10- and 20-year return level estimates from uncorrected and corrected daily precipitation are shown in Fig. 7.27 as boxplots along with the mean 95 % confidence interval of

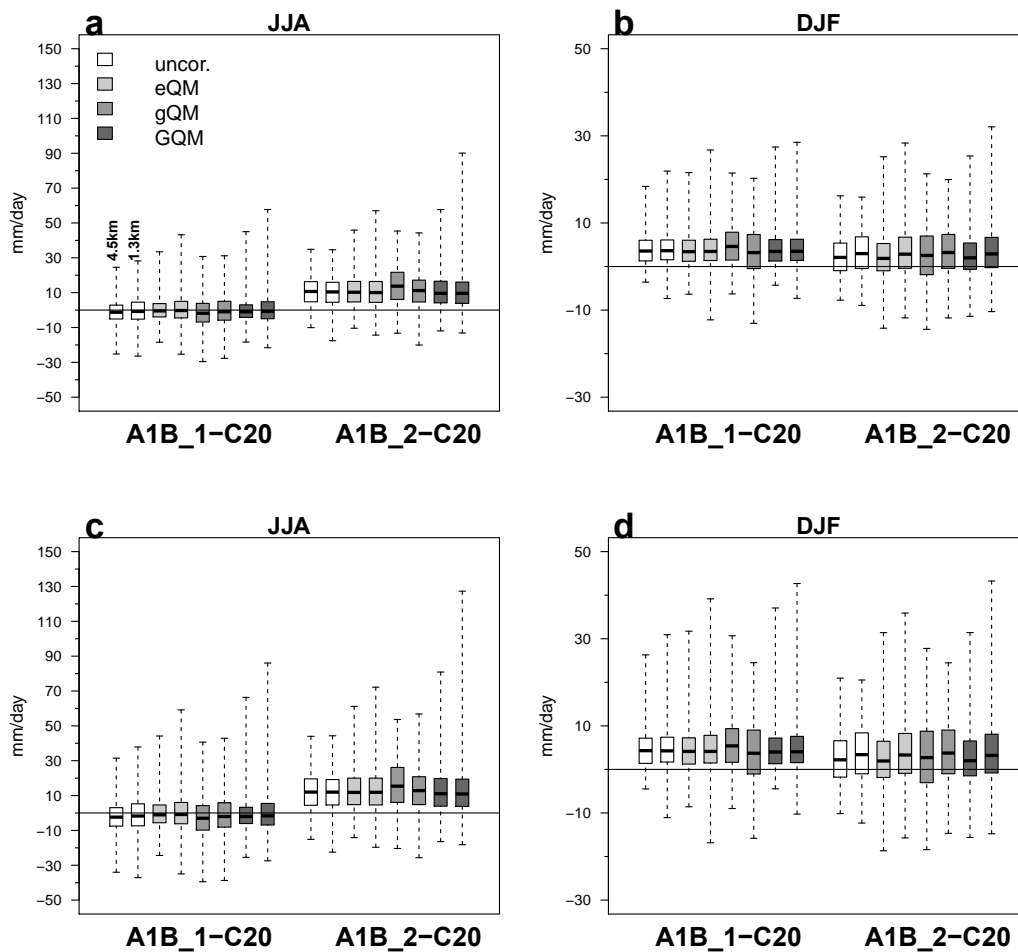


Figure 7.28 Box-and-whisker plots of the 10- and 20-year return levels of daily precipitation for the differences 2041–2050 (A1B_1) - 1991–2000 (C20) and 2091–2100 (A1B_2) - C20 from the POT models before and after the bias-corrections. The 10-year return levels are shown in a) for summer (JJA) and b) winter (DJF) and the 20-year return levels are shown in c) and d) for summer and winter, respectively. The blue lines denote the spatial mean of the 95 % confidence interval derived from the profile log-likelihood method.

the profile log-likelihood method. At some grid boxes the CI estimation failed and therefore pixels at which the upper CI exceeded 200 mm day^{-1} have been removed from the analysis. This was the case for approximately 6.5 – 8 % in the C20 but only for 1 – 2 % in the future projections. The 10-year RL from the uncorrected precipitation fields of CCLM4.5c20 and CCLM1.3c20 are about 5 mm too high in summer compared to REGNIE and their medians are about 46 mm day^{-1} . This overestimation of extremes was also found in the evaluation of the hindcast runs (chapter 4). In addition, the CI is much wider compared to that of REGNIE. The upper confidence level is at $80 - 85 \text{ mm day}^{-1}$, which is 20 mm higher than in REGNIE. After the eQM correction both the return levels and the CI reduce. However, the RL are still slightly too high because transfer functions were calibrated with the complete period and not seasonally (see section 5.2.1 (e.g. Fig. 5.6c)).

A negative impact results from the gQM correction. The RL and the CI show a large positive bias because of the large overestimation of heavy precipitation intensities. In particular in the CCLM4.5c20

estimates, whereas the bias is lower in the CCLM1.3c20 return levels. A much better effect was found for the GQM correction. The RL are comparable to those resulting from the eQM correction. Interestingly, if a correction is applied, the range of the RL differs between the 4.5 km and 1.3 km resolution. This resembles what is also evident in the observations. Thus the corrections further enhance the spatial variability in the CCLM1.3 model. In the future periods, the effects are the same as in the C20 period.

In winter (Fig. 7.27b), the return level estimates of the CCLMs are too low compared to those of REGNIE. This can again be explained by the results from section 5.2.1 (e.g. Fig. 5.6a). The bias-correction did not adjust the underestimation of winter extremes and this underestimation is transferred to the return level analysis. The deviations are about -5 mm (-12%). This indicates that at least a seasonal correction should be considered for extreme value analysis. Only the RL after gQM are close to the observations (see also Fig. 5.6e) but this is due to the systematic shift to positive extremes and therefore correct for the wrong reason. These effects were found for 20-year return levels and for the future periods as well, but even though the absolute values were affected by the bias-corrections, the relative values remain constant (Fig. 7.27), which is consistent with the findings for the extreme indices.

The spatial correlations of the return levels before and after the bias-corrections are shown in Fig. 7.29. The RL of the uncorrected model are not correlated with those of REGNIE in summer but with $r = 0.8$ fairly well in winter. After the correction with eQM and GQM the correlations improve to 0.51 and 0.48 in summer and to 0.88 in winter. In contrast, the correction with gQM enhances the correlation only slightly in summer, although the winter correlation is similarly improved as by the other methods.

These results show that the BC also alters the spatial distribution of RL or extremes. However, the summer correlations did not achieve comparably high values as in winter where already a good agreement was found. This can be explained by the non-seasonal BC. Gutjahr and Heinemann (2013) showed that if a non-seasonal correction is applied and the RL are calculated on an annual basis, then nearly perfect correlations could be achieved. However, for a seasonal analysis the transfer-functions are not flexible enough to capture the summer and winter structure equally well. The same results were found for CCLM4.5c20 and the 20-year RL.

7.3 Summary and discussion

In this chapter the performance of simulating extremes, their change signals in a warming climate and the impact of bias-correction have been analysed with the regional climate model COSMO-CLM. Because of the high horizontal resolutions of 4.5 km and 1.3 km the simulations were performed as 10-year time-slice experiments.

The evaluation of the CCLM in terms of extreme indices yielded good results with some exceptions. In general, the winter precipitation extremes are in a better agreement with the observations than the summer extremes. In particular, the location of extreme precipitation events is much better simulated in the winter season. This is due to the stronger dependency of precipitation on the large-scale dynamics. Moderate correlations were only found for indices R20 and R95PTOT. The worst correlations were found for the length of the consecutive dry days (CDD), both in summer and in winter.

Cold extremes were slightly overestimated by the CCLM, for instance the CCLM simulated 5 frost days more per year than occurred in INTERMET. On the one hand, this may be caused by the cold-bias of the CCLM but on the other hand the most frost days occur in high orographic regions where the spatial density of observational stations is sparse. This may produce uncertain interpolations in these areas and a direct comparison is difficult. Hot extremes, as tropical nights (TR20) were simulated too often in the CCLM compared to INTERMET. However, it was found that the spatial distribution of hot extremes extracted from INTERMET is not realistic, whereas the results of the CCLM models seem much more plausible.

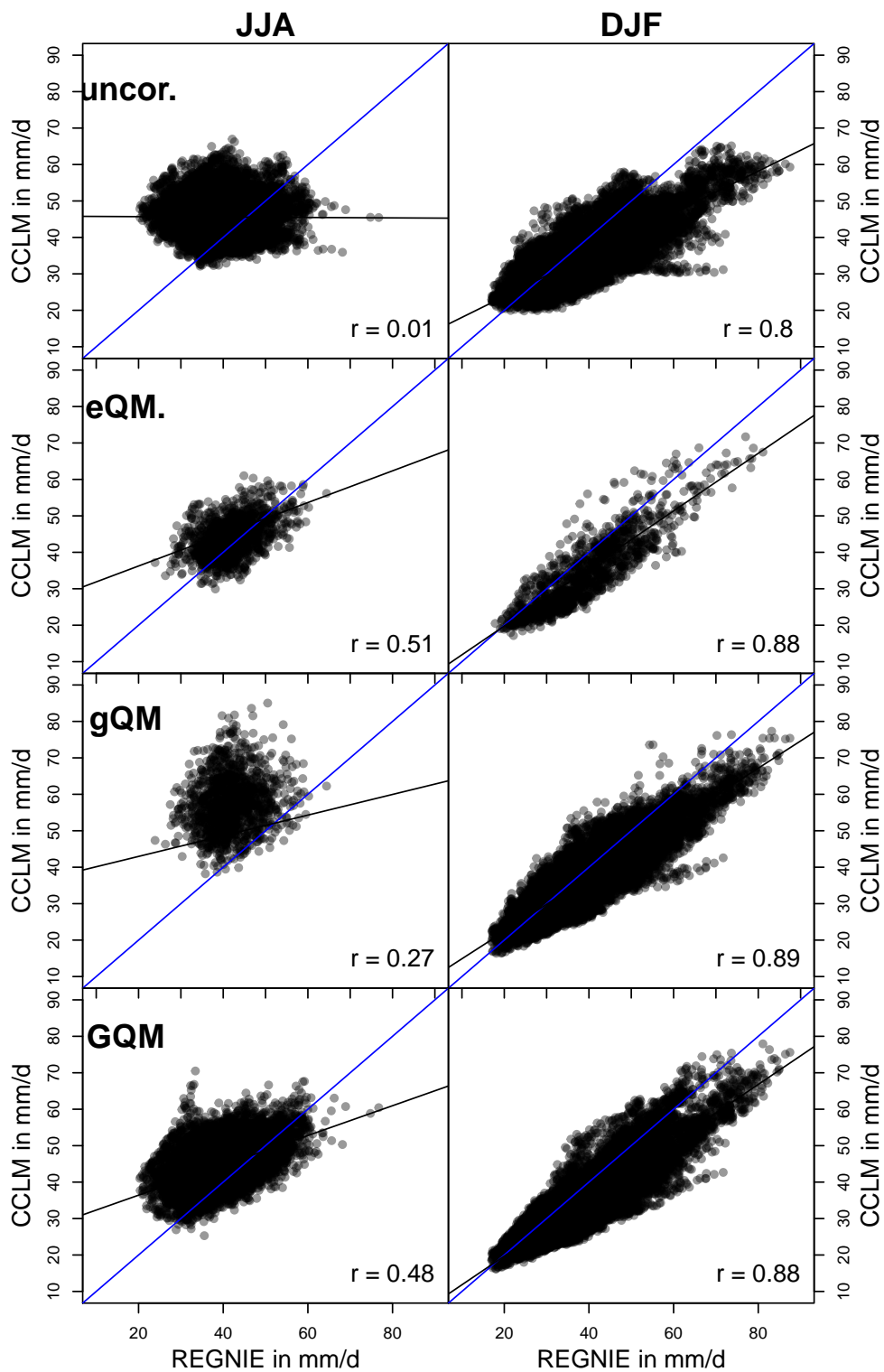


Figure 7.29 Scatterplots of the 10-year return levels (POT models) of daily precipitation from REGNIE and CCLM1.3c20 before and after the bias-corrections in summer (left column) and winter (right column). The $x = y$ line is in blue and the regression lines are in black.

By comparing the results of the CCLM4.5hc and CCLM1.3hc, a clear added value of the 1.3 km resolution was found for temperature indices. Since temperature extremes are strongly connected to orography, a better resolution improves the representation of orography and hence the spatial distribution of temperature extremes. Precipitation extremes depend less on orography in summer because of the convective nature. This explains the low correlation coefficients in summer. Although there were no large differences found for the precipitation extremes, except that CCLM1.3hc overestimates the proportion of extremes on the total sum, the better resolution may still add information if such a data set is used to force impact models, for example in hydrological applications such as catchment modelling.

The climate simulations show a significant increase for almost all temperature based hot indices until the end of the 21st century and a decrease of cold indices. The increase of hot indices until the mid of the century is only moderate and not significant but rises significantly until the end of the century. Cold indices warm already until 2041-2050. This indicates that the warming of hot extremes is delayed compared to cold extremes. For instance, heat-waves occurred every two summer in the present climate but there are about 2-3 simulated per summer in the future. A stronger signal was even found for the Rhine Valley. Here 5-7 heat-waves were simulated per summer but until 2091–2100 a signal of +10 heat-waves was found. Moreover, the duration of heat-waves increases from 16 to 46 days in the domain-average, whereas the duration of warm-spells increases from 8 to 60 days. Likewise, tropical nights increase from 1 to 13 per summer until 2091-2100.

In contrast, cold-waves were simulated every two winters, but are reduced to 1 in ten winters and their duration reduces from 16 to 6 days. Frost days reduce significantly from 46 to 14 days per winter. These results fit to results of other studies, where it was found that in summer the changes of warm extremes are larger than for "cold" extremes and in winter the changes of cold extremes are larger than for "warm" extremes (e.g. Kjellström et al., 2007; Fischer and Schär, 2010; Fischer et al., 2012). In the first case this leads to an increased variance, whereas in the second case the variance is reduced. Although the changes of summer cold and winter warm extremes were not investigated here explicitly, the increase in summer variance and decrease in winter variance was found in section 6.1.2 (see Tab. 6.4) as well. It was also found that cold extremes reduce strongest in regions with high orography while hot extremes increase in flat terrain, in particular the Rhine Valley.

The same pattern was confirmed by the analysis of return levels from 'peak-over-threshold' models. The 10- and 20-year RL of daily summer precipitation revealed an increase of about +10 mm for the end of the century. In winter, there was no change detected. The return levels of daily maximum temperature in summer increase about +6 °C in average and the winter RL warm about +8 °C. The highest return levels in the present climate were estimated for the Rhine Valley and for the Cologne and Mainz Basin with about 44 °C and they cluster at location of large cities. The advantage of a higher resolution of 1.3 km could be demonstrated for the temperature RL. For example, the Moselle Valley was not represented in the RL field from the CCLM4.5 but was clearly pronounced in the CCLM1.3.

Significant increases of precipitation indices were only found for the length and frequency of consecutive dry days (CDD, CDDn) in summer and the proportion of extremes on the total precipitation sum (R95PTOT) in summer and winter. This was also found by Sillmann and Roeckner (2008) for A1B scenario runs and by Sillmann et al. (2013a) for the new RCP scenarios. Here it was simulated that the CDD increase by +6 days in the domain-average and that 1 additional dry period occurs in summer at the end of the century. The R95PTOT increases in summer about 11 – 13 % and in winter about +7 %. However, there was no significant increase of very heavy precipitation days (R20). Thus the results indicate that the increase in R95PTOT is caused by less severe extreme events. Although the POT models show an increase of 10-year and 20-year events, the empirical RL of the models do no change until 2091–2100. But RL with return periods of ≤ 2 years increase in intensity and frequency and therefore cause the rise of R95PTOT.

The uncertainty of RL was found to be much larger in summer than in winter. This is probably because in summer the variability of extremes is larger than in winter due to the convective nature of extreme events. In contrast, in winter the highest RL were found in mountain regions. However, the simulation results do not fully confirm the expectation that the intensity of extremes increases and that they occur more often in a warmer climate. At least this was not found for the most extreme events (10-year events) in the direct model output, whereas the fitted POT models produce an increase in 10-year or 20-year events. However, these results have to be treated with caution because of the short time intervals of the simulations. In addition, only one GCM-RCM chain with only one anthropogenic scenario (A1B) was used. Therefore, the analyses should be repeated for longer simulation periods.

The bias-correction of daily precipitation caused an increase in the indices CDD and CDDn because of the reduction of the drizzle error and which mainly affects low intensities. This directly influences the dry indices because more days will have less than 1 mm of precipitation per day. However, in summer and in winter the CDD and CDDn were still underestimated in the corrected CCLM output. The BC did not influence the frequency of very heavy precipitation days (R20) but reduced the amount of RX5DAY in summer but had no effect in winter. An exception represents the correction with gQM, which caused an increase of extreme precipitation amount and frequency. But this can be explained by the large overestimation of extremes, which was already found in the verification of the hindcasts (section 5.2.1). All BC methods produced a higher R95PTOT index. In summer, the CCLM1.3c20 was already close to the observations, whereas the index was underestimated by CCLM4.5c20. After the correction, the index from the 4.5 km was close to the observations but slightly too high with regard to CCLM1.3c20. In winter, the corrections improved this index, but the proportion of extremes on the total precipitation is still underestimated. The reason is that winter extremes were not adjusted adequately by the BC methods and were still too low.

The climate change signals were less affected and only minor deviations occurred. The largest influences were found for the CCLM4.5 indices CDDn and R20. However, in most cases the signal remained constant. The analyses of the return levels resulted in the same findings. Furthermore, the spatial correlation of the RL from the CCLM and the observations did not agree in summer but well in winter. The BC methods eQM and GQM improved the summer correlation moderately, whereas the enhancement of gQM was fairly low. The discrepancies of the summer and winter correlation is due to the non-seasonal BC, which produces transfer functions that are not flexible enough to correct the spatial structure in both seasons. Thus for extreme value analyses of model output from regional climate models it is recommended to use at least seasonal corrections. In conclusion, the bias-correction does have a large impact on absolute extreme values but not on the relative change signals.

Note however that these rather short periods may hamper the statistical analysis of extremes because they are rare by definition and even more rare within 10 years. Therefore, significance tests may not be accurate because decadal variability could not be removed adequately. Nevertheless, the results can be interpreted as tendencies in a possible future climate. Further restrictions arise because only one GCM-RCM chain was used to simulate the potential future climate with only one SRES-scenario. This means that only one potential climate path is followed and conclusions cannot be drawn with regard to model variability or anthropogenic uncertainty as for example in the Coupled Model Inter-comparisons Project Phase 3 (CMIP3) (Meehl et al., 2007) and Phase 5 (CMIP5) (Taylor et al., 2012). Longer time series may improve the statistics and a censoring of the shape-parameter in the POT model fitting will become unnecessary. However, this study of the climate change of extremes with consecutive is the first one with very high-resolved continuous 10-year time-slices for the Saar-Lor-Lux and RLP area with a dynamically downscaled regional climate model.

Conclusion and outlook

The study was set out to investigate a potential future climate development in the Saar-Lor-Lux and RLP region over the next hundred years by the use of numerical simulations of the regional climate model COSMO-CLM. The focus was laid on two key-variables: precipitation and 2 m temperature, which are most often used in climate change impact studies. Climate change will have considerable implications on ecological and socio-economical systems and was therefore quantified within this study along with its uncertainty. In particular, it was explored whether the frequency and intensity of extreme events change in a future warming. Extreme events have the most severe impact on the natural environment and on society, e.g. high intensive precipitation may cause floods, extended dry periods lead to water stress or the presence of heat-waves and tropical nights constitute enhanced stress on human health. This study has further sought to answer whether an increased horizontal resolution of the climate simulations improves the results and adds new information. Finally, since regional climate models are affected with biases, it was aimed to correct the model errors and to analyse the impact of the bias-correction on the results. The literature on those subjects concerning the SLL&RLP region are scarce and only one study exists which has used identical resolved climate simulations but only for the summer months (see ?). The study sought to answer the following vital questions:

1. How does the future climate develop in the SLL&RLP region with respect to precipitation and temperature (means and extremes)?
2. Is there any added value by increasing the horizontal resolution from 4.5 km to 1.3 km?
3. Does a bias-correction improve the output of the CCLM and which method performs best?
4. What is the impact of the bias-correction on the climate change signals (means and extremes)?

To answer these questions, 10-year long time-slices have been simulated by the CCLM, which were performed at two horizontal resolutions: 4.5 km and 1.3 km for the present or control period (1991–2000), the near future (2041–2050) and the far future (2091–2100). The restriction of 10 years arose due to the high computational cost that is required by such high resolutions.

8.1 Future predictions

8.1.1 Temperature

The results showed that the annual mean of the near-surface temperature undergoes a warming of about $+1.7^{\circ}\text{C}$ until the period 2041–2050 with respect to the present climate. Furthermore, this warming increases to $+3.7^{\circ}\text{C}$ until 2091–2100. Even though these annual signals are widely homogeneous in the SLL&RLP region, several locations are expected to warm considerably stronger in summer. First and foremost, the Rhine Valley experiences a warming of up to $+5^{\circ}\text{C}$ at the end of the century. Similar signals were also found for Luxemburg, Saarland and Lorraine. Along with this increase in mean temperature, the simulations revealed an increase of the daily standard variation in summer but a decrease in winter.

These results confirm the findings of other studies and led to the hypothesis that the hot extremes in summer and the cold extremes in winter warm stronger than the mean values, which could be quantified to be about $+6^{\circ}\text{C}$ and $+8^{\circ}\text{C}$. The reason for these expected differences is that a warmer mean value shifts the whole probability density function to warmer temperatures but does not change its shape. In contrast, an increase in standard deviation does not shift the PDF but changes the shape; in this case, the PDF widens. Their joint effect produces a warming of the extremes that is larger than the warming of the mean value. A negative change of the standard deviation in winter along with a shift of the mean to warmer temperatures produces less cold extremes. These effects are consistent with the results of Knote et al. (2010), Schär et al. (2004) and Vidale et al. (2007).

According to the extreme value analysis and the results of the extreme indices, different warming rates were found for the summer hot and winter cold extremes. Hot extremes warm only moderately until 2041–2050 but then strongly until 2091–2100. Cold extremes, however, warm continuously until the end of the century. The characteristics of extremes are likely to change severely in both seasons. In particular, the Rhine Valley is likely to be affected by much more frequent heat-waves or tropical nights. In general, regions with flat terrain are mostly affected by increased summer hot extremes. In contrast, the reduction of cold extremes is simulated to be strongest in high orographic regions. A similar result was found by Knote et al. (2010), although for the summer minimum temperatures.

8.1.2 Precipitation

The results did not show significant changes in the annual mean precipitation but significant increases were found in autumn until 2041–2050 and decreases in summer until 2091–2100. Drier summers were also found in previous studies for mid-Europe (Solomon et al., 2007b) and in the SLL&RLP region (Knote et al., 2010). The variability of daily precipitation is enhanced in all seasons, though. Consequently, more frequent and intense precipitation extremes can be expected from a statistical point of view. Physically, a warming increases the potential water vapour content in the atmosphere according to the Clausius–Claperyon equation. Particularly in summer, it is expected that the frequency and intensity of extreme precipitation events occur more frequent and with higher intensities. This question could not be answered sufficiently without arising large uncertainties and the results depend on the applied analysis technique. According to the simulation results, there were no significant increases in frequency or intensity found for extreme events with intensities $> 20 \text{ mm day}^{-1}$. Moreover, the empirically assigned 10-year return levels from the simulations did not change in the future climate. If the change is calculated from the estimated values of the peaks-over-threshold models, though, then an increase could be identified. However, extreme events with shorter reoccurrence periods intensify with regard to both the empirical return level calculation and the POT models. Along with the finding of a higher proportion of extreme precipitation on the total in the future climate, it can be concluded

that this is caused by events with frequencies ≤ 2 years rather than by more severe events with longer reoccurrence periods. This also implies that events which have return periods of ≤ 2 years in the present climate will occur more often in the future.

8.2 Resolution effect

In general, the simulations with a resolution of 4.5 km and 1.3 km produced similar results. However, added value was found for the analyses of temperature with regard to the higher resolution. In particular, the simulation of extremes revealed much more small-scale features and thus showed a much more heterogeneous field as produced by the 4.5 km simulation. Furthermore, hot extremes are spatially more local than in the coarser simulation. This added value is a direct effect of the better representation of orography in the 1.3 km simulation.

The effect on the precipitation results are not as clear. The onset and intensity of convective precipitation in the afternoon/evening could be improved by explicit treatment of deep-convection, although only for intensities $> 2 \text{ mm hour}^{-1}$, which were overestimated by the model, though. The error of too frequent light rain ("drizzle") is reduced if a resolution of 1.3 km is used, but intensities of $< 10 \text{ mm day}^{-1}$ were underestimated in summer. The reason for this dry-bias could not be revealed completely so that a more detailed investigation is required. Unfortunately, this error conflicts the improvement in convective precipitation. Improved diurnal cycles, onset of convection and a reduction of persistent light rain have also been found by Kendon et al. (2012), who used a horizontal resolution of 1.5 km compared to 12 km in Great Britain. Most recently, Prein et al. (2013) showed similar improvements for the Alps-region by using also the CCLM but in a resolution of about 3 km compared to 10 km. They also concluded that the better resolved orography completely accounts for the improvement of the temperature fields, whereas the improvements in the precipitation fields are caused by the explicit treatment of deep-convection and the improved dynamics. These findings could not be entirely demonstrated as in their study due to problems in the model configuration in summer, but at least the results point to a better representation of convective precipitation. This is also indicated by higher proportions of extremes on the total precipitation in summer in the 1.3 km simulation.

8.3 Bias-correction effect

The bias-correction improved the daily precipitation and temperature fields. In particular, the empirical quantile-matching and the newly tested GQM method adjusted not only the mean but also the extreme values. Thus, it could be demonstrated that additional parameters are needed to correct extremes accurately, if a parametric quantile-matching is applied. It was found that the bias-correction had a minor to moderate impact on the climate change signals, which was also stated by e.g. Hagemann et al. (2011). The BC amplified positive change signals of precipitation but dampened negative signals so that all signals were shifted to more positive values. However, the temperature signals were almost not affected, except that the signal of daily variability in summer was considerably reduced.

With respect to the extremes, the bias-correction noticeably altered the absolute values but had nearly no impact on the change signals. Consequently, a bias-correction improves the model output fields of regional climate models, which is useful if they are intended to provide forcing data for impact models. Even though the BC corrects model errors, it is stressed here that analyses which are directly based on absolute values will considerably be affected by bias-corrections. For this reason, bias-correction should be seen as a new layer of uncertainty in the analysis chain of climate change and its impacts. On the other hand, if only relative values, such as climate change signals, are of interest then the BC is not mandatory but the detection power may be reduced, especially in case of precipitation. However, these conclusions

are only valid for precipitation and temperature but may also change with the use of other RCMs, BC methods or model domains. Therefore, the impact should be investigated for every climate change study.

8.4 Limitations and future suggestions

The major limitation of this study arose from the length of the simulation periods, which covered 10-year time-slices. This restriction emerged because of the high computational costs that are required for running regional climate simulations at a 1.3 km resolution in a domain of about $286 \text{ km} \times 286 \text{ km}$ with 40 vertical levels. As a consequence, the periods of the present and future climate were shorter than recommended by the WMO (WMO, 2011) so that the results depend stronger on individual decades than in 30-year periods, in which decadal variability would be averaged. Consequently, the results may change by using different decades. A future suggestion would therefore be to extend the simulation periods to at least 30 years. This would particularly increase the robustness and power of the climate change signal detection, the robustness of the extreme value analysis and the parametric quantile-matching method. However, the results of this study are in good agreement with published results which relied on longer time periods. In this context, a future suggestions would be to perform ensemble experiments at this high resolution to analyse the robustness of the results and the uncertainty due to the RCMs and anthropogenic scenarios.

The configuration of the 1.3 km simulation was adapted from an older model version of the CCLM and from the defaults of the numerical weather prediction configuration for a resolution of 2.8 km. Even though smaller test runs were performed to find an optimal configuration, problems with the summer precipitation occurred, which led to a considerably underestimation. It is therefore necessary to explore the reasons for this dry-bias by a sensitivity analysis in further studies. Furthermore, a detailed investigation of the physical processes that lead to the formation of convective precipitation would help to remove this systematic error. This may be also analysed with regard to differences between both resolutions to reveal any improvements in the dynamics as was found by Prein et al. (2013).

The here used methods can also be applied to other variables, such as humidity, radiation or wind, although the correction is more complicated than for precipitation or temperature, but possible as shown by Haddeland et al. (2012). An improvement may also emerge from the joint correction of temperature and precipitation as proposed by Piani and Haerter (2012). This would have the benefit that a consistent relationship is maintained between both variables, which is the main point of criticism in the application of bias-correction. However, so far this kind of correction have not been applied elsewhere in the climate community. Further possible extensions for the bias-correction could be, for instance, a seasonal correction of precipitation that may improve particularly the correction of extremes, which was not sufficient in the winter season. In addition, the transition from a gamma to a Pareto distribution in the GQM method was set to the 95% percentile and represents a rather abrupt change in the underlying PDF. A smoother transition may be established by making the transition dependent on two additional parameters as in the mixture model from Vrac and Naveau (2007), which is based on the work of Frigessi et al. (2003). However, in their study the transition was still very fast and the introduction of additional parameters is questionable then.

To model extreme events, the univariate extreme value theory was applied by fitting peak-over-threshold models to every grid-point separately. Thereby, any spatial dependency has been neglected. To account for spatial dependency, it is possible to fit the POT models in a multivariate way, as shown by Vrac et al. (2007). An alternative would be the regional frequency analysis (RFA) method (Fowler et al., 2005; Ekström et al., 2005).

Finally, a further suggestion to analyse the effect of resolution on precipitation objects would be the application of the SAL (Structure, Amplitude and Location) approach by Wernli and Paulat (2008).

The SAL is a measure for the size, shape and location of precipitation objects. With an increasing horizontal resolution and an explicit treatment of deep convection, it can be expected that the simulation of precipitation objects improves.

List of Tables

2.1	Overview of the regional climate model runs with CCLM.	22
2.2	Overview of the observational data sets.	24
3.1	ETCCDI indices	34
4.1	Comparison of domain-averaged annual precipitation and 2 m temperature.	44
4.2	Seasonal bias	44
5.1	Distribution moments of precipitation	65
6.1	Seasonal climate change signal of precipitation	74
6.2	Power analysis CCLM4.5: precipitation	77
6.3	Power analysis CCLM1.3: precipitation	77
6.4	Seasonal climate change signal of 2 m temperature	78
6.5	Power analysis CCLM4.5: temperature	81
6.6	Power analysis CCLM1.3: temperature	81
6.7	Seasonal climate change signal of precipitation after BC	83
6.8	Number of significant grid boxes after BC	85
6.9	Power analysis CCLM4.5 after BC: precipitation	86
6.10	Power analysis CCLM1.3 after BC: precipitation	86
6.11	Seasonal climate change signal of temperature after BC	87
7.1	Extreme indices hindcast	91
7.2	Extreme indices of hindcasts	99
A.1	Climatological data sets	141
A.2	Soil levels in the CCLM	142
A.3	Soil types in the CCLM	142
A.4	Initial and boundary data in the CCLM	143
A.5	Vertical levels in the CCLM	144
A.6	Namelist configuration for the CCLM4.5 and CCLM1.3 model.	145
C.1	Power analysis CCLM4.5 after BC: temperature	150
C.2	Power analysis CCLM1.3 after BC: temperature	150

List of Figures

1.1	Global warming of the surface temperature	3
1.2	Influence of anthropogenic forcing	5
1.3	Changes of extreme values	12
2.1	Sketch of a CCLM gridbox	16
2.2	Vertical staggering of grid boxes	16
2.3	Model domain of the CLM18	18
2.4	Model domain of the CCLM4.5	19
2.5	Model domain of the CCLM1.3	21
2.6	Nesting Chain	23
2.7	CLM18hc vs. E-OBS: precipitation	25
2.8	CLM18hc vs. E-OBS: 2 m temperature	25
2.9	CLM18hc vs. E-OBS: pressure	26
3.1	FDR-procedure	28
3.2	climatological monthly means of 2 m temperature	30
3.3	Schematic illustration of the running length declustering	37
3.4	Beta-prior	39
4.1	Differences: REGNIE - INTERMET	42
4.2	Annual mean precipitation	43
4.3	Annual mean of 2 m temperature	45
4.4	Monthly precipitation time series	46
4.5	climatological monthly precipitation sums	47
4.6	Monthly 2 m temperature time series	48
4.7	climatological monthly means of 2 m temperature	48
4.8	Taylor diagrams for precipitation	49
4.9	Taylor diagrams for 2 m temperature	50
4.10	Standard deviations of daily precipitation	51
4.11	Standard deviation of daily temperature	52
4.12	Histograms of precipitation	53
4.13	Histograms of 2 m temperature	54
4.14	QQ-plots of daily precipitation	55
4.15	Diurnal cycle of precipitation	55
4.16	Diurnal cycle of precipitation exceeding threshold	56
4.17	Diurnal cycle of 2 °Cm temperature	56
5.1	MAE _x from cross-validation of the hindcast	60

5.2	MAE _x of cross-validation from C20 run	62
5.3	Precipitation fields after bias-correction	63
5.4	Precipitation fields after bias correction	64
5.5	QQ-plots of BC precipitation	66
5.6	Seasonal bias-correction of precipitation	67
5.7	Temperature fields after bias-correction	68
5.8	Seasonal bias-correction of temperature	69
6.1	CCS of annual precipitation	72
6.2	CCS of annual temperature	73
6.3	CCS of DJF precipitation	75
6.4	CCS of JJA precipitation	76
6.5	CCS of DJF temperature	79
6.6	CCS of JJA temperature	80
6.7	CCS of JJA precipitation after bias-correction	82
6.8	CCS of SON precipitation after bias-correction	84
7.1	R20 index	92
7.2	Annual mean of RX5DAY	94
7.3	Frost days (FD) 1993–2000	95
7.4	Tropical nights (TR20) 1993–2000	96
7.5	Heat-wave duration 1993–2000	97
7.6	Number of heat-waves 1993–2000	98
7.7	ETCCDI precipitation indices	100
7.8	Number of CDD events	101
7.9	Number of heat-waves 1993–2000	102
7.10	ETCCDI temperature indices	103
7.11	CCS of FD events	105
7.12	CCS of TR20	106
7.13	CCS of CWDIn	107
7.14	CCS of HWDIn	108
7.15	CCS of GSL.d	109
7.16	CCS of GSL.l	110
7.17	ETCCDI precipitation indices	112
7.18	ETCCDI precipitation indices	112
7.19	Precipitation return levels from climate runs	114
7.20	Return levels of T_{\min} and T_{\max} from climate runs	115
7.21	Return levels: precipitation	116
7.22	Return levels of area-mean precipitation	117
7.23	Return levels: maximum temperature	118
7.24	Return levels: minimum temperature	119
7.25	Return levels of area-mean maximum temperature	121
7.26	Return levels of area-mean minimum temperature	122
7.27	Return levels after BC correction	123
7.28	Return levels signals after BC correction	124
7.29	Return levels scatterplot	126
B.1	Diurnal cycle of TKE and HBPL	148

C.1	Seasonal mean precipitation of CLM18	149
D.1	CDD events: hindcast vs. observations	152
D.2	RX5DAY summer events in climate simulations	153
D.3	RP95PTOT in summer	154

List of Abbreviations and Symbols

Abbreviations

A1B	SRES-scenario for future climate
ACF	Auto-correlation function
BATS	Biosphere-Atmosphere Transfer Scheme
BC	Bias-correction
C20	SRES-scenario for present climate
CCLM	see COSMO-CLM
CCS	Climate change signal
CDF	Cumulative distribution function
CI	Confidence interval
CORDEX	COordinated Regional climate Downscaling EXperiment
COSMO-CLM	Consortium for SSmall scale Modelling - Climate Limited area Mode
DWD	Deutscher Wetterdienst (German Weather Service)
E-OBS	Gridded observational data set of the ENSEMBLES project
ECHAM5/MPI-OM	General Circulation Model of the Max-Planck-Institut Hamburg coupled to the Max Planck ocean model
ECOCLIMAP	Global database of land surface parameters
ENSEMBLES	Ensemble-based Predictions of Climate Changes and their Impacts
eQM	Empirical quantile-matching
ES	Effect-size
EVT	Extreme value theory
FDR	False discovery rate
GCM	General Circulation Model
GPD	General Pareto Distribution
GQM	Quantile-matching with mixture of gamma and general Pareto distribution
gQM	Quantile-matching with gamma-distribution
INTERMET	Interpolation Software for meteorological variables of the LUWG
IPCC	Intergovernmental Panel on Climate Change
KS	Kolmogorov-Smirnov-test

LUWG	Landesamt für Umwelt, Wasserwirtschaft und Gewerbeaufsicht
MAE _x	Mean absolute error of quantiles in interval x
MA	Moving-average
MLE	Maximum-likelihood estimation
NWP	Numerical weather prediction
PDF	Probability density function
PMSL	Pressure at mean sea-level
POT	Peak-over-threshold model
PRUDENCE	Prediction of Regional scenarios and Uncertainties for Defining European Climate change risks and Effects
PTF	Parametric transfer function after Piani et al. (2010b)
QM	Quantile-matching
RCM	Regional Climate Model
REGNIE	Regionalisierung von Niederschlagshöhen (Regionalisation of precipitation)
REMO	REgional MOdel
RL	Return level
RLP	Rhineland-Palatinate
RMSE	Root-mean-squared error
SLL	Saarland-Lorraine-Luxemburg
TERRA_ML	Soil-Vegetation-Atmosphere model of CCLM
UTC	Universal time code

Symbols

α	Significance level	%
α_g	Shape-parameter of gamma-distribution	-
a	Additive parameter of PTF	-
β	Beta-error	-
β_g	Scale-parameter of gamma distribution	-
β_p	2.0×10^{-3}	s ⁻¹
b	Multiplicative parameter of PTF	-
χ^2	Chi ² -distribution	-
C_α	Confidence region	-
c_α	α -quantile of Chi ² -distribution	-
dt	Time step	s
D	Absolute climate change signal	*
d	Cohen's d	-
e_d	Evaporation of precipitation in downdraft	kg m ²
e_p	Evaporation of precipitation below cloud base	kg m ²
Γ	Gamma function	-
g_p	Convective precipitation amount	kg m ²
H_0	Null-hypothesis	-
K_p	Height-dependent conversion function	-
ℓ	log-likelihood function	-

ℓ_p	Profile-loglikelihood	-
M	N-year return period	-
n_d	Number of observed dry-days	-
N	Normal-distribution	-
n	Number of p-values	-
p_i	p-value on location i	-
p_k	Largest local p-value in FDR-procedure	-
q_u^c	Cloud water content in updraft	kg m^2
q	False discovery rate	%
r	Pearson's correlation coefficient	-
r_l	Run length	-
σ	Scale-parameter of GPD	-
σ	Standard deviation	*
σ_{xy}	Within-population standard deviation	*
$\tilde{\sigma}$	Reparameterized scale-parameter of GPD	-
$\hat{\theta}_0$	Maximum-likelihood estimates of the parameters	-
$\hat{\theta}_i$	Individual parameter estimate	-
$\hat{\theta}_{mle}$	Maximum-likelihood estimate	-
Θ	Parameter space	-
θ	Parameter set	-
θ_0	True values of parameters	-
θ_i	Individual parameter	-
θ_{-i}	All parameters except θ_i	-
u	Threshold of GPD	*
τ	Asymptotic rate of PTF	-
ξ	Shape-parameter of GPD	-
x_0	Dry-day correction parameter of PTF	-
x_d	Dry-day threshold	mm day^{-1}
$\tilde{\psi}_{i,j}$	Inversion of observed information matrix	-
ζ_u	Exceedance probability	-
z	Height	m
z_α	α -quantile of standard normal distribution	-
z_b	Height of cloud base	m
z_c	Height levels of suppressed convective precipitation	m
z_m	Return-level	-

*: same unit as output variable

Appendix A

CCLM configurations

Table A.1 Climatological data sets which were used in the CCLM4.5 and CCLM1.3. See more details in Smiatek et al. (2008).

Variable	Description	Resolution	Source
HSURF	orography	1 × 1 km	GLOBE
FR_LAND	land/sea fraction	1 × 1 km	GLCC2.0
SOILTYP	soil type	1/12°	FAO
Z0	roughness length	1 × 1 km	GLOBE
PLCOV_MN	vegetation cover resting season	1 × 1 km	GLCC2.0/CORINE
PLCOV_MX	vegetation cover growing season	1 × 1 km	GLCC2.0/CORINE
LAI_MN	leaf area index resting season	1 × 1 km	GLCC2.0/CORINE
LAI_MX	leaf area index growing season	1 × 1 km	GLCC2.0/CORINE
ROOTDP	root depth	1 × 1 km	GLCC2.0/CORINE
FOR_E	fraction evergreen forest	1 × 1 km	GLCC2.0/CORINE
FOR_D	fraction deciduous forest	1 × 1 km	GLCC2.0/CORINE
T_CL	climatological deep soil temperature	10'	CRU (2 m temperature)

Table A.2 Main levels of the soil layers which were used in the CCLM4.5 and CCLM1.3.

Level	(m)	(cm)
0:		0
1:	0.01	1
2:	0.04	4
3:	0.10	10
4:	0.22	22
5:	0.46	46
6:	0.94	94
7:	1.90	190
8:	3.82	382
9:	7.66	766
10:	15.34	1534

Table A.3 Soil types and parameters in the CCLM4.5 and CCLM1.3.

soil type		ice	rock	sand	sandy loam	loam	loamy clay	clay	peat
volume of voids	(-)	-	-	0.364	0.445	0.455	0.475	0.507	0.863
field capacity	(-)	-	-	0.196	0.260	0.340	0.370	0.463	0.763
permanent wilting point	(-)	-	-	0.042	0.100	0.110	0.185	0.257	0.265
air dryness point	(-)	-	-	0.012	0.030	0.035	0.060	0.065	0.098
minimum infiltration rate	(kg m ² s)	-	-	0.0035	0.0023	0.0010	0.0006	0.0001	0.0002
hydraulic diffusivity D_0	(10 ⁻⁹ m ² s ⁻¹)	-	-	18400	3460	3570	1180	442	106
hydraulic diffusivity D_1	(-)	-	-	-8.45	-9.47	-7.44	-7.76	-6.74	-5.97
hydraulic conductivity K_0	(10 ⁻⁹ ms ⁻¹)	-	-	47900	9430	5310	764	17	58
hydraulic conductivity K_1	(-)	-	-	-19.27	-20.86	-19.66	-18.52	-16.32	-16.48
heat capacity	(10 ⁶ J/(m ³ K))	1.92	2.10	1.28	1.35	1.42	1.50	1.63	0.58
heat conductivity λ_0	(W/(K m))	2.26	2.41	0.30	0.28	0.25	0.21	0.18	0.06
heat conductivity $\Delta\lambda$	(W/(K m))	0.0	0.0	2.40	1.58	1.55	1.50	1.50	0.50
exponent B	(-)	1.0	1.0	3.5	4.8	6.1	8.6	10.0	9.0

Table A.4 Initial and boundary data fields for CCLM4.5 and CCLM1.3.

Variable	Description	Initial	Boundary
HSURF	orography	x	
FR_LAND	land/sea fraction	x	
Z0	roughness length	x	
SOILTYP	soil type	x	
PLCOV	vegetation cover	x	x
LAI	leaf area index	x	x
ROOTDP	root depth	x	x
VIO3	vertical integrated ozone content	x	x
HMO3	ozone maximum	x	x
U	zonal wind speed	x	x
V	meridional wind speed	x	x
W	vertical wind speed	x	
T	temperature	x	x
QV	specific water vapor content	x	x
QC	specific cloud water content	x	x
PP	deviation from reference pressure	x	x
T_SNOW	temperature of snow surface	x	x
W_I	water content of interception water	x	
QV_S	specific water vapor content at the surface	x	x
W_SNOW	water content of snow	x	x
T_S	temperature at the boundary soil-atmosphere	x	x
T_SO	soil temperature at different layers	x	
W_SO	soil moisture of different layers	x	
FRESHSNW	weighting function indicating "freshness" of snow	x	
FOR_E	fraction evergreen forest	x	
FOR_D	fraction deciduous forest	x	
QI	specific cloud ice content	x	x
QR	specific rain content	x	x
QS	specific snow content	x	x
QG	specific graupel content	x	x

Table A.5 The vertical model layers (k), vertical coordinate (VCOORD(k)), the corresponding height in meters (Z(k)) and the reference pressure ($P_0(k)$).

k	VCOORD(k)	Z(k)	$P_0(k)$
1	0.0200	23588.5	20.00
2	0.0400	20780.5	40.00
3	0.0604	18834.3	60.36
4	0.0814	17289.8	81.39
5	0.1033	15978.3	103.32
6	0.1264	14817.2	126.36
7	0.1506	13763.2	150.63
8	0.1762	12789.1	176.24
9	0.2032	11878.9	203.23
10	0.2316	11021.9	231.61
11	0.2614	10211.2	261.35
12	0.2924	9441.8	292.40
13	0.3246	8711.0	324.64
14	0.3580	8016.0	357.97
15	0.3922	7355.4	392.23
16	0.4272	6728.0	427.25
17	0.4629	6132.5	462.86
18	0.4988	5568.6	498.84
19	0.5350	5035.6	534.97
20	0.5710	4532.7	571.05
21	0.6068	4059.6	606.84
22	0.6421	3615.8	642.11
23	0.6766	3200.9	676.63
24	0.7102	2814.7	710.17
25	0.7425	2456.4	742.53
26	0.7735	2125.8	773.49
27	0.8029	1822.5	802.85
28	0.8304	1545.8	830.45
29	0.8561	1295.3	856.13
30	0.8798	1070.4	879.76
31	0.9012	870.5	901.23
32	0.9205	694.7	920.48
33	0.9374	542.3	937.44
34	0.9521	412.2	952.14
35	0.9646	303.3	964.59
36	0.9749	214.3	974.87
37	0.9831	143.4	983.12
38	0.9895	89.0	989.50
39	0.9942	48.7	994.24
40	0.9976	20.0	997.63
41	1.0000	0.0	1000.00

Boundary for change from z to terrain-following: $vcoord = 0.22$

Half-level index where levels become flat:
 $k = 9$

Linear vertical gradient with respect to $\ln p$.

Table A.6 Namelist configuration for the CCLM4.5 and CCLM1.3 model.

Namelist	Variable	CCLM4.5	CCLM1.3	
lmgrid	pollon	180	180	
	pollat	40	40	
	polgam	0.0000	0.0000	
	dlon	0.0400	0.0118	
	dlat	0.0400	0.0118	
	startlon_tot	-1.2500	3.2012	
	startlat_tot	-4.7500	-1.2348	
	ie_tot	254	220	
	je_tot	254	220	
	ke_tot	40	40	
	runcpl	dt	45.0	12.0
		ydate_ini	1991010100	1991010100
		ydate_end	2000123123	2000123123
ydate_bd		1991010100	1991010100	
itype_calenda		0	0	
hincmxt		24.0	24.0	
hincmxu		24.0	24.0	
leps		F	F	
lphys		T	T	
ldiagnos		T	T	
ldfi		F	F	
luseobs		F	F	
luse_rttov		F	F	
ldump_ascii		T	T	
lreproduce		T	T	
lreorder		F	F	
lartif_data		F	F	
lperi		F	F	
l2dim		F	F	
llm		F	F	
nprocx		16	24	
nprocy		8	16	
nprocio		0	0	
nboundlines		3	3	
ncomm_type		3	3	
ldatatypes		F	F	
ltime_barrier		F	F	
itype_timing		4	4	
idbg_level		21	21	
ldebug_dyn		F	F	
ldebug_gsp		F	F	
ldebug_rad		F	F	
ldebug_tur		F	F	
ldebug_con		F	F	
ldebug_soi		F	F	
ldebug_io		F	F	
ldebug_dia	F	F		
ldebug_ass	F	F		
ldebug_lhn	F	F		
lprintdeb_all	F	F		
linit_fields	F	F		
tuning	crsmin	150.0	150.0	
	rat_lam	1.0	1.0	
	tkesmot	0.15	0.15	
	wichfakt	0.15	0.15	
	securi	0.85	0.85	
	tkhmin	1.0	1.0	
	tkmmin	1.0	1.0	
	rlam_mom	0.0	0.0	
	rlam_heat	1.0	1.0	
	rat_sea	20.0	20.0	
	rat_can	1.0	1.0	
	c_lnd	2.0	2.0	
	c_sea	1.5	1.5	
	c_soil	1.0	1.0	
	e_surf	1.0	1.0	
	pat_len	500.0	500.0	
	tur_len	500.0	500.0	
	z0m_dia	0.20	0.20	
	a_heat	0.74	0.74	
	a_mom	0.92	0.92	
	dynctl	d_heat	10.1	10.1
d_mom		16.6	16.6	
c_diff		0.2	0.2	
cle_diag		0.5	0.5	
q_crit		4.0	4.0	
qc0		0.0	0.0	
qi0		0.0	0.0	
gkdrag		0.075	0.075	
gkwake		0.5	0.5	
mu_rain		0.5	0.5	
cloud_num		5.00E+08	5.00E+08	
entr_sc		0.0003	0.0003	
epsass		0.15	0.15	
alphaass		1.0	1.0	
betasw		0.4	0.4	
betagw		0.4	0.4	
beta2sw		0.4	0.4	
beta2gw		0.4	0.4	
lrubc		F	F	
lspubc		T	T	
lexpl_lbc		T	T	
lradlbc	F	F		
lcond	T	T		
ldiabf_lh	F	F		
lw_freeslip	T	T		
l2tls	T	T		
lsl_adv_qx	T	T		
lva_impl_dyn	T	T		
lsemi_imp	F	F		
lhordiff	T	T		
lhdiff_mask	T	T		
ldyn_bbc	T	T		
lcori	T	T		
lmetr	T	T		
lcori_deep	F	F		
ladv_deep	F	F		
yef_adv_qx	Bott_2	Bott_2		
rdheight	12000.0	12000.0		
crltau	1.0	1.0		
riwidth	50000	13000		
relax_fac	0.01	0.01		
nrdtau	5	5		
maxit_si	200	200		
ikrylow_si	20	20		
iprint_si	0	0		
irunge_kutta	1	1		
irk_order	3	3		
iadv_order	5	5		
ieva_order	3	3		
itheta_adv	0	0		
inter_max	1	1		
itype_outflow_qrsg	1	1		
itype_lbc_qrsg	2	1		
itype_spubc	1	1		
nfi_spubc2	10	10		
itype_hdiff	2	2		
hd_corr_u	0.75	0.75		
hd_corr_t	0.75	0.75		
hd_corr_q	0.0	0.0		
hd_dhmax	250	250		
xkd	0.1	0.1		
eps_si	0.0	0.0		
lspecnudge	F	F		
phyctl	lgsp	T	T	
	lprogprec	T	T	
	ltrans_prec	T	T	
	ldiniiprec	F	F	
	lrad	T	T	
	lforest	T	T	
	ltur	T	T	
	l3dturb	F	F	
	l3dturb_metr	T	T	
	lfreeslip_sfc	F	F	
	lprog_tke	F	F	
	lconv	T	T	
	itype_conv	0	3	
	lconv_inst	T	T	
	lsoil	T	T	

llake	F	F
lmfk_start_co	T	F
lso	F	F
lmelt	T	T
lmelt_var	T	T
lmulti_layer	T	T
ke_soil	7	7
czbot_w_so	2.5	2.5
lturhor	F	F
lexpcor	T	T
lmpcor	F	F
lprfcor	F	F
lnonloc	F	F
lcpfluc	F	F
lconf_avg	T	T
lradf_avg	F	F
lcape	F	F
letke	F	F
hincerad	1.0	1.0
nincerad	80	300
nradcoarse	1	1
ninctura	1	1
nineconv	1	1
nincso	5	5
itype_trvg	2	2
itype_evsl	2	2
itype_gscp	4	4
itype_wclld	1	1
itype_tran	2	2
itype_turb	3	3
itype_synd	1	1
imode_tran	1	1
imode_turb	0	0
ico2_rad	0	0
icldm_rad	4	4
icldm_tran	0	0
icldm_turb	1	1
nlgw	2	2
lradtopo	F	F
nhorl	24	24

Appendix B

Additional information for chapter 4

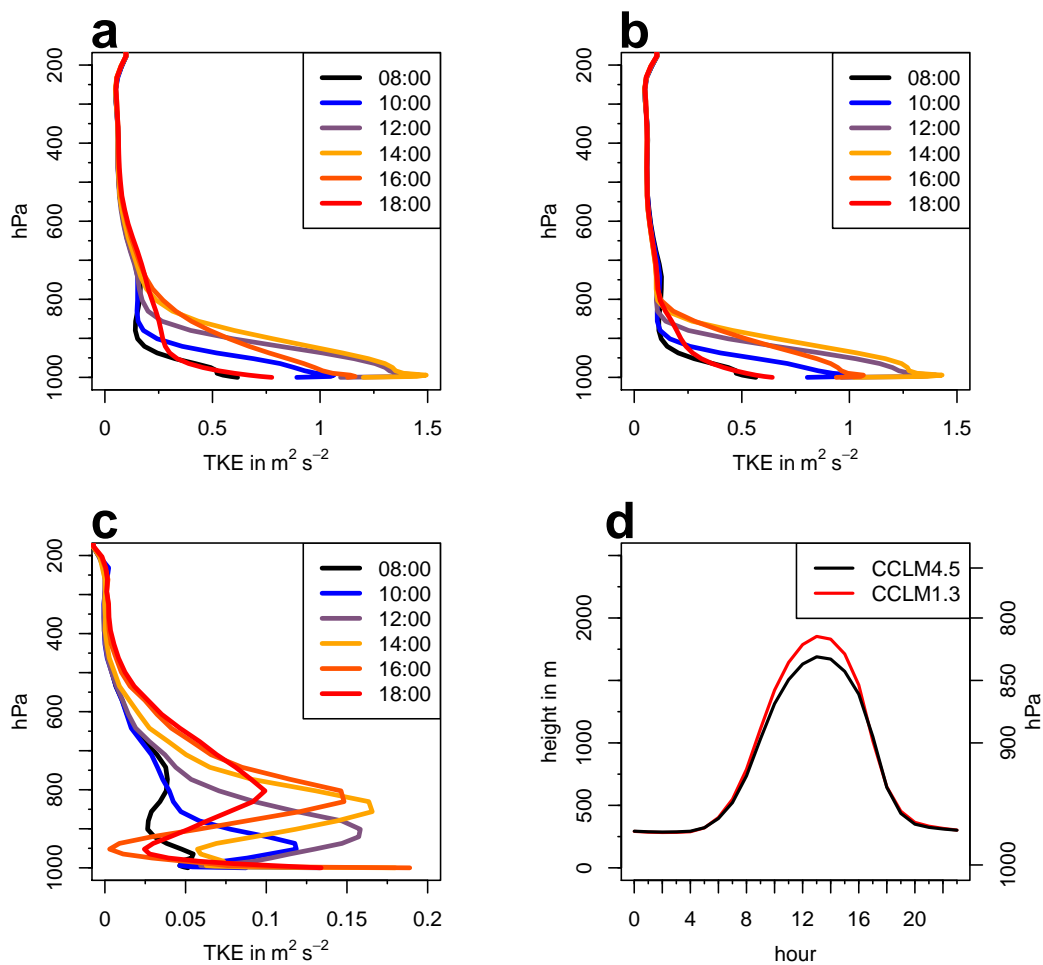


Figure B.1 Diurnal cycles of turbulent kinetic energy (TKE) of the period 1993–2000 for a) CCLM4.5hc, b) CCLM1.3hc and in c) the difference: CCLM4.5hc - CCLM1.3hc and in d) the height of boundary layer (HPBL). The time specifications in the legends are UTC.

Appendix C

Additional information for chapter 6

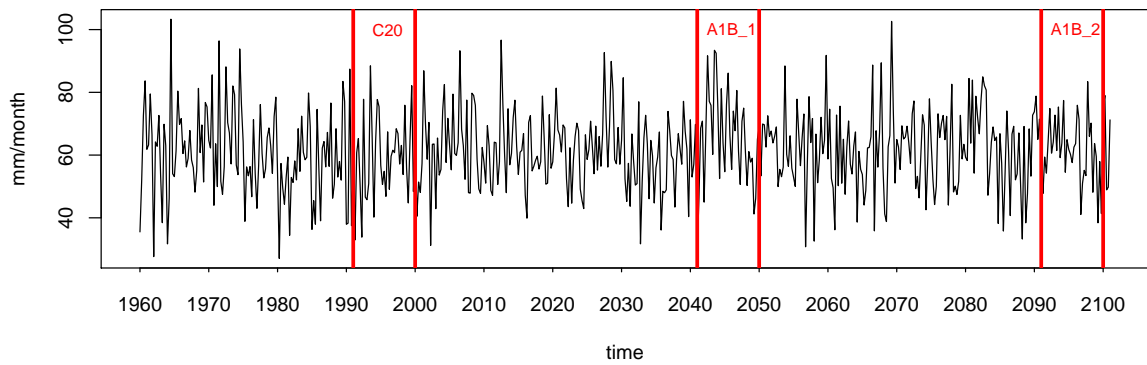


Figure C.1 Seasonal mean precipitation from CLM18 for the period 1960–2100. The red vertical lines indicate the periods of the time-slice experiments with CCLM4.5 and CCLM1.3 (they mark the 1st January of the start and end year).

Table C.1 Power analysis of the t-tests of the seasonal CCLM4.5 2 m temperature climate change signals before and after BC.

	power	n(p=0.8) ¹	σ_{xy} ²	D
$\Delta 2050$				
DJF				
CCLM	0.95	6.62	1.17	1.98
eQM	0.94	6.71	1.19	1.99
MAM				
CCLM	0.58	16.12	1.53	1.56
eQM	0.56	16.70	1.56	1.56
JJA				
CCLM	0.47	20.86	1.61	1.43
eQM	0.48	20.50	1.67	1.50
SON				
CCLM	0.98	5.63	0.96	1.80
eQM	0.98	5.53	0.97	1.83
$\Delta 2100$				
DJF				
CCLM	1.00	2.28	0.96	4.19
eQM	1.00	2.29	0.97	4.21
MAM				
CCLM	0.90	7.82	1.47	2.24
eQM	0.88	8.21	1.50	2.23
JJA				
CCLM	1.00	2.91	1.34	4.27
eQM	1.00	2.80	1.27	4.20
SON				
CCLM	1.00	2.63	1.15	4.10
eQM	1.00	2.60	1.17	4.22

¹ Number of years needed to achieve a power of 80 %.

² σ_{xy} denotes the common variance.

Table C.2 Power analysis of the t-tests of the seasonal CCLM1.3 2 m temperature climate change signals before and after BC.

	power	n(p=0.8) ¹	σ_{xy} ²	D
$\Delta 2050$				
DJF				
CCLM	0.95	6.62	1.18	1.99
eQM	0.94	6.71	1.19	1.99
MAM				
CCLM	0.57	16.57	1.52	1.53
eQM	0.55	17.07	1.52	1.50
JJA				
CCLM	0.45	22.08	1.53	1.32
eQM	0.45	21.76	1.57	1.36
SON				
CCLM	0.97	5.73	0.97	1.80
eQM	0.98	5.64	0.96	1.80
$\Delta 2100$				
DJF				
CCLM	1.00	2.28	0.97	4.24
eQM	1.00	2.30	0.97	4.20
MAM				
CCLM	0.90	7.85	1.45	2.22
eQM	0.88	8.20	1.46	2.17
JJA				
CCLM	1.00	2.97	1.29	4.00
eQM	1.00	2.88	1.21	3.89
SON				
CCLM	1.00	2.70	1.16	4.01
eQM	1.00	2.67	1.16	4.05

Appendix D

Additional information for chapter 7

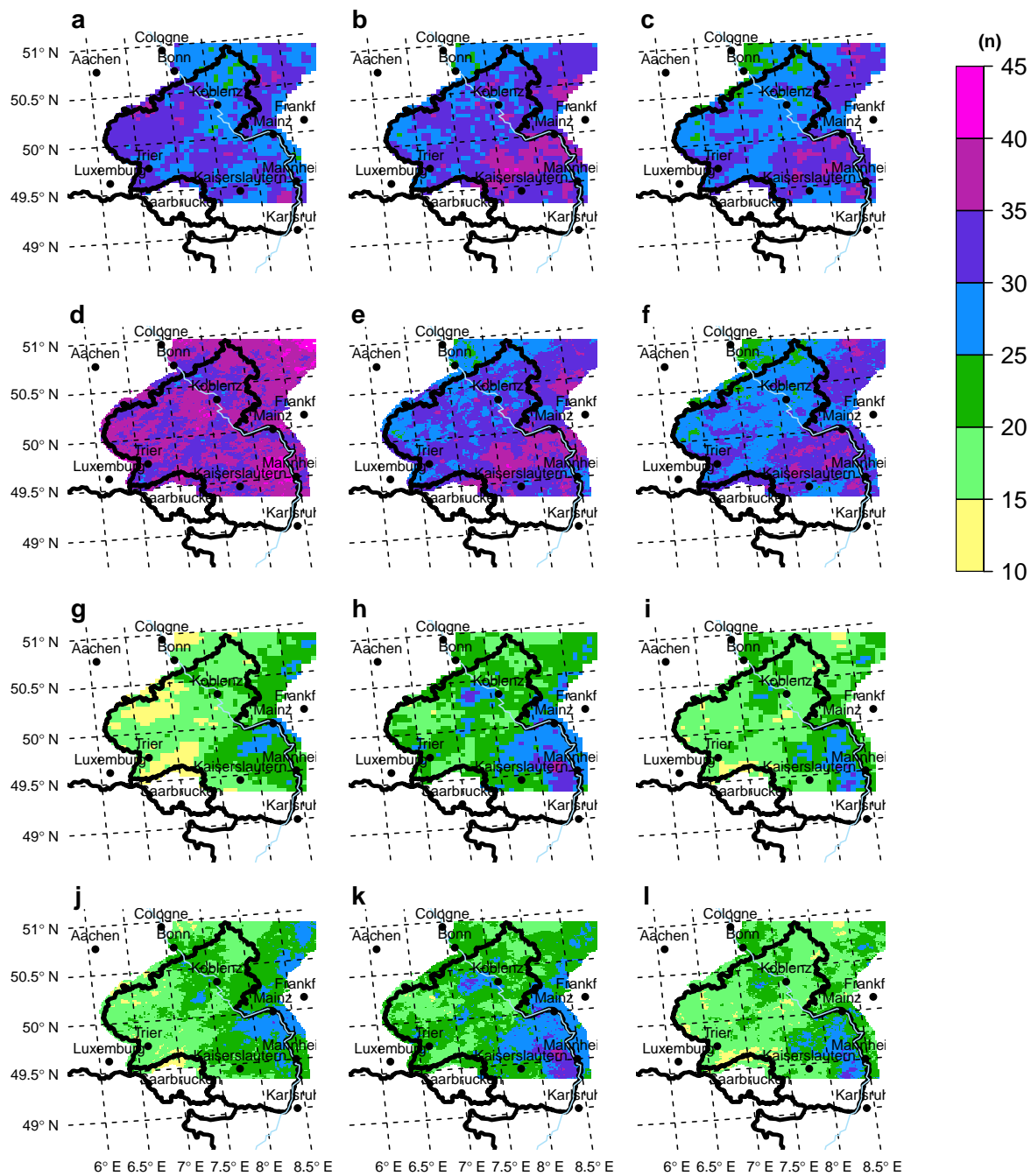


Figure D.1 Total number of consecutive dry day (CDD) events within the period 1993–2000 for summer (a-f) and winter (g-l) calculated from a) CCLM4.5, b) INTERMET4.5, c) REGNIE4.5, d) CCLM1.3, e) INTERMET1.3, f) REGNIE1.3 and g),l) from CCLM, h),k) from INTERMET and i),l)from REGNIE, respectively but for winter.

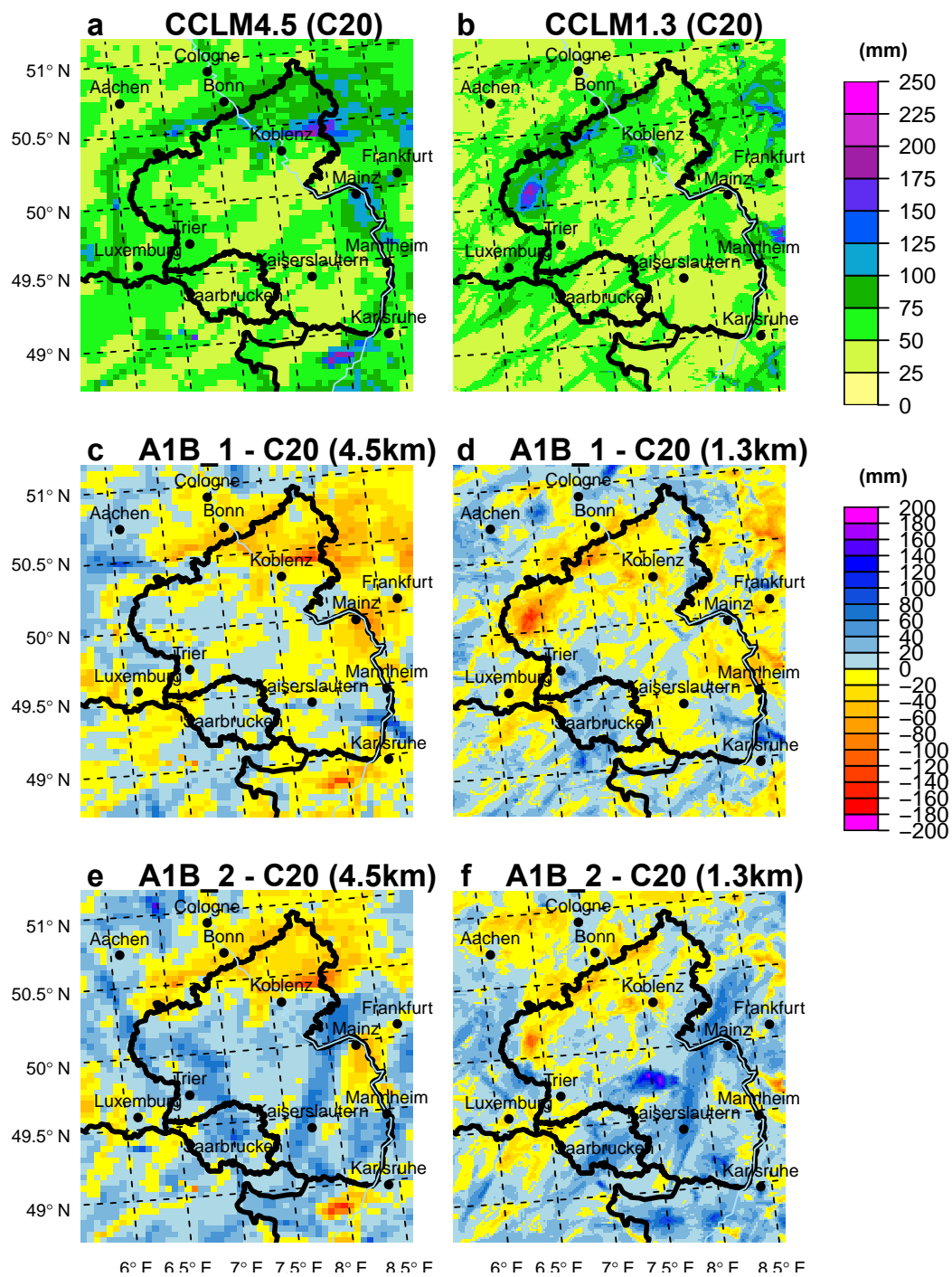


Figure D.2 Maximum precipitation sum within 5 consecutive days (RX5DAY) in summer (JJA) simulated for the present climate (1991–2000, C20) by a) CCLM4.5, b) CCLM1.3, the change until 2041–2050 (A1B1) c) 4.5 km, d) 1.3 km and the change until 2100 e) 4.5 km, f) 1.3 km.

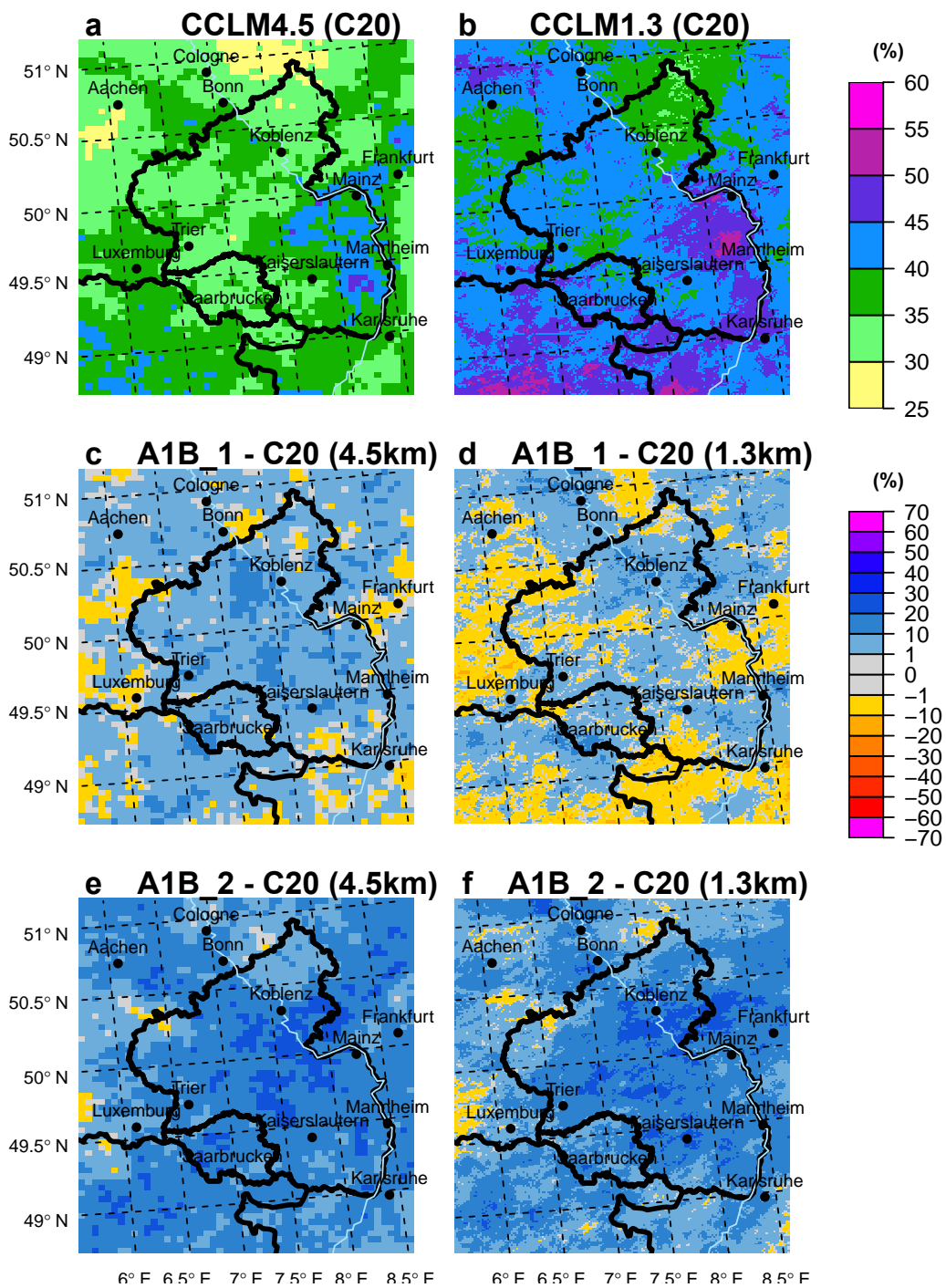


Figure D.3 Mean proportion of precipitation days > 95 % percentile on the total sum (R95PTOT) in summer (JJA) simulated for the present climate (1991–2000, C20) by a) CCLM4.5, b) CCLM1.3, the change until 2041–2050 (A1B1) c) 4.5 km, d) 1.3 km and the change until 2100 e) 4.5 km, f) 1.3 km.

Bibliography

- Allan RP, Soden BJ (2008) Atmospheric warming and the amplification of precipitation extremes. *Science*, 321(5895):1481–1484, DOI 10.1126/science.1160787
- Ament F, Simmer C (2006) Improved representation of land-surface heterogeneity in a non-hydrostatic numerical weather prediction model. *Boundary-Layer Meteorology*, 121(1):153–174, DOI 10.1007/s10546-006-9066-4
- Bachner S, Kapala A, Simmer C (2008) Evaluation of daily precipitation characteristics in the CLM and their sensitivity to parameterizations. *Meteorologische Zeitschrift*, 17/4:407–419
- Ballester J, Giorgi F, Rodó X (2010) Changes in european temperature extremes can be predicted from changes in PDF central statistics. *Climatic Change*, 98(1-2):277–284, DOI 10.1007/s10584-009-9758-0
- Bartels H, Dietzer B, Malitz G, Albrecht F, Guttenberger J (2006) KOSTRA-DWD-2000 Starkniederschlagshöhen für Deutschland (1951–2000). Fortschreibungsbericht, Deutscher Wetterdienst, Selbstverlag
- Beirlant J, Goegebeur Y, Teugels J, Segers J, De Waal D, Ferro C (2004) Statistics of Extremes, 1st edn. John Wiley & Sons, Ltd
- Bellprat O, Kotlarski S, Lüthi D, Schär C (2012) Exploring perturbed physics ensembles in a regional climate model. *Journal of Climate*, 25(13):4582–4599, DOI 10.1175/JCLI-D-11-00275.1
- Benestad R, Hanssen Bauer I, Chen D (2008) Empirical-statistical downscaling. World Scientific Publishing, New Jersey
- Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B (Methodological)*, 57(1):289–300
- Berg M, Wieland R, Mirschel W, Wenkel KO (2008) Modellierung und Simulation von Ökosystemen., Shaker Verlag Aachen, chap LandCaRe 2020 - ein Entscheidungsunterstützungssystem zur Vorhersage und Beurteilung der Potentiale ländlicher Gebiete unter dem Einfluss regionalen Klimawandels., pp 214–229
- Berg P, Feldmann H, Panitz HJ (2012) Bias correction of high resolution regional climate model data. *Journal of Hydrology*, 448–449:80–92, DOI 10.1016/j.jhydrol.2012.04.026
- Boberg F, Berg P, Thejll P, Gutowski WJ, Christensen JH (2009) Improved confidence in climate change projections of precipitation evaluated using daily statistics from the PRUDENCE ensemble. *Climate Dynamics*, 32(7-8):1097–1106, DOI 10.1007/s00382-008-0446-y

- Boé J, Terray L, Habets F, Martin E (2007) Statistical and dynamical downscaling of the seine basin climate for hydro-meteorological studies. *International Journal of Climatology*, 27(12):1643–1655, DOI 10.1002/joc.1602
- Buser CM, Künsch HR, Lüthi D, Wild M, Schär C (2009) Bayesian multi-model projection of climate: bias assumptions and interannual variability. *Climate Dynamics*, 33:849–868, DOI 10.1007/s00382-009-0588-6
- Cannon AJ (2008) Probabilistic multisite precipitation downscaling by an expanded bernoulli-gamma density network. *Journal of Hydrometeorology*, 9:1284–1300, DOI 10.1175/2008JHM960.1
- Casper MC, Grigoryan G, Gronz O, Gutjahr O, Heinemann G, Ley R, Rock A (2012) Analysis of projected hydrological behavior of catchments based on signature indices. *Hydrol Earth Syst Sci*, 16(2):409–421, DOI 10.5194/hess-16-409-2012
- Champeaux JL, Masson V, Chauvin F (2005) ECOCLIMAP: a global database of land surface parameters at 1 km resolution. *Meteorological Applications*, 12(1):29–32, DOI 10.1017/S1350482705001519
- Chen C, Haerter JO, Hagemann S, Piani C (2011) On the contribution of statistical bias correction to the uncertainty in the projected hydrological cycle. *Geophysical Research Letters*, 38(20), DOI 10.1029/2011GL049318
- Christensen JH, Christensen OB (2007) A summary of the PRUDENCE model projections of changes in european climate by the end of this century. *Climatic Change*, 81(1):7–30, DOI 10.1007/s10584-006-9210-7
- Christensen JH, Boberg F, Christensen OB, Lucas-Picher P (2008) On the need for bias correction of regional climate change projections of temperature and precipitation. *Geophysical Research Letters*, 35(20):L20,709, DOI 10.1029/2008GL035694
- Cohen J (1988) *Statistical power analysis for the behavioral sciences*, 2nd edn. Lawrence Erlbaum Associates, Inc., pp. 590
- Cohen J (1992) Statistical power analysis. *Current Directions in Psychological Science*, 1(3):98–101
- Coles S (2001) *An Introduction to Statistical Modeling of Extreme Values*, 1st edn. Springer, Berlin, pp. 224
- Coles SG, Dixon MJ (1999) Likelihood-based inference for extreme value models. *Extremes*, 2(1):5–23, DOI 10.1023/A:1009905222644
- Collins M (2007) Ensembles and probabilities: a new era in the prediction of climate change. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 365(1857):1957–1970, DOI 10.1098/rsta.2007.2068
- Cooley D (2009) Extreme value analysis and the study of climate change. *Climatic Change*, 97(1):77–83, DOI 10.1007/s10584-009-9627-x
- De Haan L, Ferreira A (2006) *Extreme Value Theory: An Introduction*. Springer, New York
- Denis B, Laprise R (2003) Sensitivity of a regional climate model to the resolution of the lateral boundary conditions. , 20(2):107–126, DOI 10.1007/s00382-002-0264-6

- Denis B, Laprise R, Caya D, J C (2002) Downscaling ability of one-way nested regional climate models: the big-brother experiment. *Climate Dynamics*, 18(8):627–646, DOI 10.1007/s00382-001-0201-0
- Deser C, Phillips A, Bourdette V, Teng H (2012) Uncertainty in climate change projections: the role of internal variability. *Climate Dynamics*, 38:527–546, DOI 10.1007/s00382-010-0977-x
- Dickinson R, Henderson-Sellers A, Kennedy P (1993) Biosphere-atmosphere transfer scheme (bats) version 1e as coupled to the near community climate model. Tech. rep., NCAR Technical Note NCAR/TN-387+STR
- Dobler A, Ahrens B (2008) Precipitation by a regional climate model and bias correction in europe and south asia. *Meteorologische Zeitschrift*, 17(4):499–509, DOI 10.1127/0941-2948/2008/0306
- Dobler L, Hinterding A, Gerlach N (2004) INTERMET-Interpolation stündlicher und tagesbasierter meteorologischer Parameter. Gesamtdokumentation. Institut für Geoinformatik der Westfälischen Wilhelms-Universität Münster i.A. Landesamt für Wasserwirtschaft Rheinland-Pfalz
- Doms G, Förster J, Heise E, Herzog H, Raschendorfer M, Reinhardt T, Ritter B, Schrodin R, Schulz JP, Vogel G (2007) A Description of the Nonhydrostatic Regional COSMO-Model Part II: Physical Parameterization. Consortium for Small-scale Modelling, Deutscher Wetterdienst, Offenbach, Germany
- Déqué M (2007) Frequency of precipitation and temperature extremes over france in an anthropogenic scenario: Model results and statistical correction according to observed values. *Global Planet Change*, 57:16–26
- Ehret U, Zehe E, Wulfmeyer V, Warrach-Sagi K, Liebert J (2012) HESS opinions "Should we apply bias correction to global and regional climate model data?". *Hydrol Earth Syst Sci*, 16(9):3391–3404, DOI 10.5194/hess-16-3391-2012
- Ekström M, Fowler H, Kilsby C, Jones P (2005) New estimates of future changes in extreme rainfall across the UK using regional climate model integrations. 2. future estimates and use in impact studies. *Journal of Hydrology*, 300(1–4):234–251, DOI 10.1016/j.jhydrol.2004.06.019
- Fedderson H, Andersen U (2005) A method for statistical downscaling of seasonal ensemble predictions. *Tellus A*, 57(3), DOI 10.3402/tellusa.v57i3.14656
- Fischer EM, Schär C (2010) Consistent geographical patterns of changes in high-impact european heatwaves. *Nature Geoscience*, 3(6):398–403, DOI 10.1038/ngeo866
- Fischer EM, Rajczak J, Schär C (2012) Changes in european summer temperature variability revisited. *Geophysical Research Letters*, 39(19):L19,702, DOI 10.1029/2012GL052730
- Fisher RA, Tippett LHC (1928) Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Mathematical Proceedings of the Cambridge Philosophical Society*, 24(02):180–190, DOI 10.1017/S0305004100015681
- Foley AM (2010) Uncertainty in regional climate modelling: A review. *Progress in Physical Geography*, 34(5):647–670, DOI 10.1177/0309133310375654

- Fowler H, Ekström M, Kilsby C, Jones P (2005) New estimates of future changes in extreme rainfall across the UK using regional climate model integrations. 1. assessment of control climate. *Journal of Hydrology*, 300(1–4):212–233, DOI 10.1016/j.jhydrol.2004.06.017
- Fowler H, Blenkinsop S, Tebaldi C (2007) Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. *International Journal of Climatology*, 27(12):1547–1578, DOI 10.1002/joc
- Fowler HJ, Ekström M (2009) Multi-model ensemble estimates of climate change impacts on UK seasonal precipitation extremes. *International Journal of Climatology*, 29(3):385–416, DOI 10.1002/joc.1827
- Fowler HJ, Kilsby CG (2007) Using regional climate model data to simulate historical and future river flows in northwest England. *Climatic Change*, 80(3–4):337–367, DOI 10.1007/s10584-006-9117-3
- Fox-Rabinovitz M, Cote J, Dugas B, Deque M, McGregor JL, Belochitski A (2008) Stretched-grid model intercomparison project: decadal regional climate simulations with enhanced variable and uniform-resolution GCMs. *Meteorology and Atmospheric Physics*, 100(1–4):159–178, DOI 10.1007/s00703-008-0301-z
- Frei C, Christensen JH, Déqué M, Jacob D, Jones RG, Vidale PL (2003) Daily precipitation statistics in regional climate models: Evaluation and intercomparison for the European Alps. *Journal of Geophysical Research*, 108:19 pp., DOI 10.1029/2002JD002287
- Frei C, Schöll R, Fukutome S, Schmidli J, Vidale P (2006) Future change of precipitation extremes in Europe: Intercomparison of scenarios from regional climate models. *Journal of Geophysical Research: Atmospheres*, 111(D6), DOI 10.1029/2005JD005965
- Frich P, Alexander LV, DellaMarta P, Gleason B, Haylock M, Tank AMGK, Peterson T (2002) Observed coherent changes in climatic extremes during the second half of the twentieth century. *Climate Research*, 19(3):193–212, DOI 10.3354/cr019193
- Frigessi A, Haug O, Rue H (2003) A dynamic mixture model for unsupervised tail estimation without threshold selection. *Extremes*, 5(3):219–235
- Ghil M, Yiou P, Hallegatte S, Malamud BD, Naveau P, Soloviev A, Friederichs P, Keilis-Borok V, Kondrashov D, Kossobokov V, Mestre O, Nicolis C, Rust HW, Shebalin P, Vrac M, Witt A, Zaliapin I (2011) Extreme events: dynamics, statistics and prediction. *Nonlin Processes Geophys*, 18(3):295–350
- Giorgi F (1990) Simulation of regional climate using a limited area model nested in a general circulation model. *Journal of Climate*, 3(9):941–963, DOI 10.1175/1520-0442(1990)003<0941:SORCUA>2.0.CO;2
- Giorgi F, Mearns LO (1991) Approaches to the simulation of regional climate change: A review. *Reviews of Geophysics*, 29(2):191–216, DOI 10.1029/90RG02636
- Giorgi F, Jones C, Asrar GR (2009) Addressing climate information needs at the regional level: the CORDEX framework. *Journal of Climate*, 22(3):175–183
- Gnedenko B (1943) Sur la distribution limite du terme maximum d'une série aléatoire. *Annals of Mathematics*, 44(3):423–453, DOI 10.2307/1968974

- Graham LP, Andréasson J, Carlsson B (2007) Assessing climate change impacts on hydrology from an ensemble of regional climate models, model scales and linking methods – a case study on the lule river basin. *Climatic Change*, 81(S1):293–307, DOI 10.1007/s10584-006-9215-2
- Grasselt R, Schüttemeyer D, Warrach-Sagi K, Ament F, Simmer C (2008) Validation of TERRA-ML with discharge measurements. *Meteorologische Zeitschrift*, 17(6):763–773, DOI 10.1127/0941-2948/2008/0334
- Gudmundsson L, Bremnes JB, Haugen JE, Engen-Skaugen T (2012) Technical note: Downscaling RCM precipitation to the station scale using statistical transformations – a comparison of methods. *Hydrol Earth Syst Sci*, 16(9):3383–3390, DOI 10.5194/hess-16-3383-2012
- Guichard F, Petch JC, Redelsperger JL, Bechtold P, Chaboureau JP, Cheinet S, Grabowski W, Grenier H, Jones CG, Köhler M M, Piriou JM, Tailleux R, Tomasini M (2004) Modelling the diurnal cycle of deep precipitating convection over land with cloud-resolving models and single-column models. *Quarterly Journal of the Royal Meteorological Society*, 130(604):3139–3172, DOI 10.1256/qj.03.145
- Gumbel E (1958) *Statistics of Extremes*. Columbia University Press, New York, p. 375
- Gutjahr O, Heinemann G (2013) Comparing precipitation bias correction methods for high-resolution regional climate simulations using COSMO-CLM. *Theoretical and Applied Climatology*, 114(3):511–529, DOI 10.1007/s00704-013-0834-z
- Haddeland I, Heinke J, Voß F, Eisner S, Chen C, Hagemann S, Ludwig F (2012) Effects of climate model radiation, humidity and wind estimates on hydrological simulations. *Hydrology and Earth System Sciences*, 16(4):305–318, DOI 10.5194/hessd-8-7919-2011
- Haerter JO, Hagemann S, Moseley C, Piani C (2010) Climate model bias correction and the role of timescales. *Hydrology and Earth System Sciences Discussions*, 7(5):7863–7898, DOI 10.5194/hessd-7-7863-2010
- Hagemann S, Machenhauer B, Jones R, Christensen OB, Déqué M, Jacob D, Vidale PL (2004) Evaluation of water and energy budgets in regional climate models applied over europe. *Climate Dynamics*, 23(5):547–567, DOI 10.1007/s00382-004-0444-7
- Hagemann S, Chen C, Haerter JO, Heinke J, Gerten D, Piani C (2011) Impact of a statistical bias correction on the projected hydrological changes obtained from three GCMs and two hydrology models. *Journal of Hydrometeorology*, 12(4):556–578, DOI 10.1175/2011JHM1336.1
- Hansen JW, Challinor A, Ines A, Wheeler T, Moron V (2006) Translating climate forecasts into agricultural terms: advances and challenges. *Climate Research*, 33(1):27–41, DOI 10.3354/cr033027
- Hasselmann K (1976) Stochastic climate models part I. Theory. *Tellus*, 28(6):473–485, DOI 10.1111/j.2153-3490.1976.tb00696.x
- Hasselmann K (1993) Optimal fingerprints for the detection of time-dependent climate change. *Journal of Climate; (United States)*, 6:1957–1971
- Hawkins E, Sutton R (2009) The potential to narrow uncertainty in regional climate predictions. *Bulletin of the American Meteorological Society*, 90(8):1095–1107, DOI 10.1175/2009BAMS2607.1

- Hay LE, Clark MP (2003) Use of statistically and dynamically downscaled atmospheric model output for hydrologic simulations in three mountainous basins in the western united states. *Journal of Hydrology*, 282:56–75, DOI 10.1016/S0022-1694(03)00252-X
- Hay LE, Wilby RL, Leavesley GH (2000) A comparison of delta change and downscaled GCM scenarios for three mountainous basins in the united states. *Journal of the American Water Resources Association*, 36(2):387–397, DOI 10.1111/j.1752-1688.2000.tb04276.x
- Haylock MR, Hofstra N, Tank AMGK, Klok EJ, Jones PD, New M (2008) A european daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006. *Journal of Geophysical Research*, 113(D20):D20,119, DOI 10.1029/2008JD010201
- Heffernan JE, Tawn JA (2004) A conditional approach for multivariate extreme values (with discussion). *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 66(3):497–546, DOI 10.1111/j.1467-9868.2004.02050.x
- Hegerl G, Zwiers F (2011) Use of models in detection and attribution of climate change. *Wiley Interdisciplinary Reviews: Climate Change*, 2(4):570–591, DOI 10.1002/wcc.121
- Hegerl G, Zwiers F, Allen M, Marengo J, Barros V, Berliner M, Boer G, Crowley T, Folland C, Free M, Gillett N, Groisman P, Haigh J, Hasselmann K, Jones P, Knadlikar M, Kharin V, H K, Knutson T, MacCracken M, Mann M, North G, Risbey J, Robock A, Santer B, Schnur R, Schönwiese C, Sexton D, Stott P, Tett S, Vinnikov K, Wigley T (2001) Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, chap Detection of Climate Change and Attribution of Causes
- Hegerl G, Zwiers F, Braconnot P, Gillett N, Luo Y, Marengo J, Nicholls N, Penner J, Stott P (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, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, chap Understanding and attributing climate change
- Hegerl GC, von Storch H, Hasselmann K, Santer BD, Cubasch U, Jones PD (1996) Detecting greenhouse-gas-induced climate change with an optimal fingerprint method. *Journal of Climate*, 9(10):2281–2306, DOI 10.1175/1520-0442(1996)009<2281:DGGICC>2.0.CO;2
- Heise E, Ritter B, Schrodin R (2006) Operational implementation of the multilayer soil mode. Technical Report No. 9, Consortium for Small-Scale Modelling (COSMO), Deutscher Wetterdienst, Offenbach, Germany
- Hohenegger C, Brockhaus P, Schär C (2008) Towards climate simulations at cloud-resolving scales. *Meteorologische Zeitschrift*, 17(4):383–394, DOI 10.1127/0941-2948/2008/0303
- Hollweg H, Böhm U, Fast I, Hennemuth B, Keuler K, Keup-Thiel E, Lautenschlager M, Legutke S, Radtke K, Rockel B, Schubert M, Will A, Woldt M, Wunram C (2008) Ensemble simulations over europe with the regional climate model CLM forced with IPCC AR4 global scenarios. Tech. Report 3, Max-Planck-Institut für Meteorologie, Gruppe: Modelle & Daten

- Hosking JRM (1990) L-moments: Analysis and estimation of distributions using linear combinations of order statistics. *Journal of the Royal Statistical Society Series B (Methodological)*, 52(1):105–124, DOI 10.2307/2345653
- Ines AV, Hansen JW (2006) Bias correction of daily GCM rainfall for crop simulation studies. *Agricultural and Forest Meteorology*, 138(1-4):44–53, DOI 10.1016/j.agrformet.2006.03.009
- 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, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, chap Summary for Policymakers.
- IPCC (2012) Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation. A Special Report of Working Group I and II of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK and New York, USA
- Jones PW (1999) First- and second-order conservative remapping schemes for grids in spherical coordinates. *Monthly Weather Review*, 127(9):2204–2210, DOI 10.1175/1520-0493(1999)127<2204:FASOCR>2.0.CO;2
- Jungclaus JH, Keenlyside N, Botzet M, Haak H, Luo JJ, Latif M, Marotzke J, Mikolajewicz U, Roeckner E (2006) Ocean circulation and tropical variability in the coupled model ECHAM5/MPI-OM. *Journal of Climate*, 19 (16):3952–3972
- Kantz H, Altmann EG, Hallerberg S, Holstein D, Riegert A (2006) Dynamical interpretation of extreme events: Predictability and predictions. In: Albeverio PDS, Jentsch DV, Kantz PDH (eds) *Extreme Events in Nature and Society*, The Frontiers Collection, Springer Berlin Heidelberg, pp 69–93
- Katz R (2010) Statistics of extremes in climate change. *Climatic Change*, 100:71–76, DOI 10.1007/s10584-010-9834-5
- Katz RW, Brown BG (1992) Extreme events in a changing climate: Variability is more important than averages. *Climatic Change*, 21(3):289–302, DOI 10.1007/BF00139728
- Kendon EJ, Roberts NM, Senior CA, Roberts MJ (2012) Realism of rainfall in a very high-resolution regional climate model. *Journal of Climate*, 25(17):5791–5806, DOI 10.1175/JCLI-D-11-00562.1
- Kessler (1969) On the distribution and continuity of water substance in the atmospheric circulations, *Meteorological Monographs*, vol 10. American Meteorological Society, Boston
- Keuler K, Radtke K, Georgievski G (2012) Summary of evaluation results for cosmo-clm version 4.8.clm13 (clm17): Comparison of three different configurations over europe driven by ecmwf reanalysis data era40 for the period 1979–2000. Tech. rep., Brandenburg University of Technology, Cottbus
- Kharin VV, Zwiers FW (2000) Changes in the extremes in an ensemble of transient climate simulations with a coupled atmosphere–ocean GCM. *Journal of Climate*, 13:3760–3788
- Kharin VV, Zwiers FW (2005) Estimating extremes in transient climate change simulations. *Journal of Climate*, 18(8):1156, DOI 10.1175/JCLI3320.1

- Kjellström E, Barring L, Jacob D, Jones R, Lenderink G, Schär C (2007) Modelling daily temperature extremes: recent climate and future changes over europe. *Climatic Change*, 81(1):249–265, DOI 10.1007/s10584-006-9220-5
- Klein Tank AMG, Können GP (2003) Trends in indices of daily temperature and precipitation extremes in europe, 1946-99. *Journal of Climate*, 16:3665–3680
- Knote C, Heinemann G, Rockel B (2010) Changes in weather extremes: Assessment of return values using high resolution climate simulations at convection-resolving scale. *Meteorologische Zeitschrift*, 19:11–23
- Lafon T, Dadson S, Buys G, Prudhomme C (2013) Bias correction of daily precipitation simulated by a regional climate model: a comparison of methods. *International Journal of Climatology*, 33(6):1367–1381, DOI 10.1002/joc.3518
- Lang M, Ouarda T, Bobée B (1999) Towards operational guidelines for over-threshold modeling. *Journal of Hydrology*, 225:103–117, DOI 10.1016/S0022-1694(99)00167-5
- Leadbetter MR, Weissmann I, de Haan L, Rootzén H (1989) On clustering of high values in statistically stationary series. Tech. Rep. 253, Center for Stochastic Processes, University of North Carolina, Chapel Hill
- Leander R, Buishand T (2007) Resampling of regional climate model output for the simulation of extreme river flows. *Journal of Hydrology*, 332(3-4):487–496
- Lenderink G, Buishand A, van Deursen W (2007) Estimates of future discharges of the river rhine using two scenario methodologies: direct versus delta approach. *Hydrol Earth Syst Sci*, 11(3):1145–1159, DOI 10.5194/hess-11-1145-2007
- Li H, Sheffield J, Wood EF (2010) Bias correction of monthly precipitation and temperature fields from intergovernmental panel on climate change AR4 models using equidistant quantile matching. *Journal of Geophysical Research: Atmospheres*, 115(D10), DOI 10.1029/2009JD012882
- van der Linden P, Mitchell J (eds) (2009) ENSEMBLES: Climate Change and its Impacts: Summary of research and results from the ENSEMBLES project. Met Office Hadley Centre, FitzRoy Road, Exeter EX1 3PB, UK
- Livezey R, Chen W (1983) Statistical field significance and its determination by monte carlo techniques. *Monthly Weather Review*, 111:46–59
- Lorenz P, Jacob D (2005) Influence of regional scale information on the global circulation: A two-way nesting climate simulation. *Geophysical Research Letters*, 32(18):L18,706, DOI 10.1029/2005GL023351
- Luterbacher J, Dietrich D, Xoplaki E, Grosjean M, Wanner H (2004) European seasonal and annual temperature variability, trends, and extremes since 1500. *Science*, 303(5663):1499–1503, DOI 10.1126/science.1093877
- Madden RA (1976) Estimates of the natural variability of time-averaged sea-level pressure. *Monthly Weather Review*, 104(7):942–952, DOI 10.1175/1520-0493(1976)104<0942:EOTNVO>2.0.CO;2

- Maraun D (2013) Bias correction, quantile mapping, and downscaling: Revisiting the inflation issue. *Journal of Climate*, 26(6):2137–2143, DOI 10.1175/JCLI-D-12-00821.1
- Maraun D, Wetterhall F, Ireson A, Chandler R, Kendon E, Widmann M, Brienen S, Rust H, Sauter T, Themeßl M, Venema V, Chun K, Goodess C, Jones C RGand Onof, Vrac M, Thiele-Eich I (2010) Precipitation downscaling under climate change. recent developments to bridge the gap between dynamical models and the end user. *Reviews of Geophysics*, 48 (RG3003):1–34, DOI 10.1029/2009RG000314.
- Martins ES, Stedinger JR (2000) Generalized maximum-likelihood generalized extreme-value quantile estimators for hydrologic data. *Water Resources Research*, 36(3):737–744, DOI 200010.1029/1999WR900330
- Masson V, Champeaux JL, Chauvin F, Meriguet C, Lacaze R (2003) A global database of land surface parameters at 1-km resolution in meteorological and climate models. *Journal of Climate*, 16(9):1261–1282, DOI 10.1175/1520-0442-16.9.1261
- Maurer EP, Brekke L, Pruitt T, Duffy PB (2007) Fine-resolution climate projections enhance regional climate change impact studies. *Eos, Transactions American Geophysical Union*, 88(47):504–504, DOI 10.1029/2007EO470006
- McKee T, Doeskin N, Kleist J (1993) The relationship of drought frequency and duration to time scales. pp 179–184
- Meehl GA, Covey C, Taylor KE, Delworth T, Stouffer RJ, Latif M, McAvaney B, Mitchell JFB (2007) THE WCRP CMIP3 multimodel dataset: A new era in climate change research. *Bulletin of the American Meteorological Society*, 88(9):1383–1394, DOI 10.1175/BAMS-88-9-1383
- Mellor GL, Yamada T (1974) A hierarchy of turbulence closure models for planetary boundary layers. *Journal of the Atmospheric Sciences*, 31(7):1791–1806, DOI 10.1175/1520-0469(1974)031<1791:AHOTCM>2.0.CO;2
- Michelangeli PA, Vrac M, Loukos H (2009) Probabilistic downscaling approaches: Application to wind cumulative distribution function. *Geophysical Research Letters*, 36:L11,708, DOI doi:10.1029/2009GL038401
- Mizuta R, Kazuyoshi O, Hiromasa Y, Akira N, Keiichi K, Seiji Y, Masahiro H, Shoji K, Hideaki K, Masayuki N (2006) 20-km-mesh global climate simulations using JMA-GSM model-mean climate states. *J Meteorol Soc Jpn*, 84(1):165–185
- Mudelsee M (2011) In *Extremis*, Springer Berlin Heidelberg, chap The Bootstrap in Climate Risk Analysis, pp 44–58. DOI 10.1007/978-3-642-14863-7_2
- Muerth MJ, Gauvin St-Denis B, Ricard S, Velázquez JA, Schmid J, Minville M, Caya D, Chaumont D, Ludwig R, Turcotte R (2013) On the need for bias correction in regional climate scenarios to assess climate change impacts on river runoff. *Hydrol Earth Syst Sci*, 17(3):1189–1204, DOI 10.5194/hess-17-1189-2013
- O’Gorman PA, Muller CJ (2010) How closely do changes in surface and column water vapor follow Clausius–Clapeyron scaling in climate change simulations? *Environmental Research Letters*, 5(2):025,207, DOI 10.1088/1748-9326/5/2/025207

- Pandey MD, van Gelder PHAJM, K VJ (2001) The Use of L-Moments in the Peak Over Threshold Approach for Estimating Extreme Quantiles of Wind Velocity
- Panofsky H, Brier G (1968) Some applications of statistics to meteorology. *The Pennsylvania State University Press*, p 224
- Piani C, Haerter JO (2012) Two dimensional bias correction of temperature and precipitation copulas in climate models. *Geophysical Research Letters*, 39(20):L20,401, DOI 10.1029/2012GL053839
- Piani C, Haerter J, Coppola E (2010a) Statistical bias correction for daily precipitation in regional climate models over europe. *Theor Appl Climatol*, 99:187–192
- Piani C, Weedon G, Best M, Gomes S, Viterbo P, Hagemann S, Haerter J (2010b) Statistical bias correction of global simulated daily precipitation and temperature for the application of hydrological models. *Journal of Hydrology*, 395(3-4):199–215, DOI 10.1016/j.jhydrol.2010.10.024
- Pickands J (1975) Statistical inference using extreme order statistics. *The Annals of Statistics*, 3(1):119–131, DOI 10.2307/2958083
- Prein AF, Gobiet A, Suklitsch M, Truhetz H, Awan NK, Keuler K, Georgievski G (2013) Added value of convection permitting seasonal simulations. *Climate Dynamics*, DOI 10.1007/s00382-013-1744-6
- Rasch D, Guiard V (2004) The robustness of parametric statistical methods. *Psychology Science*, 46:175–208
- Rauscher S, Coppola E, Piani C, Giorgi F (2010) Resolution effects on regional climate model simulations of seasonal precipitation over europe. *Climate Dynamics*, 35(4):685–711, DOI 10.1007/s00382-009-0607-7
- Rauthe M, Steiner H, Riediger U, Mazurkiewicz A, Gratzki A (2013) A central european precipitation climatology - part i: Generation and validation of a high-resolution gridded daily data set (HYRAS). *Meteorologische Zeitschrift*, in print
- Resnick S (1987) Extreme values, regular variation, and Point Processes, Applied Probability, vol 4. Springer, New York, pp 334
- Ribes A, Azaïs JM, Planton S (2009) Adaptation of the optimal fingerprint method for climate change detection using a well-conditioned covariance matrix estimate. *Climate Dynamics*, 33(5):707–722, DOI 10.1007/s00382-009-0561-4
- Richter D (ed) (1995) Ergebnisse methodischer Untersuchungen zur Korrektur des systematischen Meßfehlers des Hellmann-Niederschlagsmessers, Berichte des Deutschen Wetterdienstes, vol 194, Deutscher Wetterdienst, Offenbach
- Ritter B, Geleyn JF (1992) A comprehensive radiation scheme for numerical weather prediction models with potential applications in climate simulations. *Monthly Weather Review*, 120(2):303–325, DOI 10.1175/1520-0493(1992)120<0303:ACRSFN>2.0.CO;2
- Rockel B, Will A, Hense A (2008) The regional climate model COSMO-CLM (CCLM). *Meteorologische Zeitschrift*, 17(4):347–348

- Roeckner E, Bäuml G, Bonaventura L, Esch M, Giorgetta M, Hagemann S, Kirchner I, Kornblüeh L, Manzini E, Rhodin A, Schlese U, Schulzweida U, Tompkins A (2003) The atmospheric general circulation model echam 5. part i: Model description. Tech. Rep. 349, Max-Planck-Institut für Meteorologie
- Rojas R, Feyen L, Dosio A, Bavera D (2011) Improving pan-European hydrological simulation of extreme events through statistical bias correction of RCM-driven climate simulations. *Hydrol Earth Syst Sci*, 15(8):2599–2620, DOI 10.5194/hess-15-2599-2011
- van Roosmalen L, Sonnenborg TO, Jensen KH, Christensen J (2011) Comparison of hydrological simulations of climate change using perturbation of observations and Distribution-Based scaling. *Vadose Zone Journal*, 10(1):136–150, DOI 10.2136/vzj2010.0112
- Rummukainen M (2010) State-of-the-art with regional climate models. *Wiley Interdisciplinary Reviews: Climate Change*, 1(1):82–96, DOI 10.1002/wcc.8
- Schmidli J, Frei C, Vidale PL (2006) Downscaling from GCM precipitation: a benchmark for dynamical and statistical downscaling methods. *International Journal of Climatology*, 26(5):679–689, DOI 10.1002/joc.1287
- Schneider EK, Iii JLK (1994) An examination of internally generated variability in long climate simulations. *Climate Dynamics*, 10(4-5):181–204, DOI 10.1007/BF00208987
- Schoetter R, Hoffmann P, Rechid D, Schlünzen KH (2012) Evaluation and bias correction of regional climate model results using model evaluation measures. *Journal of Applied Meteorology and Climatology*, 51(9):1670–1684, DOI 10.1175/JAMC-D-11-0161.1
- Schär C, Vidale PL, Lüthi D, Frei C, Häberli C, Liniger MA, Appenzeller C (2004) The role of increasing temperature variability in european summer heatwaves. *Nature*, 427(6972):332–336, DOI 10.1038/nature02300
- Schrodin R, Heise E (2001) The Multi-Layer version of the DWD soil model TERRA-ML. Technical Report No. 2, Consortium for Small-Scale Modelling (COSMO), Deutscher Wetterdienst, Offenbach, Germany
- Schättler U, Doms G, Schraff C (2009) A Description of the Nonhydrostatic Regional COSMO-Model Part VII: User's Guide. Consortium for Small-Scale Modelling, cosmo Modell 4.11
- Seneviratne SI, Nicholls N, Easterling D, Goodess CM, Kanae S, Kossin J, Luo Y, Marengo J, McInnes K, Rahimi M, Reichstein M, Sorteberg A, Vera C, Zhang X (2012) Changes in climate extremes and their impacts on the natural physical environment: An overview of the IPCC SREX report. In: Field BV CB, Stocker T, Qin D, Dokken D, Ebi K, Mastrandrea M, Mach K, Plattner GK, Allen S, Tignor M, Midgley P (eds) *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation, A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change (IPCC)*., Cambridge University Press, Cambridge/New York, pp 109–230
- Sennikovs J, Bethers U (2009) Statistical downscaling method of regional climate model results for hydrological modelling. In: 18th World IMACS/MODSIM Congress, Australia
- Sharma D, Das Gupta A, Babel MS (2007) Spatial disaggregation of bias-corrected GCM precipitation for improved hydrologic simulation: Ping river basin, thailand. *Hydrol Earth Syst Sci*, 11(4):1373–1390

- Sillmann J, Roeckner E (2008) Indices for extreme events in projections of anthropogenic climate change. *Climatic Change*, 86(1-2):83–104, DOI 10.1007/s10584-007-9308-6
- Sillmann J, Kharin VV, Zhang X, Zwiers FW, Bronaugh D (2013a) Climate extremes indices in the CMIP5 multimodel ensemble: Part 1. model evaluation in the present climate. *Journal of Geophysical Research: Atmospheres*, 118(4):1716–1733, DOI 10.1002/jgrd.50203
- Sillmann J, Kharin VV, Zwiers FW, Zhang X, Bronaugh D (2013b) Climate extremes indices in the CMIP5 multimodel ensemble: Part 2. future climate projections. *Journal of Geophysical Research: Atmospheres*, 118(6):2473–2493, DOI 10.1002/jgrd.50188
- Simiu E, Heckert NA (1996) Extreme wind distribution tails: A "peaks over threshold" approach. *Journal of Structural Engineering*, 122(5):539–547, DOI 10.1061/(ASCE)0733-9445(1996)122:5(539)
- Smiatek G, Rockel B, Schättler U (2008) Time invariant data preprocessor for the climate version of the COSMO model (COSMO-CLM). *Meteorologische Zeitschrift*, 17(4):395–405, DOI 10.1127/0941-2948/2008/0302
- Solomon S, Qin D, Manning M, Alley R, Berntsen T, Bindoff N, Chen Z, Chidthaisong A, Gregory J, Hegerl G, Heimann M, Hewitson B, Hoskins B, Joos F, Jouzel J, Kattsov V, Lohmann U, Matsuno T, Molina M, Nicholls N, Overpeck J, Raga G, Ramaswamy V, Ren J, Rusticucci M, Somerville R, Stocker T, P W, Wood R, Wratt D (2007a) Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, chap Technical Summary
- Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt K, Tignor M, Miller H (eds) (2007b) 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. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA
- Stensrud DJ (2007) Parameterization schemes. Keys to Understanding Numerical Weather Prediction Models. Cambridge University Press, Cambridge, pp. 459
- Stephenson DB (2008) Definition, diagnosis, and origin of extreme weather and climate events. In: Climate Extremes and Society, Cambridge University Press
- Stappeler J, Doms G, Schättler U, Bitzer HW, Gassmann A, Damrath U, Gregoric G (2003) Mesogamma scale forecasts using the nonhydrostatic model LM. *Meteorology and Atmospheric Physics*, 82(1-4):75–96, DOI 10.1007/s00703-001-0592-9
- Stone DA, Allen MR, Selten F, Kliphuis M, Stott PA (2007) The detection and attribution of climate change using an ensemble of opportunity. *Journal of Climate*, 20(3):504–516, DOI 10.1175/JCLI3966.1
- von Storch H, Navarra A (eds) (1999) Analysis of Climate Variability: Applications of Statistical Techniques, 2nd edn. Springer Berlin
- Stott PA, Tett SFB, Jones GS, Allen MR, Ingram WJ, Mitchell JFB (2001) Attribution of twentieth century temperature change to natural and anthropogenic causes. *Climate Dynamics*, 17(1):1–21, DOI 10.1007/PL00007924

- Stott PA, Mitchell JFB, Allen MR, Delworth TL, Gregory JM, Meehl GA, Santer BD (2006) Observational constraints on past attributable warming and predictions of future global warming. *Journal of Climate*, 19(13):3055–3069, DOI 10.1175/JCLI3802.1
- Taylor KE (2001) Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research: Atmospheres*, 106(D7):7183–7192, DOI 10.1029/2000JD900719
- Taylor KE, Stouffer RJ, Meehl GA (2012) An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society*, 93(4):485–498, DOI 10.1175/BAMS-D-11-00094.1
- Tebaldi C, Knutti R (2007) The use of the multi-model ensemble in probabilistic climate projections. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 365(1857):2053–2075, DOI 10.1098/rsta.2007.2076
- Tebaldi C, Hayhoe K, Arblaster JM, Meehl GA (2006) Going to the extremes. an intercomparison of model-simulated historical and future changes in extreme events. *Climatic Change*, 79(3-4):185–211, DOI 10.1007/s10584-006-9051-4
- Tett SFB, Stott PA, Allen MR, Ingram WJ, Mitchell JFB (1999) Causes of twentieth-century temperature change near the earth's surface. *Nature*, 399(6736):569–572, DOI 10.1038/21164
- Teutschbein C, Seibert J (2010) Regional climate models for hydrological impact studies at the catchment scale: A review of recent modeling strategies. *Geography Compass*, 4(7):834–860, DOI 10.1111/j.1749-8198.2010.00357.x
- Teutschbein C, Seibert J (2012) Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods. *Journal of Hydrology*, 456–457:12–29, DOI 10.1016/j.jhydrol.2012.05.052
- Thiemeßl M, Gobiet A, Heinrich G (2012) Empirical-statistical downscaling and error correction of regional climate models and its impact on the climate change signal. *Climatic Change*, 112(2):449–468
- Thiemeßl MJ, Gobiet A, Leuprecht A (2011) Empirical-statistical downscaling and error correction of daily precipitation from regional climate models. *International Journal of Climatology*, DOI 10.1002/joc.2168
- Thom HCS (1968) Approximate convolution of the gamma and mixed gamma distributions. *Monthly Weather Review*, 96(12):883–886, DOI 10.1175/1520-0493(1968)096<0883:ACOTGA>2.0.CO;2
- Tiedtke M (1989) A comprehensive mass flux scheme for cumulus parameterization in large-scale models. *Monthly Weather Review*, 117:1779–1800
- Trenberth K, Dai A, Rasmussen R, Parsons D (2003) The changing character of precipitation. *Bull Amer Meteor Soc*, 84:1205–1217, DOI 10.1175/BAMS-84-9-1205
- Trenberth K, Jones P, Ambenje R, Bojariu R, Easterling D, Klein Tank A, Parker D, Rahimzadeh F, Renwick J, Rusticucci M, Soden B, Zhai P (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, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, chap Observations: Surface and Atmospheric Climate Change

- Trömel S, Schönwiese CD (2005) A generalized method of time series decomposition into significant components including probability assessments of extreme events and application to observational german precipitation data. *Meteorologische Zeitschrift*, 14(3):417–427, DOI 10.1127/0941-2948/2005/0039
- Uppala SM, Kållberg PW, Simmons AJ, Andrae U, Bechtold VDC, Fiorino M, Gibson JK, Haseler J, Hernandez A, Kelly GA, Li X, Onogi K, Saarinen S, Sokka N, Allan RP, Andersson E, Arpe K, Balmaseda MA, Beljaars ACM, Berg LVD, Bidlot J, Bormann N, Caires S, Chevallier F, Dethof A, Dragosavac M, Fisher M, Fuentes M, Hagemann S, Hólm E, Hoskins BJ, Isaksen L, Janssen PaEM, Jenne R, McNally AP, Mahfouf JF, Morcrette JJ, Rayner NA, Saunders RW, Simon P, Sterl A, Trenberth KE, Untch A, Vasiljevic D, Viterbo P, Woollen J (2005) The ERA-40 re-analysis. *Quarterly Journal of the Royal Meteorological Society*, 131(612):2961–3012, DOI 10.1256/qj.04.176
- Ventura V, Paciorek C, Risbey J (2004) Controlling the proportion of falsely rejected hypotheses when conducting multiple tests with climatological data. *Journal of Climate*, 17:4343–4356
- Venzon DJ, Moolgavkar SH (1988) A method for computing profile-likelihood-based confidence intervals. *Journal of the Royal Statistical Society Series C (Applied Statistics)*, 37(1):87–94, DOI 10.2307/2347496
- Vicente-Serrano SM, Beguería S, López-Moreno JI (2010) A multiscale drought index sensitive to global warming: The standardized precipitation evapotranspiration index. *Journal of Climate*, 23(7):1696–1718, DOI 10.1175/2009JCLI2909.1
- Vidale PL, Lüthi D, Wegmann R, Schär C (2007) European summer climate variability in a heterogeneous multi-model ensemble. *Climatic Change*, 81(1):209–232, DOI 10.1007/s10584-006-9218-z
- Vlcek O, Huth R (2009) Is daily precipitation gamma-distributed?: Adverse effects of an incorrect use of the Kolmogorov-Smirnov test. *Atmospheric Research*, 93(4):759–766, DOI 16/j.atmosres.2009.03.005
- Vrac M, Naveau P (2007) Stochastic downscaling of precipitation: From dry events to heavy rainfalls. *Water Resources Research*, 43:13 PP., DOI 200710.1029/2006WR005308
- Vrac M, Naveau P, Drobinski P (2007) Modeling pairwise dependencies in precipitation intensities. *Nonlin Processes Geophys*, 14(6):789–797
- Wehner M (2010) Sources of uncertainty in the extreme value statistics of climate data. *Extremes*, 13(2):205–217, DOI 10.1007/s10687-010-0105-7
- Weigel AP, Liniger MA, Appenzeller C (2008) Can multi-model combination really enhance the prediction skill of probabilistic ensemble forecasts? *Quarterly Journal of the Royal Meteorological Society*, 134(630):241–260, DOI 10.1002/qj.210
- Wernli H, Paulat M (2008) Sal - a novel quality measure for the verification of quantitative precipitation forecast. *Monthly Weather Review*, 136:4470–4487
- Widmann M, Bretherton CS, Salathe EPJ (2003) Statistical precipitation downscaling over the north-western united states using numerically simulated precipitation as a predictor. *Journal of Climate*,
- Wilks DS (2006) On field significance and the false discovery rate. *Journal of Applied Meteorology and Climatology*, 45:1181–1189

- Wilks DS (2011) *Statistical Methods in the Atmospheric Sciences*, vol 100, 3rd edn. Academic Press, ISBN:0123850223
- WMO (2011) *Guide to Climatological Practices*. World Meteorological Organization, Geneva, Switzerland, 3rd edn
- Wood A, Leung L, Sridhar V, Lettenmaier D (2004) Hydrologic implications of dynamical and statistical approaches to downscaling climate outputs. *Climatic Change*, 62:189–216
- Yang W, Andréasson J, Graham LP, Olsson J, Rosberg J, Wetterhall F (2010) Distribution-based scaling to improve usability of regional climate model projections for hydrological climate change impacts studies. *Hydrology Research*, 41(3-4):211–229, DOI 10.2166/nh.2010.004
- Zhang X, Zwiers FW (2013) Statistical indices for the diagnosing and detecting changes in extremes. In: AghaKouchak A, Easterling D, Hsu K, Schubert S, Sorooshian S (eds) *Extremes in a Changing Climate*, no. 65 in Water Science and Technology Library, Springer Netherlands, pp 1–14
- Zhang X, Hegerl G, Zwiers FW, Kenyon J (2005) Avoiding inhomogeneity in percentile-based indices of temperature extremes. *Journal of Climate*, 18(11):1641–1651, DOI 10.1175/JCLI3366.1
- Zhang X, Alexander L, Hegerl GC, Jones P, Tank AK, Peterson TC, Trewin B, Zwiers FW (2011) Indices for monitoring changes in extremes based on daily temperature and precipitation data. *Wiley Interdisciplinary Reviews: Climate Change*, 2(6):851–870, DOI 10.1002/wcc.147
- Zwiers FW, Kharin VV (1998) Changes in the extremes of the climate simulated by CCC GCM2 under CO₂ doubling. *Journal of Climate*, 11(9):2200–2222, DOI 10.1175/1520-0442(1998)011<2200:CITEOT>2.0.CO;2
- Zwiers FW, Zhang X (2003) Toward regional-scale climate change detection. *Journal of Climate*, 16(5):793–797, DOI 10.1175/1520-0442(2003)016<0793:TRSCCD>2.0.CO;2

Acknowledgements

Without the help of many people, this thesis would not lay in front of you. My first thanks are directed to my supervisor Prof. Dr. Günther Heinemann, head of the Environmental Meteorology department, whom I am very grateful for offering me this possibility and whose expertise and guidance ensured a profound framework for my research. In a same way I thank my second supervisor Prof. Dr. Markus Casper for his helpful input and discussions, in particular concerning intersections of meteorological and hydrological modelling, which allowed me to get a deeper insight into interdisciplinary research.

I also like to thank all of my colleagues in the department of Environmental Meteorology. In particular, Lukas Schefczyk who was mainly involved in the configuration and running of the COSMO-CLM and the technical implementation of the analysis code, Clemens Drüe and Uwe Baltes for the maintenance of the servers on which my programs have been running and Anne Menkhaus for enduring my absent-minded tendencies and my "no more than 5-word" phases. I am also thankful to Philipp Reiter for the long discussions on the methodologies and results, which helped me to stick to a red thread while writing.

Most of all I would like to thank my lifetime partner Anne Schneibel who always watched my back while I was working on this thesis. I know this was not at all a matter of course and therefore my thank is truly heartfelt. I am also deeply grateful for the support from my parents Norbert and Notburga Gutjahr and from Hermann Schneibel.

Special thanks go to the CLM-Community for the general help on the CCLM model and for compensating parts of my travel costs during the last three years. I thank Hans-Jürgen Panitz, Burkhardt Rockel and Andreas Will for their advices on the model configuration, Klaus Keuler und Kai Radtke for providing the CLM18 data sets, the CRCS-group of the CLM-COMmunity for intensive discussions on cloud-permitting climate simulations and statistics and Merja Tölle for providing the REGNIE data set. Furthermore, I like to thank Oliver Gronz for providing the INTERMET data set.

Finally, I acknowledge the E-OBS dataset from the EU-FP6 project ENSEMBLES (<http://ensembles-eu.metoffice.com>) and the data providers in the ECA&D project (<http://www.ecad.eu>) as well as the DKRZ in Hamburg for providing the required computing time and storage capacities.

Curriculum Vitae

Oliver Gutjahr
gutjahr@uni-trier.de
geboren am 3. Oktober 1983
in Worms, Deutschland

Ruwerstraße 3a
54317 Gutweiler
Deutschland

Berufstätigkeit

- | | |
|------------|---|
| 2010–2013 | Wissenschaftlicher Mitarbeiter an der Universität Trier im Projekt Global Change, Fach Umweltmeteorologie . |
| 2013–heute | Wissenschaftlicher Mitarbeiter an der Universität Trier im Projekt "BMBF-Projekt Laptew-See Transdrift", Fach Umweltmeteorologie. |

Hochschulausbildung

- | | |
|-----------------|---|
| 10/2010–08/2013 | Doktorant an der Universität Trier im Fach Umweltmeteorologie unter Betreuung von Univ.-Prof. Dr. Günther Heinemann. |
| 10/2005–10/2010 | Studium: Angewandte Physische Geographie (Diplom) an der Universität Trier. Diplomarbeit wurde im Fach Umweltmeteorologie geschrieben; Betreuende: Prof. Dr. Günther Heinemann und Prof. Dr. Markus Casper. |
| 10/2004–10/2005 | Studium: Geographie (Diplom) and der Universät Frankfurt a.M., Deutschland. |

Schulbildung

- | | |
|---------|---|
| 07/2003 | Abitur, Lessing Gymnasium Lampertheim, Deutschland. |
|---------|---|

Lehrtätigkeit

- | | |
|---------|---|
| SS 2013 | Numerical Modelling in Meteorology I: Dynamics |
| WS 2011 | Numerical Modelling in Meteorology II: Applications |
| SS 2011 | Numerical Modelling in Meteorology I: Dynamics |

Konferenzbeiträge

- | | |
|---------|---|
| 09/2013 | DACH 2013 (Vortrag), Innsbruck, Österreich |
| 03/2013 | COSMO user seminar (Poster), Offenbach, Deutschland |
| 02/2013 | DKRZ user workshop (Poster), Hamburg, Deutschland |
| 09/2012 | 7th CLM Assembly (Vortrag), Leuven, Belgien |
| 09/2012 | 2nd Conference on Modelling Hydrology, Climate and Land Surface Processes (Vortrag), Oslo, Norwegen |
| 04/2012 | European Geoscience Union (EGU) (Poster), Wien, Österreich |
| 03/2012 | COSMO user seminar (Vortrag), Offenbach, Deutschland |

04/2011	European Geoscience Union (EGU) (Poster), Wien, Österreich
09/2011	6th CLM Assembly (Poster), Cava de Tirreni, Italien
03/2011	COSMO user seminar (Poster), Langen, Deutschland
09/2010	5th CLM Assembly (Poster), Berlin, Deutschland

Workshops

SS 2011	Academic english certificate, Universität Trier, Deutschland
02/2010	COSMO/CLM Spring School (workshop), Langen, Deutschland

Begutachtete Publikationen

2014	Tölle M.H., Gutjahr O., Thiele G. & Busch G. (2014): Potential impact of bioenergy plantations on the future regional summer climate in central Germany. <i>J. Geophys. Res.</i> , 119, 2711–2724, DOI:10.1002/2013JD020877.
2013	Gutjahr O. & Heinemann G. (2013): Comparing precipitation bias correction methods for high-resolution regional climate simulations using COSMO-CLM. <i>Theor. and Appl. Climatol.</i> , 1–19, DOI: 10.1007/s00704-013-0834-z
2012	Casper M. C., Grigoryan G., Gronz O., Gutjahr O., Heinemann G., Ley R. & Rock A. (2012): Analysis of projected hydrological behavior of catchments based on signature indices. <i>Hydrol. Earth Syst. Sci.</i> , 16, 409–421, DOI: 10.5194/hess-16-409-2012.