

---

# Downscaling of Precipitation in the Upper Danube Catchment Area

Janus Willem Schipper

---



München 2005



---

# **Downscaling of Precipitation in the Upper Danube Catchment Area**

**Janus Willem Schipper**

---

Dissertation  
an der Fakultät für Physik  
der Ludwig-Maximilians-Universität  
München

vorgelegt von  
Janus Willem Schipper  
aus Amersfoort, Niederlande

München, den 10. Mai 2005

Erstgutachter: Prof. Dr. J. Egger

Zweitgutachter: Priv. Doz. Dr. G. Zängl

Tag der mündlichen Prüfung: 8. Juli 2005

*“Anyone who says sunshine brings happiness has never danced in the rain.”*  
(anonymous)



# Contents

<b>Summary</b>	<b>xiii</b>
<b>1 Introduction</b>	<b>1</b>
<b>2 Literature review</b>	<b>5</b>
<b>3 Sensitivity of MM5 precipitation to various physical parameterizations</b>	<b>9</b>
3.1 Introduction . . . . .	9
3.2 Model description, data and methodology . . . . .	9
3.2.1 Description of the model configurations . . . . .	10
3.2.2 Observational data and selected time periods . . . . .	12
3.3 Results . . . . .	14
3.4 Conclusions . . . . .	16
<b>4 Description of rainfall variability</b>	<b>19</b>
4.1 Introduction . . . . .	19
4.2 Linear average . . . . .	22
4.3 Fourier analysis . . . . .	23
4.3.1 Introduction . . . . .	23
4.3.2 MM5 simulations . . . . .	25
4.3.3 Observations . . . . .	26
4.3.4 Comparing MM5 with observations . . . . .	26
4.4 Running average . . . . .	28
4.4.1 Introduction . . . . .	28
4.4.2 MM5 simulations . . . . .	29
4.4.3 Observations . . . . .	32
4.4.4 Comparing MM5 with observations . . . . .	34
4.5 Spline interpolation . . . . .	36
4.5.1 Introduction . . . . .	36
4.5.2 MM5 simulations . . . . .	37
4.5.3 Observations . . . . .	39
4.6 Intercomparison . . . . .	39
<b>5 Empirical downscaling</b>	<b>43</b>
5.1 Introduction . . . . .	43
5.2 Small orographic details . . . . .	44
5.2.1 Introduction . . . . .	44

5.2.2	Correction . . . . .	46
5.2.3	Example . . . . .	47
5.3	General rainfall distribution . . . . .	50
5.3.1	Introduction . . . . .	50
5.3.2	Correction . . . . .	51
5.3.3	Example . . . . .	51
5.4	Total . . . . .	52
5.4.1	Introduction . . . . .	52
5.4.2	Correction . . . . .	53
5.4.3	Example . . . . .	55
5.5	Results . . . . .	56
<b>6</b>	<b>Dynamical empirical downscaling</b>	<b>59</b>
6.1	Introduction . . . . .	59
6.2	Wind distribution . . . . .	61
6.3	Rainfall distribution . . . . .	61
6.3.1	Rainfall classes . . . . .	61
6.3.2	Rainfall groups . . . . .	64
6.4	Results . . . . .	66
<b>7</b>	<b>Conclusions and outlook</b>	<b>69</b>
7.1	Conclusions . . . . .	69
7.2	Outlook . . . . .	71
<b>A</b>	<b>MM5 model description</b>	<b>73</b>
<b>B</b>	<b>Observations</b>	<b>77</b>
<b>C</b>	<b>Statistics</b>	<b>79</b>
C.1	Mean ( $\mu$ ) . . . . .	79
C.2	Variance ( $\sigma^2$ ) . . . . .	79
C.3	Skewness ( $\gamma_1$ ) . . . . .	79
C.4	Kurtosis ( $\gamma_2$ ) . . . . .	80
<b>D</b>	<b>Correlations and Heidke skill score</b>	<b>81</b>
D.1	Covariance . . . . .	81
D.2	Pearson's correlation . . . . .	81
D.3	Spearman rank-order correlation . . . . .	81
D.4	Heidke Skill Score . . . . .	82
	<b>References</b>	<b>83</b>
	<b>Acknowledgments</b>	<b>89</b>
	<b>Curriculum vitae</b>	<b>91</b>



# List of Figures

3.1	Orography of the research area . . . . .	10
3.2	MM5 model domain and the measurement stations . . . . .	13
3.3	Observed and simulated total rainfall for May 2000 and February 1998 . .	14
3.4	Daily observed and simulated rainfall for May 2000 and February 1998 . .	16
4.1	Areal averaged simulated daily rainfall rates . . . . .	20
4.2	Probability distribution of the simulated daily rainfall rates . . . . .	20
4.3	Probability distribution of the simulated monthly rainfall rates . . . . .	21
4.4	Areal averaged observed daily rainfall rates . . . . .	22
4.5	Climatological simulated rainfall rates by linear averaging . . . . .	23
4.6	Climatological observed rainfall rates by linear averaging . . . . .	23
4.7	Power spectrum of the simulated rainfall rates . . . . .	25
4.8	Power spectrum of the observed rainfall rates . . . . .	26
4.9	Size distribution of the Fourier coefficients . . . . .	27
4.10	Climatological annual cycle of rainfall rates by Fourier analysis . . . . .	27
4.11	Examples of a running average . . . . .	29
4.12	Standard deviation of differences between simulations and running average	30
4.13	Skewness ( $\gamma_1$ ) and kurtosis ( $\gamma_2$ ) of the averaged MM5 model simulations .	30
4.14	Frequency distributions of the simulations . . . . .	31
4.15	Climatological year using different running average periods . . . . .	32
4.16	Skewness and kurtosis of the averaged observational data . . . . .	33
4.17	Climatological year using a running average period of 91 days . . . . .	33
4.18	Correlation between the power spectra of the simulations and observations	34
4.19	Rainfall distributions at a running average period of 31 days . . . . .	35
4.20	Simulated annual rainfall cycle with a running average period of 31 days .	35
4.21	Mean simulated monthly rainfall for the period 1991–2000 . . . . .	37
4.22	Mean simulated monthly rainfalls (1991) and the standard deviations . . .	38
4.23	Mean simulated monthly rainfalls (1991) and the spline interpolated values	38
4.24	Simulated climatological rainfall with spline interpolation . . . . .	39
4.25	Observed climatological rainfall with spline interpolation . . . . .	40
4.26	Simulated rainfall (Fourier, running average and Spline interpolation) . . .	40
4.27	Observed rainfall (Fourier, running average and Spline interpolation) . . .	41
5.1	Averaged areal rainfall distribution from observations and simulations . . .	44
5.2	Vertical cross section of the 1 km and the 45 km orography . . . . .	45
5.3	Observations (1 km) and observations (45 km) and the results from eq. 5.1	47
5.4	Rainfall distribution with $f_{\text{details}}$ . . . . .	48

5.5	Average simulated and observed rainfall in east-west direction . . . . .	50
5.6	Observations (45 km) and simulations (45 km) and the results from eq. 5.3 . . . . .	52
5.7	Rainfall distribution with $f_{bias}$ . . . . .	53
5.8	Simulations (45 km) and observations (1 km) and the results from eq. 5.5 . . . . .	54
5.9	Rainfall distribution with $f_{total}$ . . . . .	55
5.10	Results of the Heidke Skill Score . . . . .	57
6.1	Research area with MM5 model grid cells . . . . .	59
6.2	Example of observed accumulation effect of precipitation . . . . .	60
6.3	The distribution of wind direction against wind speed . . . . .	62
6.4	The wind direction dependent observed precipitation patterns . . . . .	63
6.5	The wind direction dependent simulated precipitation patterns . . . . .	64
6.6	The distribution of wind direction against wind speed with classification . . . . .	65
6.7	Group I–VI for the observed precipitation distributions . . . . .	66
6.8	Group I–VI for the simulated precipitation distributions . . . . .	67
6.9	Rainfall distribution with wind direction dependent factors . . . . .	68
A.1	Total model area with research area . . . . .	73
A.2	Sigma-layers as defined in the MM5 model configuration . . . . .	74
A.3	The total MM5 model area. Outside the box model simulations are nudged . . . . .	75
B.1	The 10x10 grid cells MM5 model domain with national borders (1) . . . . .	77
B.2	The 10x10 grid cells MM5 model domain with national borders (2) . . . . .	78

# List of Tables

3.1	Sensitivity experiments and their corresponding parameterizations . . . . .	12
4.1	Statistical quantities of Fig.4.1 and Fig.4.4 . . . . .	21
4.2	Some statistical quantities with different running average periods . . . . .	31
5.1	The Spearman rank-order correlations ( $r$ ) . . . . .	56
5.2	Heidke Skill Score . . . . .	58
6.1	Number of days in each climatological seasonal wind direction group . . . . .	65
6.2	Correlations between the corrected simulations and the observations . . . . .	68
D.1	Basis for Heidke skill score. . . . .	82



# Summary

This work has been carried out in the framework of the project GLOWA-Danube (GLObal WAter cycle) where a joint effort is made by several groups to model the interaction of the water cycle and society in the Upper Danube catchment area. In particular regional climate models are used to simulate and eventually predict precipitation in this research area, while other groups convert this information into river runoff estimates and groundwater fluxes.

It has been agreed in the project that precipitation data and other meteorological data must be handed over to the hydrological groups with a spatial resolution of 1 km. Long term runs with regional climate models are, however, not feasible at 1 km resolution, because they would exceed available computer resources by far. Therefore, a pragmatic downscaling method for precipitation must be implemented which provides data of 1 km resolution on the basis of model results of fairly coarse resolution.

This downscaling uses extensively climatological precipitation observations where such downscaling relations can be derived. These observed rates are then adapted to the model results. The data are provided by the German Weather Service (DWD) and the Austrian Weather Service (ZAMG). The years 1991–2000 have been chosen as a reference period for the analysis.

The climate simulation is carried out by the mesoscale model MM5 at a resolution of 45 km. The model MM5 offers a wide range of parameterizations with respect to convective processes, the boundary layer, cloud microphysics, and the radiation balance, all directly or indirectly responsible for generating precipitation. Sensitivity studies are performed to find the best configuration for the research area and reference period.

A variety of methods is tested to generate observed and simulated climatological time series of precipitation. In particular, a linear average, a running average, a Fourier analysis, and spline interpolation are intercompared. In the end, spline interpolation between monthly values showed the best results for both time series and is used as a basis for the downscaling method.

The downscaling method has to correct two major discrepancies between the observed precipitation distribution at the 1 km resolution and the simulated distribution at the 45 km resolution. First, these are small scale details related to topography in the rainfall distribution at the 1 km resolution, which lack in the 45 km resolution. Second, there is an unrealistic southward shift of the rainfall maximum at the northern rim of the Alps in the simulations, which needs to be corrected. A specific correction factor is introduced for each problem. The correlation between the spatial distribution of observed and simulated distributions increases after using the correction factors. Due to the climatological relationships, the results time periods of 10 days and longer are superior to those for periods shorter than 10 days.

The precipitation distribution depends, of course, on the wind direction in particular

so near the Alps. Wind direction and wind speed are simulated by the MM5 model and combined with the correction factors described above. The correlation between the spatial distribution of observed and simulated precipitation increases more if the wind direction dependent correction factors are introduced. These improved correction factors depend less on climatological relationships and perform therefore better for shorter time periods. Additionally, they will be able to respond better on changes in the weather regime in future climates.

Altogether, this investigation provides a new pragmatic method to downscale model simulations on the basis of observations. This method will be used in the project GLOWA-Danube.

# Chapter 1

## Introduction

In recent years, modeling of regional climates became a research topic of considerable interest. Given the fact that global climate models are able to reproduce reasonably well today's climate and even to indicate directions of future climate change, it is an obvious next step to model also regional climates and to find out what can be learned about local future developments. See for example the Intergovernmental Panel on Climate Change – IPCC (Watson and Coauthors 2001). Specific regional models have to be used for this step, because the spatial resolution of global climate models is not sufficiently fine to resolve the details of regional climatology, at least within the next decade or so. On the other hand, the details of regional climate change have a much larger impact on daily life than the grand features of global climate change.

Precipitation is a key feature of regional climatology, and a reasonably accurate simulation of the hydrological cycle is an important aspect of a successful climate simulation. So far, it has been found that regional climate models are capable of reproducing seasonal mean values of precipitation satisfactorily, at least in general, if the moisture inflow through the boundaries is specified according to observations (Beljaars *et al.* 1996; Betts and Miller 1993). Although the mean precipitation is an important feature of the regional climate, precipitation should also be distributed realistically between different types of flow regimes. It is important for the hydrology of a region how much rain is caused by frontal passages, where most of the moisture is imported, and how much is contributed by convective situations where the advection of moisture often is less important. In addition, the need emerges to investigate and model the impact of regional climate on society as well as the feedback of human activity on the climate (see e.g. Changnon *et al.* 2000; Warner *et al.* 2000). Such interaction depends strongly on the prevalent type of precipitation and, in particular, on the occurrence of extreme events.

Many research studies focus on the difficulties in climate modeling with respect to regional influences, as well as studying the effect of regional processes to climate modeling (e.g. Giorgi *et al.* 1992). The work in the present study grew out of a project called GLOWA-Danube (GLOBal Water cycle) in which a joint effort is made to model the hydrological cycle and its linkage to society in part of the catchment area of the river Danube (Ludwig *et al.* 2003). In the framework of GLOWA-Danube it is desired to model precipitation in the catchment area of the river Danube until where it leaves Germany at the city of Passau. The choice of this area is motivated by several reasons. First, the catchment contains parts of the Alps as well as a fraction of the Alpine foreland to

the north. Therefore, the catchment contains parts with high precipitation rates from where water is exported to the lowlands. Such a situation is prone to flooding, many water-related problems (like up- and downstream conflicts, water quality, environmental protection, tourism, etc.), and vulnerability due to climatic change. Second, the Alps and their foreland enforce large gradients in climate and therefore a wide range of vegetation and land use. Third, the catchment is divided between many nations (Southern Germany and parts of Austria, Switzerland, Italy and the Czech Republic) with its borders roughly between  $46^{\circ}$ – $50^{\circ}$ N and  $8^{\circ}$ – $14^{\circ}$ E (the box in Fig. A.1). The differences in the water budget within a rather small area, along with good data coverage in both natural and social sciences, makes the Upper Danube an excellent prototype for integrative research.

There exist many regional (climate) models. An intercomparison between different models has been performed by Cox *et al.* (1998), who compared the mesoscale model MM5, the Colorado State University Regional Atmospheric Modeling System (RAMS), the Navy Operational Regional Prediction System Version 6 (NORAPS6) and the model of the United States Air Force (USAF). Although each model had its difficulties in simulating forecast values over different terrains scattered over the whole world, statistically the two best models resulting from their studies were RAMS and the MM5-model. The regional model used in this work is MM5 (see Appendix A), because of its wide-spread use and the ability to choose optimized adapted configurations for many weather situations. Among other applications, this model has also been tested extensively in the past for regional climate studies (see e.g. Leung and Ghan 1999; Giorgi *et al.* 1993). For example, a study using a relatively short integration is provided by Zhang *et al.* (2003), who simulated a 30-day period in order to capture the passage of several mesoscale weather events. Although the authors conclude that the model depends sensitively on the used configuration, they found a good correspondence between the simulated and observed rainfall in the total monthly rainfall as well as on a daily basis. Leung *et al.* (2003) used the MM5-model to study the interaction between atmospheric circulation and orography in a 20-year regional climate simulation. They found a good correspondence between the frequency and distribution of the simulated and observed rainfall as well. The total simulation area in this work is shown in Fig. A.1.

The problem of modeling a correct attribution of rainfall to weather events is far from trivial. After all, even modern numerical forecast models with their high spatial resolution and with their access to all information available face substantial difficulties when predicting rainfall rates (Pall and Eltahir 2001). In particular, the result of a forecast depends heavily on the parameterization chosen for the representation of convective processes (Colle *et al.* 2003; Braun and Tao 2000; Warner *et al.* 2000). The onset of convective rainfall, the exact location, as well as other processes that accompany convective rainfall, such as downdrafts, depend heavily on the choice of the cumulus parameterization (Gallus and Segal 2000; Gallus 1999; Wang and Seaman 1997). Altogether, then, there is the problem which parameterization to choose in a regional climate model in order to achieve reasonably realistic rain rates in a variety of characteristic synoptic situations (see e.g. Pall and Eltahir 2001; Betts 2000; Beljaars *et al.* 1996; Betts *et al.* 1996). Therefore, the first chapter of this study will be dedicated completely to a sensitivity study concerning different parameterization schemes available within the MM5 model. These parameterizations influence many variables within the model, but the main focus will be precipitation rates.



---

Many different scientific groups, like meteorologists, hydrologists, and social scientist, are involved in the GLOWA-project and use their specific models on different spatial and temporal scales. To ensure good interaction between those models an accord was made between all groups to implement a spatial resolution of 1 km. The accord about the temporal scale was 1 hour. With a resolution of 1 km, overlap exists between large-scale processes, like synoptic weather systems, and smaller scale processes, like infiltration rate and land use. Yet, long term runs at 1 km resolution exceed the available computer resources by far. Therefore, a coarser resolution of 45 km was chosen for the simulations. In order to provide reasonable rainfall rates at 1 km resolution to the other scientific groups within the GLOWA-project, a downscaling method has to be developed. Many methods exist to downscale and describe differences in rainfall rates between two different resolutions (Buishand *et al.* 2004; Widmann *et al.* 2003; Murphy 1999; Zorita and von Storch 1999; Wilby and Wigley 1997). Most methods use the relationship between predictors, like large-scale structures, and predictands, like precipitation patterns, at higher resolutions to correct a model at a coarse resolution. The method presented in this study is based on climatological relationships between simulated and observed precipitation rates for the reference period 1991–2000. Therefore, Chapter 4 describes the rainfall variability within the research area during 1991–2000. Four methods will be compared, which are linearly averaging, running average, Fourier analysis, and spline interpolation, in order to find an annual precipitation cycle representing the reference period. Special attention will be given to the differences and the accordances between the simulations and the observations. The downscaling method is then presented in Chapter 5.

One of the major goals within the GLOWA-project is to simulate future scenarios. The presented downscaling method in Chapter 5 is based on the simulated and observed precipitation distributions during 1991–2000. These distributions may change in a future climate. For this reason, a refinement of the downscaling method is presented in Chapter 6. In this chapter, it is assumed that a universal relationship exists between a precipitation distribution and a prevailing wind direction. The prevailing wind direction was defined as in the weather classification scheme by the German Weather Service (Bissolli and Dittmann 2003).

The downscaling method as presented in Chapter 5 is already implemented in the GLOWA-project and gives satisfying results. The refinement in Chapter 6 is not yet implemented, but it is planned to do this in the near future.



# Chapter 2

## Literature review

### Climatology

The quality of simulated climatologies depends, of course, strongly on the quality of the model used. However, the assessment of the models skill is also affected by the selection of the time period for which an intercomparison with the data is made. Moreover, availability of data is a crucial issue, in particular so in mountainous terrain. These problems become the more pressing the smaller the domain for which an intercomparison is conducted.

One of the earliest examples for such an intercomparison was provided by Karl *et al.* (1990), who selected two periods of 10 years (1966–1976 and 1976–1986) to statistically compare the output of a General Circulation Model (GCM) with observations for daily precipitation at five observational sites spreaded throughout the United States. An even shorter period was used by Giorgi *et al.* (1992) and Giorgi *et al.* (1993). They simulated a period of 2 years (01/01/1982–01/01/1983) and of 1.5 years (01/01/88–25/04/89) at a resolution of 60 km with an early version of the MM5 model (MM4). They found that these for that time, ‘high’ resolution simulations performed rather well when compared to observations. The difference between the length of the observed and the simulated climatological periods is mainly due to the insufficient computer power to simulate longer time periods at reasonable computational costs at that time. Another reason for choosing short periods in the last decades is the lack of sufficient long-term observations by e.g. satellites. This forced, for example, Negri *et al.* (2000) to use a 10 years period for generating a climatology of the northern part of South America.

Climatological episodes of observed meteorological variables like temperature and precipitation over the Alpine region were already described by Fliri (1975) using the 30 years period of 1931–1960. He describes a history of climatology going back to the 19th century as well. In the 1990s, many meteorologists became aware of the necessity to develop climatological precipitation fields on a regular grid, because meteorological models were increasingly linked to ecological and hydrological models (Changnon *et al.* 2000; Warner *et al.* 2000). Therefore, Daly *et al.* (1994) developed PRISM (Precipitation-elevation Regressions on Independent Slopes Model) at a resolution of 6 km using a digital elevation model (DEM) over Oregon and Washington, USA. A few years later, Widmann and Bretherton (2000) used the model PRISM to find out how useful its precipitation fields are for estimating temporal variability in local precipitation. They suggest that GCM precipitation might be a good predictor for statistical downscaling techniques. Finally, the PhD-thesis by Schwarb (2001) applied PRISM to the Alpine region ( $2^{\circ}$ – $18^{\circ}$ E/ $42.75^{\circ}$ – $48^{\circ}$ N) and modi-

fied it to a resolution of 1.25' ( $\sim 2$  km). He bases his work on the precipitation climatology data set of the Alps by Frei and Schär (1998), which was generated using high-resolution rain-gauge observations (9,546 daily precipitation observations) over the period 1971–1990. A description of this data set is given by Schmidli *et al.* (2001) as well. They state that the quality of such a climatology is highly dependent on the density of available observations, but nevertheless shows a good agreement with the mesoscale precipitation patterns and multiyear precipitation anomalies.

The feedback between atmospheric and hydrological variables like e.g. soil moisture was recognized by Schär *et al.* (1999) as well. They used a similar setup as Schär *et al.* (1996) and found that the potential for convective activity increased if soil moisture was increased, i.e. with small Bowen ratios.

As a result of the awareness that interdisciplinary research was needed to accurately simulate and understand many atmospheric processes, research projects like BayFOR-KLIM (Bavarian Cooperation for Research on Regional Climate Changes) were founded (BayFORKLIM 1999a,b, 1996). Its goal was to investigate climate changes in Bavaria, Germany, and their effects on microorganisms, plants, animals and man. The work in the present study grew out of a project called GLOWA-Danube (GLObal WAter cycle) where a joint effort is made to model the hydrological cycle and its linkage to society in part of the catchment area of the river Danube (Ludwig *et al.* 2003). In the framework of GLOWA-Danube it is desired to model precipitation mainly in the German catchment area of the river Danube.

It should be mentioned that instead of using daily data sets of precipitation, many scientists prefer monthly data (see e.g. Huffman *et al.* 1997). A reason for this is that they work on a global rather than on a regional scale.

In many applications there is an interest in a correct attribution of rainfall to weather events instead of establishing a long term climatology. This is a nontrivial task. After all, even modern numerical forecast models with their high spatial resolution and with their access to all information available face substantial difficulties when predicting rainfall rates (Pall and Eltahir 2001). Local variability of atmospheric variables at sub-grid scales cannot be simulated explicitly by numerical models and therefore needs to be parameterized (Pielke *et al.* 1991). In particular, the result of a forecast depends heavily on the parameterization chosen for the representation of convective processes (Colle *et al.* 2003; Braun and Tao 2000; Warner *et al.* 2000). The onset of convective rainfall, the exact location, as well as other effects that accompany convective rainfall, such as downdrafts, depend heavily on the choice of the cumulus parameterization (Gallus and Segal 2000; Gallus 1999; Wang and Seaman 1997). Altogether, then, there is the problem of which parameterization to choose in a regional climate model in order to achieve reasonably realistic rain rates in a variety of characteristic synoptic situations (see e.g. Pall and Eltahir 2001; Betts 2000; Beljaars *et al.* 1996; Betts *et al.* 1996).

The climate simulation in this study is carried out by the mesoscale model MM5, which is the Fifth-Generation NCAR Penn State Mesoscale Model version 3.5 (Grell *et al.* 1995; Dudhia 1993). Among other applications, this model has been used and tested extensively in the past for regional climate studies (see e.g. Leung and Ghan 1999; Giorgi *et al.* 1993). For example, a study using a relatively short integration is provided by Zhang *et al.* (2003), in which they simulated a 30-day period in order to capture the passage of several mesoscale weather events. Although they conclude that the model depends sensitively on

---

the configuration chosen, they found a good correspondence between the simulated and observed rainfall in the total monthly rainfall as well as on a daily basis. Leung *et al.* (2003) used the MM5-model to study the interaction between atmospheric circulation and orography in a 20-year regional climate simulation. They found a good correspondence between the frequency and distribution of the simulated and observed rainfall as well.

## Downscaling

General circulation models (GCMs) often fail to simulate high-resolution meteorological fields at reasonable computational costs. This plays an even more important role in climate studies for which usually many years or even decades need to be simulated. However, using a GCM at a lower resolution implies a coarser representation of e.g. topography, resulting in local height biases between the simulated and the real world. This again results to biases in the simulated and observed meteorological variables, like precipitation (Daly *et al.* 1994). One way to reduce these biases is using a downscaling technique.

Generally, downscaling techniques are subdivided into four different classes, which are stochastic weather generators, limited-area (climate) models, regression methods, and weather pattern based approaches (Zorita and von Storch 1999). A good overview of different downscaling techniques is given by Wilby and Wigley (1997). In reality, most downscaling techniques do not focus on just one class, but are a combination of two or more of them. Some authors have developed downscaling techniques for monthly data (e.g. Murphy 2000, 1999; Kilsby *et al.* 1998), whereas others have considered daily data (e.g. Beckmann and Buishand 2002; Wilby *et al.* 1998). The stochastic weather generators are not of interest for this work and will therefore not be discussed. The limited-area model approach is actually already applied by using the MM5 regional model. Although this is a downscaling technique, this work bases its downscaling on using MM5. This implies that when downscaling is mentioned the regional model results are interpreted as GCM results. This has no effect on the ideas behind downscaling, of course.

The regression method is one of the earliest downscaling approaches. It is based on linear or nonlinear relationships between sub-grid scale parameters called predictands and coarser resolution predictor variables. Kim *et al.* (1984) were the first to publish about a regression method in the climate change context. They tried to find a relationship between predictands and the anomalies in the monthly-averaged surface temperature and monthly total precipitation for the state of Oregon, USA. They found encouraging results and mention that the regression method may be of practical importance to local climate impacts. Also, most methods tested by Widmann *et al.* (2003) were based on finding coupled anomaly patterns between the predictor and the predictand. These predictor variables are usually provided by a GCM, which implies that the predictor should be realistically simulated by the GCM. Widmann *et al.* (2003) showed that the simulated precipitation by a GCM can be a good predictor for regional precipitation. They used the precipitation data set provided by Widmann and Bretherton (2000), which is consistent with the Parameter-Elevation Regressions on Independent Slopes Model precipitation climatologies (PRISM; Daly *et al.* 1994, 1997).

A simple downscaling method was introduced by Salathé (2003). He focused more on temporal downscaling and used raw large-scale model data, downscaled data, and observations of precipitation fields as input for a hydrological model of the small Yakima River in central Washington, USA. The basis of his method is the ratio between long-term observed

and simulated precipitation rates, both at a 50 km resolution. He concludes that the statistical downscaling method is very efficient and able to capture the essential precipitation features required for accurate simulation of flow in the Yakima River.

Besides using the same variable as a predictor and predictand, many approaches focus on the relationship to a different large-scale predictor variable than the predictand (Buishand *et al.* 2004; Pandey *et al.* 2000; Murphy 1999; von Storch *et al.* 1993). For example, Buishand *et al.* (2004) compared different regression methods for daily and monthly rainfall amounts using rainfall occurrence as well as rainfall amount models. They used specific humidity  $q_s$  near the surface, sea level pressure, and the west as well as the south component of the geostrophic flow ( $u$  and  $v$ , respectively) as predictor values. von Storch *et al.* (1993) used a canonical correlation technique to relate winter rainfall in the Iberian Peninsula to sea-level pressure patterns in the North Atlantic.

Weather pattern based approaches focus typically on the statistical relationship between local observations and spatial averaged meteorological data depending on a given weather classification scheme. This was already recognized by Fliri (1984). Zorita and von Storch (1999) give an overview of existing classification schemes. For example, a scheme developed by the German Weather Service (DWD) distinguishes between 40 different weather situations based on four physical criteria (Bissolli and Dittmann 2003). Bárdossy and Plate (1992) used the classification of the DWD for downscaling purposes. Salathé (2003) found that weather pattern dependent downscaling of precipitation hardly improved the results of his hydrological model, because the small improvements in the spatial distribution of precipitation had a minimal effect on the total water availability in the catchment area.

As many groups of scientists tried to find rainfall distributions that correspond well with observations, this work adds to that. However, the goal of this work is more to develop a robust and fast method to gain rainfall distributions, which can be used by the different groups involved in the GLOWA-Danube project.

# Chapter 3

## Sensitivity of MM5 precipitation to various physical parameterizations

### 3.1 Introduction

The mesoscale model MM5 (see Appendix A) offers the choice of a wide range of parameterizations with respect to convective processes, the boundary layer, cloud microphysics, and the radiation balance, to be addressed to in more detail below. Considering the extensive interaction between the different parameterizations, it is desired to find out what combination performs best with respect to the total amount of precipitation and the evolution in time of precipitation rates in the selected area where data are available for validation. A study similar to that presented here has been performed by Zängl (2004), using a relatively small area containing part of the Alps with a high resolution during a short period of a few days. The test area chosen in this study is that part of the catchment area of the Upper Danube, which is the target area of the GLOWA project. The tests are performed for two months selected from the decade January 1991 to December 2000, which has high priority to GLOWA-Danube as a reference period. In February 1998, almost all rain fell in connection with large-scale systems moving through the area of southern Germany. As is typical of winter months, there was hardly any convective precipitation. The other month selected, May 2000, had a substantial contribution of rain falling due to local convection. Just around 50% of the total rain in May 2000 fell in connection with large-scale systems. Details will be discussed below. This choice of these relatively clear-cut cases will allow an assessment of the relative merits of various schemes.

### 3.2 Model description, data and methodology

In this sensitivity study 32 experiments are performed (16 different configurations  $\times$  2 months) with a resolution of 45 km. This means that the actual orography of the Alps (see Fig. 3.1) is crudely represented, since the terrain height of a grid cell is approximately the linear average of the height of the area it covers. Therefore, local differences in height between the model and actual orography can be up to many hundreds of meters (e.g. the maximum height of the actual orography is over 3,500 meters, while the 45 km resolution has a maximum height of a little over 2,200 meters). This difference can significantly change the amount and the distribution of precipitation in the Alps and their forelands,

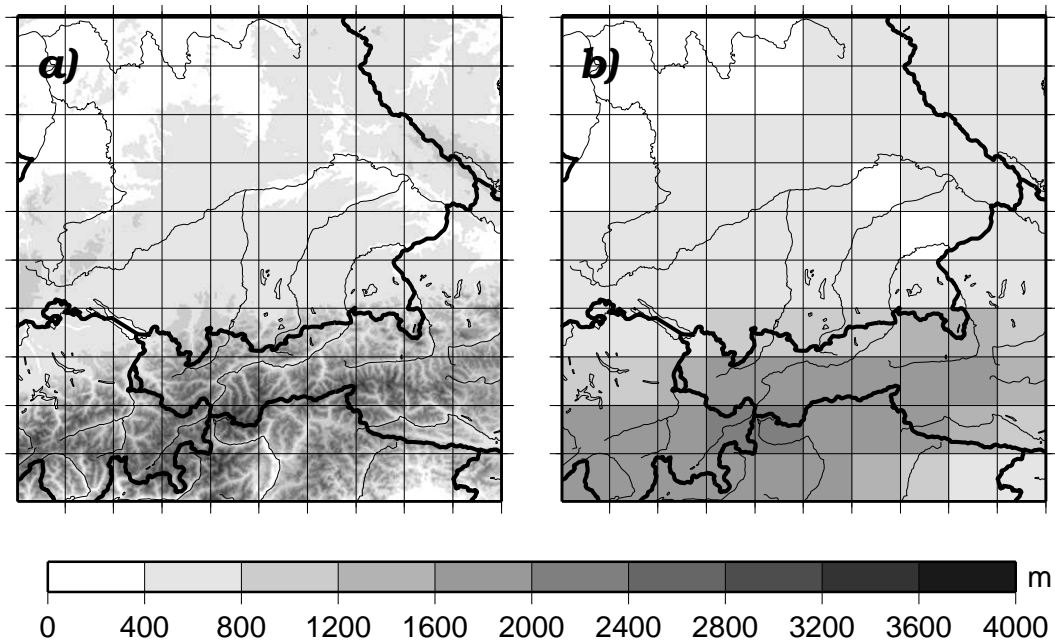


Figure 3.1: Orography of the research area with national boundaries and rivers. Panel *a)* shows the topography at a 1 km resolution and panel *b)* at a 45 km resolution.

which is the focus here. In particular, it is not possible to correctly simulate the orographic impact of rainfall variability at a 45 km resolution in the Alps. This makes a comparison with observational data very difficult. It is expected that the situation in the more flat area in the north of the research area can be simulated more accurately with respect to observations (Cosma *et al.* 2002; Daly *et al.* 1994).

### 3.2.1 Description of the model configurations

The mesoscale model MM5 should not be seen as one single model as well, but as a powerful package of model routines to be combined. Many atmospheric processes are responsible for the amount and distribution of rainfall and are calculated by different model routines. Additionally, a user can choose between different options of model routines for the same type of process to be parameterized that is suitable for the kind of model run he wants to perform. Depending on the option chosen, the results change and will be more or less realistic for the simulated situation (Zhang *et al.* 2003). Atmospheric processes are not separate, but interact and so do the model routines. This leads to a fairly large number of combinations to be tested.

Many schemes are available in the model. In this sensitivity study, four parameterization schemes are examined. For each of these four schemes, two options available in the model are considered, which differ not only in physics but also in complexity and are most applicable to the models resolution and precipitation events investigated herein. They deal with cumulus convection, boundary layer processes, cloud microphysics, and the radiation balance and either directly or indirectly affect rainfall. A brief description of each of these options is now given.



Cumulus parameterization schemes are important for precipitation simulations. They calculate the amount of convective rainfall within a grid cell, which cannot be simulated explicitly. Here, two well-known schemes are tested, the one developed by Grell (Grell 1993, hereafter referred to as GR) and that developed by Betts and Miller (Betts and Miller 1993, hereafter referred to as BM). GR is based on a quasi-equilibrium, simple, single-cloud scheme and accounts for convectively induced downdrafts in the atmosphere. BM is based on the relaxation adjustment to a reference post-convective thermodynamic profile and does not consider downdrafts explicitly (Cohen 2002; Betts 1986). Therefore, BM is probably less suited for handling severe convection. Of course, the success of a cumulus parameterization depends on the situation that is simulated. For example, Cohen (2002) tested cumulus parameterizations in idealized sea-breeze simulations. He found that GR gives unrealistically high rainfall rates, while BM produced more realistic rainfall rates. Wang and Seaman (1997) simulated a set of six rainfall events in cold and warm seasons and found that the amount of rain during light rain events in the warm seasons was overestimated by GR and more realistically modelled by BM. Differences existed between the temporal distribution of rainfall, as well. They conclude that generally GR underestimates the amount of convective rainfall. Gochis *et al.* (2002) performed a sensitivity study with respect to the evolution of the north American monsoon system and also concluded that there exists a high sensitivity of regional models precipitation with respect to cumulus parameterization schemes. Different cumulus parameterizations contain different assumptions and parameter specifications, which makes each them more appropriate in some regions than in others (Gochis *et al.* 2002).

The parameterization of the boundary layer deals with processes like local and non-local mixing of the boundary layer, vertical diffusion, and entrainment. The convective schemes described above are driven internally by the distribution of humidity and buoyancy in the lower layers of the atmosphere. Sensible and latent heat fluxes are calculated within the boundary-layer scheme and are therefore crucial for simulating precipitation. Two boundary layer parameterization schemes are selected, namely, the Mellor-Yamada scheme as used in the Eta model (Janjić 1994, 1990; hereafter referred to as the Eta-scheme), and the MRF scheme, which is part of the Medium-Range Forecast (MRF) model developed by the National Centers for Environmental Prediction (NCEP) (Hong and Pan 1996). The Eta-scheme predicts the turbulent kinetic energy budget equation to decide whether the boundary layer will become more turbulent and has local vertical mixing. The MRF scheme uses a representation of the counter-gradient term the first-order closure of the K theory. This K theory requires a well-mixed boundary layer and will most likely perform less satisfactory during situations with deep convection (Stull 1999).

Besides convective precipitation, which usually takes place on sub-grid level, much rain falls in connection with large-scale systems. These mesoscale cloudy frontal regions usually cover more than a single grid box. In addition, the precipitation is usually not as heavy as in a convective situation, and has to be handled differently in an explicit moisture parameterization, also called a micro-physical scheme. In such a parameterization, the amount of cloudwater/-ice and the processes changing the conditions of aggregation within the grid cell determine whether rain will be predicted or not. The parameterizations tested here are the ‘Simple Ice’ parameterization (Dudhia 1989) and the ‘Reisner1’ parameterization (Reisner *et al.* 1998). The difference between the parameterizations is that ‘Reisner1’ carries separate prognostic variables for cloud ice and snow. Thus, it allows for the co-

existence of cloud ice and super-cooled cloud water below 0°C and gradual melting of snow above 0°C. The ‘Simple Ice’ scheme does not consider this.

The last two schemes tested here parameterize radiation that interacts with clouds and reaches the surface. Input parameters such as surface emissivity and albedo affect the balance between incoming and outgoing radiation, which is again an important input in the boundary layer scheme. The two parameterizations used here are ‘Cloud Radiation’ (hereafter referred to as CR) and the radiation scheme from the Community Climate Model developed by the National Center for Atmospheric Research (NCAR) (hereafter referred to as CCM) (Kiehl *et al.* 1998; Hack *et al.* 1993). CR accounts explicitly for the interaction of longwave and shortwave radiation within moist and cloudy air. The CCM scheme is similar to CR but in addition considers radiation with multiple spectral bands. It uses the resolved clouds from the micro-physical scheme.

A list of the experiments and the parameterizations used in each are shown in Table 3.1.

	Cumulus	PBL	Moisture	Radiation
exp01	Grell	Eta	Simple Ice	Cloud-radiation
exp02	Grell	Eta	Simple Ice	CCM
exp03	Grell	Eta	Reisner1	Cloud-radiation
exp04	Grell	Eta	Reisner1	CCM
exp05	Grell	MRF	Simple Ice	Cloud-radiation
exp06	Grell	MRF	Simple Ice	CCM
exp07	Grell	MRF	Reisner1	Cloud-radiation
exp08	Grell	MRF	Reisner1	CCM
exp09	Betts-Miller	Eta	Simple Ice	Cloud-radiation
exp10	Betts-Miller	Eta	Simple Ice	CCM
exp11	Betts-Miller	Eta	Reisner1	Cloud-radiation
exp12	Betts-Miller	Eta	Reisner1	CCM
exp13	Betts-Miller	MRF	Simple Ice	Cloud-radiation
exp14	Betts-Miller	MRF	Simple Ice	CCM
exp15	Betts-Miller	MRF	Reisner1	Cloud-radiation
exp16	Betts-Miller	MRF	Reisner1	CCM

Table 3.1: Sensitivity experiments and their corresponding parameterizations, which deal with cumulus convection (Cumulus), boundary layer processes (PBL), cloud microphysics (Moisture), and the radiation balance (Radiation).

### 3.2.2 Observational data and selected time periods

Observations for the research area were available from the observations network of the German and Austrian Weather Service (DWD and ZAMG, respectively) covering the period January 1991 to December 2000. The location of the measurement stations for this period are shown in Fig. 3.2. The amount and distribution of the stations differed just slightly for each day due to data gaps. Figure 3.2 also shows the research area divided into 45×45 km grid cells used in MM5. The number of stations within each cell varies across the model domain shown. Hatched MM5 cells in Fig. 3.2 represent cells without a station.

For this study, cells containing five or less stations are excluded, simply because fewer than five stations per grid cell would not provide sufficient rainfall coverage across a grid cell. Additionally, grid cells where the stations were not equally spread (approximately less than a quarter of the MM5 cell covered) and were located in the relatively flat area in the north were excluded. Unfortunately, sparse observations were available in Austria, resulting in excluding Austrian measurement stations completely in this sensitivity study. Although this filters out nearly half of the grid cells in MM5 (shaded cells in Fig. 3.2), the number of measurement stations is only slightly reduced (97% of the original remain). On average each grid cell contains 16 stations. For this study the observed rainfall data for each station within the reduced research area is linearly averaged.

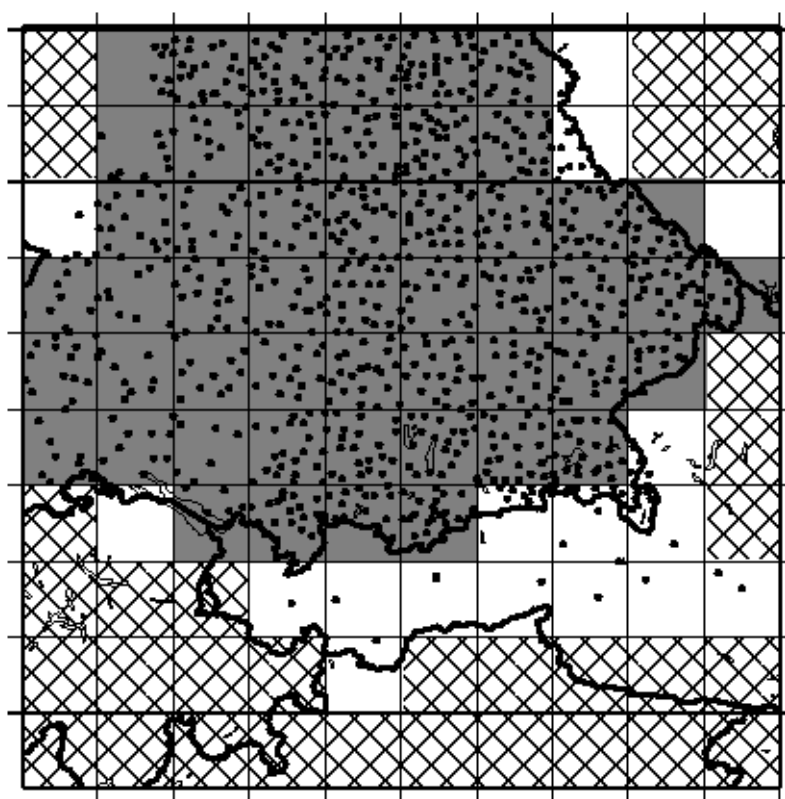


Figure 3.2: The 10x10 grid cells of the MM5 model covering the research area with national borders and the measurement stations represented as black dots; hatched cells: contain no stations at all; white cells: contain five or less, or not equally spread stations; shaded cells: contain six or more stations and are used in this study for comparison (see text for explanation).

In order to test the parameterizations under different rainfall regimes, two time periods having different rainfall characteristics are examined here. The time periods chosen for analysis from the dataset are February 1998 and May 2000. The first time period represents a typical winter month where most rain fell in association with synoptic-scale systems. The second time period represents a typical summer month where the rainfall was dominated by localized convective systems.

### 3.3 Results

The total amount of resolved scale rainfall and parameterized rainfall for each of the model experiments listed in Table 3.1 is shown in Fig. 3.3 for May 2000 (left panel) and for February 1998 (right panel). Also shown in both panels is the observed rainfall. The observation bar is shaded differently, because the observations fail to distinguish between resolved scale rainfall and convective rainfall. The observation bar shows the linear average of the observed precipitation rates at the measurement stations within the shaded area of Fig. 3.2.

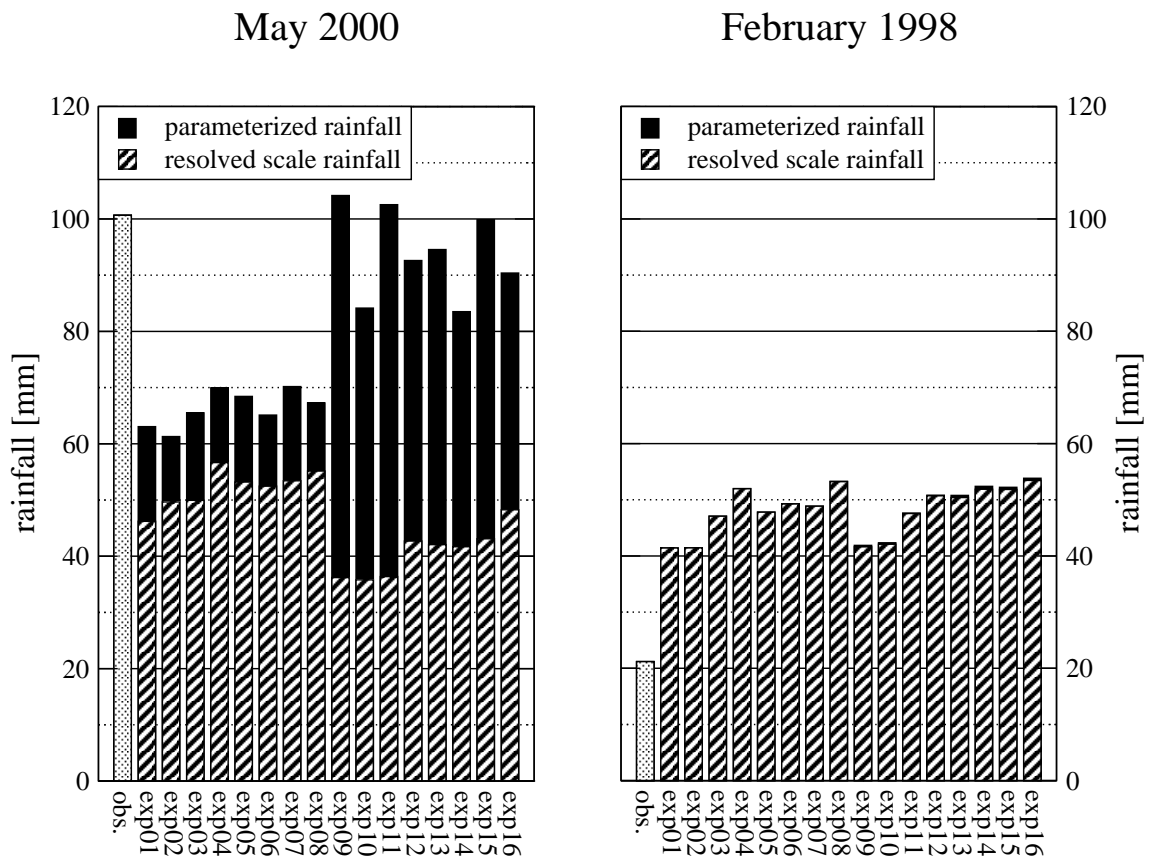


Figure 3.3: Observed and simulated total rainfall for May 2000 (left panel) and February 1998 (right panel). All 16 model experiments are split up between resolved scale rain and convective rain. The observations are shown as a separate bar.

For the May 2000 case (Fig. 3.3, left panel), the total amount of rainfall for experiments 9 to 16 is similar to the observed rainfall total. However, experiments 1 to 8 underestimate the observed total by 30–40%. This underestimate is related to the failure to produce enough ‘CON’ in the latter 8 experiments. For these latter experiments, parameterized rainfall contributes between 50 and 65% of the total rainfall. Further examining each experiment for each day in May 2000 (not shown) shows that during the first 16 days, which were especially dominated by localized convection, the first 8 experiments produce

less parameterized rain per day than the last 8 experiments. Wang and Seaman (1997) also showed a too low convective to total rainfall ratio using GR and remark that this could be improved by changing the precipitation efficiency parameter in the parameterization. This is not verified within this study and because the total amount of rainfall in experiments 9, 11, and 15 for May 2000 represents the observed rainfall best, the rest of this study will focus on these three experiments.

Figure 3.4 (upper panel) shows the total daily rainfall for these experiments along with the observed daily rainfall for May 2000. The three model runs are very similar for each day, and generally represent the observed rainfall well. On a few days the model lags the observed rainfall by 1 day (e.g. 8, 17, and 21 May). However, on 2, 4 and 30 May, the difference between the simulated and observed rainfall amount is quite large. The difference on the 30th is due to a fast moving low pressure area bringing a considerable amount of rain to southwestern Germany. Local maxima were observed of up to 128 mm on 30 May 2000. This event was not adequately represented in any of the model experiments, denoted by the small vertical extension of the gray bar at 30 May 2000. Despite the failure of experiments 9, 11 and 15 for the May 30th rain event, these experiments are believed to have more skill than the other experiments due to their success during the rest of the month.

Examining now the February 1998 case (Fig. 3.3, right panel), the total amount of rainfall for all model runs is between 2 and 2.5 times greater than the observed rainfall total. None of the runs produces a considerable amount of convective rain, which was expected for February 1998.

Figure 3.4 (lower panel) shows the daily simulated rainfall for the three model experiments for February 1998. Although the maximum amount of observed rain during the first 21 days was 1.3 mm, all 3 model experiments simulate the 3 rainfall events with totals between 1 and 5mm per day (3, 8, 13, and 17 February). On 22 February a cold front passed over the research area, and all 3 model experiments simulate the timing and amount of rain well. However, on the following days, the difference between the daily rain totals in each experiment and the observed rain total varies by as much as 10 mm (25 February). One likely reason for this difference is the rather coarse resolution of 45 km in the model leading to an underestimation of the height of the Alps by the model. This was already described in section 3.2. In the model, southerly winds from the Mediterranean sea transported moist air to the research area. This resulted in the occurrence of rain producing clouds over the Alps in the days after the cold front passage. In reality however, much of the humid air was blocked on the southern side of the mountains, and so produced less rain on the northern side in the research area. Besides the problems of simulating rainfall during winter months because of snow (Leung *et al.* 2003) this could have caused an additional overestimate of rain in the model. Note that all experiments suffer from these problems, so that an intercomparison with respect to the parameterization is still possible.

Of the three experiments chosen for analysis here (experiments 9, 11 and 15), one needs to be chosen for further work in the GLOWA-Danube project. Since all three experiments perform similarly well for the cases examined here, a selection will be based on the type of parameterizations used. Since the goal of this study is to find a configuration that will perform well in a variety of situations, the more sophisticated parameterizations will be preferred. Considering first the moisture parameterization, experiment 9 differs from experiments 11 and 15 in that the latter 2 use the more complex ‘Reisner1’ parameterization.

## May 2000 (upper panel) &amp; February 1998 (lower panel)

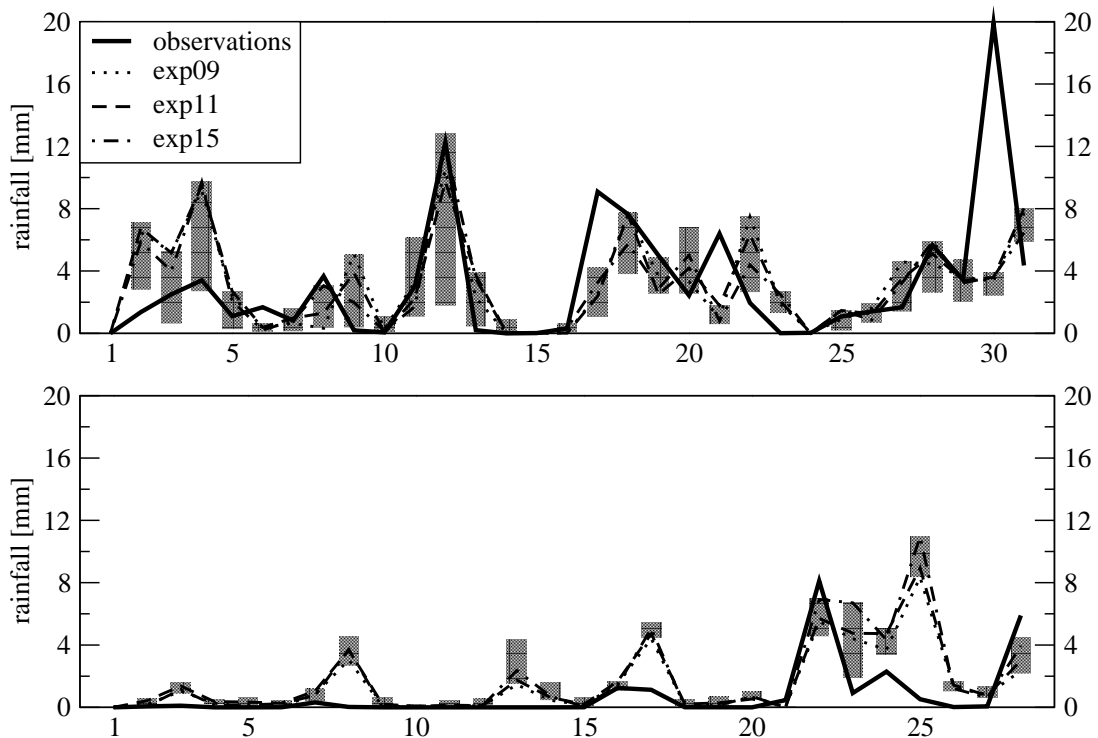


Figure 3.4: Daily observed and simulated (exp09, exp11 and exp15) rainfall for May 2000 (upper panel) and February 1998 (lower panel). The gray bars give an impression of the results of the other experiments, where the upper limit of each bar denotes the maximum rainfall of all 16 experiments on a certain day and the lower limit the minimum rainfall.

The choice for the ‘Reisner1’ parameterization over the ‘Simple Ice’ parameterization is in accordance with the results found by Colle and Mass (2000). Furthermore, considering the boundary layer parameterization, experiment 11 differs from experiment 15 in that the former uses the ‘Eta’ parameterization. Since this parameterization is more suited to higher vertical resolutions, which may in future be increased in the model, this will be the preferred boundary layer parameterization scheme. In summary this means that the configuration used in experiment 11 will be the preferred MM5 configuration for future studies for the GLOWA-Danube project.

### 3.4 Conclusions

Sensitivity tests with MM5 were performed with a grid spacing of 45 km to find the configuration of parameterizations which produces the best precipitation amounts and distribution for the Upper Danube catchment area in the framework of project GLOWA-

Danube. The sensitivity tests are carried out choosing two different parameterizations for four atmospheric processes: cumulus convection, boundary layer processes, cloud microphysics and radiation. The sensitivity study was performed using rainfall data from the German and Austrian Weather Service from February 1998 and May 2000 which were months dominated by different rainfall regimes. February 1998 was a typical winter month dominated by large-scale rainfall, while May 2000 was dominated by localized convective rainfall of horizontal scale less than the model resolution.

It was found that for May 2000 the experiments using the ‘Grell’ cumulus parameterization failed to simulate a sufficient amount of parameterized rainfall, resulting in too low total rainfall amount compared to the observations. From the experiments with the cumulus parameterization of ‘Betts-Miller’, experiments 9, 11, and 15 are focused on, because of their good representation of total rainfall. The daily simulated rainfall for these three experiments are almost identical to the observations. A considerable amount of rainfall observed during a fast moving low pressure area in the southwest of Germany on 30 May was not captured by the model. Despite this inadequacy, experiments 9, 11, and 15 are believed to be the best, because of their skill during the rest of the months.

The over-estimation of rainfall in all sensitivity runs during February 1998 was mainly caused by the synoptic situation during the days following a frontal passage on 22 February. The mountains as represented by the model were too low to block moist air coming from the Mediterranean sea, which in turn resulted in the model overestimating the amount of precipitation.

Since no final decision could be made which of the experiments 9, 11, or 15 performed best regarding the total and daily amount of rainfall, the final decision was based on the type of parameterizations used in each experiment. The ‘Reisner1’ parameterization for moisture was preferred over ‘Simple Ice’, because it is more sophisticated in handling water and ice in clouds, which will probably give more realistic results when applying this parameterization to model runs in future. Because it is possible that future model runs will simulate the research area with a higher vertical resolution and the ‘Eta’ boundary layer parameterization can handle such a situation better than the ‘MRF’ parameterization, the ‘Eta’ parameterization is preferred.

This leaves the configuration of experiment 11 with a cumulus parameterization of ‘Betts-Miller’, a boundary layer parameterization of ‘Eta’, a moisture scheme of ‘Reisner1’, and a radiation scheme called ‘Cloud-radiation’ to be most skillful in the research area of GLOWA-Danube. Due to the known difficulties in the ‘Grell’ cumulus parameterization in simulating large amounts of convective precipitation, ‘Betts-Miller’ was the logical choice. The ‘Eta’ boundary layer parameterization handled the deep convection better during the summer months, because of prediction the turbulent kinetic energy budget equation rather than using the K theory. The latter would provide better results during months with less deep convection. Only slight difference could be found between the two tested microphysical schemes. The more sophisticated ‘Reisner1’ is believed to handle better future climates.

The choice for this configuration does not mean that this configuration can simulate the total and daily amount of rainfall best at all times. But the configuration of experiment 11 showed most skill in two different months and it is likely that the configuration will do well in other situations during future model runs within GLOWA-Danube.





# Chapter 4

## Description of rainfall variability

### 4.1 Introduction

This chapter focuses on the variability of the averaged simulated and observed rainfall rates over the whole research area during a ten year integration period. This longterm integration will be used as input for the relationships, which will be derived between simulations and observations in Chapter 5 and 6. The research area equals the area in Chapter 3. The observations used here are described in Appendix B. The model simulations in this chapter are based on the configuration determined by the sensitivity study from Chapter 3.

In the next four sections, 10 years of modeled and observed daily rainfall data are discussed. Each section describes an approach to generate an averaged annual rainfall cycle from the 10 years reference period, which covers the period 1991–2000. This averaged annual rainfall cycle will be referred to as climatological year. The main motivation of each approach is to minimize fluctuations in both the modeled and observed rainfall data, such that they display the highest agreement.

First in section 4.2, daily rainfall rates during the reference period are linearly averaged. The deficiencies are revealed of a short averaging period of ten years and the necessity for more sophisticated methods. In section 4.3, Fourier analysis is applied in order to filter out insignificant frequencies in the annual rainfall cycle. The next approach tries to find a running average period for which most of the noise is reduced. Last, in section 4.5, monthly means are defined as fixed data points to apply spline interpolation. An intercomparison of all four approaches is described in section 4.6.

Note, that the reference period 1991–2000 contains three leap years (1992, 1996, and 2000) and therefore  $10 \cdot 365 + 3 = 3653$  days. These three extra days do not play a crucial role in a climatological sense, and are omitted for this reason. When 29 February occurs in a testing period, it will be handled the same as 28 February.

### MM5 simulations

Fig. 4.1 shows the daily simulated precipitation, averaged over the 10x10 MM5 grid cells covering the research area for the reference period 1991–2000. The variability between individual days is obvious. The statistical quantities (mean:  $\mu$ ; standard deviation:  $\sigma$ ; skewness:  $\gamma_1$ ; kurtosis:  $\gamma_2$ , for explanation see Appendix C) are given in Table 4.1. The variability in daily rainfall rates is expressed in a standard deviation,  $\sigma$ , of 3.51 mm/day

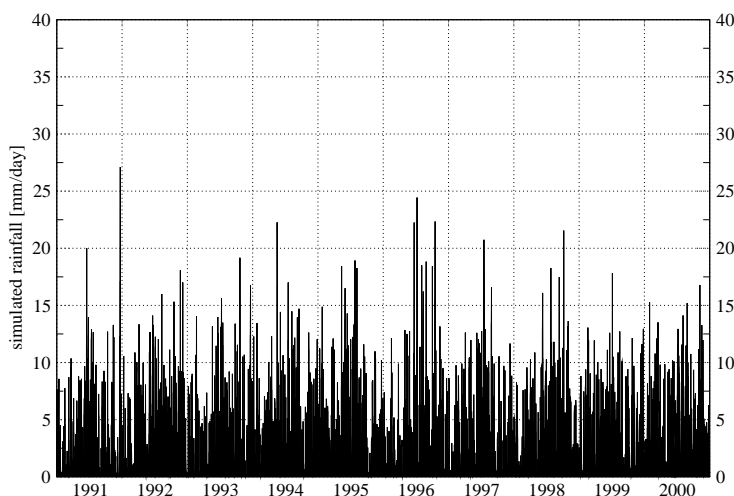


Figure 4.1: Spatial averaged simulated daily rainfall rates for the research area over the period 1991–2000.

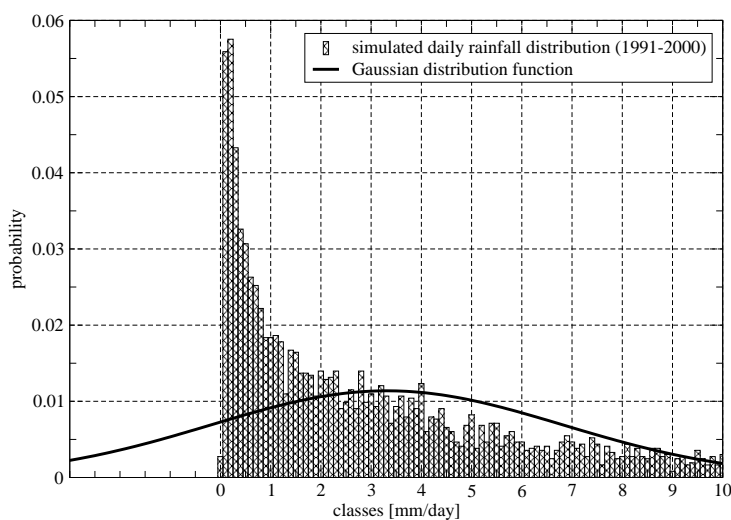


Figure 4.2: Probability distribution of the simulated daily rainfall rates during the period 1991–2000. The bold line corresponds to the Gaussian curve with the same mean value ( $\mu$ ) and standard deviation ( $\sigma$ ) as for the simulated rainfall rates (3.31 and 3.51, respectively). The classes on the x-axis are millimeters and show roughly the range of 2 standard deviations of the Gaussian curve.

with an average value,  $\mu$ , of 3.31 mm/day. Assuming a normal distribution function would mean that 68% of the values can be found between approximately 1.86 ( $\approx 3.51 - \frac{1}{2} \cdot 3.31$ ) mm and 5.17 ( $\approx 3.51 + \frac{1}{2} \cdot 3.31$ ) mm. Yet, daily precipitation does not behave like a normal

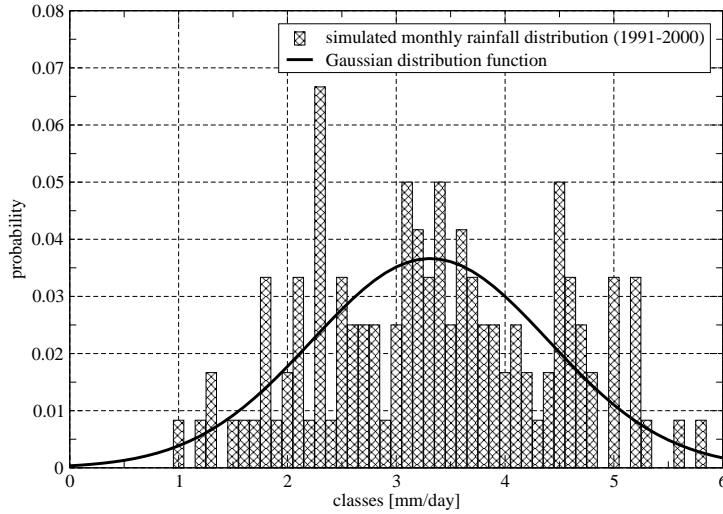


Figure 4.3: Probability distribution of the simulated monthly rainfall rates during the period 1991-2000. The bold line corresponds to the Gaussian curve with the same mean value ( $\mu$ ) and standard deviation ( $\sigma$ ) as for the simulated rainfall rates (3.31 and 1.09, respectively). The classes on the x-axis are millimeters and show roughly the range of 2 standard deviations of the Gaussian curve.

<i>data</i>	<i>n</i>	$\mu$	$\sigma$	$\gamma_1$	$\gamma_2$
MM5	3,650	3.31	3.51	1.66	3.50
OBS	3,650	3.07	4.34	2.27	6.82

Table 4.1: Some statistical quantities of the simulated time series shown in Fig.4.1 and observed time series shown in Fig.4.4.

distribution function, but is positively skewed and leptokurtic, due to its natural lower boundary of zero (e.g. Lettenmaier 1995; von Storch and Zwiers 1995).

Fig. 4.2 shows the probability distribution of simulated daily rainfall data with a class size of 0.1 mm. The bold line in Fig. 4.2 represents a normal distribution function. It shows that the distribution is strongly skewed to the left and has a steeper peak than a normal distribution, so that it is leptokurtic (Table 4.1).

A longer accumulation time of precipitation, of for instance a month, removes many of the short term precipitation peaks and has a distribution closer to the normal distribution function. This is demonstrated in Fig. 4.3, where the probability distribution of the simulated monthly rainfall rates during the period 1991–2000 is shown. Again, the bold line represents the normal distribution function.

## Observations

The observed daily rainfall rates are shown in Fig. 4.4. The observed variation between individual days seems higher than for the simulations (Fig 4.1). The statistical quantities

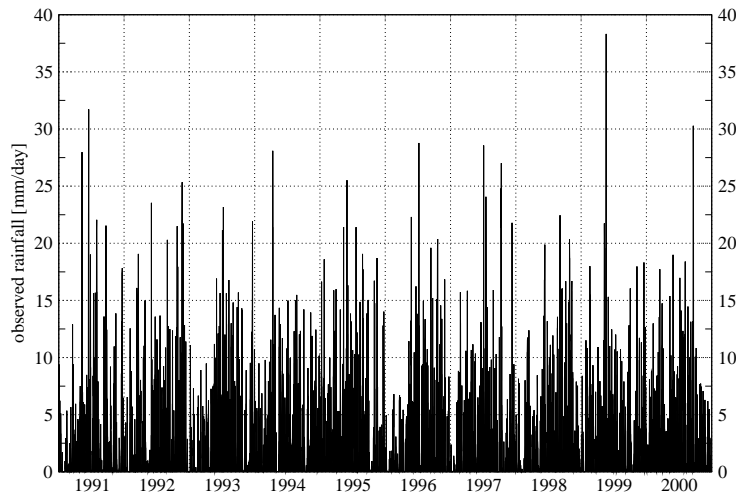


Figure 4.4: Spatial averaged observed daily rainfall rates for the research area over the period 1991–2000.

are shown in Table 4.1 and confirm this. Although the mean value,  $\mu$ , for the observations is smaller than for the simulations, the standard deviation,  $\sigma$ , is higher, indicating more and higher peaks during 1991–2000. The natural lower boundary for precipitation is zero, which gives a larger  $\gamma_1$  for the observations than for the MM5 simulations as well.

Note that the observations have a resolution of 1 km as described in Appendix B and the MM5 simulations have a resolution of 45 km. The Alps are hardly represented at a resolution of 45 km resulting in major differences in the distribution and the amount of rainfall rates. This chapter uses spatial averaged rainfall rates only and these differences are neglected for now. However, they will be discussed in Chapter 5.

## 4.2 Linear average

In order to build a climatological year over a 10 years period, the easiest approach is to build the sum for each day over all years and divide the result by 10 (Figure 4.5 and 4.6 for the MM5 simulations and the observations, respectively). The advantage is that each day of the climatological year is an average of 10 days at exactly the same time of the year. A disadvantage is that there were grid cells in which no rain fell on a certain day for each year. These days would have a climatological rainfall rate of zero. This effect increases with decreasing resolution. Figs. 4.5 and 4.6 do not show this, because they only show the areal average rainfall rate, but it is visible in the areal rainfall distribution. The temporal distribution is very noisy as well, which makes it very hard to compare both data sets.

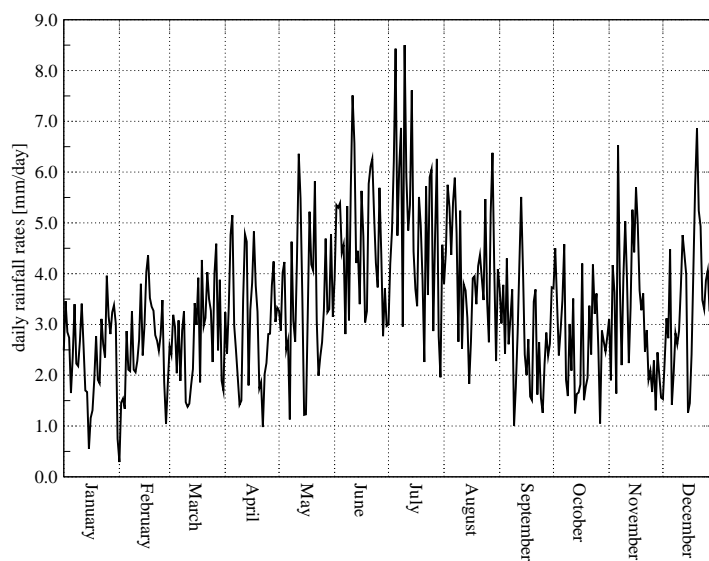


Figure 4.5: Climatological simulated rainfall rates by linear averaging.

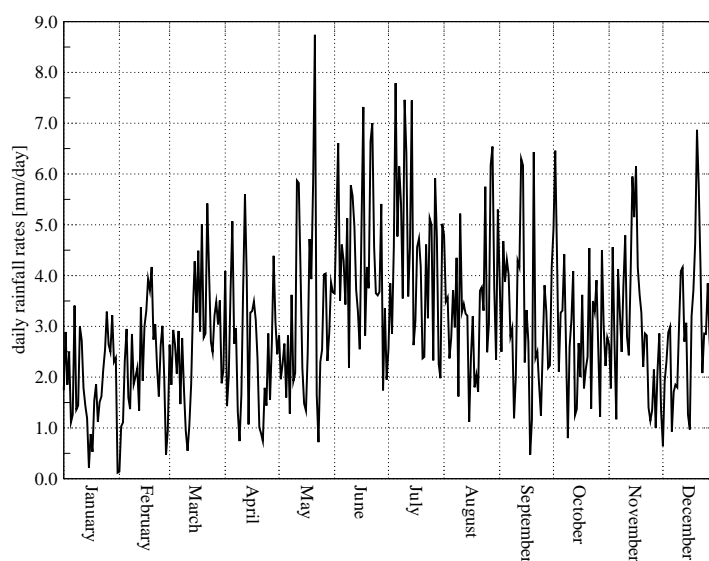


Figure 4.6: Climatological observed rainfall rates by linear averaging.

## 4.3 Fourier analysis

### 4.3.1 Introduction

A periodic function  $f(t)$  can be described as an infinite sum of harmonic functions. Each function has a period  $T_p$  and is a multiple of the fundamental frequency  $\omega_0 = \frac{2\pi}{T}$ , where  $T$  is the length of the complete data set. The sum is called a Fourier series and is given by

(Press *et al.* 2001, chap. 13)

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos n\omega_0 t + b_n \sin n\omega_0 t) \quad (4.1)$$

The constants  $a_0$ ,  $a_n = a(n\omega_0)$ , and  $b_n = b(n\omega_0)$  are called Fourier coefficients. The Fourier coefficients are defined as

$$\begin{aligned} a_0 &= \frac{2}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} f(t) dt \\ a_n &= \frac{2}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} f(t) \cos n\omega_0 t dt \quad n = 1, 2, 3, \dots \\ b_n &= \frac{2}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} f(t) \sin n\omega_0 t dt \quad n = 1, 2, 3, \dots \end{aligned}$$

A discrete function at equidistant time points (separated by  $\Delta t$ ) has a sampling frequency of  $\omega_a = \frac{2\pi}{\Delta t}$ . A Fourier representation is a continuous function and is an approximation on a limited number of points of the discrete function. The Fourier coefficients for the representation of the discrete function  $y_i$  ( $i = 1, \dots, 2m-1$ ) for  $2m$  equidistant measurement points are

$$\begin{aligned} a_0 &= \frac{1}{m} \sum_{i=0}^{2m-1} y_i \\ a_n &= \frac{1}{m} \sum_{i=0}^{2m-1} y_i \cos \frac{2\pi n i}{T} \quad n = 1, 2, \dots, m \\ b_n &= \frac{1}{m} \sum_{i=0}^{2m-1} y_i \sin \frac{2\pi n i}{T} \quad n = 1, 2, \dots, m-1 \end{aligned}$$

Now, equation 4.1 turns into

$$f(t) = \frac{a_0}{2} + \sum_{i=1}^{2m-1} \left[ a_n \cos \left( \frac{2\pi n i}{T} \right) + b_n \sin \left( \frac{2\pi n i}{T} \right) \right] \quad (4.2)$$

This results in  $m$  coefficients  $a_n$ ,  $(m-1)$  coefficients  $b_n$  and a constant  $a_0$ . The latter one represents the average value of the complete data set. In this case, it holds a value of 3.31 mm and 3.07 mm for the simulated and the observed daily precipitation rates over the period 1991–2000, respectively. The sum of the frequency components is  $m + (m-1) + 1 = 2m$ .

The highest frequency to determine a harmonic signal  $f(t)$  from a discrete data set is defined as  $\omega_g$ . In this study with 3,650 data points, it would be  $\omega_g = \frac{1}{3650}$ . However, this representation of  $f(t)$  leads to false results (aliasing effects), because a sine wave needs at least two data points per frequency (Press *et al.* 2001, chap. 12). The highest possible frequency,  $\omega_a$ , to determine  $f(t)$  most accurately is given by the Nyquist theorem and is defined as,

$$\omega_a > 2\omega_g \quad (4.3)$$

For the 10 years reference period 1991–2000 with 3,650 daily data points available, the highest possible frequency  $\omega_a$  is  $\frac{1}{1825}$ .

### 4.3.2 MM5 simulations

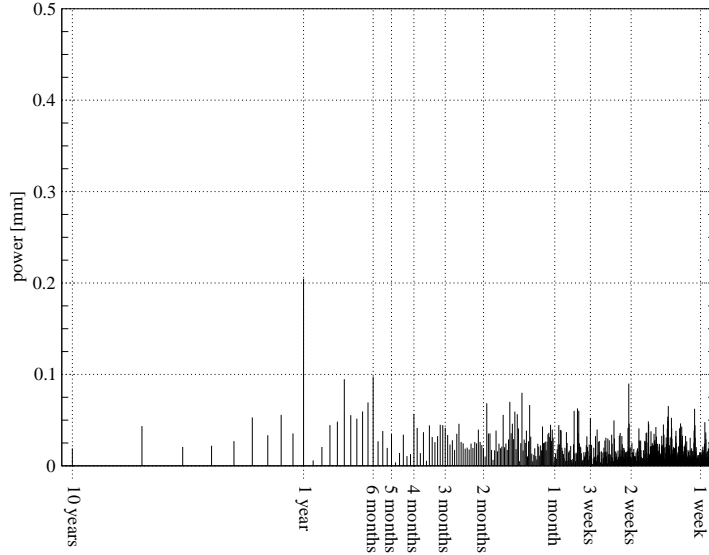


Figure 4.7: Power spectrum of the simulated rainfall rates (x-axis is logarithmic). This estimate is derived from the auto-covariance.

Equation 4.2 sums the harmonics with frequencies from 1 (with a corresponding period of 10 years, due to the length of the data set) to 1,825 (not 3,650, because of eq. 4.3). Each added frequency estimates the time series from Fig. 4.1 more accurately. The time series is estimated exactly when using 1825 frequencies. The amplitude, or power, of each frequency is determined by

$$power = \sqrt{a^2 + b^2} \quad (4.4)$$

where  $a$  and  $b$  correspond to equation 4.2. Fig. 4.7 shows the power spectrum of the simulated spatial averaged rainfall rates using the auto-covariance. The auto-covariance function is described as

$$G(\tau, P) = \frac{1}{n - \tau} \sum_{t=1}^{n-\tau} (P_{t+\tau} - \bar{P})(P_t - \bar{P}) \quad 0 \leq \tau \leq n - 1 \quad (4.5)$$

where  $P$  denotes the areal averaged simulated rainfall rate on each day during 1991–2000,  $\bar{P}$  denotes the temporal averaged rainfall rate over 1991–2000,  $t$  denotes each time step,  $\tau$  denotes the time period over which the auto-covariance is generated, and  $n$  denotes the length of the time series, which in this case is 3,650. The maximum power of 0.20 is located at the yearly frequency. This means that the daily rainfall rates follow an annual cycle, as expected. The power of the first frequency, with a period of 10 years, is small (0.019) and comparable to the power of the majority of the frequencies corresponding to periods less than a year. Except for the annual frequency, the power of the frequencies belonging to periods between 10 years and a week indicate white noise. The power of the frequencies smaller than a week decrease only slightly compared to the powers shown in Fig. 4.7. Pink noise, as suggested by Antal *et al.* (2001) and Yano *et al.* (2001) as a character of daily rainfall rates could not be found.

### 4.3.3 Observations

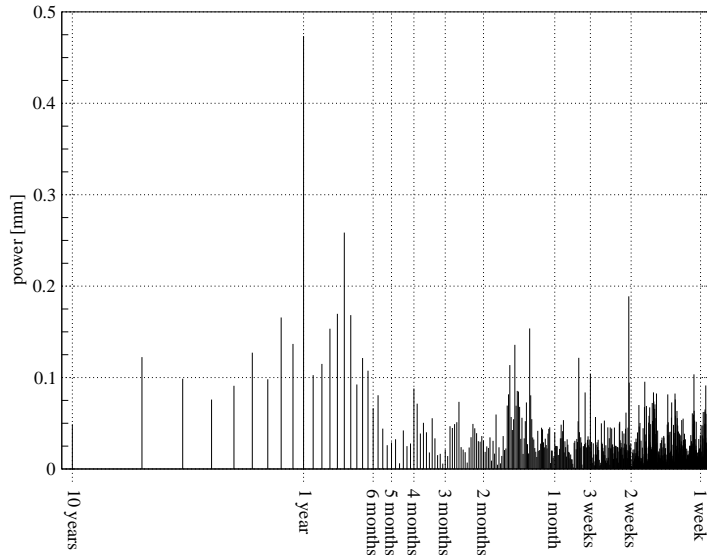


Figure 4.8: Power spectrum of the observed rainfall rates (x-axis is logarithmic). This estimate is derived from the auto-covariance.

The power spectrum of the auto-covariance of the areal averaged observed rainfall rates in the research area for the period 1991–2000 is shown in Fig. 4.8. As for the simulated rainfall rates, the annual frequency is clearly visible with a power 1.14. The powers in Fig. 4.8 are slightly higher than for the simulations in Fig. 4.7. This was already visible by the slightly higher differences in observed rainfall rates between individual days in Fig. 4.4 than in the simulated rainfall rates in Fig. 4.1. Except for the annual cycle, the power of the periods between 10 years and 1 week indicate white noise again.

### 4.3.4 Comparing MM5 with observations

Let us compare the daily rainfall rates of the MM5 simulations with those of the observations for the time period 1991–2000. Except for the annual frequency, in the previous sections no frequency contributed significantly more to determine the annual rainfall cycle than any other one in both data sets. In order to determine how well two data sets coincide, the covariance is introduced here and is shown in Fig. 4.9. According to Fig. 4.9 frequencies up to one month contribute considerably to the agreement between the simulations and the observations. However, sensitivity tests show that with frequencies of one month and higher major rainfall oscillations exist, especially during the winter months. The next peak in the covariance is at frequencies up to 45 days. Frequencies lower than one year do not contribute considerably to the power spectra of the simulations and the observations and have small values for the covariance as well. This suggests no considerable trend between 1991–2000 and these frequencies are neglected. This reduces the spectrum to frequencies between 1 year and 45 days.



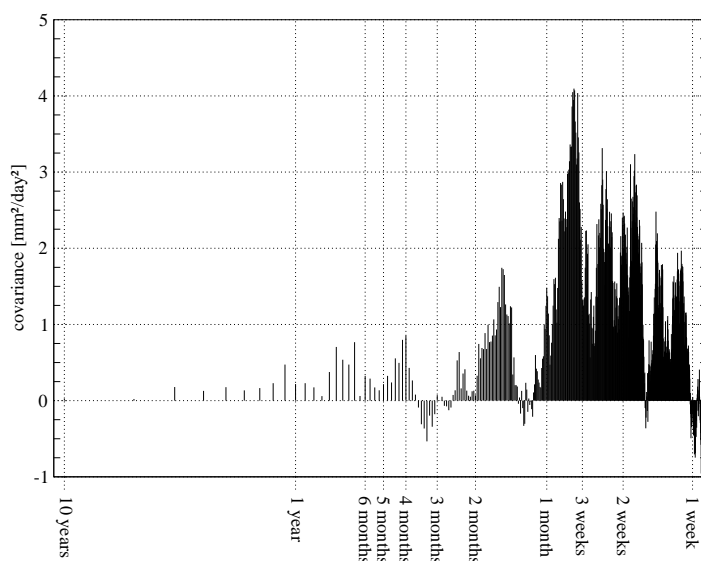


Figure 4.9: Covariance between simulation and observations for increasing frequencies (decreasing periods).

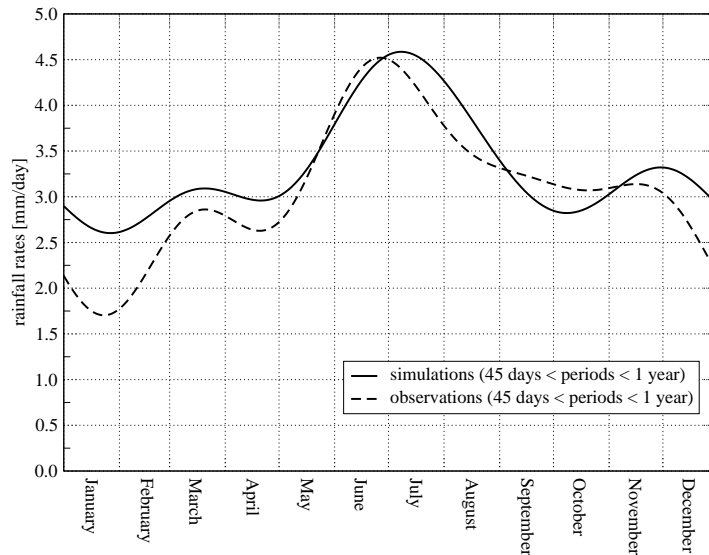


Figure 4.10: Climatological annual cycle of rainfall rates by Fourier analysis using periods between 1 year and 45 days for the MM5 model output and the observations.

The result of the Fourier analysis of the simulations and observations using frequencies between one year and 45 days is shown in Fig. 4.10. Both curves show maximal rainfall rates at approximately 4.5 mm/day during summer (JJA). In winter (DJF) and spring (MAM) the simulated rainfall rates are higher than observed. Especially at the end of January these differences are as large as almost 1 mm/day.

## 4.4 Running average

### 4.4.1 Introduction

Another method to reduce apparently insignificant oscillations in a 10-year data set of daily rainfall rates is to use a running average. The running average method defines a value at  $t_0$  as the linearly averaged value over a chosen amount of days before and after  $t_0$  (for example,  $t_{-5\text{ days}}$  and  $t_{+5\text{ days}}$  for 5 days before and after  $t_0$ , respectively) for each day of a certain time series. An increasing length of the time period over which is averaged, increasingly flattens a time series. The goal of this section is to find a smoothed time series without removing too many fluctuations, like an annual cycle.

The running average method is defined by

$$R_0^{ra} = \frac{1}{n} \sum_{i=0}^{n-1} R_{i-\frac{n-1}{2}} \quad (4.6)$$

where  $R$  denotes the parameter to be averaged (in this case the daily rainfall rate).  $R_0^{ra}$  is the averaged rainfall rate over  $n$  time steps at time step  $t = 0$ . Since days are discrete,  $n$  should be an odd number.

A disadvantage of the running average method is that it requires adjustments at the start and at the end of a data set. In order to calculate the first and the last  $R_0'$  of the data set, data exceeding the limits of a data set are needed. This problem has an effect on a larger part of a data set with increasing integration periods, i.e. a larger  $n$ .

A solution would be to decrease the integration period,  $n$ , at the edges of the data set. This means that the first and last  $R_0'$  would equal  $R_0$  (see equation 4.6). It is a disadvantage of this method that not every part of the data set is an average over the same integration period.

Another solution is to neglect such a number of data points at the start and at the end of the data set for which additional data are needed from outside the data set. The goal of this study is to build a yearly course of rainfall rates from a relatively short period of 10 years. Neglecting data would shorten the period even more and therefore not suitable for this study.

As a final solution it is assumed that there is no trend during the 10-year period. This was already assumed in section 4.3. Consequently, the data set may assumed cyclic and the start of the data set is defined as part of the start of the data set and part of the end. This provides a value using an equally long averaging period for each time step in the data set. Now, every value in the climatological year can be seen as an average over  $n \times 10$  years, where  $n$  is the running average period.

Independent of the length of the period, the individual days are linearly averaged. This implies that all days in the averaging period contribute with the same weight to the averaged value. This may lead to unrealistic jumps in precipitation if the data varies a lot between individual days. This effect can be weakened by reducing the influence of days farther away from the middle of the averaging period. For example, this can be done by multiplying the averaging period with the Gaussian distribution. An advantage is that extreme events at the edges of the averaging period do not have a great influence on the averaged value as values in the middle. Consequently, by weakening the influence of days farther from the middle, days closer to middle have a stronger influence, which results in

more fluctuation between individual days. A much longer averaging period would be needed to remove of these daily variations. This was confirmed by some additional tests, which are not shown here. Although in most cases, days far from the middle of the averaging period do not have anything to do with days in the middle, it is the goal of the running average method to smoothen the annual rainfall cycle, removing any insignificant daily variations. Therefore, linearly averaged periods are used.

#### 4.4.2 MM5 simulations

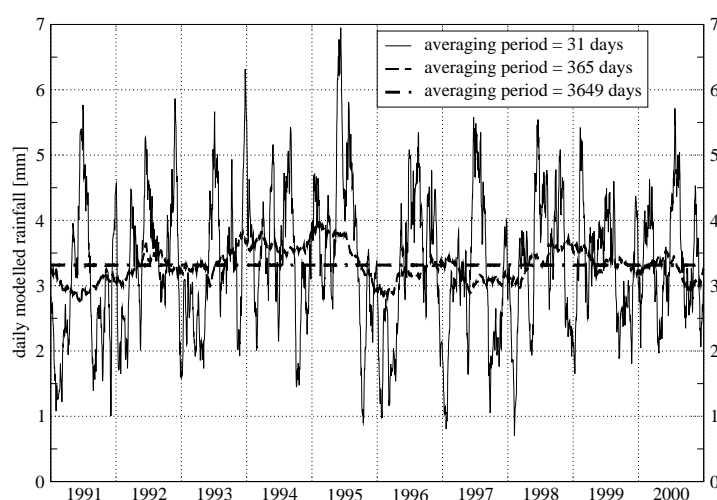


Figure 4.11: Examples of a running average over a period of 31, 365 and 3649 days using the data shown in Fig. 4.1.

The data set used to build a running average time series for the MM5 model simulations is the same as used in section 4.1 and shown in Fig. 4.1. It is obvious that using an  $n$  of 3649 days, and using the assumption of a cyclic data set described in section 4.4.1, a 10-years averaged value is constant and equals  $a_0$  from equation 4.2 and 4.2. The other limit is  $n = 1$  and results in the exact representation of the time series. To show the effectiveness of the running average method, the results with a running average period of 31, 365 and 3649 days are shown in Fig. 4.11 by example.

Averaging the rainfall over three days already filters out most of the high peaks and results in a large difference in standard deviation. Except for running average up to 11 days, no running average period contributes more to the representation of the time series than another according to Fig. 4.12.

Figure 4.13 shows the skewness,  $\gamma_1$ , and the kurtosis,  $\gamma_2$ , depending on the length of the averaging period  $n$ . The skewness is symmetrical about a running average of 1825 days and the kurtosis is rotated by  $180^\circ$  around that point. The symmetrical character of  $\gamma_1$  and  $\gamma_2$  in Fig. 4.13 is explained by the way of averaging. The assumption of a cyclic data set in section 4.4.1 means that if a running average period exceeds the temporal boundaries

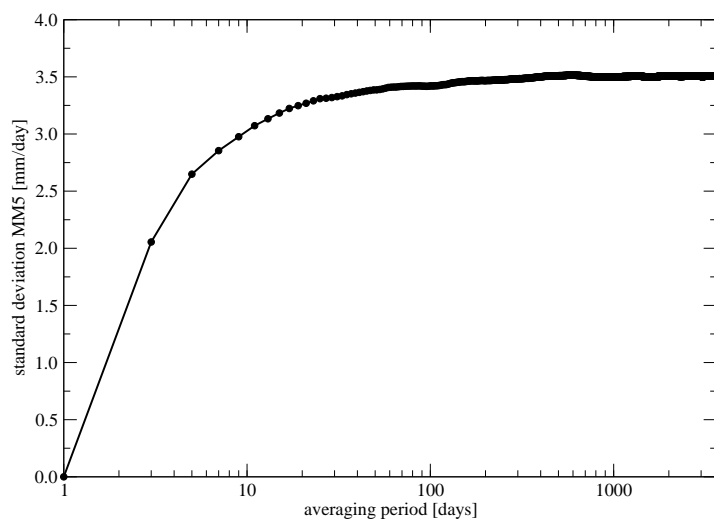


Figure 4.12: Standard deviation of the difference between the original MM5 simulations and the averaged data with a period shown on the x-axis (logarithmically).

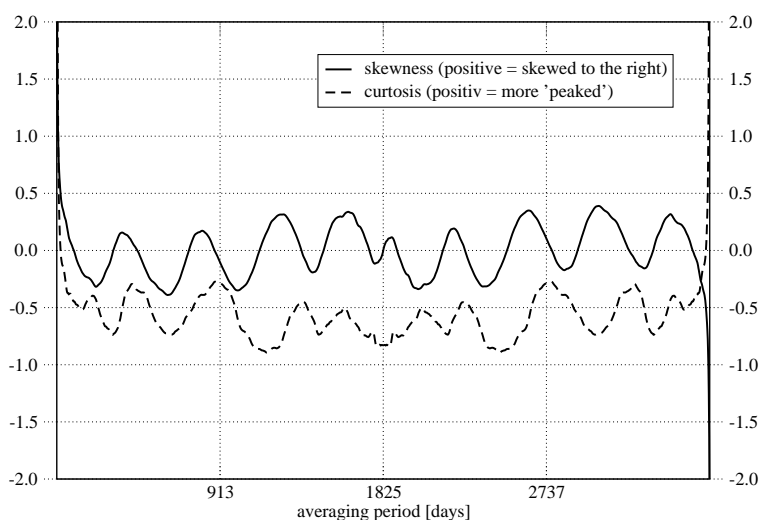


Figure 4.13: Skewness ( $\gamma_1$ ) and kurtosis ( $\gamma_2$ ) of the averaged MM5 model simulations with the periods shown on the x-axis.

at the end of the data set, the exceeding days will be taken from the start of the data set. The same holds for a running average period at the beginning of the data set for which data from the end of the data set are taken. As a result, the skewness,  $\gamma_1$ , of the  $n$ th running average period is negatively correlated to the  $(3650 - n)$ th running average period. Because the frequency distributions of the  $n$ th and  $(3650 - n)$ th running average period are

mirrored, the kurtosis,  $\gamma_2$ , of the  $n$ th running average period is equal to the  $(3650 - n)$ th running average period. Additionally, a running average period of more than 1825 days (5 years) is obviously too long to capture the interannual rainfall variation and is therefore be neglected from now on. Fig. 4.14 shows the frequency distributions of the data with a running average of 31 and 3,619 (3,650-31) days (also marked in Fig. 4.13). The statistical quantities belonging to these two running average periods are shown in Table 4.2.

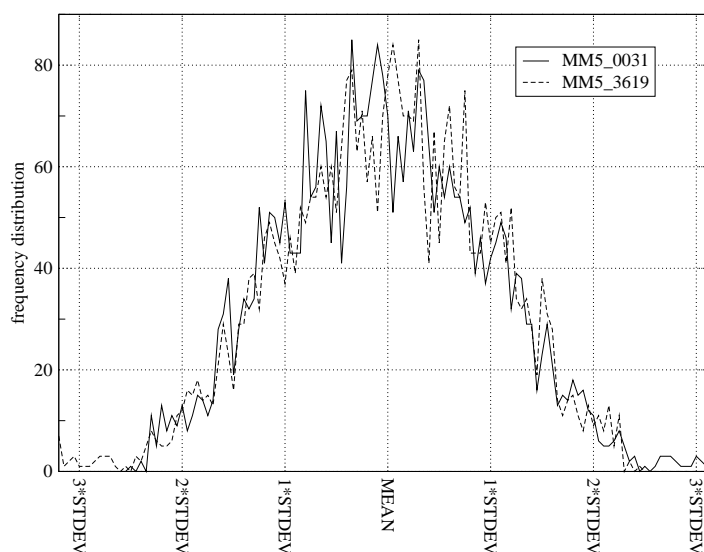


Figure 4.14: Frequency distributions of the simulations with a running average of 31 and 3,619 (3,650-31) days.

<i>av.per.</i>	$\mu$	$\sigma$	$\gamma_1$	$\gamma_2$
0011	3.315	1.573	0.402	-0.026
0031	3.315	1.048	-0.231	0.176
0091	3.315	0.687	0.002	-0.344
0295	3.315	0.311	-0.012	-0.042
3619	3.315	0.009	-0.231	-0.176

Table 4.2: Some statistical quantities of the simulated data with running average periods of 11, 31, 91, 295, and 3619 days.

The skewness,  $\gamma_1$ , and the kurtosis,  $\gamma_2$ , of a standard normal distribution equals zero (see section 4.1). The distribution of rainfall has a  $\gamma_1$  larger than 0 due to the many small rainfall rates, which are bounded on the left by 0. The  $\gamma_2$  for daily rainfall rates is usually larger than 0 as well, due to the large amount of small rainfall rates (see e.g. Table 4.1). Smoothing extreme rainfall rates from the data set would change the distribution of rainfall to a more standard-like distribution. From Fig. 4.13, three running average periods of interest are found. The first one has a  $\gamma_2$  of (almost) 0, corresponding to a running average period of 11 days. The second one has a  $\gamma_1$  of (almost) 0, corresponding to a running average period of 91 days (roughly 3 months). The third and last point of interest has

both a  $\gamma_1$  and  $\gamma_2$  of (almost) 0 and corresponds to a running average period of 295 days (almost one year). Using the running average period of 11, 91 and 295 days, a climatological time series is generated and shown in Fig. 4.15. This means that a 10-years time series is created using these running average periods and subsequently the average of each day is built.

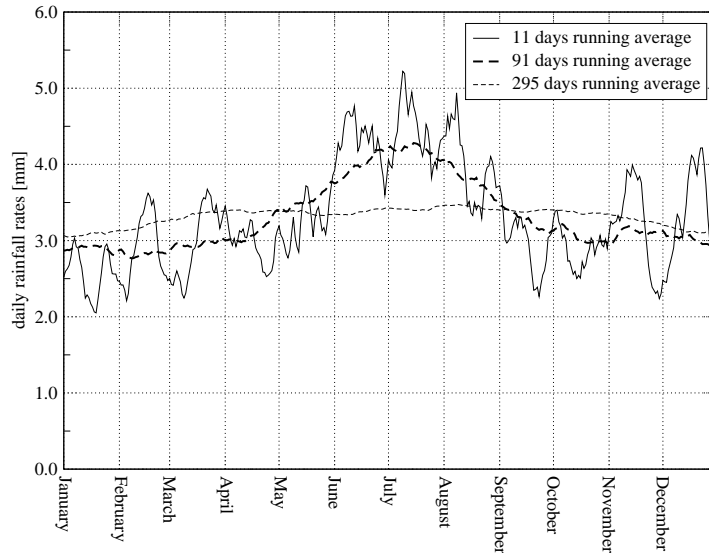


Figure 4.15: Climatological year using a running average period of 11, 91 and 295 days.

The time series with a running average period of 295 days has a  $\gamma_1$  and  $\gamma_2$  of 0 and is therefore distributed as a standard distribution, but almost lacks to have an annual cycle. The time series with a running average period of 11 days shows a strong annual cycle, but also many unrealistic jumps in rainfall rates between individual days. This is the result of only the skewness,  $\gamma_2$ , equaling zero, like with a normal distribution, and the kurtosis,  $\gamma_1$ , larger than zero. The high value for  $\gamma_1$  indicates that the time series is strongly skewed to the right. The time series with a running average period of 91 days shows a rather smooth annual cycle. It shows rather small differences in precipitation rates between individual days, due to its mean over  $91 \times 10$  years ( $= 910$ ) values.

### 4.4.3 Observations

As for the MM5 model simulations in section 4.4.2, the same tests are performed for the observations. The used observations are shown in Fig. 4.4.

Like in Fig. 4.13, the skewness and kurtosis of the observations are shown in Fig. 4.16. Again, the skewness is symmetrical about the middle and the kurtosis is rotated by  $180^\circ$  around a point in the middle. Of course, the reasons are the same as well.

In section 4.4.2, it was found that the best annual cycle with the least daily variation was produced using an running average period for which the skewness equalled (almost) 0. The first time this occurs in Fig. 4.16 is at a running average period of 95 days, which almost equals the running average period from section 4.4.2 for the model simulations.

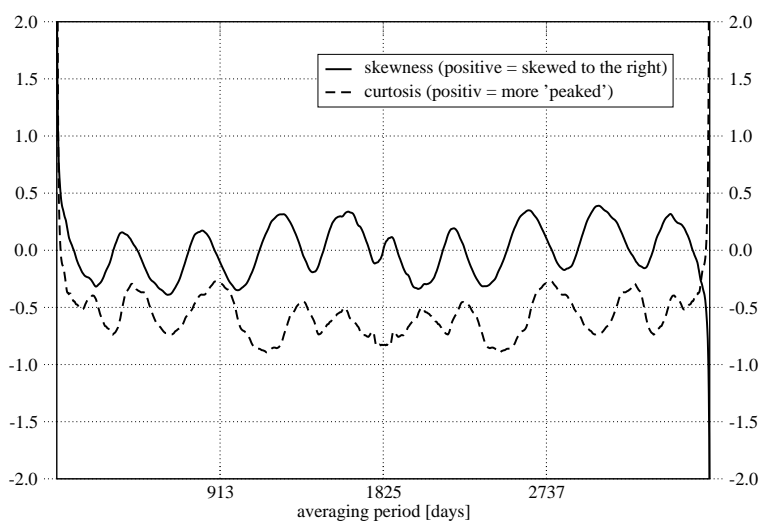


Figure 4.16: Skewness and kurtosis of the averaged observational data with the periods shown on the x-axis.

The skewness corresponding to a running average period of 91 days for the observations is also very small and therefore a running average period of 91 days is used to generate the climatology for the observations and is shown in Fig. 4.17. The climatologies of the MM5

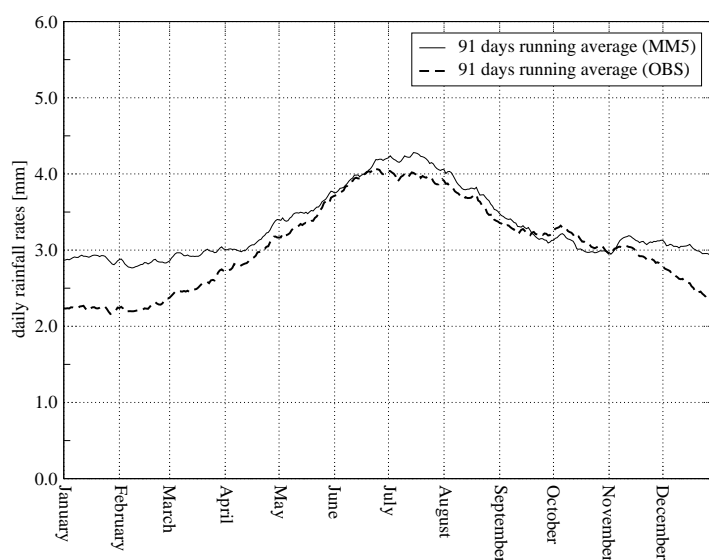


Figure 4.17: Climatological year using a running average period of 91 days for the MM5 model simulations (MM5) and the observations (OBS).

model simulations and the observations correspond quite well. Both climatologies show a maximum of around 4 mm in July and a minimum in the winter months (DJF). The

lower rainfall rates during the winter months for the observations compared to the model simulations confirm the results found in Fig. 4.10. Any conclusions drawn from Fig. 4.17 is discussed in section 4.4.4.

#### 4.4.4 Comparing MM5 with observations

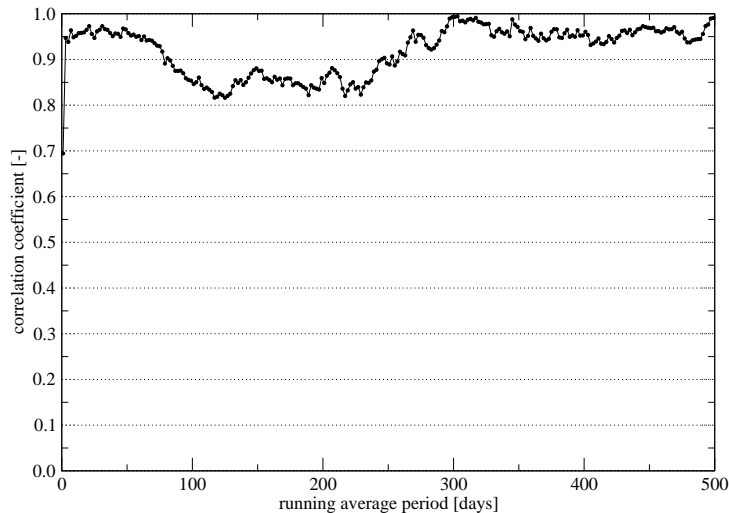


Figure 4.18: Correlation coefficients between the power spectra of the rainfall distribution from the MM5 model simulations and the observations.

In sections 4.4.2 and 4.4.3, a smooth annual rainfall cycle was found for the MM5 model simulations and the observations, respectively. For both the simulations as well as for the observations, an almost equal running average period was found of a bit over 90 days. This period describes the seasonal cycle within each year and does not necessarily mean a good representation of the annual precipitation cycle during 1991–2000. Short periods of about five days usually do not represent the annual precipitation cycle very well either, since many differences between the individual years still exist, resulting in many fluctuations in a climatological year. No specific time-scale seems to be more important than another, except for the annual time-scale (see section 4.3.2). Better results could be gained using a compromise between these two time-scales (Wirth 2005, pers. comm.). This section will therefore focus on the analogy between the MM5 model simulations and the observations, more than on the individual time series.

The goal is to find for which running average period the MM5 model simulations and the observations coincide best. It is not only the goal to find a smooth annual distribution of rainfall, although many small daily variations are not significant representations of the 10 years reference period from 1991–2000 and will most likely differ between the simulations and the observations. Generally, daily rainfall rates in summer are higher than in winter and vary more due to more heavy summer rainfall showers.

The correlation between the power spectra of both data sets is regarded to find this annual cycle. It is assumed that the correlation between the power spectra is high, when



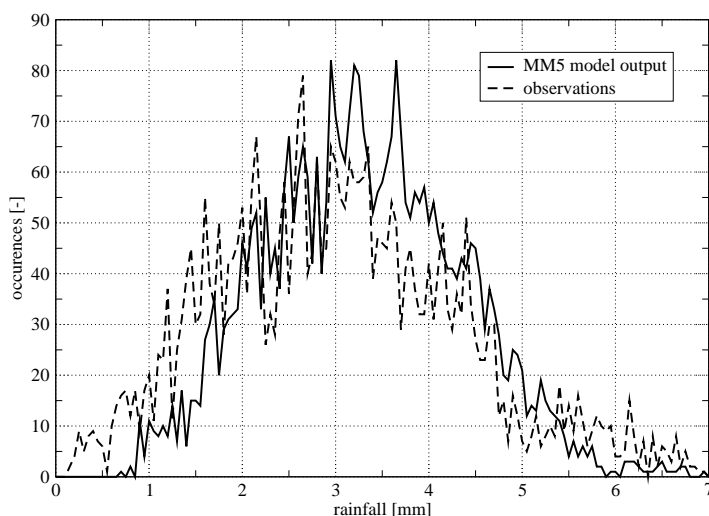


Figure 4.19: Rainfall distributions of the MM5 model simulation and the observations at a running average period of 31 days.

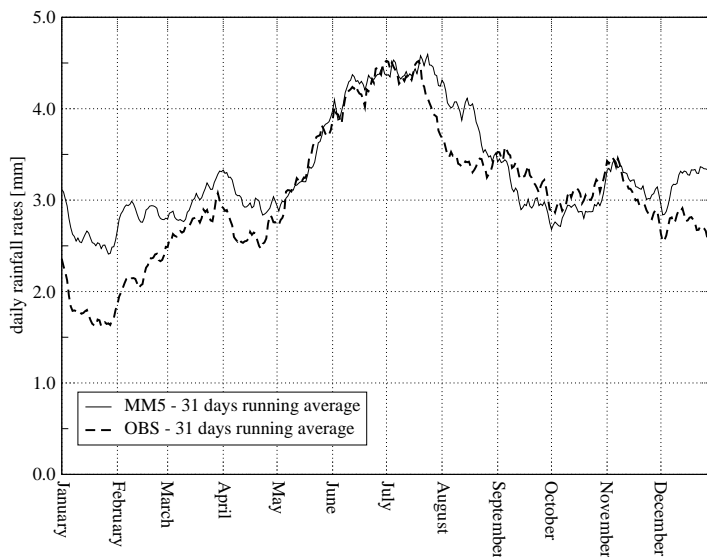


Figure 4.20: The climatological annual cycle of rainfall from the MM5 model simulations and the observations using a running average period of 31 days.

the correlation between the rainfall distributions is high as well. The benefit of this correlation coefficient against the coefficient of the data itself is that every rainfall rate is treated the same. Days with high rainfall rates do not have more influence on the correlation coefficient than days with low rainfall rates have. The height of the correlation coefficient for running average period between 1 and 500 days is shown in Fig. 4.18. To avoid

statistically insignificant data playing a too important role in generating the correlation coefficient, only data within a bin width of two standard deviations are used. This means that 5% of the data at the edges of the distribution is neglected. In Fig. 4.18, a running average period of 1 day (which actually means raw data), has a correlation coefficient of 0.69. The coefficient is then rising until a maximum of 0.97 at a running average period of 31 days. Naturally, for longer integration times, the correlation coefficient is high as well. However, this is an artifact of the averaging; both rainfall curves will eventually be constant. In section 4.4.2 it was already mentioned that a running average period exceeding 295 days fails to represent the natural annual rainfall distribution. The two distributions of the model and the observations with a running average period of 31 days are shown in Fig. 4.19. The climatological rainfall curves for the MM5 simulations and the observations using a running average period of 31 days are shown in Fig. 4.20. The running average period of 31 days seems to be a good compromise between seasonal and daily rainfall cycles described above and is preferred over the 91 days of section 4.4.2 and 4.4.3.

## 4.5 Spline interpolation

### 4.5.1 Introduction

The third method to generate a climatological year uses fixed data points. The days in between are mathematically estimated by spline interpolation (Press *et al.* 2001). An advantage of this method is the removal of unwanted sub-monthly variations. A disadvantage is the inflexibility of the data points through which is interpolated. Spline interpolation is able to generate maxima between two data points, but these will always be mathematically generated and not in any sense directly based on the information between the data points.

The equation describing the interpolation curve is (Press *et al.* 2001),

$$y = Ay_j + By_{j+1} + Cy_j'' + Dy_{j+1}'' \quad (4.7)$$

where

$$\begin{aligned} A &\equiv \frac{x_{j+1} - x}{x_{j+1} - x_j} \\ B &\equiv 1 - A = \frac{x - x_j}{x_{j+1} - x_j} \\ C &\equiv \frac{1}{6}(A^3 - A)(x_{j+1} - x_j)^2 \\ D &\equiv \frac{1}{6}(B^3 - B)(x_{j+1} - x_j)^2 \end{aligned}$$

and  $y_j$  and  $y_{j+1}$  are two data points with  $y_j''$  and  $y_{j+1}''$  their second derivatives, respectively. The second derivatives are determined as follows. Taking the first derivative of eq. 4.7 with respect to  $x$  using the definitions of  $A$ ,  $B$ ,  $C$ , and  $D$  gives

$$\frac{dy}{dx} = \frac{y_{j+1} - y_j}{x_{j+1} - x_j} - \frac{3A^2 - 1}{6}(x_{j+1} - x_j)y_j'' + \frac{3B^2 - 1}{6}(x_{j+1} - x_j)y_{j+1}'' \quad (4.8)$$

and the second derivative gives

$$\frac{d^2y}{dx^2} = Ay_j'' + By_{j+1}'' \quad (4.9)$$

Since  $A = 1$  at  $x_j$  and  $A = 0$  at  $x_{j+1}$ , while  $B$  is just the opposite of  $A$ , it shows that the second derivative is continuous. Requiring that the first derivative should be continuous as well, allows to calculate the still unknown second derivative. After some rearranging of eq. 4.8 it gives

$$\frac{x_j - x_{j-1}}{6} y''_{j-1} + \frac{x_{j-1} - x_{j+1}}{3} y''_j + \frac{x_{j+1} - x_j}{6} y''_{j+1} = \frac{y_{j+1} - y_j}{x_{j+1} - x_j} - \frac{y_j - y_{j-1}}{x_j - x_{j-1}} \quad (4.10)$$

For a unique solution,  $y''_{j-1}$  and  $y''_{j+1}$  are set to zero, giving a so-called natural cubic spline (Press *et al.* 2001).

### 4.5.2 MM5 simulations

Again, like in section 4.3.2 and 4.4.2, the basis to analyse the daily rainfall rates of the MM5 model output is shown in Fig. 4.1. In order to be able to represent the annual cycle sufficiently monthly mean values will act as the data points described above. In section 4.4.4 it was already found that a monthly time period is a good compromise between seasonal and daily rainfall cycles. Figure 4.21 shows the monthly means for the period 1991–2000. Every monthly mean is defined at the middle of each month, where it is assumed to be most representative for that particular month. The monthly means, shown in Fig. 4.21,

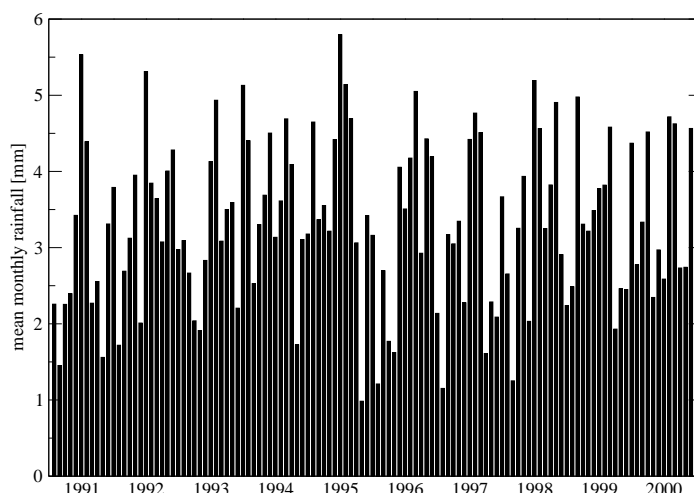


Figure 4.21: Mean simulated monthly rainfall for the period 1991–2000.

are averages over periods of 28-31 days and much variability is found throughout the 10 years reference period. Nevertheless, an annual rainfall cycle is recognized. Since the daily variation is large, the standard deviation of each monthly mean is large as well. Figure 4.22 shows the monthly means with their standard deviations for 1991. The standard deviation is in particular applicable to single-peaked symmetrical distributions (Press *et al.* 2001). This is roughly approximated by monthly rainfall data, as shown in Fig. 4.19 by a running average period of 31 days. In section 4.4 the problem of how to handle boundary conditions

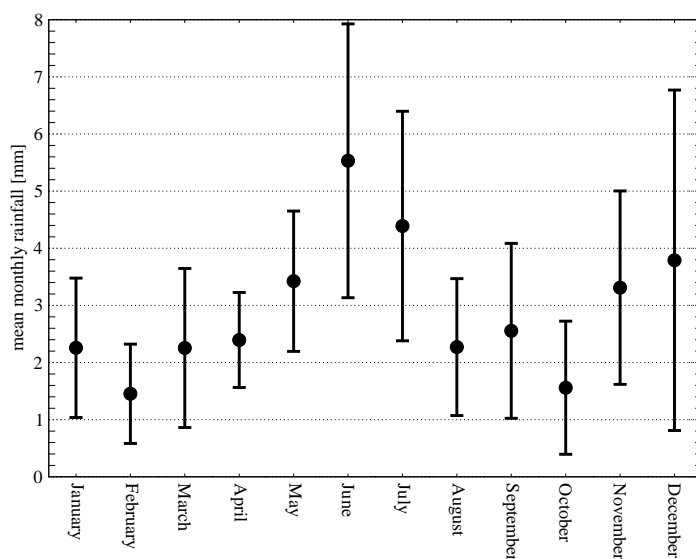


Figure 4.22: Mean monthly rainfall for the year 1991 and the individual standard deviations.

was solved by assuming a cyclic data set. This assumption is used here as well. In order to interpolate values later than 16 December 2000, it is assumed that the mean of January 2001 equals the mean of January 1991. Also, it is assumed that the mean of December 1990 equals December 2000. Figure 4.23 shows the monthly means for 1991 with the spline interpolated values on a daily basis. The different amount of days per month is considered. Every monthly mean in Fig. 4.23, which is defined at the middle of each month, is hit by

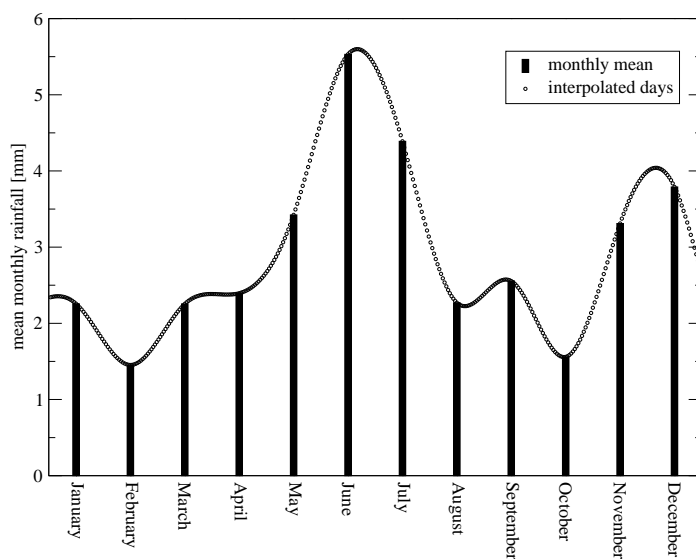


Figure 4.23: Mean monthly rainfall for the year 1991 with the spline interpolated values.

the interpolation curve. All values in between two monthly means are the result of the

mathematical equation given by equation 4.7. The apparent maximum rainfall just after 15th June and at the beginning of December is purely mathematical and not in any sense based on daily observations. Calculating an interpolation curve for the reference period 1991–2000 and building a yearly mean results in the curve shown in Fig. 4.24.

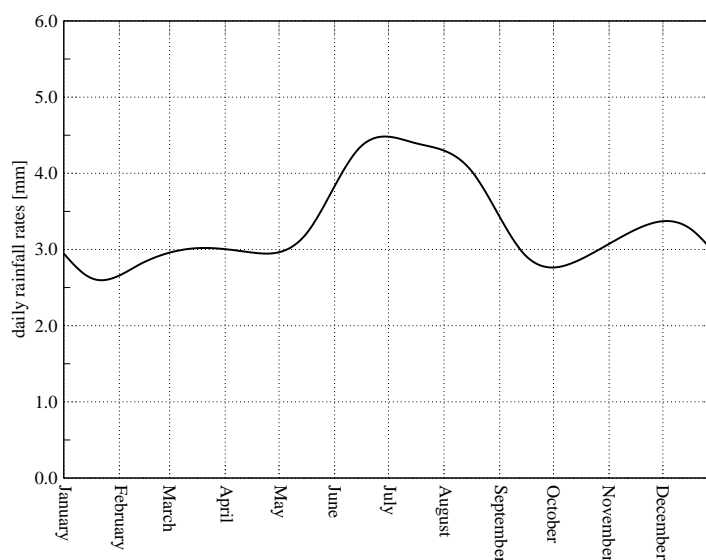


Figure 4.24: The climatological annual cycle of rainfall from the MM5 model simulations using monthly means and spline interpolation.

### 4.5.3 Observations

The data for the spline interpolation of the observations is shown in Fig. 4.4. Like in section 4.5.2, monthly means are built using the daily rainfall data. Then, using these monthly means, the days in between are estimated using spline interpolation. Climatological observed values using spline interpolation are shown in Fig. 4.25, together with the results from section 4.5.2.

## 4.6 Intercomparison

The goal of this chapter was to find a climatological annual rainfall cycle based on the reference period 1991–2000 using daily rainfall rates. Linearly averaging was tried in section 4.2, but did not remove adequately the many insignificant jumps in rainfall rates between individual days. In sections 4.3, 4.4, and 4.5, three different methods were approached, which were Fourier analysis, running average, and spline interpolation, respectively. The results of the three approaches for the MM5 model simulations is shown in Fig. 4.26. The results of the three approaches for the observations is shown in Fig. 4.27. The results of the Fourier analysis and spline interpolation are almost identical. The running average approach still shows many fluctuations throughout the climatological year. A disadvantage of spline interpolation is its rather arbitrary choice of the monthly basis values, whereas the results

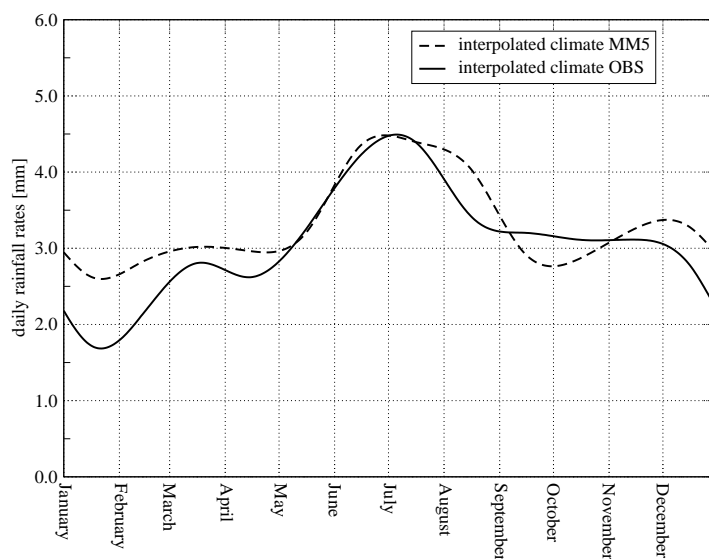


Figure 4.25: The climatological annual cycle of rainfall from the MM5 model output using monthly means and spline interpolation.

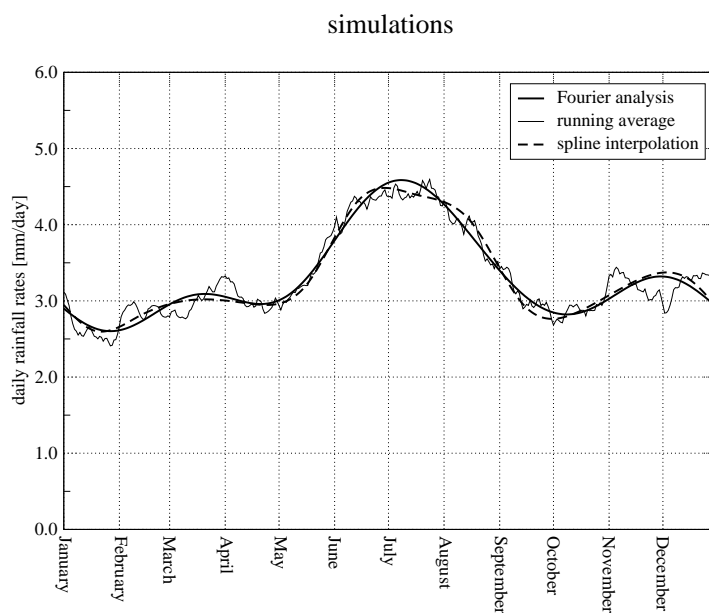


Figure 4.26: The climatological annual cycle of rainfall from the MM5 model simulations using Fourier analysis ( $45 \text{ days} \leq \text{period} \leq 1 \text{ year}$ ), running average (31 days) and spline interpolation (monthly means).

of the Fourier analysis are based on the significance of the individual frequencies in the original data base. Unfortunately, applying Fourier analysis spatially, i.e. for each 1 km grid cell, negative rainfall rates were generated, because the fluctuations in time between individual grid cells are too large. This leaves spline interpolation to be used in the next

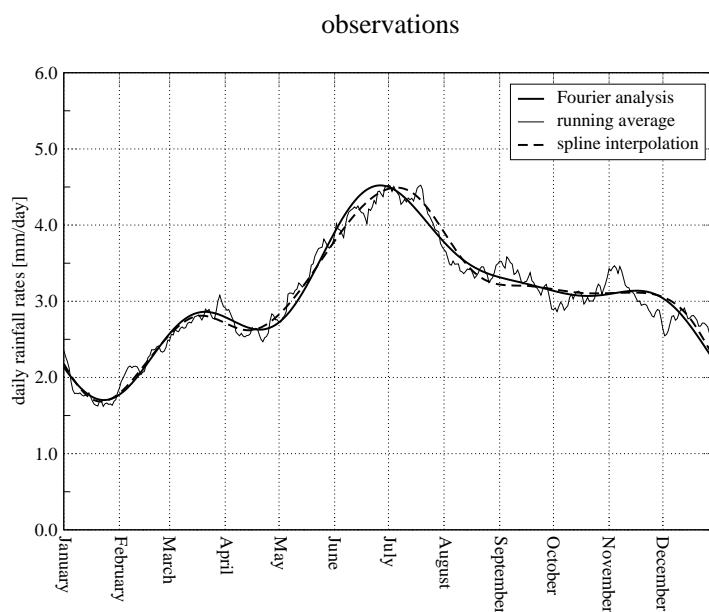


Figure 4.27: The climatological annual cycle of rainfall from the observations using Fourier analysis ( $45 \text{ days} \leq \text{period} \leq 1 \text{ year}$ ), running average (31 days) and spline interpolation (monthly means).

chapters.

The climatological annual cycle of the simulated and observed precipitation show a high agreement. This confirms that the annual cycle was captured quite well by the model. Still, a comparison with other regional climate models would confirm this even more. In the from the European Union financed project PRUDENCE (contract EVK2-CT2001-00132) 8 models simulated, among others, precipitation rates over a time period of 30 years. This time period covered 1961 until 1990. Unfortunately, this does not coincide with the reference period in this study. Simulations for the reference period 1991–2000 are planned within the framework of the project in the next half year (Rockel 2005, pers. comm.). An intercomparison with the MM5 model would be challenging.





# Chapter 5

## Empirical downscaling

### 5.1 Introduction

As has been stated above, the low resolution of the grid has consequences for the representation of the orography in the model (see Fig. 3.1). The highest peaks of the Alps at the 1 km resolution are about 3,500 meters asl and the maximum height at the 45 km resolution of the MM5 model is about 2,200 meters. The clearly visible valleys at the 1 km resolution are not visible at the 45 km resolution. In Chapter 4 the daily rainfall rates during the reference period of 1991–2000 were discussed. These observed and simulated rainfall rates extended the research area of 450×450 kilometers and were averaged spatially. However, the distribution of rainfall affect the hydrological water cycle, which is of great importance to the other disciplines, which are represented in the GLOWA-project. In this chapter typical spatial distributions of rainfall at both resolutions over the Alps and their forelands are discussed (see Fig. 5.1). Much effort will be given to find agreements and disagreements between the rainfall distributions between both resolutions.

During the last decade many techniques have been developed to reduce disagreements between scales (see e.g. Widmann *et al.* 2003; Murphy 1999; Zorita and von Storch 1999). An overview of many downscaling techniques can be found in Wilby and Wigley (1997). On the one hand so-called limited area models (LAMs), e.g. MM5, usually covering an area of around  $10^7$  km<sup>2</sup>, scale down the simulation output of general circulation models (GCMs). It is expected that the higher resolution of the LAM gives better results in regional simulations than the GCM. Still, results are very dependend on the parameterization chosen in the LAM (see Chapter 3) to account for subgrid processes, as well as indirectly dependend on the parameterizations in the GCM. On the other hand statistical downscaling methods are developed. These methods are usually relatively simple and fast compared to the LAMs. Buishand *et al.* (2004) described a method to compare downscaled simulations with observations using rainfall occurrences on a daily and a monthly basis. In contrast, Zorita and von Storch (1999) compared different methods to estimate the amount of rainfall at observation stations in the Iberian Peninsula. Though, most methods use the relationship between predictors, like large-scale structures, and predictands, like precipitation patterns, at a higher resolutions to correct the GCM. The method introduced in this study corresponds to the statistical methods, but uses the relationship between the same variables, in this case precipitation rates, at a 45 km and a 1 km resolution. Corrected precipitation rates at a resolution of 1 km are compared to spatial distributed observations

at the same resolution (Appendix B). The downscaling method uses the climatological rainfall distributions from Chapter 4 and is verified with the observations for different time periods.

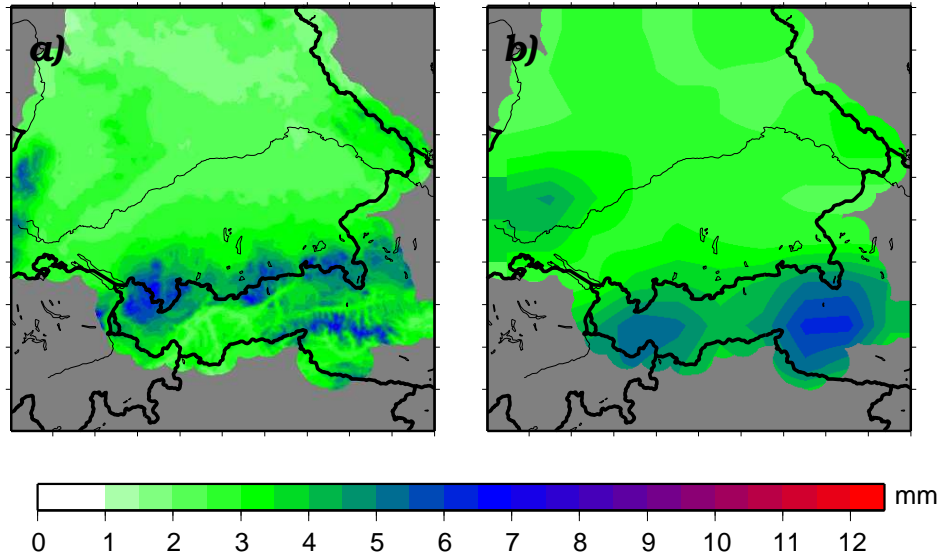


Figure 5.1: Averaged areal rainfall distribution from the observations at a resolution of 1 km (a) and the MM5 model simulations at a resolution of 45 km (b) over the period 1991–2000.

## 5.2 Small orographic details

### 5.2.1 Introduction

Local differences in rainfall rates between the 1 km resolution and the 45 km resolution are obvious (Fig. 5.1) and show strong analogies in both the amount and the distribution of precipitation with the Bavarian Climate Atlas (BayFORKLIM 1996) created during the project BayFORKLIM (BayFORKLIM 1999a,b). Figure 5.2 shows a vertical cross section of the orography of both resolutions at the highest peak of the 1 km resolution, which is located 135 km east of the western border of the research area. The slope at the 45 km resolution rises much slower than the slope at the 1 km resolution, which exhibits deep narrow valleys and high steep mountain peaks. The smooth rise between the 45 km grid cells is caused by bilinear interpolation, which assumes a linear rising between two grid cells. The reason for using this interpolation is explained by means of precipitation patterns later in this chapter. The northern part of the research area (the right side in Fig. 5.2) shows hardly any horizontal height differences. The differences in topography between both resolutions are not so pronounced in the northern part as in the southern part.

The difference in resolution affects the areal distribution of meteorological parameters like precipitation (see Fig. 5.1), which is the focus here. Figure 5.1a shows a clear distinction between relatively dry valleys and wet mountain tops. The reason for these local differences is that air is lifted by mountain ridges and due to dry adiabatic expansion cools

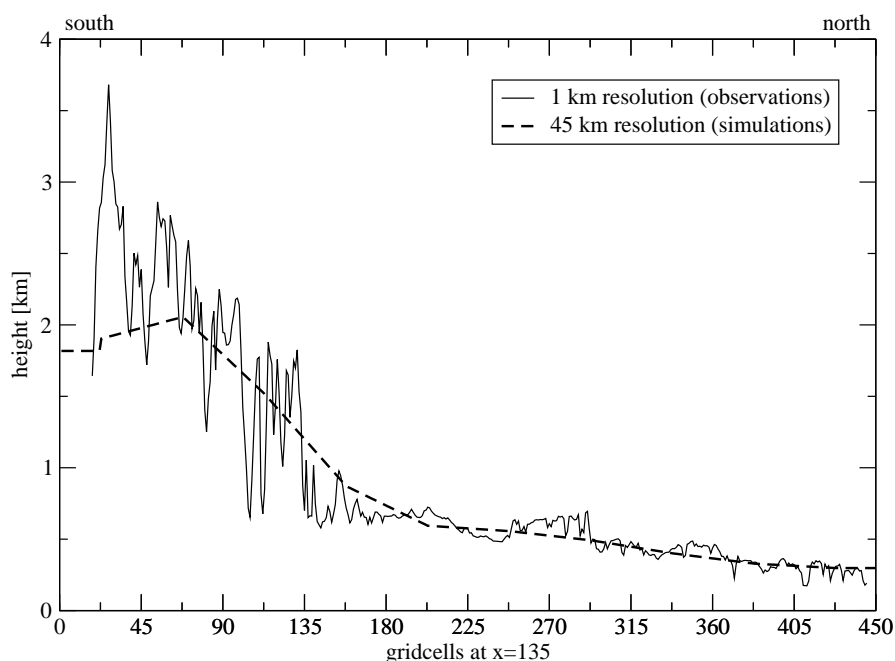


Figure 5.2: Vertical cross section of the 1 km orography on which the observations are interpolated (OBS) and the 45 km orography from the model simulations (MM5) at the highest peak in the 1 km orography.

down until it reaches the lifting condensation level (LCL). When the air rises above the LCL the available water vapor condensates, small droplets are generated and eventually precipitation is formed (see e.g. Kraus 2000; Stull 1999; McIlveen 1995; Garratt 1992). Air rises at mountain ridges by upslope winds resulting in an increasing precipitation rate at mountain slopes. It explains the long-term precipitation pattern of relatively wet mountain tops in regions like the Alps. In the models representation of the Alps no such valleys are available (Fig. 3.1b) and, consequently, the simulated precipitation pattern lacks valley-like structures as well (Fig. 5.1b).

The precipitation fields in Fig. 5.1b are bilinearly interpolated as was done with the models orography. The reason for this interpolation it that an equal distribution of rainfall rates at each 1 km cell in the  $45 \times 45$  km area would cause large differences in rainfall rates between two grid cells at the 45 km resolution. The differences between grid cells at the 45 km resolution would become differences between cells at the 1 km resolution. This would be highly unnatural and have a huge negative effect on the hydrological models in the GLOWA-project, which requests meteorological data at a 1 km resolution. A bilinear interpolation assumes a gradual linearly increasing or decreasing in precipitation amount between the grid cells at the 45 km resolution. In this case, the middle of a 45 km grid cell is assumed to be most representative for the area of  $45 \times 45$  km surrounding it. Other interpolation types, like bicubical and spline interpolation (Press *et al.* 2001), were tested, but generated negative rainfall rates during dry periods and are not mass conservative.

Bilinear interpolation is mass conservative, which means that no rain is created or depleted within the area of interpolation.

### 5.2.2 Correction

Figure 5.1 shows the average precipitation pattern for the period 1991–2000. However, daily patterns differ slightly from this climatological pattern. The climatological year was defined in Chapter 4 as the average year of the period 1991–2000 using spline interpolation. For each day in the observed climatological year, each 1 km grid cells is proportional to another. For example, the precipitation rate on a mountain top is  $x$  times higher than in the valley on a particular day. If  $x$  is known, the precipitation rate in the valley can be estimated, provided that the mountain top precipitation rate is known, or vice versa. Instead of using the proportions between 1 km grid cells, the proportion between a 1 km grid cell and a linearly averaged number of cells is used. In this case, the number of cells that is averaged is 2025, which corresponds to the amount of 1 km cells in the 45×45 km size of a grid cell in the 45 km resolution. Between each 1 km cell and the 45×45 km cell in which it is located a proportion can be derived, which is given by

$$f_{\text{details}} = \frac{\overline{\text{OBS}}_{45}}{\overline{\text{OBS}}_1} \quad (5.1)$$

In eq. 5.1, the correcting factor,  $f_{\text{details}}$ , is the proportion of the 45×45 km linearly averaged climatological observed precipitation,  $\overline{\text{OBS}}_{45}$ , to the 1 km climatological observed precipitation,  $\overline{\text{OBS}}_1$ . The 45×45 km cells correspond spatially with the cells from the 45 km resolution of the MM5 model.

Fig. 5.3 shows the results of eq. 5.1. Panel (a) shows the distribution of the averaged observed rainfall,  $\overline{\text{OBS}}_{45}$ , and panel (b) the observed rainfall with a resolution of 1 km,  $\overline{\text{OBS}}_1$ . Panel (c) shows the division of panel (a) and (b) following eq. 5.1. The red areas in panel (c) of Fig. 5.3 denote areas where  $\overline{\text{OBS}}_{45}$  is lower than  $\overline{\text{OBS}}_1$ . These areas are predominantly the higher Alpine mountains which are resolved at the 1 km resolution and have relatively high precipitation rates. The blue areas denote a higher  $\overline{\text{OBS}}_{45}$  than  $\overline{\text{OBS}}_1$ , which are mainly dryer valleys at the 1 km resolution. Instead of generating  $f_{\text{details}}$  for the whole period of 1991–2000,  $f_{\text{details}}$  is generated for every day in the climatological year.

For now, it is assumed that the proportion of the linearly averaged climatologically observed precipitation rate at the 45 km resolution to the climatologically observed precipitation at the 1 km resolution equals the proportion of the simulations between the two resolutions. Hence, equation 5.1 describes the relationship between the MM5 model simulations at the 45 km resolution and the climatological observations at the 1 km resolution as well. Now, the downscaled rainfall rates are gained by

$$\frac{\text{MM5}_1}{\text{MM5}_{45}} = \frac{\overline{\text{OBS}}_1}{\overline{\text{OBS}}_{45}} \implies \text{MM5}_1 = \text{MM5}_{45} \cdot \frac{\overline{\text{OBS}}_1}{\overline{\text{OBS}}_{45}} = \text{MM5}_{45} \cdot \frac{1}{f_{\text{details}}} \quad (5.2)$$

where  $\text{MM5}_1$  denotes the downscaled precipitation rate on a particular 1 km cell on a particular day.  $\text{MM5}_{45}$  denotes the simulated precipitation rate in a 45 km cell, in which  $\text{MM5}_1$  is located.  $f_{\text{details}}$  denotes the correction factor for that particular 1 km cell of  $\text{MM5}_1$  defined in eq. 5.1. If  $\overline{\text{OBS}}_{45} = \overline{\text{MM5}}_{45}$  and  $\text{MM5}_{45} = \overline{\text{MM5}}_{45}$ , then  $\text{MM5}_1 = \overline{\text{OBS}}_1$ .

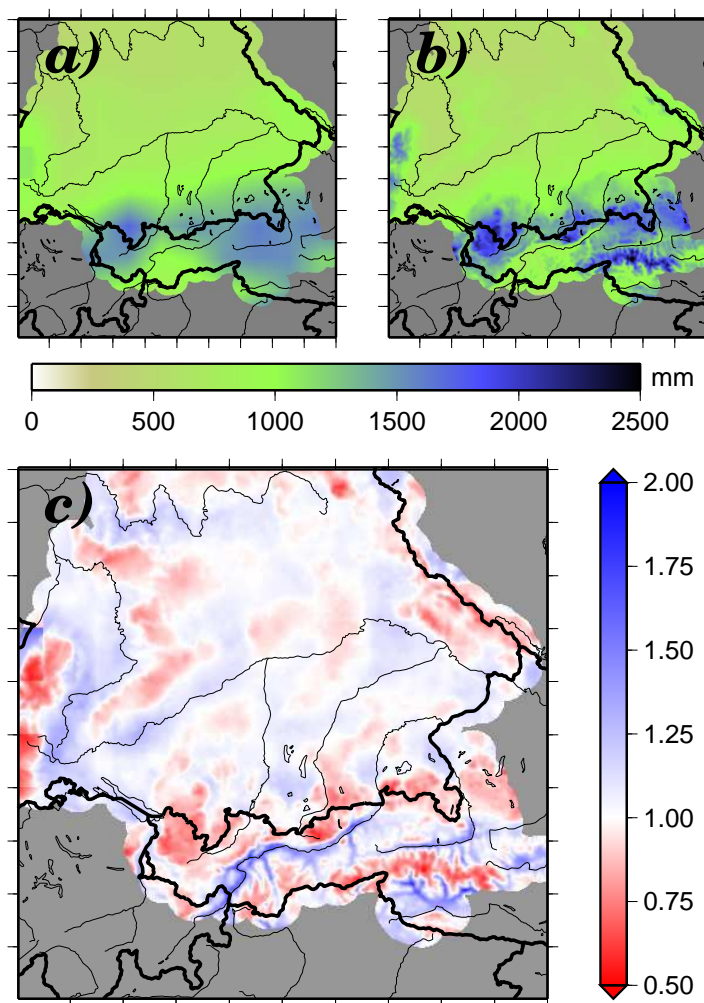


Figure 5.3: The linearly averaged climatological (1991–2000) observed rainfall distribution at the 45 km resolution (panel *a*) and the observations at a 1 km resolution (panel *b*). Panel *c* shows the spatial results of  $f_{details}$  from eq. 5.1 at a resolution of the 1 km.

The simulated  $MM5_{45}$  from eq. 5.2 is gained by a multiplicative correction. This means that in areas where the model does not simulate any rainfall, the correction does not add or subtract a fixed amount of rainfall. The great benefit of a multiplicative correction is the ability to better simulate changing climate conditions, which is one of the goals within the GLOWA-project, than a additive correction. Namely, in future climates the availability of rainfall may change, but it is unlikely that the proportion between a  $45 \times 45$  km area and a 1 km area located within is to change. An additive correction would have the disadvantage of adding a fixed amount of rainfall even under very dry climate conditions. Additionally, it may lead to negative rainfall rates if more rainfall is subtracted than simulated.

### 5.2.3 Example

The following example shows the correction for small orographic details during weather situations from the period 1991–2000. The factor  $f_{details}$  from eq. 5.1 is based on the

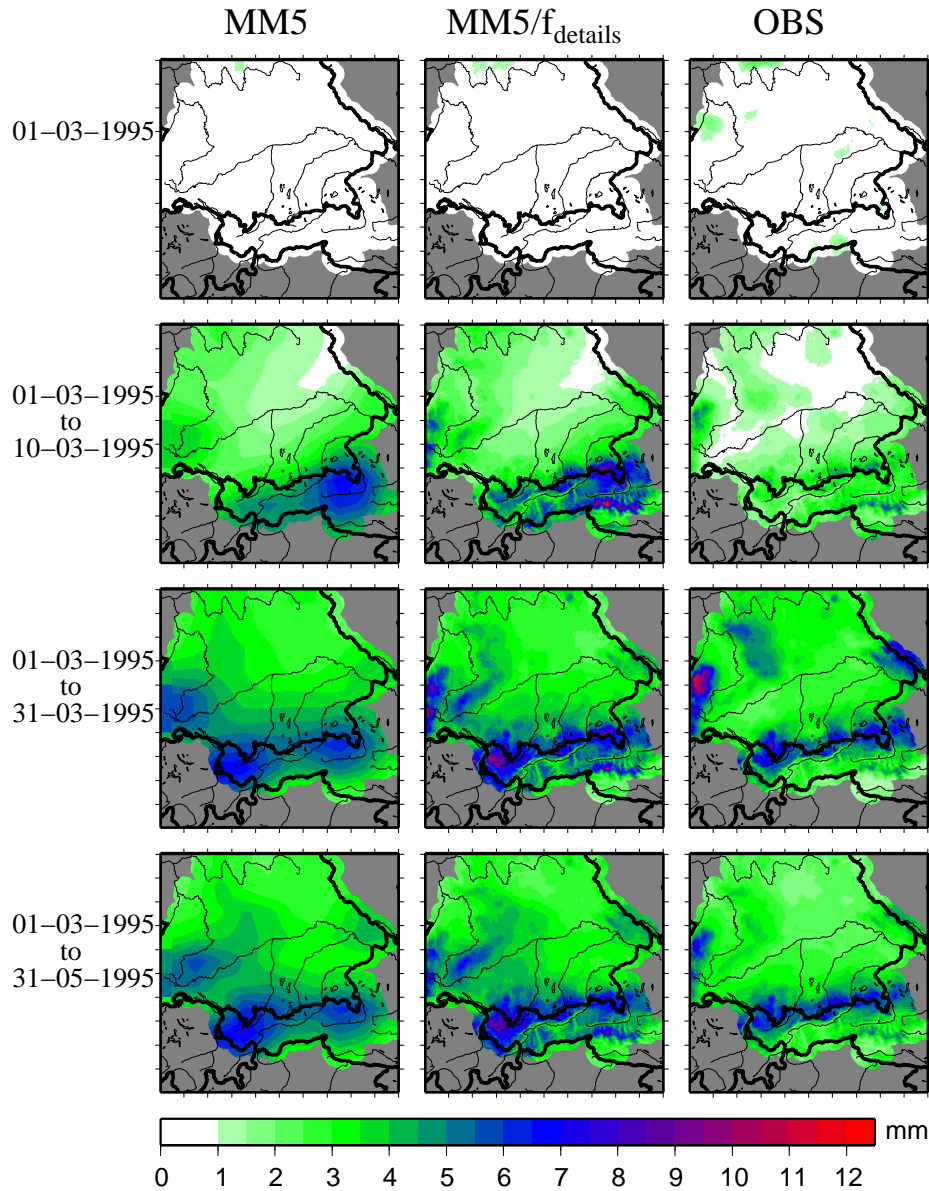


Figure 5.4: Rainfall distribution with  $f_{\text{details}}$ .

climatological relationship between the 45 km and the 1 km resolution. It is expected that better results will be gained, when using the climatological relationship for longer time periods.

The left column of images in Fig. 5.4 shows the precipitation patterns of the MM5 model simulation at a resolution of 45 km for different periods of time. The images in the middle column show the corrected simulations using eq. 5.2. The temporal-averaged precipitation distribution of these corrected model simulations is an average on a daily basis. This means that first daily fields of correction factors, similar to Fig. 5.3c, are generated for the averaging period. Secondly, these daily factors are used to correct the daily simulated precipitation fields. Last, the daily corrected rainfall rates are averaged over the length of the time period. The observations are shown in the most right column of images.

The three images in the top row show the results for 1 March 1995. This is an extreme choice, because climatological correction should not be expected to work for a single day. On this day, very little rain was simulated, especially in the mountainous area. Still some rain was observed. Of course, the multiplicative correction factor cannot produce rain in areas where no rain is simulated. Thus, the Alpine region remains dry even after the correction was applied.

The second row shows the results for the period of 1–10 March 1995. In section 5.2.2 it was assumed that the simulated rainfall rates in a 45 km grid cell correspond with the observed rainfall rates in a 45×45 km linearly averaged area. This seems not true for the mountainous area during this period. The model simulations show higher rainfall rates at the most south-eastern part of Germany than is observed. This results in too high rainfall rates after correcting. The correction factor,  $f_{\text{details}}$ , accounts for a redistribution of precipitation only and cannot correct these discrepancies. Yet, the typical small orographic details are visible in the corrected rainfall distribution.

The period of March 1995 covers 31 days. Corresponding results are shown in the third row of Fig. 5.4. There is a clear improvement with respect to the period of 10 days. Individual days do differ in the areal rainfall distribution between the model simulation and the observations (more days have been investigated, but are not shown here). An average rainfall distribution over a complete month reduces many discrepancies. Small orographic details in the spatial rainfall distribution are clearly visible. The observed rainfall rates in the Black Forest, which is located at the most western border of the research area, are underestimated more by the simulations than could be expected from the climatology (Fig. 5.3c). Equation 5.2 corrects for this climatologically expected difference only and the corrected rainfall rates draw nearer to the observations in the Black Forest than the uncorrected rainfall rates, but do not generate the observed rainfall rates completely. The difference in rainfall rates at the Bavarian Forest, which is located on the border between Germany and the Czech Republic, is larger than could be expected from both climatologies. Hence, the corrected rainfall rates still underestimate the observed rainfall rates in that area.

The last row in Fig. 5.4 shows the result for a three-months period (1 March – 31 May 1995; 92 days). Already a period of 90 days approaches climatological conditions and results in a high agreement between the simulations and the observations. Small orographic details are clearly visible. The simulated rainfall rates in the Black Forest and the Bavarian Forest show the typical spatial distributions expected from the observations. Any small difference between the corrected and the observed rainfall rates are explained by the weather conditions in the period of 1st March to 31st May 1995. According to the German Weather Service (DWD) strong north- and northwestern winds dominated March 1995 (DWD 1995a,b). These wind directions are not typical for the 10-years mean (see Chapter 6), which may have resulted in more observed rainfall rates than climatologically expected at the northern edges of the Alpine mountain ridge.

From Fig. 5.4 can be concluded that a downscaling method, which only corrects for differences in precipitation rates between mountain tops and valleys is very much dependent on the areal rainfall distribution in the model simulations. For example, in the second row of Fig. 5.4 the simulated rainfall rates vary much from the observed over an area of 45 km<sup>2</sup> in the south-eastern part of the research area. Although less visible this occurs in the third and fourth row as well. The assumption from section 5.2.2 that  $\overline{\text{OBS}}_{45} = \overline{\text{MM}}_{45}$

and  $OBS_{45} = MM5_{45}$  seems not to work in various meteorological situation and will be discussed in section 5.3.

## 5.3 General rainfall distribution

### 5.3.1 Introduction

The correction in section 5.2 was based on the assumption that the linearly averaged value of the 1 km observed precipitation fields equals, at least climatologically, the 45 km simulated value from the MM5 model. Figure 5.5 shows the zonally averaged simulated and observed climatological rainfall distribution, i.e. averaged in east-west direction. The maximum amount of simulated rainfall for the climatological period of 1991–2000 is located over the top of the Alps, whereas the maximum amount of observed rainfall is located in the forelands of the Alps. This discrepancy is an effect of the low resolution in the model simulations. Air masses are lifted too slowly, due to the gradual ascending slope of the represented mountains (see Fig. 5.2). The air saturates not until near the highest grid cells, where most rainfall will be initiated. The lower mountains at the 45 km resolution fail to have the rain shadowing effect as well, which results in drier and wetter areas north and south of the Alps depending on the prevailing wind direction.

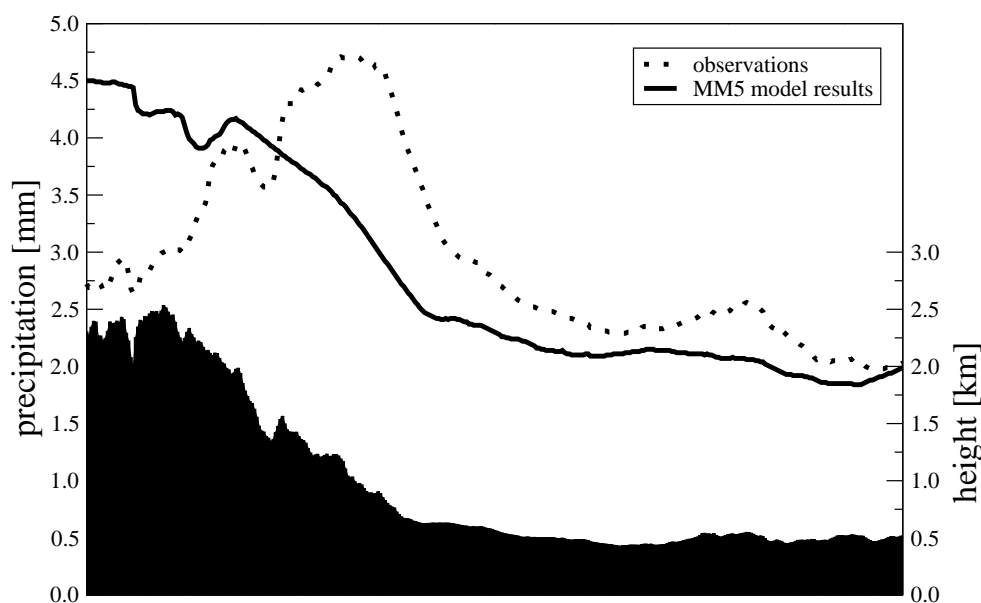


Figure 5.5: The climatological averaged rainfall distribution of the MM5 model simulations and the observations in the east-west direction in the research area. The mountains in the 1 km resolution are denoted black.



### 5.3.2 Correction

The redistribution of the simulated 45 km precipitation over the individual 1 km fields in section 5.2.2 did not correct for the total amount of precipitation in a simulated  $45 \times 45$  km grid cell. This correction is applied on a 45 km scale and is represented by a bias factor,  $f_{\text{bias}}$ , which is defined by

$$f_{\text{bias}} = \frac{\overline{\text{MM5}_{45}}}{\overline{\text{OBS}_{45}}} \quad (5.3)$$

In eq. 5.3, the correcting factor,  $f_{\text{bias}}$ , is the division of the  $45 \times 45$  km climatological simulated precipitation,  $\overline{\text{MM5}_{45}}$ , by the  $45 \times 45$  km linearly averaged climatological observed precipitation,  $\overline{\text{OBS}_{45}}$ . Equation 5.3 lowers the maximal simulated precipitation over the Alps on a 45 km scale and enhances the simulated precipitation in the Alpine forelands. Climatological results from eq. 5.3 are shown in Fig. 5.6.

The correction factor,  $f_{\text{bias}}$ , corrects the general rainfall distribution at a resolution of 45 km, only. For this reason, panel (c) in Fig. 5.6 does not show any local structures like dry valleys and wet mountain tops. Nevertheless, panel (c) shows the difference in the general rainfall distribution between the simulations and the observations. The red color in Fig. 5.6 denotes areas where less rainfall was simulated climatologically than observed, whereas the blue color denotes areas where more rainfall was simulated climatologically than observed. The areal averaged precipitation rates from Fig. 5.5 are a general phenomena along the complete northern Alpine ridge. Although partly outside the research area, this effect is visible along the southern ridge as well.

From eq. 5.3 it follows that if  $\overline{\text{OBS}_{45}} = \overline{\text{MM5}_{45}}$ ,  $f_{\text{bias}}$  turns to 1 and the assumption of section 5.2 is maintainable. These areas are white in panel (c) of Fig. 5.6. When  $\overline{\text{OBS}_{45}} = \overline{\text{MM5}_{45}}$ , the bilinear interpolation of both will equal as well, then  $\text{OBS}_1 = \text{MM5}_1$ .

In practice, equation 5.3 works as

$$\frac{\text{MM5}_{45}^*}{\overline{\text{MM5}_{45}}} = \frac{\overline{\text{OBS}_{45}}}{\overline{\text{MM5}_{45}}} \implies \text{MM5}_{45}^* = \overline{\text{MM5}_{45}} \cdot \frac{\overline{\text{OBS}_{45}}}{\overline{\text{MM5}_{45}}} = \overline{\text{MM5}_{45}} \cdot \frac{1}{f_{\text{bias}}} \quad (5.4)$$

where  $\text{MM5}_{45}^*$  denotes the corrected simulated rainfall rate at the 45 km resolution.

### 5.3.3 Example

The example is set up in the same way as in section 5.2.3, except that equation 5.3 with  $f_{\text{bias}}$  is used instead of equation 5.1 with  $f_{\text{details}}$ . The example is shown in Fig. 5.7. The first row with the results for one single day, 1 March 1995, shows again the results of a multiplicative correction: if almost no rain has fallen, almost no rain can be shifted.

The period of 10 days, 1st March until 10 March, shows the shifting of rain from the Alps center to the northern Alps. The shifting by  $f_{\text{bias}}$  acts at the 45 km scale. It does not correct for dry valleys and wet mountain tops, as done by  $f_{\text{details}}$  at the 1 km scale.

The simulated rainfall for March 1995, row 3 in Fig. 5.7, is less than for the observations. The correction with  $f_{\text{details}}$  was able to alter the rainfall distribution on a local scale, but the correction with  $f_{\text{bias}}$  shifts the maximal rainfall rates to the north.

The average observed rainfall in row 4 of Fig. 5.7 is less than in row 3. The average simulated rainfall hardly changes between the two periods, resulting in a slightly better correlation between the corrected simulated and observed rainfall.

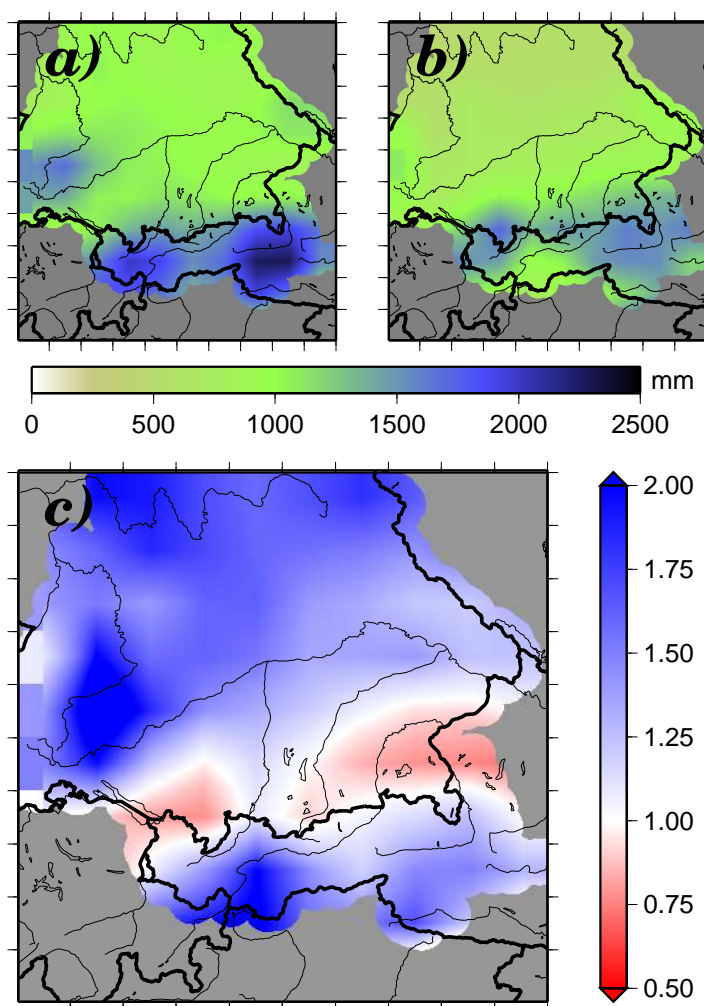
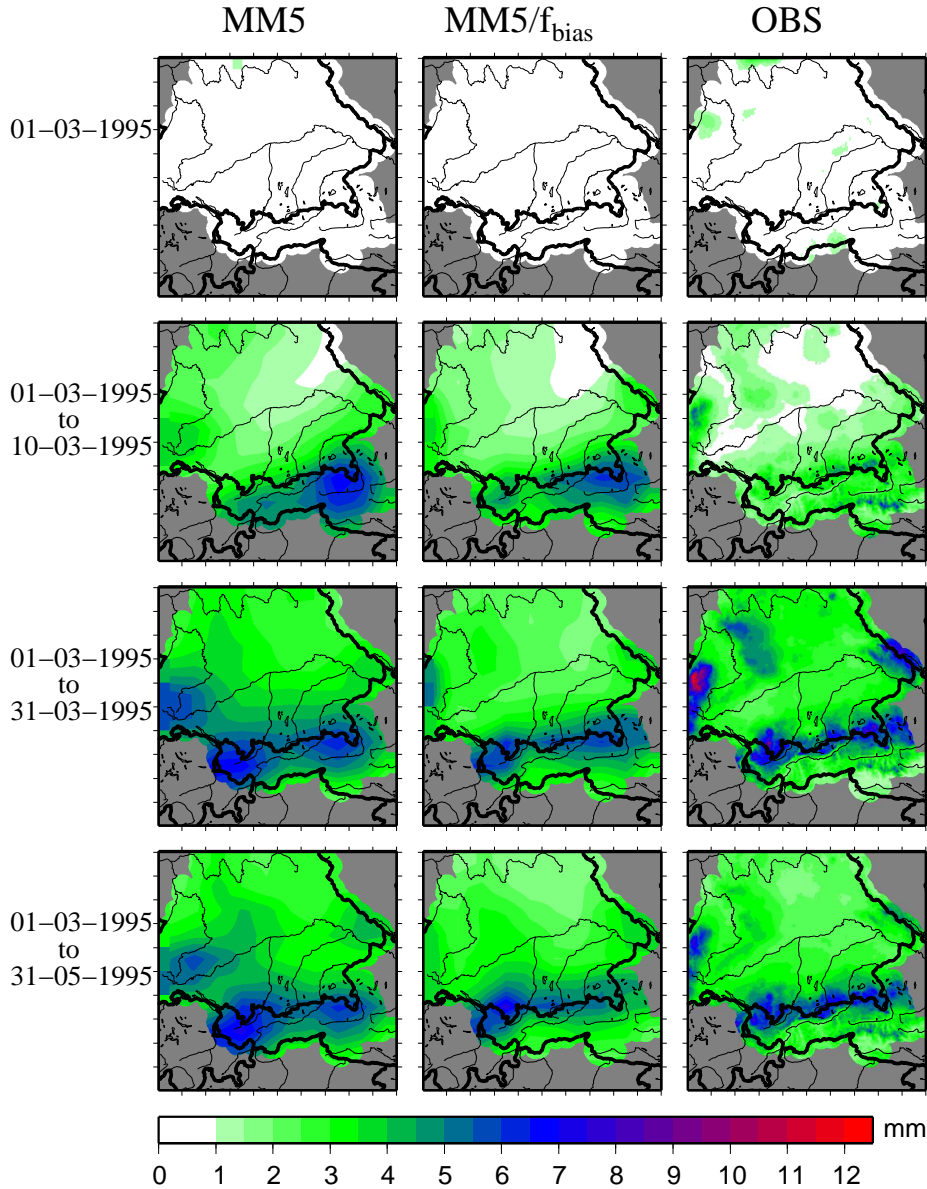


Figure 5.6: The climatological (1991–2000) rainfall distribution from the MM5-model with a 45 km resolution (panel *a*) and the linearly averaged climatological (1991–2000) observed rainfall distribution with the 45 km resolution (panel *b*). Panel *c* shows the spatial results of eq. 5.3 at the 45 km resolution. All rainfall distributions are bilinear interpolated.

## 5.4 Total

### 5.4.1 Introduction

In sections 5.2 and 5.3, simulated rainfall was corrected for small orographic details and for the general rainfall distribution in the research area, respectively. Applying a correction factor on the rainfall distribution, the results of both methods show an improvement of the simulated rainfall distribution compared to the observed rainfall distribution. Still, in both sections, just one of the two major resolution dependent deficiencies in the rainfall distribution of the MM5 model simulations is corrected. This section combines both corrections.

Figure 5.7: Rainfall distribution with  $f_{\text{bias}}$ .

### 5.4.2 Correction

The multiplication of both correction factors from sections 5.2 and 5.3 combines the correction for the small orographic details and the general rainfall distribution. The result is given by

$$f_{\text{total}} = f_{\text{details}} \cdot f_{\text{bias}} = \frac{\overline{\text{OBS}_{45}}}{\overline{\text{OBS}_1}} \cdot \frac{\overline{\text{MM5}_{45}}}{\overline{\text{OBS}_{45}}} = \frac{\overline{\text{MM5}_{45}}}{\overline{\text{OBS}_1}} \quad (5.5)$$

In eq. 5.5,  $f_{\text{total}}$  denotes the combined correction factor of  $f_{\text{details}}$  and  $f_{\text{bias}}$ . The other variables in eq. 5.5 were already described in sections 5.2 and 5.3. From eq. 5.5 it follows that the total correction of the MM5 model precipitation uses  $\overline{\text{MM5}_{45}}$  and  $\overline{\text{OBS}_1}$ , only, shown in Figure 5.8. Panel (a) and (b) were already shown in figs. 5.6 and 5.3, respectively, but not their combination, which is shown in Fig. 5.8. The spatial distribution of  $f_{\text{total}}$  in

panel (c) resolves small orographic details, like dry valleys and wet mountain tops, and additionally shifts the general rainfall distribution away from the highest peaks in the Alpine region, in order to simulate upslope precipitation.

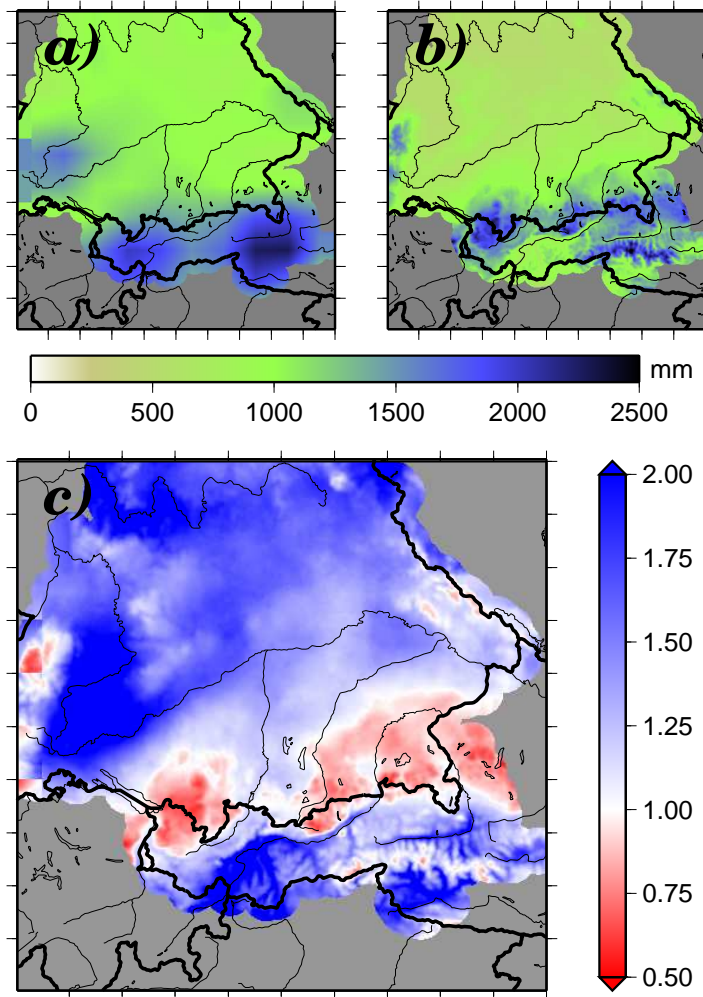


Figure 5.8: The climatological (1991–2000) rainfall distribution from the MM5-model with a 45 km resolution (panel *a*) and the observations with a 1 km resolution (panel *b*). Panel *c* shows the results of eq. 5.5.

In practice, equation 5.5 works like

$$\frac{\overline{\text{MM5}_1}}{\overline{\text{MM5}_{45}}} = \frac{\overline{\text{OBS}_1}}{\overline{\text{MM5}_{45}}} \implies \overline{\text{MM5}_1} = \overline{\text{MM5}_{45}} \cdot \frac{\overline{\text{OBS}_1}}{\overline{\text{MM5}_{45}}} = \overline{\text{MM5}_{45}} \cdot \frac{1}{f_{\text{total}}} \quad (5.6)$$

From eq. 5.6 it follows that an average observed precipitation field is not needed (i.e.  $\overline{\text{OBS}_{45}}$  is removed).  $\overline{\text{OBS}_1}$  and  $\overline{\text{MM5}_{45}}$  both correct the small orographic details and the general rainfall distribution at once.

### 5.4.3 Example

The same example like in sections 5.2.3 and 5.3.3 is discussed in this section. It compares the observed precipitation at the 1 km resolution, with both the corrected MM5 simulation at the 1 km grid and the raw MM5 simulation at the 45 km grid for different periods of time. Columns 1 and 3 in Fig. 5.9 equal columns 1 and 3 of figs. 5.4 and 5.7, respectively. Column 2 is the result of eq. 5.6. Figure 5.9 shows the orographic dependent rainfall distribution like already seen in Fig. 5.4. Although less visible, the maximal precipitation amount is moved slightly northwards, which was already shown in Fig. 5.7. Upslope precipitation in the southern part of the Alpine region should be visible as well, which is not the case, because of the extrapolation due to lacking observational precipitation in the south (see Appendix B).

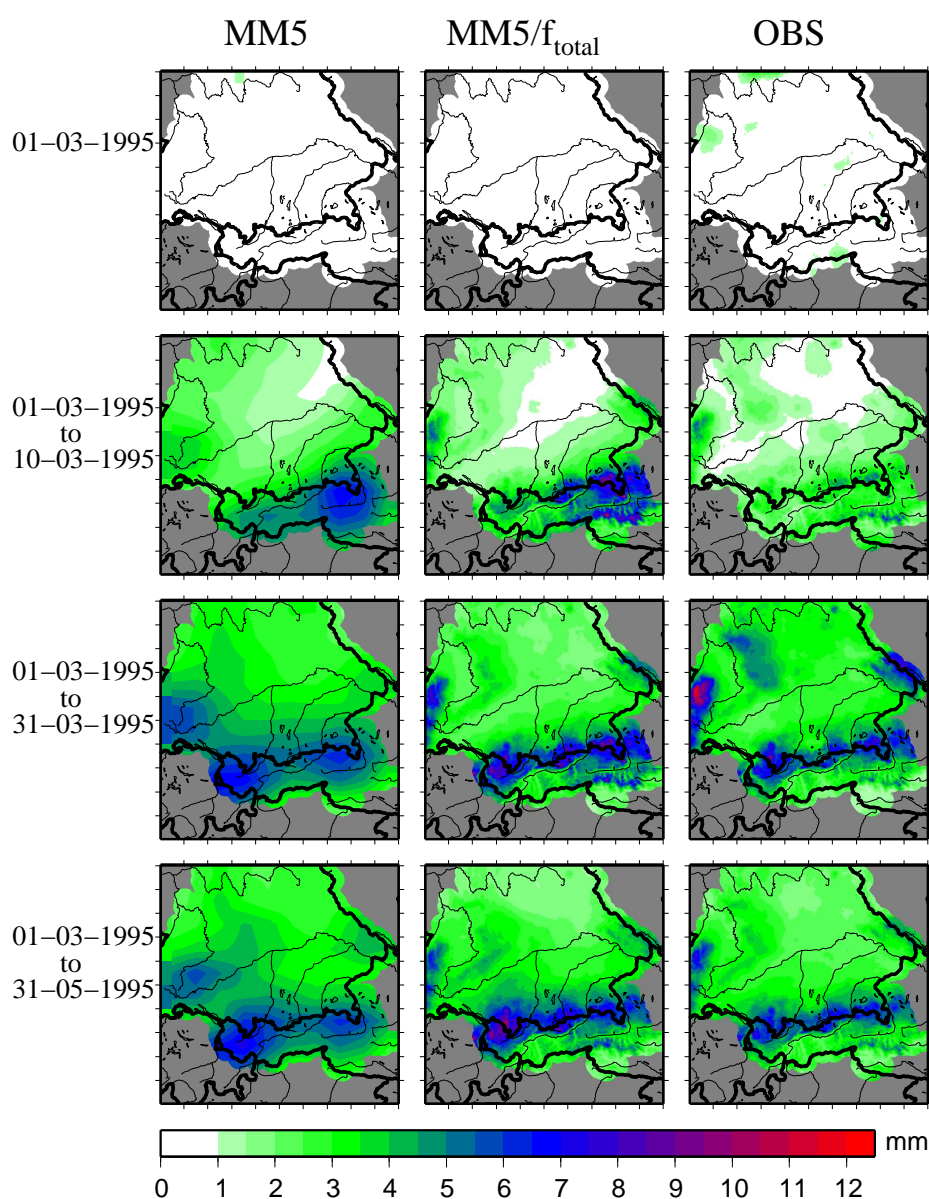


Figure 5.9: Rainfall distribution with  $f_{total}$ .

## 5.5 Results

### Correlation

In sections 5.2, 5.3, and 5.4, results were shown through figures, only. The spatial correlations between the different rainfall distribution from figs. 5.4, 5.7, and 5.9 are given by Spearman rank-order correlation in Table 5.1. The  $r$ ,  $\gamma_1$ , and  $\gamma_2$  are described in Ap-

	1st March	1st March– 10th March	1st March– 31st March	1st March– 31st May
raw	0.37	0.79	0.41	0.73
details	0.38	0.73	0.63	0.84
bias	0.37	0.85	0.57	0.86
total	0.39	0.89	0.72	0.94

Table 5.1: The Spearman rank-order correlations ( $r$ ) for the uncorrected model simulations ('raw'), the simulations corrected with  $f_{\text{details}}$  ('details'),  $f_{\text{bias}}$  ('bias'), and  $f_{\text{total}}$  ('total') with the observations for all used time periods (columns).

pendix C. A rank-order correlation instead of a 'normal' correlation (like the Pearson correlation) was used, because it does not require a normal distribution of data. However, tests showed that both correlation hardly differ for these data sets. This is most likely due to the enormous amount of data points in every data set (almost 150,000). From now on, column 1 (1st March) will be called  $T_1$  (time period 1), and column 2, 3, and 4 will be called  $T_2$ ,  $T_3$ , and  $T_4$ , respectively. The lengths of the time periods  $T_2$ ,  $T_3$ , and  $T_4$  are 10, 31, and 92 days, respectively. The first four rows in Table 5.1, which hold the correlations of different MM5 simulations with the observations, will be called  $R_{\text{raw}}$ ,  $R_{\text{details}}$ ,  $R_{\text{bias}}$ , and  $R_{\text{total}}$ , respectively. This means that  $R_{\text{raw}}(T_1)$  is the correlation between the raw MM5 precipitation and the observations for 1 March 1995.

The results from the examples in sections 5.2–5.4 showed partly an increase in spatial agreement between the rainfall distribution of the model simulations and the observations with increasing time period. This is confirmed by Table 5.1, where  $R_{\text{raw}}(T_4)$  is higher than  $R_{\text{raw}}(T_1)$ . However,  $R_{\text{raw}}(T_2)$  is higher than  $R_{\text{raw}}(T_3)$  and  $R_{\text{raw}}(T_4)$ . These values for  $R_{\text{raw}}$  are all relatively low. This is obvious if one reminds that the raw MM5 precipitation has a resolution of 45 km, without a correction for any small-scale details and general distribution in the precipitation pattern, and the observations have a resolution of 1 km. Additionally, rainfall is simulated and observed in parts of the research area, only. These parts do not necessarily coincide leading to a low correlations.  $R_{\text{raw}}(T_1)$  shows the correlation over one single day and even though the areal precipitation distribution of 1 March 1995 has a relative good agreement compared to many other individual days (not shown), a correlation of 0.37 is rather low. Even after correction of the MM5 precipitation, the correlation hardly increases ( $R_{\text{details}}(T_1)$ ,  $R_{\text{bias}}(T_1)$ , and  $R_{\text{total}}(T_1)$ ).

The correlation increases with increasing time period ( $T_2$ – $T_4$ ). The correlation between total corrected MM5 precipitation and the observations is higher than between the raw MM5 precipitation and the observations. The higher  $R_{\text{bias}}$  than  $R_{\text{details}}$  for  $T_2$  and  $T_4$  shows the importance of using a bias correction to correct for upslope precipitation. Combining the effects from small orographic details with the general rainfall distribution results in the

highest correlations ( $R_{total}(T_4)$ : 0.94). The correlation between the observations and the total corrected MM5 simulations for shorter time periods is high as well ( $R_{total}(T_2)$ : 0.89;  $R_{total}(T_3)$ : 0.72).

The result from Table 5.1 is that the simulated rainfall distribution approaches the observed distribution when applying a correction for small orographic effects and the general rainfall distribution. Longer time periods (from 10 days on) have better results than shorter time periods.

## Heidke skill score

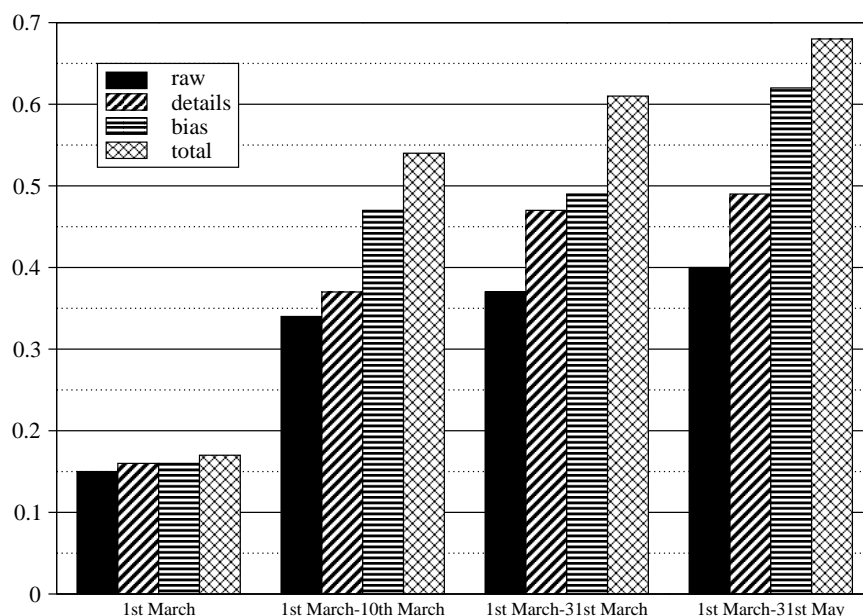


Figure 5.10: Results of the Heidke Skill Score.

The Spearman rank-correlations already gave a good impression of the improvements of the corrected model simulations compared to the uncorrected simulations. Still, an additional measure is used as an extra assurance in quantifying the differences between the MM5 simulations and observations. This measure is the Heidke skill score ( $HSS$ ), explained in Appendix C, and applied to the average rainfall distribution of  $T_1$  to  $T_4$ . Its benefit is that it is more stable concerning precipitation extremes. Due to varying distributions between  $T_1$  to  $T_4$ , the criteria that is needed for the  $HSS$  is the median of the observed precipitation and, consequently, changes with every time period (see Table 5.2).

The results of the Heidke skill score is shown in Fig. 5.10. All values in Fig. 5.10 are above zero, which means a positive correlation between the model simulations and the observations. The results for  $T_1$  are much lower than for the other periods, which confirms the results from Table 5.1. Unlike Table 5.1, the  $HSS$  in Fig. 5.10 increases

with increasing time period for all types of correction. The difference of the *HSS* between the total correction and the raw simulations increases with increasing time period as well, indicating that the presented correction methods work better for longer time periods. Like for the correlations, the total correction works best with the *HSS*.

	1st March	1st March– 10th March	1st March– 31st March	1st March– 31st May
raw	0.15	0.34	0.37	0.40
details	0.16	0.37	0.47	0.49
bias	0.16	0.47	0.49	0.62
total	0.17	0.54	0.61	0.68
median	0.01	1.56	3.34	3.15

Table 5.2: Heidke Skill Score.



# Chapter 6

## Dynamical empirical downscaling

### 6.1 Introduction

In Chapters 4 and 5 the annual precipitation cycle was given attention without taking notice of any changes in daily weather conditions, which could have played an important role in the spatial rainfall distribution. This resulted in a downscaling method, depending highly on the reference period 1991–2000. Of course, the use of climatological relationships resulted in a better performance of the downscaling method for longer time periods, but only because longer time periods approach the climatological situation more closely.

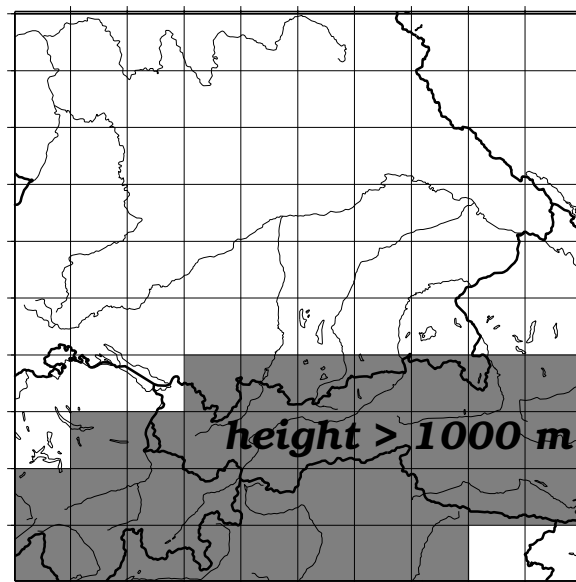


Figure 6.1: Research area with MM5 model grid cells; shaded cells denote the mountainous part containing cells exceeding 1000 meters.

The GLOWA-project requests simulated rainfall rates of current and future climates at a 1 km resolution over a relatively short period of a month at most. The downscaling method developed in Chapter 5 has difficulties to fulfill these requirements, because it uses fixed climatological precipitation patterns. Therefore, a refinement of the statistical

downscaling method from Chapter 5 is introduced in this chapter and is based on individual weather situations.

One approach to develop a weather situation dependent downscaling method is to couple local precipitation to large-scale circulation patterns (González-Rouco *et al.* 2000; Zorita and von Storch 1999; Karl *et al.* 1990). González-Rouco *et al.* (2000), for example, studied the relationship between regional precipitation and large-scale sea-level pressure.

Another approach is to use a weather classification scheme. The German Weather Service (DWD) uses a weather scheme to distinguish between 40 different weather situations. This classification is based on four physical criteria, which are the horizontal wind direction at 700 hPa (NE, SE, SW, and NW), the cyclonicity at 900 and 550 hPa (either anti-cyclonic or cyclonic), and the air humidity in the total troposphere (humid or dry). Bissolli and Dittmann (2003) describe this classification more detailed. Bárdossy and Plate (1992) used the classification of the DWD for downscaling purposes. Examples of other weather classification dependent downscaling techniques are described by Zorita and von Storch (1999).

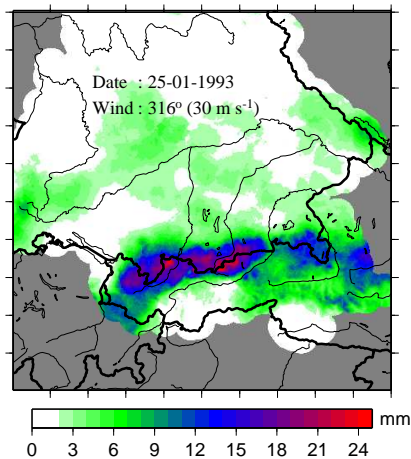


Figure 6.2: Example of observed accumulation effect of precipitation in the Alpine region for strong northerly winds ( $306^\circ$  and  $33 \text{ m s}^{-1}$ ).

The weather situation dependent downscaling method has two big advantages. First, it is able to generate weather situation dependent downscaling factors, which are capable to better react to individual weather situations in the present climate. Second, a changing weather situation in a future climate, resulting in a changing average precipitation distribution, is better captured by weather situation dependent downscaling factors.

Of all four criteria used in weather classes of the DWD the prevailing wind direction affects the precipitation pattern most in the Alpine region. The accumulation effect by upslope precipitation is strongly correlated with the prevailing wind direction. Like with the weather classification of the DWD, the prevailing wind direction in this study is defined as the average wind direction at an air pressure of 700 hPa ('free atmosphere'). This roughly corresponds to the 14th  $\sigma$ -layer (counted from above) in the used MM5 model configuration (see Appendix A). The pressure level is calculated by

$$p(x, y, \sigma) = \sigma(p_s - p_t) + p_t + p'(x, y, \sigma) \quad (6.1)$$

where  $p(x, y, \sigma)$  denotes the pressure at a certain  $\sigma$ -level at location  $(x, y)$ ,  $p_s$  and  $p_t$  denote the surface pressure and the top pressure (100 hPa), respectively.  $p'(x, y, \sigma)$  denotes the so-called pressure perturbation, which is a pressure correction depending on the current weather situation and location and is calculated by the model each time step. Tests for a winter and a summer month showed that the pressure perturbation over the Alps ranges between 675 and 725 hPa as a result of passing front. At the reference height, wind disturbances from the surface, as they are found in the planetary boundary layer, are negligibly small. Nevertheless, it is low enough to capture the prevailing wind direction important to meteorological processes like precipitation, because the highest mountains of the Alpine ridge in the model reach a height of approximately 2,200 m.

Tests showed that the research area of  $450 \times 450$  km is large enough to contain horizontal wind shears exceeding  $90^\circ$ . Therefore, only the mountainous part is used to minimize this effect. The mountainous part is chosen instead of the flatter northern part, because it influences local rainfall distribution more strongly. A lower limit of 1000 meters is chosen to define the mountainous part (see Fig. 6.1).

The Alps are roughly West-East orientated, causing northerly and southerly winds to produce upslope precipitation along the northern and southern Alps, respectively. Figure 6.2 shows an example of the upslope precipitation at northerly winds. On the contrary, westerly and easterly winds produce upslope precipitation along the western and eastern Alps, respectively. These effects are less visible in this study, due to the location of the research area. Upslope precipitation plays a significant local role in the Alpine region and should be considered in the downscaling method.

## 6.2 Wind distribution

The prevailing wind direction is defined as the average wind direction at around 700 hPa in the mountainous region of the research area. Tests showed that the 3-hourly simulated wind directions remained almost constant at this reference height during the day. Except for large weather events like the passing of a major front, they remained within an angle of  $30^\circ$ . Thus, daily averaged simulated wind directions are a good representation of the prevailing wind direction over the Alps in most cases.

The average daily wind directions in the research area are not equally distributed over  $360^\circ$ . In more than 44% of the days, winds come from the west ( $210^\circ$ – $330^\circ$ ) and almost 7.5% come from the east ( $30^\circ$ – $150^\circ$ ) between 1991–2000 (Figure 6.3). Figure 6.3 shows a clear dependency of the wind speed on the wind direction as well. The highest wind speeds are found at the most frequent wind directions (between  $240^\circ$ – $300^\circ$ ).

## 6.3 Rainfall distribution

### 6.3.1 Rainfall classes

Each wind direction results in a specific precipitation pattern. Figure 6.4 shows the observed precipitation patterns for wind direction classes of  $30^\circ$  for the period of 1991–2000. The wind directions and wind speed were simulated by the MM5 model. For the time being, the unequal distribution of wind directions as shown in Fig. 6.3 is neglected.

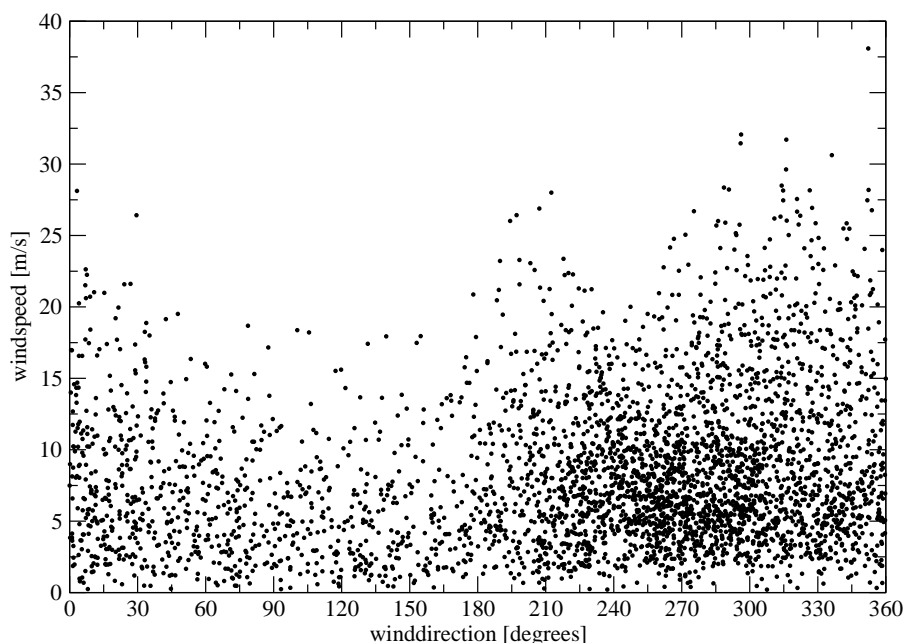


Figure 6.3: The distribution of wind direction against wind speed in the research area between 1991–2000.

The circle inside Fig. 6.4 shows the absolute and relative number of days in each wind direction class during the period of 1991–2000. For example, the wind direction class between  $270^\circ$  and  $300^\circ$  holds an average over 456 days (about 12.5% of 3650 days). Wind directions between  $210^\circ$ – $270^\circ$  carry more moisture than easterly winds, because their origin is either the Mediterranean or the Atlantic. Easterly winds, which come from the usually dryer inland, carry less moisture and are less frequent than westerly winds. Thus, westerly winds play a more dominant role in the rainfall distribution over the Alps. Still, also easterly winds are considered, because the distribution of winds as seen in Fig. 6.3 could change in a future climate. Wind directions from the north (between  $300^\circ$  and  $30^\circ$ ) clearly show an increase in total precipitation at the northern edge of the Alpine ridge, while wind directions from the south (between  $120^\circ$  and  $240^\circ$ ) show an increase at the southern edge.

Tests showed that the precipitation distributions on days with wind speeds lower than  $5 \text{ m s}^{-1}$  did not produce the typical pattern one would expect at a certain wind direction. These days are neglected and put in a separate ‘no wind’ class, which is shown in the lower left corner of Fig. 6.4. It is an average over 1099 days (about 30.1% of 3650 days). The mean precipitation pattern over all wind directions is shown in the lower right corner of Fig. 6.4 and corresponds with Fig. 5.3*b* and 5.8*b*. In general, Figure 6.4 shows that the precipitation pattern, as well as the amount of precipitation, are very much dependent on the prevailing wind direction.

The MM5 model is capable to simulate wind direction dependent precipitation patterns as well (Fig. 6.5). The simulation results show rather similar patterns as for the obser-

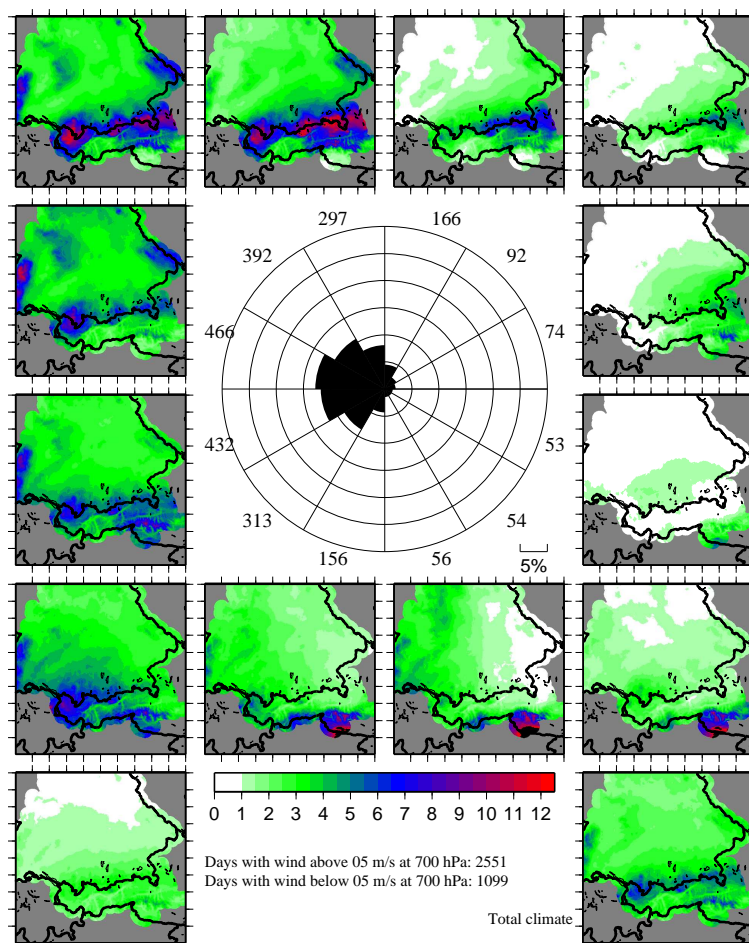


Figure 6.4: The wind direction dependent observed precipitation patterns for wind classes of  $30^\circ$  at 700 hPa in the mountainous area ( $> 1000$  m). The precipitation pattern for weak winds ( $< 5 \text{ m s}^{-1}$ ) is shown in the lower left corner. The total precipitation pattern is shown in the lower right corner.

vations. As was to be expected from Chapter 5, small orographic details are not visible at the lower resolution of the simulations as they occur at the higher resolution of the observations. Also, the observed precipitation maximum shifts to the north at northerly winds and to the south at southerly winds. This is by far less visible in the simulated precipitation maximum and confirms the need to correct the simulations for a bias at these wind directions. In the situation with winds coming from the West and the East, the bias between the simulation and the observations is less visible. These different biases confirm the need for a wind direction dependent downscaling method.

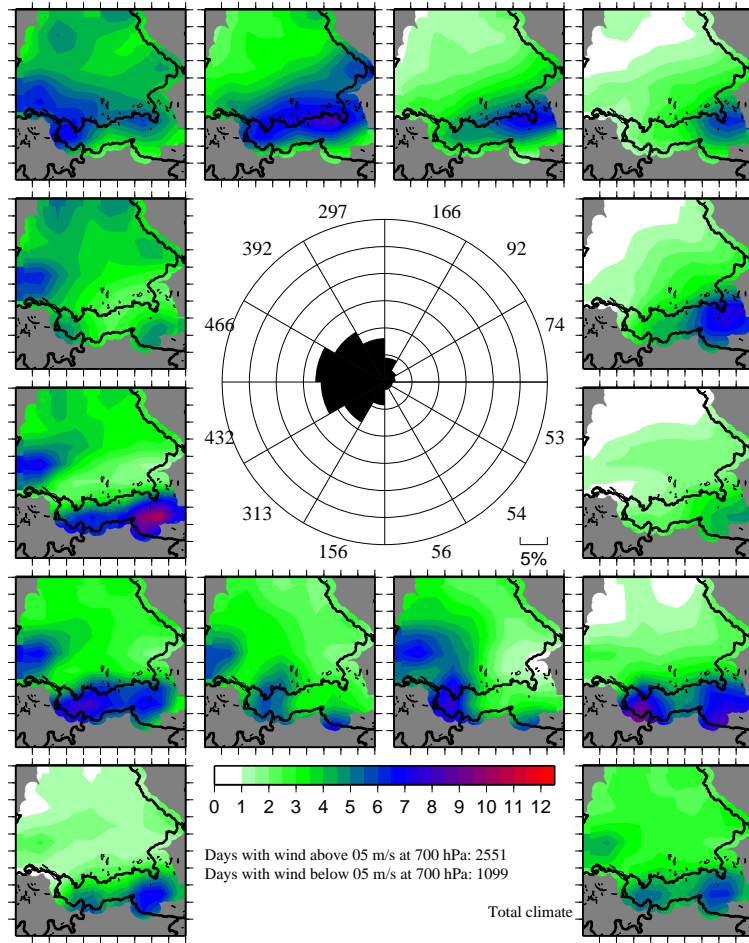


Figure 6.5: The wind direction dependent simulated precipitation patterns for wind classes of  $30^\circ$  at 700 hPa in the mountainous area ( $> 1000$  m). The precipitation pattern for weak winds ( $< 5$  m  $s^{-1}$ ) is shown in the lower left corner. The total precipitation pattern is shown in the lower right corner.

### 6.3.2 Rainfall groups

Each rainfall class in section 6.3.1 spanned  $30^\circ$  and contained a different number of days ranging from 53 days (1.5%) to 456 days (12.5%). Although each class showed a different precipitation pattern, groups of classes can be formed with more or less the same pattern. These clusters of classes enables one to build mean precipitation patterns over a larger number of days, which is statistically more significant. Especially for easterly wind directions, which carry less moisture and produce less rain, the small number of days makes it complicated to define a specific precipitation pattern. Figure 6.6 shows again wind direction against wind speed as in Fig. 6.3, except that it additionally shows the classification of the groups. Table 6.1 shows the exact amount of days per group, subdivided into seasons. In Table 6.1, the seasonal percentage of each group is shown between brackets compared

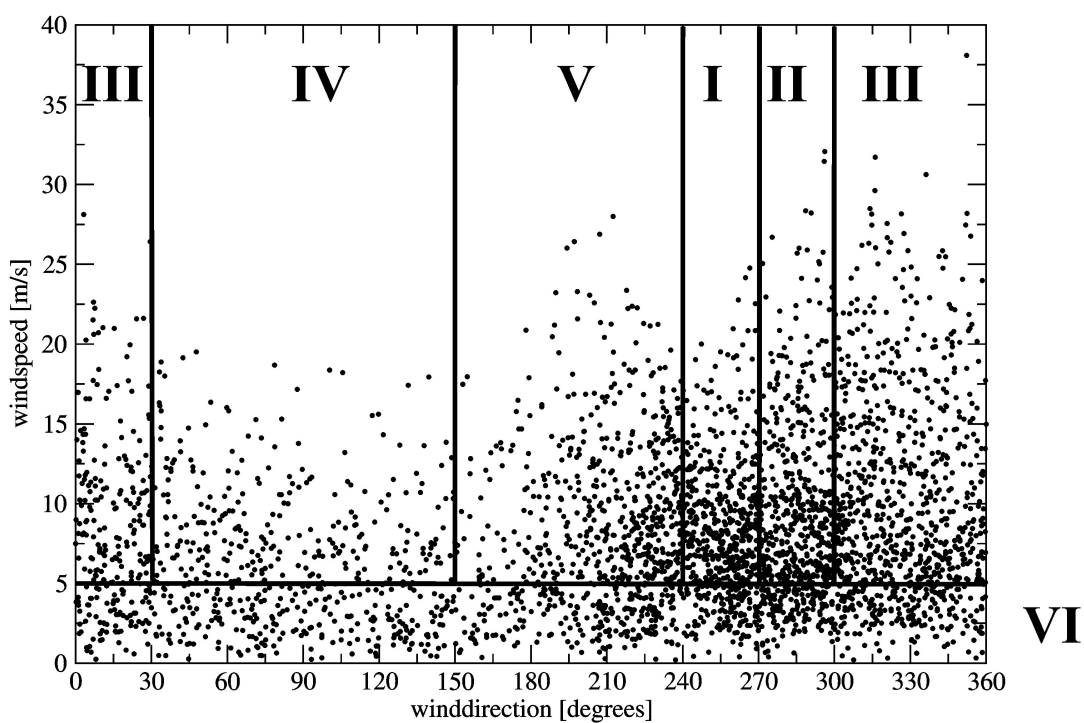


Figure 6.6: The distribution of wind direction against wind speed with classification scheme in the research area between 1991–2000.

	DJF	MAM	JJA	SON	YEAR
240°–270°(I)	95(11%)	71(8%)	138(15%)	128(14%)	432(12%)
270°–300°(II)	152(17%)	92(10%)	122(13%)	100(11%)	466(13%)
300°–30°(III)	272(30%)	257(28%)	142(15%)	184(20%)	855(23%)
30°–150°(IV)	110(12%)	84(9%)	25(3%)	54(6%)	273(7%)
150°–240°(V)	100(11%)	119(13%)	113(12%)	193(21%)	525(14%)
‘weak wind’(VI)	171(19%)	297(32%)	380(41%)	251(28%)	1099(30%)
total	900	920	920	910	3650

Table 6.1: Number of days in each climatological seasonal wind direction group.

to all days in that season. The six Roman group letters (I–VI) correspond to the letters in Fig. 6.6. The easterly winds in group IV cover the largest range of wind directions, but hold the smallest number of days for all seasons. In the summer months, this is 25 days corresponding to less than 3% of the total amount of days. This will not influence the precipitation distribution much over longer time periods, at least in the current climate, because the amount of summer precipitation in group IV is almost negligible. The small number of days at easterly wind directions forces to use seasonal precipitation distributions, instead of shorter periods. The ‘weak wind’-group (VI) is largest during the summer months.

Figures 6.7 and 6.8 show the observed and simulated precipitation distributions for groups I to VI.

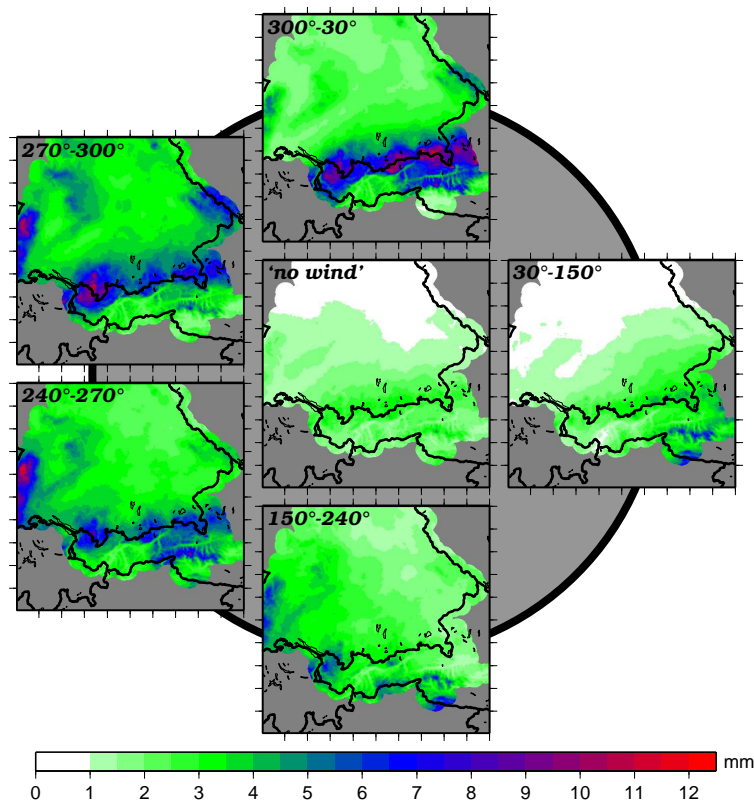


Figure 6.7: The six observed precipitation distributions (I–VI) described in Table 6.1 for the Spring months (March, April, and May) during 1991–2000.

## 6.4 Results

For every season and every wind direction group described above, a downscaling factor is generated given by eq. 5.5. Like in Chapter 5, the period from 1 March to 31 May 1995 is chosen to verify the results of the wind direction dependent (*WDD*) downscaling method.



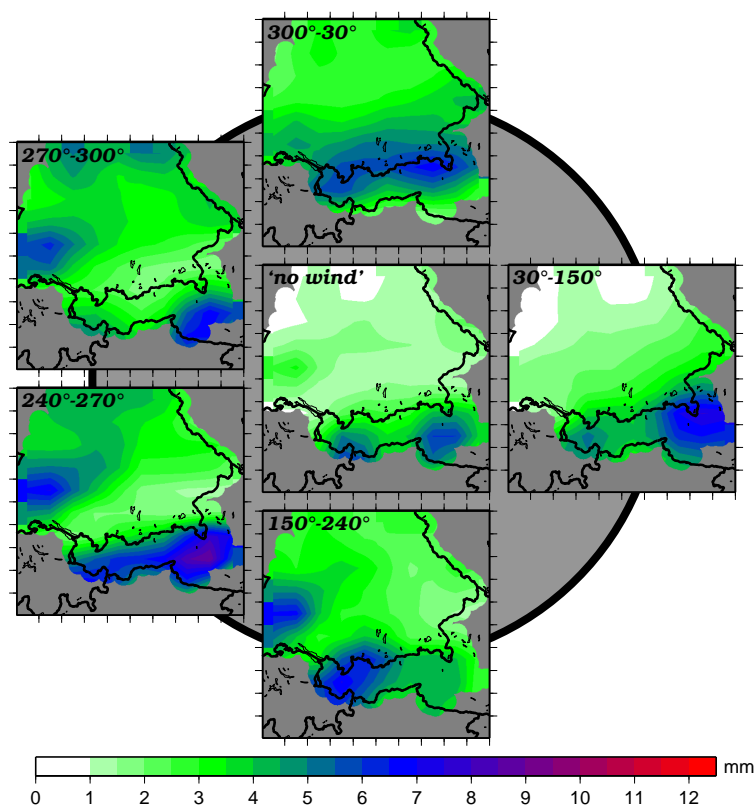


Figure 6.8: The six simulated precipitation distributions (I–VI) described in Table 6.1 for the Spring months (March, April, and May) during 1991–2000.

The results are shown in Fig. 6.9. In contrast to figs. 5.4, 5.7, and 5.9 the simulation results are not shown. Instead, the downscaled simulations using the total correction from Chapter 5 are shown in the left column. The middle column shows the results of the *WDD* correction. The right column in Fig. 6.9 shows the observed precipitation pattern.

The differences between the *WDD* downscaling method from this chapter and the general method from Chapter 5 are small. The mean precipitation pattern of the first ten days of March 1995 (Fig. 6.9, second row) shows an improvement in the *WDD* precipitation distribution compared to the general method. As already discussed in Chapter 5, March 1995 was dominated by strong north/north-west winds. In this case, the *WDD* downscaling method corrects for the bias due to upslope precipitation rates better than the general method. This results in a stronger reduction of maximum precipitation at the highest Alpine ridge.

The bias correction of the precipitation pattern for the full month March 1995 (Fig. 6.9, third row) shows similar results. Due to the strong northern winds, the maximum precipitation rates are shifted more to the northern Alps by the *WDD* method than by the general method. In this particular case, it seems that even too much precipitation is shifted north compared to the observations. This may be the result of strong northern winds combined with local showers as was observed in March 1995 (DWD 1995a,b). Small scale precipi-

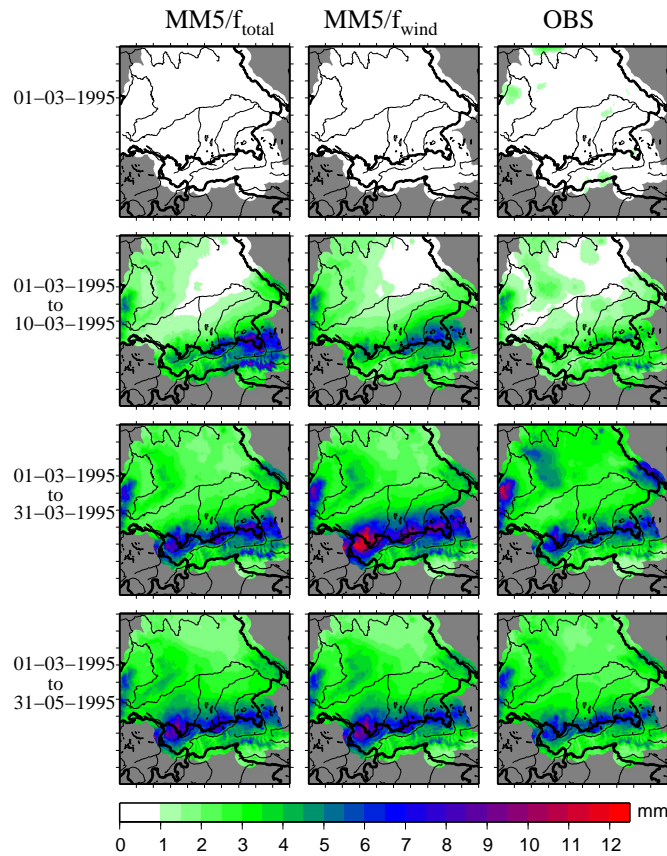


Figure 6.9: Rainfall distribution with wind direction dependent factors.

tation events like local showers cannot be explicitly resolved by the model due to its low resolution of 45 km. Consequently, every precipitation event will have a horizontal influence area of at least  $45 \times 45$  km. Due to the bilinear interpolation, this area is extended even more. As a result, the model produced too much precipitation along the Alpine ridge. The downscaling factors of the *WDD* method shifted more precipitation to the north, than the general method. That way, the worse performance of the *WDD* method is caused by the deficit of the model and not necessarily by the ability of *WDD* method to downscale precipitation rates.

	1st March	1st March– 10th March	1st March– 31st March	1st March– 31st May
total	0.39	0.89	0.72	0.94
wind	0.38	0.87	0.67	0.94

Table 6.2: Correlations between simulation corrected with the wind dependent downscaling method (*WDD*), with the general method, and the observations.

# Chapter 7

## Conclusions and outlook

### 7.1 Conclusions

Let us summarize the findings of this work. To do this, we proceed more or less along the individual chapters. The sensitivity study described in Chapter 3 aimed to test various parameterizations schemes with respect to precipitation. After finding the optimized configuration for the Upper Danube catchment area, a time period of 10 years (1991–2000) was chosen to generate a quasi-climatological year of daily precipitation rates. Various averaging methods were tested in Chapter 4. The climatological precipitation rates were used as a basis to the downscaling method described in Chapter 5. Finally, in Chapter 6 a refinement of the downscaling method from Chapter 5 is introduced, in which is tried to couple certain precipitation pattern to prevailing wind directions over the Alpine region.

With the MM5 model, one has the ability to choose different combinations of parameterizations in order to find the optimized configuration for a certain research area. The sensitivity study here was set up for four major parameterizations affecting the simulation of precipitation. These were the parameterizations that deal with cumulus convection (dealing with the simulation of the amount of convective precipitation in a grid cell), boundary layer processes (dealing with processes like local and non-local mixing of the boundary layer, vertical diffusion, and entrainment), cloud microphysics (dealing with rain that falls in connection with large-scale systems), and the radiation balance (deals with radiation that interacts with clouds and reaches the surface). Two months from the reference period 1991–2000 were examined, having different rainfall regimes. In February 1998, rain fell in connection to large-scale synoptic systems, whereas May 2000 was dominated by localized convective systems.

It was found that the runs using the cumulus parameterization 'Grell' failed to simulate a sufficient amount of parameterized rainfall during May 2000. When using the cumulus parameterization 'Betts-Miller' the model simulated a more plausible ratio between parameterized and resolved rainfall. All parameterizations had difficulties with the accurate simulation of rainfall, because the represented mountains at a resolution of 45 km are much lower than in reality. The mountains failed to block moist air from the Mediterranean at southerly winds, resulting in too much rain in the northern Alps. The boundary layer parameterization 'Eta' was preferred over 'MRF', because it is able to simulate model runs with a higher vertical resolution. The cloud microphysics parameterization 'Reisner1' was preferred over 'Simple Ice', because it is more sophisticated and is therefore able to simulate

future climate scenarios better. This resulted in the configuration of 'Betts-Miller' for the cumulus parameterization, 'Eta' for the boundary layer parameterization, 'Reisner1' for the cloud microphysics scheme, and 'Cloud-Radiation' for the radiation balance scheme.

With describing rainfall variability, as done in Chapter 4, one should remember that rainfall rates are bounded by zero and not normally distributed, but positively skewed and leptokurtic. Therefore, a rank-correlation, the Spearman correlation, is used. However, it should be noted that correlations hardly changed if using a 'normal' correlation. This is probably due to the enormous amount of small daily rainfall rates. The spatial averaged simulated rainfall rate over 1991–2000 is higher (3.31 mm) than the observed rainfall rate (3.07 mm). However, the observed rainfall rates showed more fluctuations in the spatial averaged rainfall rates over 1991–2000 than the simulations did.

Four different methods are examined in order to find a climatological year of simulated and observed rainfall data. First, a linear averaging of each day was tried. Unfortunately, generating a climatological year of spatial averaged rainfall data by linearly averaging gives many rainfall fluctuations, because of the short averaging period of ten years. Second, instead of using 10 days to average, like was done by the linear averaging method, a running average was applied on the data set. To avoid inconsistencies at the start and the end of the data set a cyclic data set was assumed, which implies no trend over the reference period 1991–2000. At a running average period of 31 days, the observed and simulated data set coincide most. However, with this relatively short period still many fluctuations exist between individual days. A longer running average period would need such a long time period that even effects like a annual cycle is weakened. The third method used Fourier analysis. Again, no trend was assumed and frequencies between 1 year and 45 days gave the most plausible representation of rainfall rates in a 10 years averaged time series. Unfortunately, applying this range of frequencies to individual grid cells, rainfall rates could become negative during dry spells. The last method used spline interpolation with monthly data as fixed points. Naturally, this interpolations smoothens the rainfall rates radically. The benefit is that it is very robust, because it is based on monthly values. Negative rainfall rates could not be found as well. Although the spline interpolation is quite inflexible due to the fixed monthly data points, it was used for the downscaling method developed in the following chapters.

The downscaling method itself is motivated by the fact that at 45 km resolution, the landsurface is represented much more crudely than at 1 km resolution. This difference in resolution has an enormous influence on the distribution of various meteorological parameters, like precipitation. Small-scale orographic effects, like dry valleys and wet mountain tops, which are visible in a 1 km resolution, are not represented at 45 km resolution. A bilinearly spatially interpolated rainfall field is preferred to compare 45 km values with 1 km values, since it does not have large discrepancies at the borders of each 45 km grid cell. If it is assumed that an averaged observed rainfall rate over  $45 \times 45 \text{ km}^2$  equals a 45 km simulated rainfall rate, a climatological relationship between the observed rainfall rate in a averaged 45 km resolution and a 1 km resolution can be used to correct for small orographic details. Due to the lower and less steep mountains in the 45 km resolution, the model simulates the maximum rainfall rates closer to the highest mountain tops and fails to simulate upslope precipitation in most cases. A climatological relationship between an averaged observed rainfall rate over  $45 \times 45 \text{ km}^2$  and a 45 km simulated rainfall rate is able to correct for this bias in the general rainfall distribution between the observations and the

simulations. A combination of the correction for small orographic details and the general rainfall distribution is able to correct for both effects at once. Depending on the length of the time period, the correlation, as well as the Heidke skill score, showed a significant improvement of the model simulations after applying the downscaling correction. Since the downscaling is based on climatological relationships, a correction over a longer integration period gives better results.

The refinement introduced in Chapter 6 tried to find a correlation between the precipitation distribution and the prevailing wind direction over the Alps. In more than 44% of the days during 1991–2000, winds came from the west ( $210^{\circ}$ – $330^{\circ}$ ) and in almost 7.5% of the days winds came from the east ( $30^{\circ}$ – $150^{\circ}$ ). At the most frequent wind directions, which were western winds, the wind speeds were highest. Rainfall distributions in the research area, especially the mountainous part, are dependent on the prevailing wind direction. The prevailing wind direction is defined as the average wind direction at around 700 hPa over the area where the orography exceeds 1000 m to minimize wind shear. A lower wind speed limit of  $5 \text{ m s}^{-1}$  filtered out days with ‘weak wind’. The wind direction dependent downscaling only slightly improved the results compared to the general downscaling method in the previous downscaling. Still, it is expected that the wind direction dependent downscaling method will be more robust as weather conditions change in future climates.

## 7.2 Outlook

The studies performed in this work were a first approach to find a fast downscaling method. It needed to be fast, because the GLOWA-project requires current and future climate scenario simulations.

The sensitivity of MM5 precipitation to various physical parameterizations was tested in Chapter 3. Although the findings from this chapter are believed to be optimal for the current climate, it is possible that the configuration found needs to change in future simulations. The sensitivity of the downscaling method to these changes should be tested as well.

Until now, the downscaling method has been verified with observations within the reference period, only. This will always give slightly better results and the method necessarily needs testing outside the reference period. Due to lacking observations during time periods other than the reference period, this has not been done so far, but is planned in near future.

This study shows just one example period (1 March – 31 May 2005). Although a few more periods have been tested with similar results, more testing needs to be performed. Especially, the limits of the downscaling method should be examined during extreme weather situations in order to draw definite conclusions about its reliability.

In Chapter 6, one of the DWD criteria to classify weather situations was used (wind direction at around 700 hPa). In future studies, other criteria or combinations of criteria should also be tested to find out what combinations give the best results.

Except for precipitation, other meteorological parameters need to be downscaled within the GLOWA-project. This has already been done for temperature at 2 m, partly using the achievements of Chapter 5. More testing on more meteorological parameters also needs to be performed to find the applicability of the method.



# Appendix A

## MM5 model description

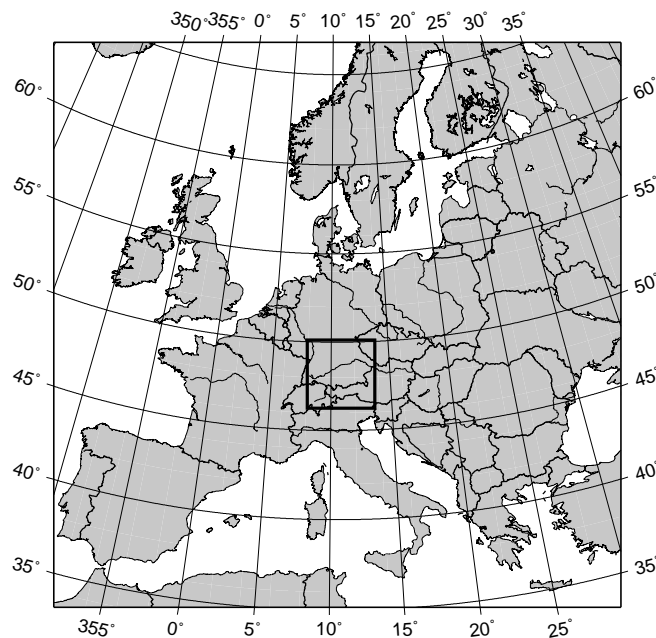


Figure A.1: Total model area with research area (box).

This study uses the Fifth-Generation NCAR Penn State Mesoscale Model version 3.5, MM5 (Grell *et al.* 1995; Dudhia 1993). The model domain covers Europe (see Fig. A.1) and contains 81 by 81 horizontal grid points at a resolution of 45 km, and has 28 vertical sigma-layers (Miller and White 1984). These layers are: 1., .998, .994, .99, .985, .98, .97, .96, .945, .93, .91, .89, .87, .85, .8, .75, .7, .65, .6, .55, .5, .45, .4, .35, .3, .2, .1, .0 (Fig. A.2). The research area which is examined in this study is located in the middle of the model domain (see box in Fig. A.1), and contains 10 by 10 grid points. This inner region will be referred to as the research area. The size of the outer model domain is chosen to minimize the effects of the boundaries on the study area. The coarse resolution used here is chosen in order to minimize computational costs, because the MM5 model will eventually be integrated in a group of models within the GLOWA-project.

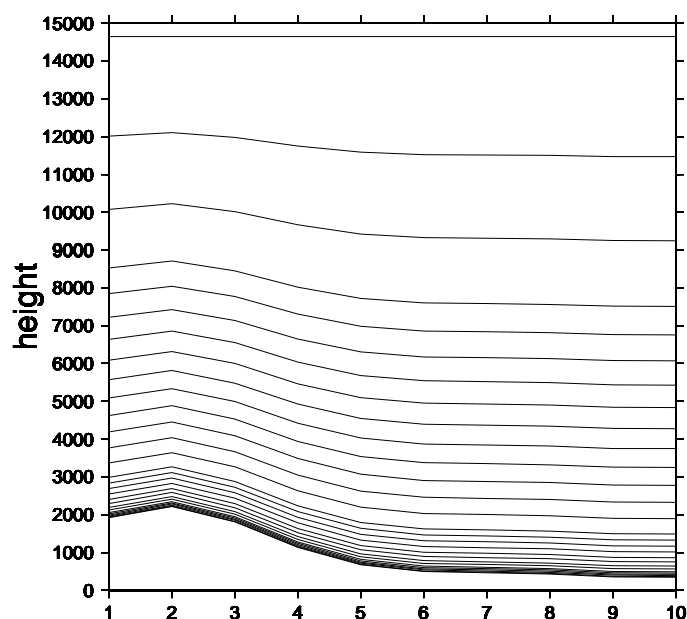


Figure A.2: Sigma-layers as defined in the MM5 model configuration.

ECMWF analysis data (Chapter 3) and re-analysis data (other chapters) were used as model input. The model was initialized at the lower boundaries (soil temperature, soil humidity, etc.) and was updated at the lateral boundaries every 6 hours. This means that soil properties were simulated by the model only. Model output was given every hour, except for the results in Chapter 3, for which it was every 3 hours. To ensure that the model could not drift to far from the observations during the 10 years reference period, so-called nudging was applied. This nudging is called FDDA (Four Dimensional Data Assimilation) and bends the model results towards the observations (in this case the re-analysis data) at every time step. The area in which nudging was applied is shown in Fig. A.3 as the area outside the box. The size of the box was chosen such that the model has the ability to develop its dynamics within the research area, without being corrected by observations.

The soil moisture and temperature in four layers (10, 30, 60, and 100 cm), as well as variables like canopy moisture and water-equivalent snow depth, are predicted by a landsurface scheme called OSU/Eta Land-Surface Model (Chen and Dudhia 2001).



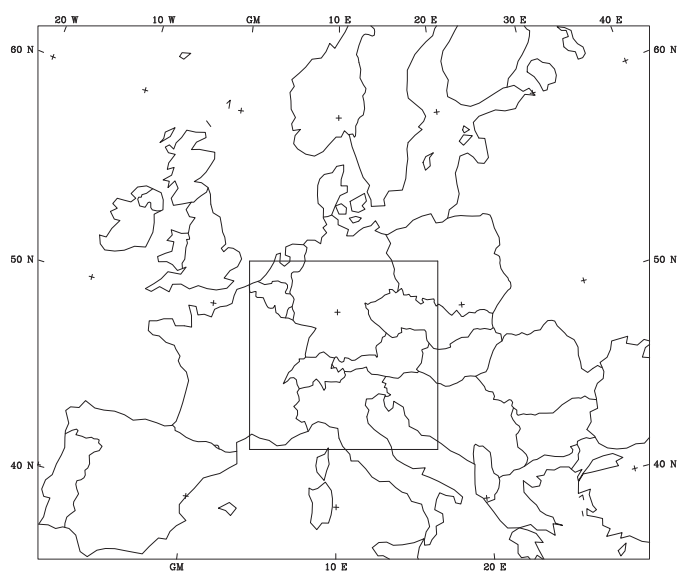


Figure A.3: The total MM5 model area. Outside the box model simulations are nudged using FDDA.



# Appendix B

## Observations

Observational data for the research area were available from the network of the German and Austrian Weather Service (DWD and ZAMG respectively) spanning January 1991 to December 2000. The measurement stations are not equally distributed in the research area, but accumulate more in the northern part of the area. Due to the sparse observations available in the Alpine region, three strategies were followed throughout this study to obtain the best environment for comparison with the model results.

First, in the Chapter 3, areal averaged model simulations are compared to areal averaged observations. To ensure using as much observations as available in the relatively short study period of one month, it was decided to only use those MM5 grid cells for comparison with observations in which at least 6 measurement stations were available. Grid cells where the stations were not equally spread (approximately less than a quarter of the MM5 cell covered) and were located in the relatively flat area in the north were excluded, as well. This gives the reduced research area shown in Fig. B.1.

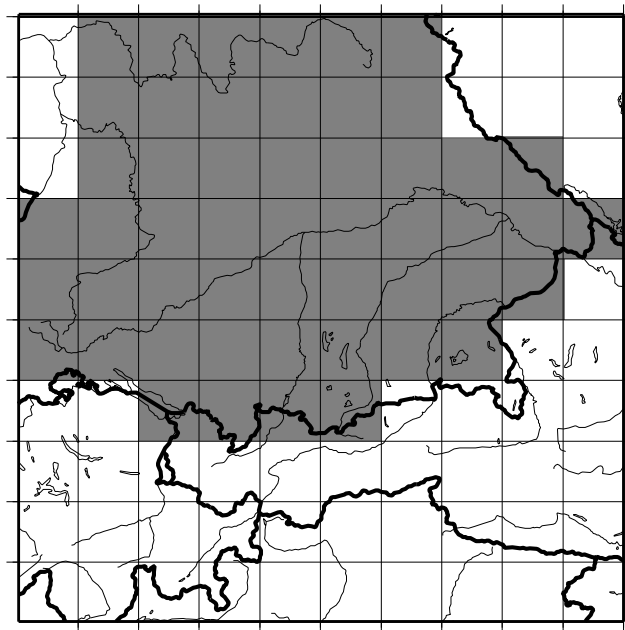


Figure B.1: The 10x10 grid cells MM5 model domain with national borders; shaded cells contain six or more stations and are used in this study for comparison (see text for explanation).

During the time of this study, additional observations from more measurement stations in the Austrian measurement network were obtained. Still, the density of the Austrian station network stayed lower than that in the German area. Therefore, no re-calculation of the results from Chapter 3 was performed.

Second, in Chapters 4 and 5, all available observations are interpolated in order to cover the complete research area. Areas without any observations were extrapolated using the climatology of the Parameter-elevation Regression on Independent Slope Model (PRISM) (Frei and Schär 1998; Daly *et al.* 1994). Schwarb (2001) and Schwarb *et al.* (2001) provided a precipitation climatology for the entire European Alps (43°N–49°N, 2°E–18°E) at 2.5 minutes (about 2 km) resolution based on 9546 station records for the time period 1971–1990.

Third, in Chapter 6, daily wind direction dependent precipitation patterns were needed. Climatological extrapolated data from PRISM would influence these patterns. Therefore, an area was cut out covering the area in which data were available for every day throughout the 10 years reference period plus a 20 km border and is shown in Fig. B.2.

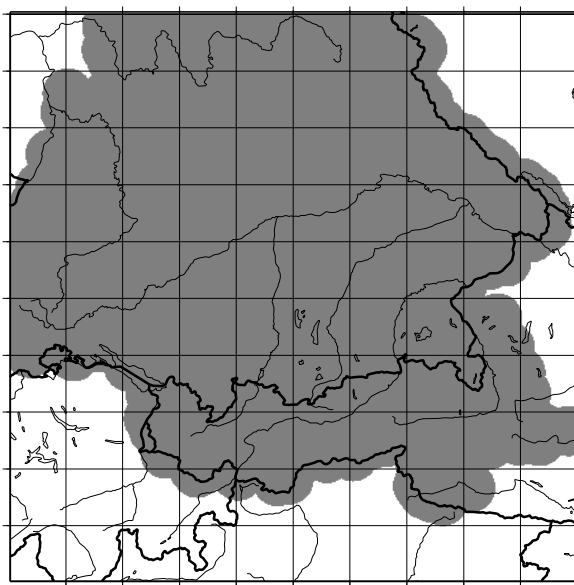


Figure B.2: The 10x10 grid cells MM5 model domain with national borders; shaded area: contains stations in which data were available during for every day throughout the 10 years reference period plus a 20 km border (see text for explanation).

# Appendix C

## Statistics

This chapter contains a description concerning the distribution of precipitation during a 10 years reference period (1991-2000). Daily rainfall has a fractional character (see e.g. Yano *et al.* 1996) and cannot be treated like temperature, as has been done by for example ?. Here an attempt is made to determine the annual distribution of rainfall during the reference period without neglecting the most significant inter-annual phenomena. The statistical quantities used in these sections are defined as follows.

### C.1 Mean ( $\mu$ )

The mean value of a data set  $x_i$  with  $i = 1, \dots, n$  is defined as,

$$\mu_x = \frac{1}{n} \sum_{i=1}^n x_i. \quad (\text{C.1})$$

In this case,  $n = 3650 \rightarrow 10 \text{ years} \times 365 \text{ days}$ .

### C.2 Variance ( $\sigma^2$ )

Values of a data set vary around  $\mu$ . This variation is defined for a data set  $x_i$  with  $i = 1, \dots, n$  as,

$$\sigma_x^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mu_x)^2. \quad (\text{C.2})$$

The root of the variance is defined as the *standard deviation* ( $\sigma_x$ ). Assuming a normal distribution, half a standard deviation away from the mean in either direction (i.e.  $-0.5 \cdot \sigma_x \leq \mu_x \leq 0.5 \cdot \sigma_x$ ) covers roughly 68% of the data.

### C.3 Skewness ( $\gamma_1$ )

Most data sets do not follow a normal distribution. Skewness is a measure of symmetry, or more precisely, the lack of symmetry of a distribution compared to a normal distribution and defined for a data set  $x_i$  with  $i = 1, \dots, n$  as,

$$\gamma_{1,x} = \frac{1}{n} \sum_{i=1}^n \left( \frac{x_i - \mu_x}{\sigma_x} \right)^3. \quad (\text{C.3})$$

If a data set is distributed symmetrically around  $\mu$ , then  $\gamma_1 = 0$ . Distributions with  $\gamma_1 < 0$  are *negatively skewed* or *skewed to the left*. Consequently, distributions with  $\gamma_1 > 0$  are *positively skewed* or *skewed to the right*. Daily rainfall distributions, bounded naturally on the left by no rainfall rates, are generally strongly skewed to the right, because of a wide 'tail' that extends far to the right.

## C.4 Kurtosis ( $\gamma_2$ )

Besides skewness, a data set can be more peaked compared to a normal distribution, or less, which is expressed by the kurtosis ( $\gamma_2$ ) and defined for a data set  $x_i$  with  $i = 1, \dots, n$  as,

$$\gamma_{2,x} = \frac{1}{n} \sum_{i=1}^n \left( \frac{x_i - \mu_x}{\sigma_x} \right)^4. \quad (\text{C.4})$$

The kurtosis of a normal distribution is three. From now on, the kurtosis is shifted by -3 in order to have  $\gamma_2 = 0$  for a normal distribution. *Platykurtic* distributions have  $\gamma_2 < 0$  and are less 'peaked' than a normal distribution. Distributions with  $\gamma_2 > 0$  are *leptokurtic* and more 'peaked' than a normal distribution.

# Appendix D

## Correlations and Heidke skill score

### D.1 Covariance

The covariance is a measure how much the deviations of two data sets match. If the variance is defined as in eq. C.2, the covariance of two data sets  $x_i$  and  $y_i$  with  $i = 1, \dots, n$  is defined as,

$$cov_{xy} = \frac{1}{n} \sum_{i=1}^n (x_i - \mu_x)(y_i - \mu_y). \quad (D.1)$$

### D.2 Pearson's correlation

The correlation is the standardized covariance of two data sets. The Pearson's correlation coefficient is calculated for two data sets  $x_i$  and  $y_i$  with  $i = 1, \dots, n$  as follows (Press *et al.* 2001, chap. 14),

$$r_{Pearson} = \frac{\sum_{i=1}^n (x_i - \mu_x)(y_i - \mu_y)}{\sqrt{\sum_{i=1}^n (x_i - \mu_x)^2} \sqrt{\sum_{i=1}^n (y_i - \mu_y)^2}}. \quad (D.2)$$

### D.3 Spearman rank-order correlation

Extreme values may have a great impact on the Pearson's correlation coefficient. The Spearman rank-order correlation coefficient is based on the position, or rank, of a certain value in a data set. The lowest value gets the ranking 1, while the highest value gets the highest ranking. The Spearman rank-order correlation coefficient is calculated for two data sets  $x_i$  and  $y_i$  with  $i = 1, \dots, n$  as follows (Press *et al.* 2001, chap. 14),

$$r_{Spearman} = \frac{\sum_{i=1}^n (R_i - \mu_R)(S_i - \mu_S)}{\sqrt{\sum_{i=1}^n (R_i - \mu_R)^2} \sqrt{\sum_{i=1}^n (S_i - \mu_S)^2}}, \quad (D.3)$$

in which  $R_i$  denotes the ranking of  $x_i$  and  $S_i$  denotes the ranking of  $y_i$ . The advantage over the Pearson's correlation is that the Spearman correlation is more robust, because of its lower dependency on extreme values in a data set.

## D.4 Heidke Skill Score

The Heidke skill score is a categorical skill score depending on the number of events for which a criterion is true (see e.g. Barnston 1992). Four groups are formed using this criterion and shown in Table D.1. Since it is the goal to find out how well the simulations do compared to the observations, the simulations provide the criterion and therefore, it is the median over the simulations for each time period (1 day, 10 days, 31 days, or 90 days). The letters a–d in Table D.1 denote a number of events. For example, the letter a denotes the amount of grid cells for which the simulated and observed precipitation amount were both lower than the median of the simulations over the whole research area. The Heidke

		observations	
		criterion true	criterion false
simulation	criterion true	a	b
	criterion false	c	d

Table D.1: Basis for Heidke skill score.

skill score ( $HSS$ ) is then calculated as:

$$C = a + d \quad (D.4)$$

$$N = a + b + c + d \quad (D.5)$$

$$E = \frac{(a + b)(a + c) + (d + b)(d + c)}{a + b + c + d} \quad (D.6)$$

$$HSS = \frac{C - E}{N - E} \quad (D.7)$$

where  $C$  is the number of correct simulations (i.e. the number of grid cells where the simulation and observation fall into the same class),  $N$  is total number of simulations (in the case of the research area described in Appendix A and B, 450 cells  $\times$  450 cells = 202500 cells), and  $E$  is the number of simulations expected to be correct by chance. If the simulations and the observations are perfectly negatively correlated, the  $HSS$  will be -1. A perfectly randomly based simulation has a  $HSS$  of 0, whereas a perfectly correlating simulation has a  $HSS$  of 1.



# Bibliography

- Antal, T., M. Droz, G. Györgyi, and Z. Rácz, 2001:  $\frac{1}{f}$ -noise and extreme value statistics. *Phys. Rev. Lett.*, **87**, nr. 24.
- Bárdossy, A., and E. J. Plate, 1992: Space-time model for daily rainfall using atmospheric circulation patterns. *Water Resour. Res.*, **28**, 1247–1259.
- Barnston, A. G., 1992: Correspondence among the correlation, rmse, and Heidke forecast verification measures; refinement of the Heidke score. *Weather and Forecasting*, **7**, nr. 4, 699–709.
- BayFORKLIM, 1996: *Klimaatlas von Bayern (German)*. Bayerischer Klimaforschungsverbund, 114pp.
- BayFORKLIM, 1999a: *Climate change in Bavaria and its impact. Final report (German)*. Bayerischer Klimaforschungsverbund, 90pp.
- BayFORKLIM, 1999b: *Regional climate modeling. Final report (German)*. Bayerischer Klimaforschungsverbund, 30pp.
- Beckmann, B. R., and T. A. Buishand, 2002: Statistical downscaling relationships for precipitation in the Netherlands and north Germany. *Int. J. Climatol.*, **22**, 15–32.
- Beljaars, A. C. M., P. Viterbo, M. J. Miller, and A. K. Betts, 1996: The anomalous rainfall over the United States during July 1993: sensitivity to land surface parameterization and soil moisture anomalies. *Mon. Wea. Rev.*, **124**, 362–383.
- Betts, A. K., 1986: A new convective adjustment scheme. Part I: Observational and theoretical basis. *Quart. J. Roy. Meteor. Soc.*, **112**, 677–691.
- Betts, A. K., 2000: Idealized model for equilibrium boundary layer over land. *J. Hydrometeorol.*, **1**, 507–523.
- Betts, A. K., J. H. Ball, A. C. M. Beljaars, M. J. Miller, and P. A. Viterbo, 1996: The land surface-atmosphere interaction: a review based on observational and global modeling perspectives. *J. Geophys. Res.*, **101**, 7209–7225.
- Betts, A. K., and M. J. Miller, 1993: Idealized model for equilibrium boundary layer over land. *Meteor. Monogr.*, **46**, 107–121.
- Bissolli, R., and E. Dittmann, 2003: *The land surface-atmosphere interaction: a review based on observational and global modeling perspectives. Objektive Wetterlagenklassen*. Klimastatusbericht des DWD 2002.

- Braun, S. A., and W. K. Tao, 2000: Sensitivity of high-resolution simulations of hurricane Bob (1991) to planetary boundary layer parameterizations. *Mon. Wea. Rev.*, **128**, 3941–3961.
- Buishand, T. A., M. V. Shabalova, and T. Brandsma, 2004: On the choice of the temporal aggregation level for statistical downscaling of precipitation. *J. Climate*, **17**, 1816–1827.
- Changnon, S. A., R. A. Pielke(Jr.), D. Changnon, R. T. Sylves, and R. Pulwarty, 2000: Human factors explain the increased losses from weather and climate extremes. *Bull. Amer. Meteor. Soc.*, **81**, nr.3, 437–442.
- Chen, F., and J. Dudhia, 2001: Coupling an advanced land surface–hydrology model with the Penn State–NCAR MM5 modeling system. Part I: Model implementation and sensitivity. *Mon. Wea. Rev.*, **129**, 569–585.
- Cohen, C., 2002: A comparison of cumulus parameterizations in idealized sea-breeze simulations. *Mon. Wea. Rev.*, **130**, 2554–2571.
- Colle, B., and C. F. Mass, 2000: The 5–9 February 1996 flooding event over the Pacific Northwest: sensitivity studies and evaluation of the MM5 precipitation forecasts. *Mon. Wea. Rev.*, **128**, 593–617.
- Colle, B., J. B. Olson, and J. S. Tongue, 2003: Multiseason verification of the MM5. Part II: Evaluation of high-resolution precipitation forecasts over the northeastern United States. *Weather and Forecasting*, **18**, nr.3, 458–480.
- Cosma, S., E. Richard, and F. Miniscloux, 2002: The role of small-scale orographic features in the spatial distribution of precipitation. *Quart. J. Roy. Meteor. Soc.*, **128**, 75–92.
- Cox, R., B. L. Bauer, and T. Smith, 1998: A mesoscale model intercomparison. *Bull. Amer. Meteor. Soc.*, **79**, nr.2, 265–283.
- Daly, C., R. P. Neilson, and D. L. Phillips, 1994: A statistical-topographic model for mapping climatological precipitation over mountainous terrain. *J. App. Meteor.*, **33**, 140–158.
- Daly, C., G. Taylor, and W. Gibson, 1997: The PRISM approach to mapping precipitation and temperature. in *10th Conf. on Applied Climatology* Reno, NV, Amer. Meteor. Soc. 10–12.
- Dudhia, J., 1989: Numerical study of convection observed during the winter monsoon experiment using a mesoscale two-dimensional model. *J. Atmos. Sci.*, **46**, 3077–3107.
- Dudhia, J., 1993: A nonhydrostatic version of the Penn State/NCAR mesoscale model: Validation tests and simulation of an Atlantic cyclone and cold front. *Mon. Wea. Rev.*, **121**, 1493–1513.
- DWD, 1995a: Wetterkarte des Deutschen Wetterdienstes 1995/i.
- DWD, 1995b: Wetterkarte des Deutschen Wetterdienstes 1995/ii.

- Fliri, F., 1975: *Das Klima der Alpen im Raume von Tirol*. Universitätsverlag Wagner, Innsbruck-München, 454pp.
- Fliri, F., 1984: *Synoptische Klimatographie der Alpen zwischen Mont Blanc und Hohen Tauern*. Universitätsverlag Wagner, Innsbruck-München, 686pp.
- Frei, C., and C. Schär, 1998: A precipitation climatology of the Alps from high-resolution raingauge observations. *Int. J. Climatol.*, **18**, 873–900.
- Gallus, W. A., 1999: Eta Simulations of three extreme precipitation events: sensitivity to resolution and convective parameterization. *Weather and Forecasting*, **14**, nr.3, 405–426.
- Gallus, W. A., and M. Segal, 2000: Sensitivity of forecast rainfall in a Texas convective system to soil moisture and convective parameterization. *Weather and Forecasting*, **15**, nr.5, 509–526.
- Garratt, J. R. (ed.), 1992: *The atmospheric boundary layer*. Cambridge University Press, 316pp.
- Giorgi, F., G. T. Bates, and S. J. Nieman, 1992: Simulation of the arid climate of the southern great basin using a regional climate model. *Bull. Amer. Meteor. Soc.*, **73**, nr.11, 1807–1822.
- Giorgi, F., G. T. Bates, and S. J. Nieman, 1993: The multiyear surface climatology of a regional atmospheric model over the western United States. *J. Climate*, **6**, 75–95.
- Gochis, D. J., W. J. Shuttleworth, and Z. L. Yang, 2002: Sensitivity of the modeled north American monsoon regional climate to convective parameterization. *Mon. Wea. Rev.*, **130**, 1282–1298.
- González-Rouco, J. F., H. Heyen, E. Zorita, and F. Valero, 2000: Agreement between observed rainfall trends and climate change simulations in the southwest of Europe. *J. Climate*, **13**, 3057–3065.
- Grell, G. A., 1993: Prognostic evaluation of assumptions used by cumulus parameterizations. *Mon. Wea. Rev.*, **121**, 764–787.
- Grell, G. A., J. Dudhia, and D. R. Stauffer, 1995: *A description of the Fifth-Generation Penn State/NCAR Mesoscale Model (MM5)*, *Tech. Note NCAR/TN-398+STR*. National Center for Atmospheric Research, Boulder, CO, 122pp.
- Hack, J. J., B. A. Boville, B. P. Briegleb, J. T. Kiehl, P. J. Rasch, and D. L. Williamson, 1993: *Description of the NCAR community Climate Model (CCM2)*, *NCAR Technical Note, NCAR/TN-382+STR*. National Center for Atmospheric Research, Boulder, CO, 120pp.
- Hong, S. Y., and H. L. Pan, 1996: Nonlocal boundary layer vertical diffusion in a medium-range forecast model. *Mon. Wea. Rev.*, **124**, 2322–2339.

- Huffman, G. J., R. F. Adler, P. Arkin, A. Chang, R. Ferraro, A. Gruber, J. Janowiak, A. McNab, B. Rudolf, and U. Schneider, 1997: The global precipitation climatology project (GPCP) combined precipitation dataset. *Bull. Amer. Meteor. Soc.*, **78**, nr. 1, 5–20.
- Janjić, Z. I., 1990: The step-mountain coordinate: Physical package. *Mon. Wea. Rev.*, **118**, 1429–1443.
- Janjić, Z. I., 1994: The step-mountain eta coordinate model: Further development of the convection, viscous sublayer, and turbulent closure schemes. *Mon. Wea. Rev.*, **122**, 927–945.
- Karl, T. R., W. C. Wang, M. E. Schlesinger, R. W. Knight, and D. Portman, 1990: A method of relating general circulation model simulated climate to the observed local climate. Part I: Seasonal statistics. *J. Climate*, **3**, 1053–1079.
- Kiehl, J. T., J. J. Hack, G. B. Bonan, B. A. Boville, D. L. Williamson, and P. J. Rasch, 1998: The National Center for Atmospheric Research Community Climate Model: CCM3. *J. Climate*, **11**, 1131–1149.
- Kilsby, C. G., P. S. P. Cowpertwait, P. E. O’Connell, and P. D. Jones, 1998: Predicting rainfall statistics in England and Wales using atmospheric circulation variables. *Int. J. Climatol.*, **18**, 523–539.
- Kim, J. W., J. T. Chang, N. L. Baker, D. S. Wilks, and W. L. Gates, 1984: The statistical problem of climate inversion: determination of the relationship between local and large-scale climate. *Mon. Wea. Rev.*, **112**, 2069–2077.
- Kraus, H., 2000: *Die Atmosphäre der Erde*. Verlag Vieweg, 470pp.
- Lettenmaier, D. P., 1995: *Stochastic modeling of precipitation with applications to climate model downscaling*. In: H. von Storch, and A. Navarra (eds.), *Analysis of Climate Variability: Applications of Statistical Techniques*. Springer Verlag, pages 197–212.
- Leung, L. R., and S. J. Ghan, 1999: Pacific Northwest climate sensitivity simulated by a regional climate model Driven by a GCM. Part I: Control simulations. *J. Climate*, **12**, 2010–2030.
- Leung, L. R., Y. Qian, and X. Bian, 2003: Hydroclimate of the Western United States based observations and regional climate simulations of 1981–2000. Part I: Seasonal statistics. *J. Climate*, **16**, 1892–1911.
- Ludwig, R., W. Mauser, S. Niemeyer, A. Coljan, R. Stolz, H. Escher-Vetter, M. Kuhn, M. Reichstein, J. Tenhunen, A. Kraus, M. Ludwig, M. Barth, *et al.*, 2003: Web-based modelling of energy, water and matter fluxes to support decision making in mesoscale catchments – the integrative perspective of GLOWA-Danube. *Physics and Chemistry of the Earth*, **28**, 621–634.
- McIlveen, R., 1995: *Fundamentals of weather and climate*. Chapman & Hal, 497pp.

- Miller, M. J., and A. A. White, 1984: On the non-hydrostatic equations in pressure and sigma coordinates. *Quart. J. Roy. Meteor. Soc.*, **110**, 515–533.
- Murphy, J., 1999: An evaluation of statistical and dynamical techniques for downscaling local climate. *J. Climate*, **12**, 2256–2284.
- Murphy, J., 2000: Predictions of climate change over Europe using statistical and dynamical downscaling techniques. *Int. J. Climatol.*, **20**, 489–501.
- Negri, A., E. N. Anagnostou, and R. F. Adler, 2000: A 10-yr climatology of amazonian rainfall derived from passive microwave satellite observations. *J. App. Meteor.*, **39**, 42–56.
- Pall, J. S., and E. A. B. Eltahir, 2001: Pathways relating soil moisture conditions to future summer rainfall within a model of the land-atmosphere system. *J. Climate*, **14**, 1227–1242.
- Pandey, G. R., D. R. Cayan, M. D. Dettinger, and K. P. Georgakakos, 2000: A hybrid orographic plus statistical model for downscaling daily precipitation in northern California. *J. Hydrometeor.*, **1**, 491–506.
- Pielke, R. A., G. A. Daly, J. S. Snook, T. J. Lee, and T. G. E. Kittel, 1991: Nonlinear influence of mesoscale land use on weather and climate. *J. Climate*, **4**, 1053–1069.
- Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, 2001: *Numerical Recipes in Fortran 77*. Chapter 13. Cambridge University Press, 933pp.: second edition.
- Reisner, J., R. M. Rasmussen, and R. T. Brintjes, 1998: Explicit forecasting of supercooled liquid water in winter storms using the MM5 mesoscale model. *Quart. J. Roy. Meteor. Soc.*, **124**, 1071–1107.
- Rockel, B., 2005: GKSS Research Centre Geesthacht: personal communication.
- Salathé, E. P., 2003: Comparison of various precipitation downscaling methods for the simulation of streamflow in a rainshadow river basin. *Int. J. Climatol.*, **23**, 887–901.
- Schär, C., C. Frei, D. Lüthi, and H. C. Davies, 1996: Surrogate climate-change scenarios for regional climate models. *Geoph. Res. Lett.*, **23**, nr. 6, 669–672.
- Schär, C., D. Lüthi, U. Beyerle, and E. Heise, 1999: The soil-precipitation feedback: A process study with a regional climate model. *J. Climate*, **12**, 722–741.
- Schmidli, J., C. Frei, and C. Schär, 2001: Reconstruction of mesoscale precipitation fields from sparse observations in complex terrain. *J. Climate*, **14**, 3289–3306.
- Schwarb, M., 2001: *The alpine precipitation climate evaluation of a high-resolution analysis scheme using comprehensive rain-gauge data*: Ph.D. thesis: Swiss Federal Institute of Technology, Zürich: ETH No. 13911.
- Schwarb, M., C. Daly, C. Frei, and C. Schär, 2001: *Mean annual and seasonal precipitation in the European Alps 1971–1990. Hydrological Atlas of Switzerland*. Federal Office for Water and Geology, Bern, Switzerland, Plates 2.6 and 2.7.

- von Storch, H., E. Zorita, and U. Cubasch, 1993: Downscaling of global climate change estimates to regional scales: An application to Iberian rainfall in wintertime. *J. Climate*, **6**, 1161–1171.
- von Storch, H., and F. W. Zwiers, 1995: *Statistical Analysis in Climate Research*. Cambridge University Press, 513pp.
- Stull, R. B., 1999: *An introduction to boundary layer meteorology*. Kluwer Academic Publishers, 670pp.
- Wang, W., and N. L. Seaman, 1997: A comparison study of convective parameterization schemes in a mesoscale model. *Mon. Wea. Rev.*, **124**, 252–278.
- Warner, T. T., D. N. Yates, and G. H. Leavesley, 2000: A community hydrometeorology laboratory for fostering collaborative research by the atmospheric and hydrologic sciences. *Bull. Amer. Meteor. Soc.*, **81**, nr.7 1499–1505.
- Watson, R. T., and Coauthors, 2001: *Climate Change 2001: Synthesis Report*: Technical Report: IPCC, Geneva, Switzerland: 184pp.
- Widmann, M., and C. S. Bretherton, 2000: Validation of mesoscale precipitation in the NCEP reanalysis using a new gridcell dataset for the northwestern United States. *J. Climate*, **13**, 1936–1950.
- Widmann, M., C. S. Bretherton, and E. P. S. Jr., 2003: Statistical precipitation downscaling over the northwestern United States using numerically simulated precipitation as a predictor. *J. Climate*, **16**, 799–816.
- Wilby, R. L., H. Hassan, and K. Hanaki, 1998: Statistical downscaling of hydrometeorological variables using general circulation model output. *J. Hydrol.*, **205**, 1–19.
- Wilby, R. L., and T. M. L. Wigley, 1997: Downscaling general circulation model output: A review of methods and limitations. *Prog. Phys. Geogr.*, **21**, 530–548.
- Wirth, V., 2005: University of Mainz, Institute for Atmospheric Physics: personal communication.
- Yano, J. I., K. Fraedrich, and R. Blender, 2001: Tropical convective variability as  $\frac{1}{f}$  noise. *J. Climate*, **14**, 3608–3616.
- Yano, J. I., J. C. McWilliams, and M. W. Moncrieff, 1996: Fractality in idealized simulations of large-scale tropical cloud systems. *Mon. Wea. Rev.*, **124**, 838–848.
- Zängl, G., 2004: The sensitivity of simulated orography precipitation to model components other than cloud microphysics. *Quart. J. Roy. Meteor. Soc.*, **130**, 1857–1875.
- Zhang, D. L., W. Z. Zheng, and Y. K. Xue, 2003: A numerical study of early summer regional climate and weather over LSA-East. Part I: Model implementation and verification. *Mon. Wea. Rev.*, **131**, 1895–1909.
- Zorita, E., and H. von Storch, 1999: The analog method as a simple statistical downscaling technique: comparison with more complicated methods. *J. Climate*, **12**, 2474–2489.

# Acknowledgments

First of all I would like to thank my supervisor Prof. J. Egger. Four years ago, he gave me the opportunity to give a seminar at the theoretical department of the meteorological institute. I am very pleased that he immediately noticed my interest in interdisciplinary scientific work and decided that I could work on the project GLOWA-Danube. In the course of my stay in Munich he gave me the opportunity to prepare own strategies and develop own ideas. Thank you very much for that.

Second, I would like to thank my co-corrector Priv. Doz. Dr. Günther Zängl. His knowledge about the MM5 model and the precipitation distribution over the Alps is unmatched and his suggestions were extremely helpful to my work. In particular, he helped me to improve the standard downscaling method to the refined method presented in Chapter 6. Thank you as well for giving me much advise about how to write a work like this, as well as about writing scientific papers.

I would also like to express many thanks to my colleague Andreas Pfeiffer. He introduced me to the GLOWA-Danube community and gave me many opportunities to present my work at numerous project meetings. Additionally, he supported me in getting acquainted with the Unix and the Linux operating systems, getting to know the model MM5, and handling my many computer problems. Getting through a four years period whilst literally facing each other every day demands either a high ability of cooperation or the development of a friendship. Without excluding the first, the second contributed considerably to a very pleasant period. Thank you very much.

My work could not have been successful without the preparation of long-term observed precipitation data. These were provided through Barbara Früh. On a continuous basis she updated the precipitation fields and made it possible for me to work with them without serious problems. Moreover, our long discussions about the downscaling method contributed considerably to many improvements during the last four years. Thank you for supporting me throughout this period.

At this point, I take the opportunity to also thank Prof. V. Wirth. His comments on my work, in particular on Chapter 4, made it a more complete work. In our discussions, he always praised my ideas with such an enthusiasm that it gave me clear insight in what I was doing and the way I want to go.

I would also like to thank the project leader of the project GLOWA-Danube Prof. W. Mauser. He gave me the possibility to work with many other disciplines in science, inside as well as outside the geosciences, and appreciate their way of thinking. His critical comments on my work and the meteorology in general allowed me to gain a better insight into the many disciplines I encountered in the project.

Additionally, I want to thank all the people at the institute for supporting me and giving me advice in many different kinds of subjects. Although I cannot mention all of

them, I want to thank Uschi Pliete in particular for helping me with administrative work, and Heinz Lösslein for having the patience to explain to me the interesting and unexpected features in a computers life. I also want to thank Markus Garhammer for providing one of his beautiful photos to put on the front page of this work and Robert Goler for proof-reading some of the chapters. In my time here I was lucky to be a part of the positive environment of the meteorological institute in Munich and I will always remember these years as a very pleasant period of my life.

Last, but not least, this work could not have originated without the utter support of my parents, my sister Marije and her husband Jeroen. Thank you all very much for your never-ending trust and the possibility to breath in fresh salty Dutch air once in a while.



# Curriculum vitae

Janus Willem (Hans) Schipper was born in Amersfoort (Netherlands) on 21 August 1975. After 7 years he moved to Oostkapelle, Zeeland, close to the North Sea coast. He attended 'De Regenboog' primary school, followed by the secondary schools MAVO, HAVO, and finally VWO. His interest in physics and weather phenomena in general made him study 'Bodem, Water en Atmosfeer' (Soil, Water and Atmosphere - BWA) at the Wageningen University in Fall 1995. There, he was active in playing volleyball and organizing a volleyball tournament with around 300 attendants in 1997. In 1998 he became chairman of the study society of BWA called 'Pyrus'. He used the university's contacts to go to the 'Deutscher Wetter Dienst' (German Weather Service - DWD) in Lindenberg (Germany) under the supervision of Dr. F. Beyrich for nearly 4 months in 1999. His work at the DWD consisted of assessing the results of the Large Aperture Scintillometer, which was located in Lindenberg. His thesis then involved the study of the energy balance components from measurements of the 'Lindenberg Inhomogeneous Terrain - Fluxes between Atmosphere and Surface - a long-term Study 1998' (LITFASS'98) experiment under the supervision of Prof. A.A.M. Holtslag. Subsequently, he went to the 'National Oceanic & Atmospheric Administration' (NOAA) in Boulder, Colorado, USA, for half a year under the supervision of Dr. W.M. Angevine. His focus was to study the ability to extract the sensible heat flux from wind profiler measurements. In 2001 he successfully finished his studies in Wageningen and moved to Munich (Germany) to do his PhD. At the Meteorological Institute of the Ludwig Maximilian University, he worked for the project GLOWA ('GLobal WAter cycle'). After 4 years he finished his PhD, resulting in a downscaling method to redistribute rainfall in the Alpine region using a numerical model presented in this work. He hopes to successfully continue to work in science at the Institute for Meteorology and Climate Research at the University of Karlsruhe (Germany).