Project no. GOCE-CT-2003-505539

Project acronym: ENSEMBLES

Project title: ENSEMBLE-based Predictions of Climate Changes and their Impacts

Instrument: Integrated Project

Thematic Priority: Global Change and Ecosystems

**Deliverable 5.32: Detailed study of heavy precipitation events in Alpine region using ERA40 driven RCMs**

Due date of deliverable: 31st August 2009
Actual submission date: 30th September 2009

Start date of project: 1 September 2004    Duration: 60 Months

Organisation name of lead contractor for this deliverable: ETH Zurich

---

<table>
<thead>
<tr>
<th>Dissemination Level</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>PU</td>
<td>Public</td>
</tr>
<tr>
<td>PP</td>
<td>Restricted to other programme participants (including the Commission Services)</td>
</tr>
<tr>
<td>RE</td>
<td>Restricted to a group specified by the consortium (including the Commission Services)</td>
</tr>
<tr>
<td>CO</td>
<td>Confidential, only for members of the Consortium (including the Commission Services)</td>
</tr>
</tbody>
</table>

Revision [Final]
Detailed study of heavy precipitation events in the Alpine region using ERA40 driven RCMs

Authors:
Julian Arnold, Pardeep Pall,
Thomas Bosshard, Sven Kotlarski, Christoph Schär

Correspondence:
pardeep.pall@env.ethz.ch

September 2009
Abstract

This report evaluates fifteen regional climate models (RCMs) from the European Commission’s ENSEMBLES project, with regard to their simulation of daily precipitation in the region of the European Alps. Model integrations, having a spatial resolution of 25km, have been driven for “present-day” (1961-1990) climate by boundary conditions from the ERA-40 re-analysis. The E-OBS gridded dataset, which has been produced as part of the ENSEMBLES project, serves as observational reference for the evaluation. Focus is placed on extreme events diagnosed via multiyear return values of daily precipitation amounts, using techniques from extreme value theory for their estimation. Most models tend to overestimate these extreme precipitation diagnostics. Domain mean biases generally exceed +10% and can reach values of up to several tens of percent, particularly in winter. In summer and in the southern part of the evaluation area, however, a general underestimation of precipitation intensity causes a deviation from this tendency, and so a notable number of models achieve a reasonable performance. Characteristic model deficiencies with regard to the spatial representation of extreme precipitation in dynamically active seasons such as autumn and winter involve, on one hand, overestimation of return values along and over the Alpine mountain rims and, on the other hand, underestimation (in part also realistic simulation) of return values in a zone of inner-alpine areas. The mechanisms underlying these error patterns are thought to be overestimation of topographic precipitation enhancement and overestimation of rain shadowing, respectively. Overall, there is a slight tendency for models to less realistically simulate extreme precipitation than average precipitation. This finding, however, may be viewed with a certain degree of reservation given known issues in E-OBS with regard to smoothing of extremes due to gridding of station data, and the relative sparsity of this data over the Alpine region. Nevertheless, in spite of substantial deviations, the models are able to capture observed mesoscale structures of extreme precipitation over this region and thus to refine climate information from General Circulation Models (GCMs) in areas of complex topography.
# Contents

1 Introduction  
  1.1 Motivation  
  1.2 Objectives and Related Tasks  
  1.3 Outline  

2 Theoretical Background  
  2.1 The Alpine Precipitation Climate  
    2.1.1 Mean Precipitation  
    2.1.2 Strong Precipitation  
    2.1.3 Trends  
  2.2 Extreme Value Theory  
    2.2.1 Definition  
    2.2.2 Block Maxima Method  

3 Data  
  3.1 Observational Data Set  
  3.2 Regional Climate Model Data  

4 Methodology  
  4.1 Climatological Diagnostics  
  4.2 Extreme Value Analysis  
  4.3 Alpine Subregions  
  4.4 Uncertainty Calculations  

5 Results  
  5.1 Domain Means for Alpine Subregions  
  5.2 Spatial Distribution of Relative Bias  
    5.2.1 Extreme Precipitation in Winter  
    5.2.2 Extreme Precipitation in Summer  
    5.2.3 Extreme Precipitation in Autumn  

6 Conclusions  

7 References
List of Figures

2.1 Geography of the European Alpine Region 5
3.1 Observed (E-OBS) mean daily precipitation 12
3.2 Observed (E-OBS) 90% quantile of precipitation on days with ≥ 1mm 12
4.1 Gumbel diagrams for grid box samples of daily precipitation extremes 17
4.2 Alpine subregions used in model evaluations of this report 18
5.1 Domain mean values and 90% bootstrap confidence intervals of mean precipitation (mea) for ENSEMBLES regional climate models and observations 22
5.2 Domain mean values and 90% bootstrap confidence intervals of wet day frequency (fre) for ENSEMBLES regional climate models and observations 23
5.3 Domain mean values and 90% bootstrap confidence intervals of wet day intensity (int) for ENSEMBLES regional climate models and observations 24
5.4 Domain mean values and 90% bootstrap confidence intervals of empirical 90% quantile of wet day precipitation amounts (q90) for ENSEMBLES regional climate models and observations 25
5.5 Domain mean values and 90% bootstrap confidence intervals of 5-year return value of 1-day precipitation extreme (x1d.5) for ENSEMBLES regional climate models and observations 26
5.6 Domain mean values and 90% bootstrap confidence intervals of 20-year return value of 1-day precipitation extreme (x1d.20) for ENSEMBLES regional climate models and observations 27
5.7 Boxplot diagrams for relative bias of ENSEMBLES regional climate models in domain mean basic (mea, fre, int, q90) and extreme precipitation (x1d, x5d) diagnostics 29
5.8 Relative bias of ENSEMBLES regional climate models in 5-year return value of 5-day precipitation extreme (x5d.5) in winter (DJF) 32
5.9 Relative bias of ENSEMBLES regional climate models in 5-year return value of 1-day precipitation extreme (x1d.5) in summer (JJA) 34
5.10 Relative bias of ENSEMBLES regional climate models in 5-year return value of 1-day precipitation extreme (x1d.5) in autumn (SON) 36
5.11 Mean autumn (SON) 5-year return value of 1-day precipitation intensity (x1d.5) for ENSEMBLES regional climate models and observations 39
## List of Tables

3.1 Overview of ENSEMBLES Regional Climate Models 13

4.1 Diagnostic quantities of daily precipitation 15

5.1 Model bias in domain mean of basic precipitation diagnostics for the northern and southern Alpine subregion 30

5.2 Model bias in domain mean of extreme precipitation diagnostics for the northern and southern Alpine subregion 30
List of Abbreviations

CTRL  Control Simulation
E-OBS  ENSEMBLES daily gridded observational data set
ERA-40 ECMWF reanalysis of the global atmosphere 1957-2002
ETHZ  Swiss Federal Institute of Technology Zurich
GCM  General Circulation Model
GEV  Generalized Extreme Value distribution
IPCC  Intergovernmental Panel on Climate Change
NA  Northern Alps
NH  Northern Hemisphere
RCM  Regional Climate Model
SA  Southern Alps
SCEN  Climate Scenario Simulation
SRES  Special Report on Emissions Scenarios
1 Introduction

1.1 Motivation

The Alps play a significant role in the hydrologic cycle of the European continent, leading to the frequently quoted expression “water tower of Europe”. Indeed, the interaction between Alpine orography and impinging weather systems produces annual precipitation amounts that are more than twice as high as neighbouring regions (Frei and Schmidli, 2006). Thanks to their abundance of precipitation the Alps act as a substantial freshwater reservoir, consisting of glaciers, snow fields, lakes, and river systems. The stored water forms an important resource for energy and food production as well as for the provision of drinking water. This holds not only for Alpine countries, but also for many European regions in their surroundings: especially in summer, when melting of snow and ice in the Alps compensates reduced runoff generation in these regions (Viviroli and Weingartner, 2004).

Analogous to total annual precipitation, the frequency of strong precipitation in the Alpine region is higher than the rest of the European continent (Schär et al, 1998). Heavy precipitation events, often associated with severe flooding, are an integral part of Alpine climatology. This is exemplified by a whole string of remarkable cases during the past 22 years: Valtellina / ValPoschia - Vallo d’Intelvi (1987), Vaucluse (1992), Upper Valais (1993), Piedmont (1994), Switzerland / Eastern Austria (1999), Valais / Ticino / Aosta Valley (2000), Central Switzerland / Vorarlberg / Bavaria (2005). The last serious event in August 2005 turned out to be extremely unusual, both in terms of intensity and spatial extent (MeteoSchweiz, 2006). In a numerical sensitivity experiment MeteoSwiss (2006) demonstrated that the influence of the Alpine orography was a very crucial factor for the observed high precipitation amounts.

The study of the spatial and temporal variability of precipitation in the Alpine region has a long history. From early on, people began to establish national observation networks, that yielded long precipitation time series. These data represent an invaluable source of information and they allowed climatologists to shed light on the Alpine precipitation climate and its evolution over the course of the 20th century. What stands out particularly among the numerous analyses and studies is the Alpine precipitation climatology by Frei and Schär (1998). It is a gridded analysis with a resolution of about 25 km and covers the entire mountain range and the adjacent foreland. The climatology is based on observations at more than 6600 rain-gauge stations from the uniquely dense networks of the Alpine countries and has been determined from daily analyses of data spanning the period 1971-1990. More recently, a new similarly high-resolution gridded data set of daily precipitation has been produced that includes the Alpine region (Haylock et al, 2008). In addition to European-wide coverage, it offers the advantage of a prolonged time span (1950-2006).

Scientific interest did not remain confined to investigating past and present precipitation climate of the Alps: the particular susceptibility of this mountain region to the impacts of a changing climate (Beniston, 2003) and the enhanced tendency for the occurrence of extreme events fostered the construction of reliable regional scenarios for the future Alpine precipitation climate. These scenarios are obtained by the development and operation of physical climate models. However, whilst conventional General Circulation Models (GCMs), are able to adequately represent global-to continental-scale features they cannot adequately represent regional details. The coarse spatial resolution of a typical GCM leads to an inadequate representation of the complex Alpine relief. In addition, it prevents modeling of mesoscale processes and phenomena, that are often caused orographically. Since strong precipitation is in particular often considerably influenced by topography and mesoscale processes, a GCM’s ability to simulate these effects in its regional-scale climate is highly restricted. Regional climate scenarios are therefore constructed through the process of downscaling or regionalization (Gyalistras et al, 1998), that is, the spatial refinement of larger-scale climate information. There exist two principal procedures: dynamical downscaling and...
empirical / statistical downscaling (Christensen et al 2007a). A widely used technique in dynamical downscaling consists of nesting a higher-resolution Regional Climate Model (RCM) into a GCM over the region of interest (Giorgi and Mearns, 1999). Usually, the flow of information occurs exclusively from the GCM to the embedded RCM (one-way nesting mode). Running such an RCM over a specific time slice requires forcing lateral boundary conditions and initial data that are derived from corresponding GCM simulations or large-scale observational data. In the context of modeling errors, it should be emphasized that nested RCMs are not expected to eliminate fundamental biases in the GCM driving fields. Rather, their role is to add more detailed regional information to the large-scale climate scenario (Giorgi, 2008).

RCMs are operated and continuously improved by a whole series of European institutes. Fifteen such models currently form part of a 5-year ambitious research programme: the ENSEMBLES project, funded by the European Commission and coordinated by the Met Office Hadley Centre in the UK (Hewitt, 2005). The project aims at the development of an ensemble climate prediction system based on state-of-the-art GCMs and RCMs with European roots. This system is designed to produce reliable and objective probabilistic scenarios of future climate for the assessment of related impacts on environment and society. The probabilistic approach includes quantification of the prediction uncertainty, which in case of regional scenarios arises from the following sources (Christensen et al, 2007a): (i) **sampling uncertainty** - Model climate represents an average over a finite number of years; (ii) **regional model uncertainty** - RCMs use different discretization techniques for basic equations and different parameterization schemes for subgrid processes such as convection, land-atmosphere exchange, radiation or clouds; (iii) **emission uncertainty** - Choice of Intergovernmental Panel on Climate Change (IPCC) special report emissions scenarios (SRES)\(^2\) (Randall et al, 2007); (iv) **boundary uncertainty** - RCMs are driven by different boundary conditions from different GCMs. In order to better understand the uncertainty inherent in RCM simulations and to reduce it finally by improving the representation of processes and feedbacks (see e.g. Jacob et al, 2007), the activities within the ENSEMBLES project include comprehensive evaluation of involved models. This can be achieved by assessing their performance in simulating “present-day climate” (Randall et al, 2007), since RCMs allow continuous forcing at their lateral boundaries with gridded observational data (Gyalistras et al, 1998).

As part of ETH Zürich’s contribution to the EU-ENSEMBLES project, this report evaluates fifteen RCMs with regard to their representation of extreme precipitation in the region of the European Alps. This mountainous zone is focused on because the complexity of its terrain poses a considerable challenge to RCMs in reproducing the observed climate of intense precipitation. The challenge is further increased by the fact that the Alps constitute a transition between two major climate zones (Schär et al, 1998): the northward mid-latitude temperate- and the southward Mediterranean-climate type. Both aspects thus make the Alps a very interesting model evaluation region that could provide valuable insights into capabilities and limitations of RCMs with regard to simulating precipitation.

---

1The term scenario is defined as plausible representation of future climate conditions, often in simplified form, for use in exploring the potential effects of anthropogenic climate change (IPCC, 2007).

2The term emission scenario is defined as plausible representation of the future evolution of emissions of potentially radiatively active substances such as greenhouse gases or aerosols (IPCC, 2007).
1.2 Objectives and Related Tasks

Two basic objectives are addressed in this report:

- **Evaluation of ENSEMBLES RCMs:**
  Fifteen high-resolution RCMs from the EU-ENSEMBLES project are to be evaluated with regard to their representation of mean and extreme daily precipitation in present-day climate (1961-1990) over the region of the European Alps. A newly-produced, quality-controlled, high-resolution gridded data set (E-OBS; Haylock et al, 2008) is used as the observational reference.

- **Comparison of ENSEMBLES RCMs with PRUDENCE RCMs:**
  Skill of the considered RCMs in representing mean and extreme precipitation is to be compared with that of a previous generation of RCMs involved into the PRUDENCE³ project (Christensen et al, 2007). Relevant evaluation results for PRUDENCE RCMs are presented by Frei et al (2003) and Frei et al (2006).

³PRUDENCE = Prediction of Regional scenarios and Uncertainties for Defining EuropeanN Climate change risks and Effects

1.3 Outline

The remainder of this report is structured as follows. Theoretical background knowledge about key topics of this report is offered in the next chapter (**Chapter2**). This comprises both a brief description of major features of the Alpine precipitation climate, and a short introduction to extreme value theory focusing on the block maximum method to be employed for analyzing extreme precipitation. The subsequent chapter (**Chapter3**) characterizes the RCM data and the observational reference data (E-OBS) to be analyzed. It is followed by a description of the methodology adopted for the evaluation (**Chapter4**) and the presentation of the corresponding model evaluation and inter-comparison results (**Chapter5**). The report ends with a discussion of the main findings along with some concluding remarks (**Chapter6**).
2 Theoretical Background

2.1 The Alpine Precipitation Climate

The European Alps are a major mountain range running in an 800 km long arc with an average width of about 200 km through Austria, France, Germany, Liechtenstein, Italy and Switzerland (Figure 1). They are exposed to oceanic, continental, polar and subtropical (in particular Mediterranean and sometimes even Saharan) influences and are situated in a comparatively warm part of the Northern Hemisphere (NH) mid-latitudes owing to the proximity of the Gulf Stream and the Mediterranean Sea, which act as significant heat reservoirs (Beniston, 2006). Their configuration and spatial dimensions enable the Alps to modify or to trigger synoptic weather systems (Schär et al., 1998), producing Alpine lee cyclogenesis for example. Not least, these dynamic interactions between mountains and weather systems have far-reaching effects on the Alpine precipitation climate, which is the subject of the following section. Mean seasonal precipitation, strong precipitation, and precipitation trends for the Alps are briefly described below, principally based on Frei and Schmidli (2006).

![Figure 2.1: Geography of the European Alpine Region. Source: National Geographic Society.](image-url)
2.1.1 Mean Precipitation

Roughly speaking, the general distribution of mean precipitation is characterized by two elongated wet zones along the northern and southern rim of the Alps and a drier inner-Alpine sector in between. This pattern is mainly the effect of orographic precipitation enhancement at the mountain rims and of leeside rain-shadowing of the inner-Alpine regions (Frei and Schär, 1998). When considering the individual seasons, the picture has to be somewhat differentiated: In winter precipitation activity occurs along the northern rim of the Alps, especially in the western part, while in summer it is not only localized in the central and eastern parts, but also much more pronounced. The particular feature of the Southern Alpine wet zone is its subdivision into two anomaly centers (Southern Switzerland, northernmost Piedmont and Lombardy - Julian and Carnic Alps north of the Adriatic Sea), which are present in all season except in winter. Of central importance to the existence of these two precipitation signatures is their position: close to the Mediterranean Sea and at the northern end of major topographic indentations that channel the advection of moist air masses from the South. (Frei and Schmidli, 2006)

2.1.2 Strong Precipitation

Basically, strong precipitation events in the Alps fall into two categories: (1) locally confined events of relatively short duration, and (2) long-lasting and widespread episodes. The former are caused by thunderstorm cells (convective precipitation), whereas the latter are connected to synoptic disturbances or persistent advection and blocking of moist air masses (stratiform precipitation with embedded convection).

Intense Daily Precipitation

Intense daily precipitation is especially found in southern parts of the Alps. Regions with the highest tendency for large daily amounts include Southern Switzerland / Northern Italy as well as areas north of the Adriatic Sea. Here, strong precipitation events appear with particular frequency in autumn, when a pool of warm moisture-laden air is available over the Mediterranean Sea. Ahead of a cold front, a deep upper-level trough or a lee-cyclone, this comparatively weakly stratified air-mass can be steered towards the Alps and, as a result of advective and convective lifting, produce intense and sustained precipitation south of the Alpine crest.

Compared to the surrounding flatland areas, the entire Alpine range shows a higher frequency of intense daily precipitation. The degree, however, varies depending on region and season. Along the northern rim of the Alps, for example, intense daily precipitation arises most frequently during summer and is generally associated with thunderstorms (Frei and Schmidli, 2006).

Multiday Precipitation Episodes

Long-lasting precipitation phases (10 days and more) are particularly common in the northern Alpine area. They are predominantly observed in winter and spring, with the exception of the central and eastern Alpine rim, where they also occur in summer. A clear difference from the North is found in the lower Rhone Valley, the Po Valley, in Eastern Austria and in the inner-Alpine valleys (southern Tirol, Engadin, Aosta Valley). These regions much less frequently experience continuous precipitation episodes (Frei and Schmidli, 2006).
Droughts

The fact that the Alpine range constitutes a major weather divide, hindering the progress of Atlantic weather systems into the Mediterranean region, shows up well in the persistence of dry conditions. Across the Alps, there is a pronounced gradient in the average length of the longest continuous drought in a year. South of the Alpine range, the corresponding duration is about twice as long as that in the North. An interesting aspect in this context is that parts of the Southern domain coincide with the typical heavy rainfall areas at the Southern rim of the Alps. While in the north persistent droughts can happen throughout the year, the south, that is, the Mediterranean Coast and the Po Valley is affected by such episodes primarily in summer (Frei and Schmidli, 2006).

2.1.3 Trends

For the period of the 20th century, Schmidli et al (2002) detected clear trends in mean winter and autumn precipitation. The long-term signal in winter indicates a statistically significant increase by 20-30 percent in a large interconnected area of the western and northern Alpine region and a statistically non-significant decrease in the same magnitude in the south-eastern Alps. In autumn the analysis revealed a statistically significant decrease in the south-west and south-east of the Alpine region, again in the order of 20-30 percent. The other seasons behaved less clearly: changes are less pronounced and in most cases statistically non-significant.

In further analyses, Schmidli and Frei (2005) found that the increase in mean winter precipitation can be ascribed on average both to an increase in precipitation frequency and precipitation intensity. Additionally, they observed a general increase in the frequency of intense winter and autumn precipitation. (Frei and Schmidli, 2006b)

2.2 Extreme Value Theory

2.2.1 Definition

As the name implies, extreme value theory represents, in a certain respect, a special discipline within statistics. Unlike "classical statistics", which basically focuses on the average stochastic behaviour of a process, extreme value statistics deals with models and techniques for quantifying the extreme behaviour of a process: that is, at the tail of a distribution. Over the last decades, extreme value theory gained considerable importance and has established itself in various applied fields (Coles, 2001). It is used, e.g., for modeling natural extremes / risks (extreme rainfall, floods, ocean waves, wind storms), for portfolio adjustment in the insurance industry or for modeling risk assessment on financial markets.

It is in the nature of extreme values that they are rare. Consequently, extreme value analyses face the difficulty of estimating the probability of process events that are so extreme that they have never been observed. This necessitates an extrapolation from observed values to unobserved, unusually large or small values, and a fundamental task of extreme value theory consists of providing models that enable such a procedure (Coles, 2001).

There are two main modeling approaches to extreme value analysis:

- **Block maxima method**: modeling the sample (block) maximum over a certain time period
- **Peak over threshold method**: modeling sample values exceeding a given threshold within a certain time period
The following subsection provides a brief introduction to the block maxima method following Coles (2001), since this report will employ this approach.

### 2.2.2 Block Maxima Method

#### Fundamentals

The following description of the block maxima approach is based on a typical example of use: modeling of extreme summertime precipitation at a given location, using daily amounts from a period of several decades. As a first step, the block maxima method requires the subdivision of the entire data set into large block sequences of equal size. For our case, this means that each summer season (June/July/August, JJA) is treated as a block sample. The second step is then to determine the block maxima $M_n$, that is, the seasonal maximum of each summer season that consists of $n = 92$ daily observations $X_i$:

$$M_n = \max\{X_1, ..., X_n\} \quad (2.1)$$

With regard to the sampling distribution of the summer maxima, a fundamental theorem from extreme value theory, the extremal type theorem, states that, as $n$ increases towards infinity, the distribution of the maxima $M_n$ after linear renormalization converges to one of the so-called extreme value distributions, independent of the parent distribution (i.e. the distribution of the observations $X_i$).

The group of extreme value distributions includes three classes of asymptotic distributions, known as Gumbel, Fréchet, and Weibull distributions. For reasons of convenience, these three classes are combined into one single family of distributions:

$$\text{GEV}(z ; \mu, \sigma, \xi) = \exp\left\{-\left[1 + \frac{\xi(z-\mu)}{\sigma}\right]^{-1/\xi}\right\}, \quad (2.2)$$

defined on the set \{z: 1 + \xi(z-\mu)/\sigma > 0\} where $z$ are the maxima. This cumulative distribution function (cdf) is termed a generalized extreme value (GEV) distribution and contains three parameters: a location parameter $\mu$, a scale parameter $\sigma$ and a shape parameter $\xi$. From this arises the third step in the procedure of the block maximum method, fitting the GEV distribution to the sample block maxima. The distribution can be fitted by estimating the corresponding parameters, using either the method of maximum likelihood or the method of L-moments.

Underlying the block maximum approach is the assumption that the observations $X_i$ (considered as random quantities) are independently and identically distributed (iid). In practice, it can be violated by a non-stationary behaviour (seasonality, trends) of the modelled process. Applying the block maxima method is, moreover, only meaningful under the condition of sufficiently large data blocks, with the critical size depending on the parent distribution of the observations $X_i$.

#### Return Values and Gumbel Diagram

It is very common to define blocks that correspond to the length of one year. Depending on whether all days or those of a specific season are considered, this results in an extreme value data set of annual maxima or seasonal maxima. Widely used in conjunction with such maxima are the concepts of return value and return period. By definition, the return value for a return period of $T=1/p$ years is the threshold $z_p$ that is exceeded by an annual or seasonal extreme with a probability of $p$. It is calculated as a quantile of the annual $l$ seasonal maxima distribution, whose estimate is obtained through inversion of Equation 2.2:
\[ z_p = \mu - \sigma \xi [1 - \{ -\log(1 - p) \}]^{\frac{1}{\xi}} \quad \text{for } \xi \neq 0, \]

\[ z_p = \mu - \sigma \log \{ -\log(1 - p) \} \quad \text{for } \xi = 0, \]  

(2.3)

where \( GEV(z_p) = 1 - p \). The interpretation of a return value \( z_p \) for a return period of \( T \) years is twofold: (1) an event with this return value is expected to occur on average in \( T \) years next time, or (2) there is on average one corresponding event within a time period of \( T \) years.

An effective graphical device for the assessment of a GEV fit to a set of annual / seasonal maxima or for the extrapolation to long return periods is the Gumbel diagram, that depicts the return value \( z_p \) as a function of the return period \( T \). The peculiarity of the Gumbel diagram is that the return period axis (bottom axis) is subject to a linear scaling with respect to the Gumbel variate \( Y \), that is defined as:

\[ Y(z_p) = -\log y_p \quad \text{with } y_p = -\log(GEV(z_p)) = -\log(1 - p). \]  

(2.4)

The result of this axis transformation is, on the one hand, a condensation of the distribution tail and thus a focus on the effect of extrapolation, and, on the other hand, the representation of the Gumbel distribution (\( \xi = 0 \)) as a straight line, as it follows from:

\[ z_p = \mu - \sigma \xi [1 - y_p^{\frac{1}{\xi}}] \quad \text{for } \xi \neq 0, \]

\[ z_p = \mu - \sigma \log y_p \quad \text{for } \xi = 0, \]  

(2.5)

where \( y_p \) is defined as in Equation 2.4. In case of \( \xi > 0 \), the GEV fit in the Gumble diagram is concave without a finite bound, and if \( \xi < 0 \), it takes on a convex form with asymptotic limit at \( \mu - \sigma \xi \). This reveals the enormous influence of the shape parameter on the extrapolation to extended return periods and at the same time explains the need for a correspondingly accurate estimation. For examples of Gumble diagrams, see Figure 4.1 in the Methodology chapter.

Quantification of Uncertainty

Estimates of model parameters are always afflicted with uncertainty due to sampling variability. In extreme value models this uncertainty propagates up to return values and can considerably increase on extrapolation (Coles, 2001). Each extreme value analysis should therefore include information about the reliability of return values. In essence, there are three possibilities to quantify the uncertainty of return values: resampling confidence intervals (parametric or non-parametric), asymptotic maximum likelihood confidence intervals or likelihood-profile confidence intervals. Each of these options has its advantages and disadvantages, and the choice depends on the actual requirements and conditions.
3 Data

3.1 Observational Data Set

The observational reference for this report is E-OBS created by Haylock et al (2008) for evaluation activities within the framework of the European Union’s (EU) ENSEMBLES project. It is a high-resolution gridded data set of daily precipitation totals and temperature variables, encompassing the European continent and covering the period from 1950 to 2006. The underlying observing station network is entirely located on land and varies both with respect to time and space. Highest station density is generally in the United Kingdom (UK), Netherlands and Switzerland. The time span of the record, the spatial extent and grid resolution, the number of contributing observation stations, and the integration of daily uncertainty estimates make the E-OBS data set unique compared to similar pre-existing data sets. In the 25 km resolution version, the E-OBS rotated pole grid conforms with the data grid of many of the ENSEMBLES RCMs to be evaluated in this report.

Haylock et al (2008) followed a three-step procedure to interpolate spatially inhomogeneous station observations to the regular grid. Interpolation was geared to generate grid box values that represent the best estimate average of the respective observations. For the evaluation of RCMs these values form a more solid base than direct point observations, since modelled processes, in particular precipitation, show a spatially averaged character (Osborn and Hulme, 1998, in Haylock et al, 2008).

On assessing the uncertainty in the E-OBS data set, Haylock et al (2008) made the assumption that interpolation clearly exceeds other potential sources of uncertainty (measurement and recording errors, data quality, homogeneity), which is usually justified in interpolated data sets. They demonstrated that average interpolation uncertainty over the entire E-OBS area is determined by the season and the number of observation stations. As this number was reduced at the start (1950-1960) and end (1990-2006) of the E-OBS period, uncertainty tends to be larger then. A disadvantageous concomitant of spatial interpolation is smoothing of extremes. Haylock et al (2008) examined this aspect in the E-OBS data set and showed that the 10-year return value of daily precipitation is on average reduced by a factor of 0.66. For some observing stations it was found to be reduced by even more than half. Given that models describe area averaged processes, Haylock et al (2008) infer that the reduction of extremes in the E-OBS data set allows direct comparison with RCMs of the same spatial resolution.

A visual inspection of the E-OBS data set is enabled by Figure 3.1 and Figure 3.2. The former depicts seasonal means of daily precipitation for each grid point and the reference period 19611990, whereas the latter presents the empirical 90% quantile of precipitation on days with \( \geq 1 \) mm, again on a grid point basis, for all seasons and the period 1961-1990. Note the Alpine ridge-like features present in many cases, with more widespread precipitation in summer and autumn. Also note the differences between mean and extreme patterns.
Figure 3.1: Observed (E-OBS) mean daily precipitation (1961-1990, mm/d) for (a) winter (DJF), (b) spring (MAM), (c) summer (JJA) and (d) autumn (SON).

Figure 3.2: Observed (E-OBS) 90% quantile of precipitation on days with ≥Amm (1961-1990, mm/d) for (a) winter (DJF), (b) spring (MAM), (c) summer (JJA) and (d) autumn (SON).
Table 3.1: Overview of ENSEMBLES Regional Climate Models from which control simulations are evaluated in this thesis.

<table>
<thead>
<tr>
<th>Model Denomination</th>
<th>Institution</th>
<th>Model Grid</th>
</tr>
</thead>
<tbody>
<tr>
<td>C4I-RCA3</td>
<td>Community Climate Change Consortium for Ireland (C4I), Met Éireann</td>
<td>Rotated lon/lat grid (0.22°)</td>
</tr>
<tr>
<td>CHMI-Aladin</td>
<td>Czech Hydrometeorological Institute (CHMI)</td>
<td>Lambert projection</td>
</tr>
<tr>
<td>CNRM-Aladin</td>
<td>Centre National de Recherches Météorologiques (CNRM), Météo-France</td>
<td>Lambert projection</td>
</tr>
<tr>
<td>DMI-HIRHAM</td>
<td>Danish Meteorological Institute (DMI)</td>
<td>Rotated lon/lat grid (0.22°)</td>
</tr>
<tr>
<td>EC-GEMLAM</td>
<td>Recherche en Prévision Numérique, Environment Canada (EC)</td>
<td>Rotated lon/lat grid (0.22°)</td>
</tr>
<tr>
<td>ETHZ-CLM</td>
<td>Swiss Federal Institute of Technology Zurich (ETHZ)</td>
<td>Rotated lon/lat grid (0.22°)</td>
</tr>
<tr>
<td>HC-HadRM3Q0</td>
<td>Hadley Centre (HC), UK Met Office</td>
<td>Rotated lon/lat grid (0.22°)</td>
</tr>
<tr>
<td>ICTP-RegCM</td>
<td>Abdus Salam International Centre for Theoretical Physics (ICTP), Trieste (Italy)</td>
<td>Lambert projection</td>
</tr>
<tr>
<td>INM-RCA3</td>
<td>Instituto Nacional de Meteorología (INM), Madrid (Spain)</td>
<td>Regular lon/lat grid (0.25°)</td>
</tr>
<tr>
<td>KNMI-RACMO</td>
<td>Royal Netherlands Meteorological Institute (KNMI)</td>
<td>Rotated lon/lat grid (0.22°)</td>
</tr>
<tr>
<td>METNO-HIRHAM</td>
<td>Norwegian Meteorological Institute (METNO)</td>
<td>Rotated lon/lat grid (0.22°)</td>
</tr>
<tr>
<td>MPI-REMO</td>
<td>Max-Planck-Institute for Meteorology (MPI), Hamburg (Germany)</td>
<td>Rotated lon/lat grid (0.22°)</td>
</tr>
<tr>
<td>OURANOS-CRCM</td>
<td>Consortium on Regional Climatology and Adaptation to Climate Change (OURANOS), Montreal (Canada)</td>
<td>Regular lon/lat grid (0.25°)</td>
</tr>
<tr>
<td>SMHI-RCA</td>
<td>Swedish Meteorological and Hydrological Institute (SMHI)</td>
<td>Rotated lon/lat grid (0.22°)</td>
</tr>
<tr>
<td>UCLM-PROMES</td>
<td>Universidad de Castilla La Mancha (UCLM), Toledo (Spain)</td>
<td>Lambert projection</td>
</tr>
</tbody>
</table>

*Model denomination is composed of the institution acronym followed by the model acronym.

*Grid resolution is indicated in brackets.
3.2 Regional Climate Model Data

This report evaluates fifteen RCMs from the ENSEMBLES project. More precisely, fifteen corresponding data sets of gridded daily precipitation totals from 40-year control simulations (CTRL) from 1961 to 2000. All experiments have been driven by boundary conditions from the ERA-40 re-analysis (Uppala et al, 2005), encompass the European continent, and have been operated at a spatial resolution of 25 km. Eight data sets were supplied on a rotated pole grid, matching that of E-OBS, with the other seven supplied on a variety of other grids. As a component of data pre-processing, the fields of these seven data sets have been bi-linearly interpolated onto the rotated pole grid. Another important step of model data preparation was to discard sea grid box values in all fifteen data sets by applying the E-OBS (land-only) land-sea mask. Table 3.1 provides an overview of the ENSEMBLES regional climate models, with model denominations used in this report, names of operating institutes, and some data grid characteristics.
4 Methodology

A basic approach to the evaluation of climate models consists of examining their ability to reproduce current climate. Here, reanalysis-driven control simulations (CTRL) of daily precipitation are evaluated against observations for the time period 1961-1990 that represents present-day climate conditions. Firstly, the characteristics of average and extreme precipitation will be analyzed in the reference data set of observations (E-OBS). Secondly, this analysis will be repeated with the CTRL integrations of all fifteen ENSEMBLES RCMs. The results of both steps will then be combined to identify model deficiencies and to assess model performance with regard to the representation of extreme precipitation events in the European Alps. Note that, although available as data sets for the time span from 1961 to 2000, RCM simulations will be evaluated for a 30-year time slice (1961-1990), since this is a conventional baseline period over which stationarity can be assumed (Frei et al, 2006). Furthermore, this report adopts key features of the methodology used in Frei et al (2006) that evaluated a previous generation of RCMs—the so-called PRUDENCE RCMs with regard to precipitation extremes in the European Alps. This facilitates a corresponding comparison between the ENSEMBLES and PRUDENCE RCMs, which is also a central objective of this report.

4.1 Climatological Diagnostics

The procedure of this report rests on a series of different climatological diagnostics that serve as statistical indices for model evaluation. A compilation of these quantities with acronyms and definitions is given in Table 4.1. The statistical analyzes, that is, the calculation of the diagnostics,

<table>
<thead>
<tr>
<th>Diagnostic</th>
<th>Definition</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>mea</td>
<td>mean precipitation</td>
<td>mm/d</td>
</tr>
<tr>
<td>fre</td>
<td>wet day frequency</td>
<td>fraction</td>
</tr>
<tr>
<td>int</td>
<td>wet day intensity, mean precipitation on wet days</td>
<td>mm/d</td>
</tr>
<tr>
<td>q90</td>
<td>empirical 90% quantile(^1) of wet day precipitation amounts</td>
<td>mm/d</td>
</tr>
<tr>
<td>x1d.TT</td>
<td>return value of 1-day precipitation intensity with return period TT=5, 10, 20, 50 years</td>
<td>mm/d</td>
</tr>
<tr>
<td>x5d.TT</td>
<td>return value of 5-day precipitation intensity(^2) with return period TT=5, 10, 20, 50 years</td>
<td>mm/d</td>
</tr>
</tbody>
</table>

Table 4.1: Diagnostic quantities of daily precipitation employed for model evaluation. Wet days are defined as days with precipitation ≥ 1mm.

\(^1\)5-day precipitation intensity results from applying a 5-day moving window average to 1-day precipitation intensity.
is performed independently for each model grid point, subdivided into seasons (winter, DJF; spring, MAM; summer, JJA; autumn, SON). In essence, two classes of diagnostics can be distinguished: those characterizing average or intense daily precipitation, termed basic diagnostics, and those reflecting strong precipitation events, termed extreme diagnostics. It is the latter to which this report devotes its main attention. Basic diagnostics are mainly used for comparison with and interpretation of extremes. As in the study of Frei et al (2006), a threshold of 1mm/d is introduced for some of the basic diagnostics in order to separate wet from dry days. Frei et al (2006) found that results are not sensitive to small changes in this threshold. The following list elaborates on the purpose of the individual diagnostics:

- **Basic diagnostics**
  - **mean precipitation (mea)**
    Quantification of climatological mean precipitation behaviour.
  - **wet day frequency (fre)**
    Quantification of average precipitation occurrence.
  - **wet day intensity (int)**
    Quantification of mean intensity of the precipitation process.
  - **empirical 90% quantile of wet day precipitation amounts (q90)**
    Quantification of precipitation intensity of “intense events”, that is, events related to the tail of the precipitation probability density function (pdf).

- **Extreme precipitation diagnostics**
  - **1-day precipitation intensity (x1d)**
    Assessment of extreme 1-day precipitation amounts, characteristic of heavy precipitation of short duration, and in general of convective formation. In the European Alps extreme events of that sort occur preferentially in summer (see Section 2.1.2). Focus will therefore be on return values of 1-day precipitation intensity in the analysis for extreme summer rainfall.
  - **5-day precipitation intensity (x5d)**
    Assessment of extreme 5-day precipitation amounts, typical of long-lasting heavy precipitation episodes caused by synoptic disturbances or persistent advection of moist air masses. In the study region such multi-day phenomena usually belong to the winter season (see Section 2.1.2). For this reason, focus will be on return values of 5-day precipitation intensity in the analysis for extreme winter rainfall.

### 4.2 Extreme Value Analysis

Return values of precipitation extremes with an average recurrence ranging from 5 to 50 years (Table 4.1) will be estimated using techniques from extreme value theory as follows:

1. Use of block maximum method: fitting of a generalized extreme value (GEV) distribution to seasonal maxima of 1-day and 5-day precipitation intensity and calculation of return values from the resulting distribution.
2. Estimation of GEV distribution parameters by method of maximum likelihood; use of a modified form of the GEV likelihood function, via incorporation of a geophysical Bayesian prior distribution for the GEV shape parameter.
Figure 4.1: Gumbel diagrams for four gridbox samples of daily precipitation extremes. Examples (a) and (b) refer to two adjacent grid boxes in southwestern Germany and the autumn season (SON), whereas examples (c) and (d) refer to two neighbouring grid boxes in eastern France and the winter season (DJF). Sample extremes (black dots) were simulated by the METNO-HIRHAM (examples (a) and (b)) and MPI-REMO (examples (c) and (d)) regional climate models. Fitted GEV distributions are shown for the case when estimated from the conventional likelihood function (pink curves) and for the case when estimated from the likelihood function with the geophysical prior (blue curves). Uncertainty of GEV estimates is indicated by asymptotic maximum likelihood 95% confidence bounds (dotted curves).
As in the study of Frei et al (2006), this report opts to take into account a prior distribution for the GEV shape parameter. This modification of the likelihood function was suggested by Martins and Stedinger (2000) in the context of geophysical applications. The proposed geophysical prior aims at minimizing the estimation of absurd values of the GEV shape parameter, a phenomenon that is often encountered when calculating the conventional maximum likelihood estimator from small samples of geophysical data (Hosking, 1985; Martins and Stedinger, 2000; Frei et al, 2006). The stabilizing influence of the geophysical prior clearly manifests itself in the spatial configuration of the shape parameter determined from grid box maxima: Frei et al (2006) demonstrated that including a prior into the classical GEV likelihood function markedly reduces the high gridbox to gridbox spatial variability of the shape estimates and leads to a much smoother distribution with virtually no anomalous values. From a physical point of view this is much more realistic.

The implications of the prior can be pursued further on the level of quantile estimates. Figure 4.1 depicts Gumbel diagrams for two samples of precipitation extremes from neighbouring ENSEMBLES model grid boxes in southwestern Germany (Figures 4.1a and 4.1b) as well as for two further samples from adjacent grid boxes in eastern France (Figures 4.1c and 4.1d). In each of the two examples of nearby grid boxes, but particularly in the second one, the use of the geophysical prior causes the respective GEV fits to approach each other, which is in line with the aforementioned smoothing of the spatial distribution of the shape parameter. From Figure 4.1 it follows, moreover, that the prior generally affects GEV quantiles at return period of 10 years and above.

4.3 Alpine Subregions

Evaluation of RCMs against observations comprises quantifying model biases in diagnostics relative to the observational reference (E-OBS). Relative biases are calculated individually for grid boxes and seasons, and the resulting spatial patterns are investigated for the European Alpine region (land-only). In addition to the grid box scale analysis, this report calculates areal means of all precipitation diagnostics across two specific Alpine subregions, and evaluates relative biases of RCM control simulations with respect to these values. The two subregions, illustrated in Figure 4.2, are the Northern Alps (NA) and the Southern Alps (SA). The definition and the use of these domains is motivated by the fact that the Alpine mountain range acts as a climatic divide, separating mid-latitude temperate and Mediterranean climate regimes (Schär et al, 1998). There is a pronounced gradient of moisture origin across this divide’s axis (Sodemann and Zübler, 2009); while in the northern Alps it is the North Atlantic and Central Europe that dominate as moisture sources, the southern Alps largely receive their moisture from the Mediterranean Sea. The latter differs from the North Atlantic moisture source by a higher variability and unsteadiness (Sodemann and Zübler, 2009). There is also the fact that air masses formed over the Mediterranean Sea have a higher moisture transport capacity than those arriving from the North Atlantic source region (Barry, 1992), which is a consequence of thermodynamic properties (warmer air has a higher moisture capacity according to the Clausius-Clapeyron relation). The NA and SA subregions share roughly the same number of grid boxes (NA: 285 grid boxes; SA: 287 grid boxes) and they bear a strong resemblance to those used in the studies of Frei et al (2006) and Sodemann and Zübler (2009). The work of Frei et al (2006) is of particular relevance here because it includes an evaluation of the PRUDENCE generation of RCMs with respect to the representation of extreme precipitation. Finally, a brief mention should be made of the southeastern French Alps, that are not attached to the NA subregion. It is true that they are situated west of the Alpine main ridge, but, due to their southern position and proximity to the Mediterranean Sea, they exhibit features of a Mediterranean climate.
4.4 Uncertainty Calculations

The analyses in this report are completed by the expression of uncertainty in the domain mean values of precipitation diagnostics. Assessment of uncertainty is carried out by calculating non-parametric bootstrap 90% confidence intervals. The procedure involves the following steps:

1. For a certain season, the 30-year time series of daily values (for mea, fre, int, q90) and of seasonal maxima (for x1d, x5d) respectively, are resampled for each grid box of an Alpine subregion. The condition is made that all grid boxes are sampled from the same days (for mea, fre, int, q90) and years (for x1d, x5d), respectively, to preserve the geophysical spatial correlation of the data (Wilks, 1997; Frei et al, 2006).

2. Diagnostics are then determined for each resampled grid box data set and averaged over the subdomain, which yields one bootstrap object.

3. Fifty bootstrap samples of domain mean diagnostics are generated by repeating steps 1 and 2.

4. Finally, it remains to calculate empirical 5% and 95% quantiles of the bootstrap sampling distributions to obtain the 90% confidence intervals.
5 Results

This chapter presents the results of the model evaluation and is structured into two parts. The first one is devoted to the spatial averages of basic and extreme precipitation diagnostics, the associated uncertainty, as well as the respective relative bias in model simulations (Section 5.1). The second part considers the spatial distribution of the model bias, with the focus on extreme precipitation diagnostics (Section 5.2).

5.1 Domain Means for Alpine Subregions

Figures 5.1 to 5.6 depict in seasonal order the basic and extreme (x1d.5, x1d.20) diagnostics averaged over the northern (NA, subfigures a) and southern (SA, subfigures b) Alpine subregion. A qualitative comparison (increase or decrease) of the simulated with the observed interseason variations reveals a generally reasonable agreement. Still, each RCM deviates from the observations for several of these quantities. With regard to individual diagnostics, it is found that less RCMs deviate in the NA than in the SA subregion. While mea is excluded from this finding, it holds most notably for int. In qualitative terms (increase or decrease) models reproduce the observed north-south gradient across the Alpine ridge fairly well. However, the majority of models exhibit deviations in summer: CNRM, HC, INM, KNMI, MPI and OURANOS as well as partly also C4I, DMI, EC and SMHI do not produce an increase for int, q90, x1d.5 and x1d.20 from north to south. Thus, the summertime intensity process south of the Alps during summer appears to be a significant source of error in models. As has already been observed by Frei et al (2006) for PRUDENCE models, extreme diagnostics (x1d.5, x1d.20) and int evoke comparison with respect to intermodel pattern and seasonal variation of biases, which is not the case between extreme diagnostics (x1d.5, x1d.20) and fre. This implies that model inadequacies in the representation of extreme precipitation might be ascribed in the first place to failings in the intensity process.

In addition to domain mean values, Figures 5.1 to 5.6 depict the pertaining 90% confidence intervals, indicating that uncertainty increases from mea and q90 up to extreme precipitation diagnostics. Superimposed on this is a tendency of the uncertainty to increase from the NA to the SA subregion, which is the result of a higher precipitation variability south of the Alps (see Section 2.1). In the SA subregion uncertainty tends to be largest in autumn (SON). By contrast, no corresponding season can be identified in the NA subregion. Finally, an important aspect to point out is that a host of model biases exceeds confidence bounds of observations, suggesting that the considered model biases are not directly related to random errors as a consequence of the limited integration period (cf. Frei et al, 2006).

Figure 5.7 gives, in the form of boxplot diagrams, a comprehensive overview of the relative bias of models in domain mean diagnostics and Tables 5.1 (basic diagnostics) and 5.2 (extreme diagnostics) list the corresponding numbers separately for individual models:
Figure 5.1: Domain mean values (symbols) and 90% bootstrap confidence intervals (vertical lines) of mean precipitation (mea, mm/d) for the Northern Alps (NA, Figure 5.1a) and Southern Alps (SA, Figure 5.1b) subregion. Results are depicted for ENSEMBLES regional climate models (black) and observations (blue (NA), orange (SA)).
Figure 5.2: Domain mean values (symbols) and 90% bootstrap confidence intervals (vertical lines) of wet day frequency (fre, fraction) for the Northern Alps (NA, Figure 5.2a) and Southern Alps (SA, Figure 5.2b) subregion. Results are depicted for ENSEMBLES regional climate models (black) and observations (blue (NA), orange (SA)).
Figure 5.3: Domain mean values (symbols) and 90% bootstrap confidence intervals (vertical lines) of wet day intensity (int, mm/d) for the Northern Alps (NA, Figure 5.3a) and Southern Alps (SA, Figure 5.3b) subregion. Results are depicted for ENSEMBLES regional climate models (black) and observations (blue (NA), orange (SA)).
Figure 5.4: Domain mean values (symbols) and 90% bootstrap confidence intervals (vertical lines) of empirical 90% quantile of wet day precipitation amounts (q90, mm/d) for the Northern Alps (NA, Figure 5.4a) and Southern Alps (SA, Figure 5.4b) subregion. Results are depicted for ENSEMBLES regional climate models (black) and observations (blue (NA), orange (SA)).
Figure 5.5: Domain mean values (symbols) and 90% bootstrap confidence intervals (vertical lines) of 5-year return value of 1-day precipitation extreme (x1d.5, mm/d) for the Northern Alps (NA, Figure 5.5a) and Southern Alps (SA, Figure 5.5b) subregion. Results are depicted for ENSEMBLES regional climate models (black) and observations (blue (NA), orange (SA)).
Figure 5.6: Domain mean values (symbols) and 90% bootstrap confidence intervals (vertical lines) of 20-year return value of 1-day precipitation extreme (x1d.20, mm/d) for the Northern Alps (NA, Figure 5.6a) and Southern Alps (SA, Figure 5.6b) subregion. Results are depicted for ENSEMBLES regional climate models (black) and observations (blue (NA), orange (SA)).
mea/fre/int in NA subregion (Figure 5.7a) In **winter** and **spring** positive biases in general predominate. Extremely (partly also exceptionally) high positive biases in mea are observed for DMI (+85% in DJF, +49% in MAM) and ICTP (+83% in DJF, +58% in MAM). These biases originate both from an overestimation of fre and int. While DMI shows larger deficiencies in the precipitation intensity process, deficiencies of ICTP are localized more in the precipitation occurrence process. The only models with significant negative biases for mea and fre are CNRM (-33% / -34% for mea / fre in DJF) and UCLM (-12% / -18% for fre / int in MAM). To characterize performance of models in **summer** they are assigned to four distinct categories: Category 1 (EC, OURANOS, UCLM) - small positive bias in mea, mutual compensation of bias in fre and int, Category 2 (CHMI, CNRM, ICTP) - positive bias in mea, partial compensation of positive bias in fre by negative bias in int, Category 3 (HC) - positive bias in mea, combination of positive bias in fre and int, Category 4 (C4I, DMI, ETHZ, INM, KNMI, METNO, MPI, SMHI) - small to moderate negative or positive bias in mea, inclusion of precipitation frequency distribution necessary for a plausible interpretation of bias in mea (cf. Frei et al, 2003). In **autumn** model bias varies between -10% and +27% for mea and between -7% and +28% for fre. An exception to this is CNRM. As in winter, it shows significant negative biases in mea (-42%) and fre (-29%).

mea/fre/int in SA subregion (Figure 5.7b) In **winter** and **spring** most models show small to large positive biases for mea and fre. Some of the models underestimate int that, however, is (almost) compensated or overcompensated by the overestimation of fre. For the description of their performance in **summer** models are assigned to four distinct categories: Category 1 (ETHZ, ICTP) – small bias in mea, mutual compensation of bias in fre and int, Category 2 (CHMI, CNRM, OURANOS) – positive bias in mea, partial compensation of extremely large positive bias in fre by negative bias in int, Category 3 (EC, KNMI, UCLM) – positive (negative) bias in mea, combination of positive (negative) bias in fre and int, Category 4 (C4I, DMI, HC, INM, METNO, MPI, SMHI) – small to moderate negative or positive bias in mea, reference to differences to the observed precipitation frequency distribution necessary for a plausible interpretation of bias in mea (cf. Frei et al, 2003). In **autumn** C4I, DMI, ETHZ, HC, INM and SMHI all show a good performance for mea, which is the consequence of a mutual (partial) compensation of biases in fre and int. As in the NA subregion, CNRM strongly underestimates mea in autumn due to large negative biases in fre and int.

def in NA and SA subregion (Figures 5.7a-b) Distribution of relative bias (as shown by boxplots) and corresponding numbers for individual models are very similar between def and int, which is in line with the aforementioned observation that model errors in extreme precipitation primarily depend on shortcomings in the intensity process. The most noticeable feature in the SA subregion is that nearly all models underestimate def and int in summer. The only exceptions with respect to def are ETHZ (+5%), METNO (+5%) and UCLM (+12%) and in case of int ETHZ (+3%) and UCLM (+5%).

Extreme diagnostics (Figures 5.7c-f) With regard to extreme diagnostics there is generally a tendency for relative bias to increase from small (5-year) to large (50-year) return values, and a somewhat less pronounced tendency for the intermodal variability to increase also. At the same time there is a tendency, especially for x1d, that the model bias decreases / increases from north to south in summer / autumn. An increase of the intermodal variability from north to south can be observed for x1d in all seasons. For x5d this increase is confined to the transition seasons. In the **NA subregion** most models show positive biases. In the case of x1d models tend to have their largest bias in summer and, somewhat less
Figure 5.7: Boxplot diagrams for relative bias (%) of ENSEMBLES regional climate models in domain mean basic (mea, fre, int, q90; panels a-b) and extreme precipitation (x1d, x5d; panels c-f) diagnostics. Boxplots are presented for the Northern Alps (panels a, c and e) and Southern Alps (panels b, d and f) subregion. Each boxplot displays five sample quantiles of a corresponding data group: sample minimum (lower end of lower whisker), lower quartile (lower end of box), median (bar inside box), upper quartile (upper end of box), sample maximum (upper end of upper whisker). Data values being far removed from the others in the data group are indicated as outliers (black dots).
Table 5.1: Model bias in domain mean of basic precipitation diagnostics for the Alpine subregions, expressed relative to (that is in percent) of observational data.

<table>
<thead>
<tr>
<th></th>
<th>DEF</th>
<th>MAM</th>
<th>JJA</th>
<th>SON</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>x18.5</td>
<td>x23.2</td>
<td>x18.5</td>
<td>x23.2</td>
</tr>
<tr>
<td></td>
<td>x18.5</td>
<td>x23.2</td>
<td>x18.5</td>
<td>x23.2</td>
</tr>
<tr>
<td></td>
<td>x18.5</td>
<td>x23.2</td>
<td>x18.5</td>
<td>x23.2</td>
</tr>
<tr>
<td></td>
<td>x18.5</td>
<td>x23.2</td>
<td>x18.5</td>
<td>x23.2</td>
</tr>
<tr>
<td></td>
<td>x18.5</td>
<td>x23.2</td>
<td>x18.5</td>
<td>x23.2</td>
</tr>
<tr>
<td></td>
<td>x18.5</td>
<td>x23.2</td>
<td>x18.5</td>
<td>x23.2</td>
</tr>
<tr>
<td></td>
<td>x18.5</td>
<td>x23.2</td>
<td>x18.5</td>
<td>x23.2</td>
</tr>
<tr>
<td></td>
<td>x18.5</td>
<td>x23.2</td>
<td>x18.5</td>
<td>x23.2</td>
</tr>
<tr>
<td></td>
<td>x18.5</td>
<td>x23.2</td>
<td>x18.5</td>
<td>x23.2</td>
</tr>
<tr>
<td></td>
<td>x18.5</td>
<td>x23.2</td>
<td>x18.5</td>
<td>x23.2</td>
</tr>
<tr>
<td></td>
<td>x18.5</td>
<td>x23.2</td>
<td>x18.5</td>
<td>x23.2</td>
</tr>
<tr>
<td></td>
<td>x18.5</td>
<td>x23.2</td>
<td>x18.5</td>
<td>x23.2</td>
</tr>
<tr>
<td></td>
<td>x18.5</td>
<td>x23.2</td>
<td>x18.5</td>
<td>x23.2</td>
</tr>
<tr>
<td></td>
<td>x18.5</td>
<td>x23.2</td>
<td>x18.5</td>
<td>x23.2</td>
</tr>
<tr>
<td></td>
<td>x18.5</td>
<td>x23.2</td>
<td>x18.5</td>
<td>x23.2</td>
</tr>
<tr>
<td></td>
<td>x18.5</td>
<td>x23.2</td>
<td>x18.5</td>
<td>x23.2</td>
</tr>
<tr>
<td></td>
<td>x18.5</td>
<td>x23.2</td>
<td>x18.5</td>
<td>x23.2</td>
</tr>
<tr>
<td></td>
<td>x18.5</td>
<td>x23.2</td>
<td>x18.5</td>
<td>x23.2</td>
</tr>
<tr>
<td></td>
<td>x18.5</td>
<td>x23.2</td>
<td>x18.5</td>
<td>x23.2</td>
</tr>
<tr>
<td></td>
<td>x18.5</td>
<td>x23.2</td>
<td>x18.5</td>
<td>x23.2</td>
</tr>
<tr>
<td></td>
<td>x18.5</td>
<td>x23.2</td>
<td>x18.5</td>
<td>x23.2</td>
</tr>
<tr>
<td></td>
<td>x18.5</td>
<td>x23.2</td>
<td>x18.5</td>
<td>x23.2</td>
</tr>
<tr>
<td></td>
<td>x18.5</td>
<td>x23.2</td>
<td>x18.5</td>
<td>x23.2</td>
</tr>
</tbody>
</table>

Table 5.2: Model bias in domain mean of extreme precipitation diagnostics for the Alpine subregions, expressed relative to (that is in percent) of observational data.

<table>
<thead>
<tr>
<th></th>
<th>DEF</th>
<th>MAM</th>
<th>JJA</th>
<th>SON</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>x18.5</td>
<td>x23.2</td>
<td>x18.5</td>
<td>x23.2</td>
</tr>
<tr>
<td></td>
<td>x18.5</td>
<td>x23.2</td>
<td>x18.5</td>
<td>x23.2</td>
</tr>
<tr>
<td></td>
<td>x18.5</td>
<td>x23.2</td>
<td>x18.5</td>
<td>x23.2</td>
</tr>
<tr>
<td></td>
<td>x18.5</td>
<td>x23.2</td>
<td>x18.5</td>
<td>x23.2</td>
</tr>
<tr>
<td></td>
<td>x18.5</td>
<td>x23.2</td>
<td>x18.5</td>
<td>x23.2</td>
</tr>
<tr>
<td></td>
<td>x18.5</td>
<td>x23.2</td>
<td>x18.5</td>
<td>x23.2</td>
</tr>
<tr>
<td></td>
<td>x18.5</td>
<td>x23.2</td>
<td>x18.5</td>
<td>x23.2</td>
</tr>
<tr>
<td></td>
<td>x18.5</td>
<td>x23.2</td>
<td>x18.5</td>
<td>x23.2</td>
</tr>
<tr>
<td></td>
<td>x18.5</td>
<td>x23.2</td>
<td>x18.5</td>
<td>x23.2</td>
</tr>
<tr>
<td></td>
<td>x18.5</td>
<td>x23.2</td>
<td>x18.5</td>
<td>x23.2</td>
</tr>
<tr>
<td></td>
<td>x18.5</td>
<td>x23.2</td>
<td>x18.5</td>
<td>x23.2</td>
</tr>
<tr>
<td></td>
<td>x18.5</td>
<td>x23.2</td>
<td>x18.5</td>
<td>x23.2</td>
</tr>
<tr>
<td></td>
<td>x18.5</td>
<td>x23.2</td>
<td>x18.5</td>
<td>x23.2</td>
</tr>
<tr>
<td></td>
<td>x18.5</td>
<td>x23.2</td>
<td>x18.5</td>
<td>x23.2</td>
</tr>
<tr>
<td></td>
<td>x18.5</td>
<td>x23.2</td>
<td>x18.5</td>
<td>x23.2</td>
</tr>
<tr>
<td></td>
<td>x18.5</td>
<td>x23.2</td>
<td>x18.5</td>
<td>x23.2</td>
</tr>
<tr>
<td></td>
<td>x18.5</td>
<td>x23.2</td>
<td>x18.5</td>
<td>x23.2</td>
</tr>
<tr>
<td></td>
<td>x18.5</td>
<td>x23.2</td>
<td>x18.5</td>
<td>x23.2</td>
</tr>
<tr>
<td></td>
<td>x18.5</td>
<td>x23.2</td>
<td>x18.5</td>
<td>x23.2</td>
</tr>
<tr>
<td></td>
<td>x18.5</td>
<td>x23.2</td>
<td>x18.5</td>
<td>x23.2</td>
</tr>
<tr>
<td></td>
<td>x18.5</td>
<td>x23.2</td>
<td>x18.5</td>
<td>x23.2</td>
</tr>
</tbody>
</table>

30
pronounced, in winter. On the other hand, there is a tendency for models to have their largest bias in winter and spring with respect to x5d. These maxima have a clear climatological background: north of the Alps intense precipitation events of short duration occur most frequently during summer, whereas long-lasting strong rainfall episodes are most common in winter and spring (see Section 2.1.2). In the SA subregion most models also exhibit moderate to large positive biases in winter, spring and autumn. As in the case of int and q90, the distribution of model biases in summer clearly differs from those in the other seasons. A first group of models (CNRM, INM, OURANOS) has large negative biases, a second and at the same time the largest group (C4I, CHMI, DMI, EC, HC, ICTP, KNMI, MPI, SMHI) shows relatively small biases, while a third group (ETHZ, METNO, UCLM) has large positive biases. It is interesting to note that biases in int vary between -37\% and -22\% with the first group, between -24\% and -7\% with the second and between -3\% and +5\% with the third group. Remember that models of the third group have been specially mentioned in the case of summertime q90, since they are the only ones with positive bias in this diagnostic. Two special cases with regard to both Alpine subregions are DMI and OURANOS. The latter model systematically shows large positive biases (exception: JJA in NA and SA). DMI and OURANOS both suffer from significant deficiencies in the intensity process (over- and underestimation, respectively).

5.2 Spatial Distribution of Relative Bias

The two typical forms of extreme precipitation in the Alpine region, relatively short-lived events of convective origin and long-lasting episodes associated with synoptic-scale disturbances, occur with particular frequency in summer and winter, respectively (see Section 2.1.2). It is for this reason that the spatial analysis of the relative bias concentrates on wintertime x5d (Section 5.2.1) and summertime x1d (Section 5.2.2). In addition a detailed look is taken at the spatial bias pattern of x1d in autumn (Section 5.2.3). This season coincides with a dynamically particularly active phase in the climate of the Southern Alpine region, with the frequency of heavy precipitation reaching its maximum (see Section 2.1.2). In each of the three considered seasons, spatial analysis is extended to the relative bias in fre and int, in order to elucidate the quality of the simulated precipitation occurrence and intensity process also from a spatial perspective. Both processes, especially the latter, are potential sources of error for the representation of extreme precipitation.

5.2.1 Extreme Precipitation in Winter x5d.5 (Figure 5.8) An essential common feature of models is the overestimation of the 5-year return value of 5-day precipitation extremes in the northern (exceptions: CNRM, INM) and southern Alpine rim zone as well as along the northwestern rim of the Jura Mountains (exception: CNRM). On the other hand, models show either quite realistic or too low return values in a narrow stripe of inner-alpine areas and in the Swiss Plateau, that is, to the lee of the Alpine marginal chains and the Jura Mountains, respectively. In the Po Valley, the majority of models are found to have low to large negative biases. It is interesting to note that there is a remarkable similarity between the spatial bias distributions of C4I, INM, and SMHI, all sharing the same dynamical core (RCA). Both CHMI and CNRM as well as DMI and METNO also use an equivalent dynamical core (Aladin and HIRHAM, respectively). However, a resemblance regarding the spatial error structure is discernible only to a much lesser extent, especially between CHMI and CNRM.
Figure 5.8: Visualization of relative bias of ENSEMBLES regional climate models in 5-year return value of 5-day precipitation extreme ($\Delta 5d.5$) in winter (DJF), expressed in percent of observations (E-OBS, mm/day, top left panel) and calculated for the reference period 1961-1990.
Figure 5.8: (continued) Visualization of relative bias of ENSEMBLES regional climate models in 5-year return value of 5-day precipitation extreme (x5d.5) in winter (DJF), expressed in percent of observations (E-OBS, mm/d, top left panel on previous page) and calculated for the reference period 1961-1990.
Figure 5.9: Visualization of relative bias of ENSEMBLES regional climate models in 5-year return value of 1-day precipitation extreme (x1d.5) in summer (JJA), expressed in percent of observations (E-OBS, mm/d, top left panel) and calculated for the reference period 1961-1990.
Figure 5.9: (continued) Visualization of relative bias of ENSEMBLES regional climate models in 5-year return value of 1-day precipitation extreme (x1d.5) in summer (JJA), expressed in percent of observations (E-OBS, mm/d, top left panel on previous page) and calculated for the reference period 1961-1990.
Figure 5.10: Visualization of relative bias of ENSEMBLES regional climate models in 5-year return value of 1-day precipitation extreme (x1d.5) in autumn (SON), expressed in percent of observations (E-OBS, mm/d, top left panel) and calculated for the reference period 1961-1990.
Figure 5.10: (continued) Visualization of relative bias of ENSEMBLES regional climate models in 5-year return value of 1-day precipitation extreme (x1d.5) in autumn (SON), expressed in percent of observations (E-OBS, mm/d, top left panel on previous page) and calculated for the reference period 1961-1990.
Over the Alpine range and the Jura Mountains, all models (exception: CNRM) tend to overestimate wet-day frequency. Conversely, models exhibit too low wet-day frequency over the northwestern Italian Alps and the Ticino (exception: CNRM) as well as in the Po Valley (exceptions: EC, ICTP, OURANOS).

As for x5d.5 all models show positive biases in the northern (exceptions: INM, OURANOS) and southern (exception: OURANOS) rim zone of the Alpine range and along the northwestern rim of the Jura Mountains. In the inner-alpine domain and the Swiss Plateau models exhibit quite realistic values or large negative biases, again as for x5d.5. Performance of the simulated precipitation intensity over the Po Valley varies from model to model, with the magnitude of the relative bias being generally limited.

5.2.2 Extreme Precipitation in Summer

Models can be subdivided into three groups of similar spatial bias distribution. **Group 1** (C4I, CHMI, CNRM, DMI, EC, HC, INM, KNMI, MPI): An irregular pattern of negative and positive biases is found over the Alpine chain, with some of the models having a slight tendency towards positive/negative biases along the mountain rim/inner-alpine zone. While return values in general are too high north and northwest of the Alps, models tend to underestimate them in the Mediterranean region and in the Po Valley. **Group 2** (ETHZ, ICTP, METNO, UCLM): Apart from the inner-alpine zone and the main crest of the Apennine, positive biases dominate the entire domain. **Group 3** (OURANOS): There is an area-wide underestimation of return values, most pronounced over the main ridge of the Alps, the Apennine and the Dinaric Alps.

Two groups of models with similar spatial bias distribution can be formed. **Group 1** (C4I, DMI, EC, ETHZ, HC, INM, KNMI, METNO, MPI, SMHI, UCLM): Over the Alpine range models in general exhibit an irregular pattern of positive and isolated negative biases, with maximum values tending to lie over the southwestern Alps and south of western Austria. Conversely, they generally show negative biases north and east of the Alps and in the Po and Rhone Valley. **Group 2** (CHMI, CNRM, ICTP, OURANOS): There is a considerable large-scale overestimation of wet day frequency by models. In southern France and Italy, relative bias exceeds the value of +50% across the whole area. Deviating from the other models, ICTP quite realistically reproduces wet day frequency in Austria, southern Germany, eastern Switzerland and northernmost Italy.

Models can be subdivided into two groups of similar spatial error structure. **Group 1** (CHMI, CNRM, OURANOS): Precipitation intensity is widely underestimated, in case of CHMI and CNRM most pronouncedly in the Mediterranean region and in case of OURANOS most markedly over the main ridge of the Alps, the Apennine and the Dinaric Alps. **Group 2** (C4I, DMI, EC, ETHZ, HC, ICTP, INM, KNMI, METNO, MPI, SMHI, UCLM): Over the Alpine mountain range, a first model subgroup (C4I, DMI, HC, INM, KNMI, METNO, SMHI) shows an irregular pattern of positive and negative biases, a second subgroup (EC, ICTP) exhibits predominantly negative biases, and a third subgroup (ETHZ, MPI, UCLM) shows positive/negative biases in marginal/inner areas. In the Mediterranean region and in the Po Valley, a first model subgroup (C4I, DMI, HC, INM, KNMI, MPI, SMHI) underestimates precipitation intensity, while a second subgroup (EC, ETHZ, ICTP, METNO, UCLM) shows positive or negative but, in general, small biases. North and northwest of the Alps, models of Group 2 overall depict quite realistic values of precipitation intensity.
Figure 5.11: The mean autumn (SON) 5-year return value of 1-day precipitation intensity (x10.5, mm/d) for ENSEMBLES regional climate models and observations (E-OBS, top left panel), calculated for the reference period 1961-1990.
Figure 5.11: (continued) The mean autumn (SON) 5-year return value of 1-day precipitation intensity (x1d.5, mm/d) for ENSEMBLES regional climate models and observations (E-OBS, top left panel on previous page), calculated for the reference period 1961-1990.
In summer a large number of models have a spatial error distribution in common that reveals an irregular, randomly appearing pattern over the Alpine mountain ridge. This applies to all diagnostics considered and is likely associated with the parameterization of convective precipitation, whose activity over the Alpine topography is largest during the summer season.

5.2.3 Extreme Precipitation in Autumn

**x1d.5 (Figure 5.10)** There is a clear correspondence between the spatial bias distribution for x1d.5 in autumn and that observed for wintertime x5d.5 (see Section 5.2.1). Depending on model, however, magnitude and spatial extent of the southern rim pattern in x1d.5 (SON) is either less or more pronounced. Moreover, with most models, the positive bias pattern along the northern Alpine rim is distinctly less pronounced in autumn.

In order to gain a direct insight into the models’ ability to reproduce mesoscale structures of observed precipitation extremes, Figure 5.11 depicts directly – as a complement to Figure 5.10 – the spatial distribution of x1d.5 in SON. Over low mountain ranges such as the Jura, Black Forest or Vosges, the observations (Figure 5.11, top left panel) reveal distinct regional patterns that are related to mechanisms of orographic precipitation. In almost all model simulations these features can be clearly recognized. Yet, there are significant deviations in terms of magnitude, spatial extent and location. For example, the narrow ribbon over the Jura mountains is shifted to the west and overemphasized in many models. Furthermore, the observations reveal two prominent mesoscale patterns that are, in addition to orographic precipitation mechanisms, related to channeling effects on moist airflows and thus more complex. One is located over Southern Switzerland (Ticino) and the other has its position in the area where Austria, Italy and Slovenia meet. Nearly all models are found to reproduce these structures. In spite of that there are major deviations that in general involve a shift towards the west and, especially in case of the pattern over Southern Switzerland, an extension along the mountain rim to the southwest. Finally, a brief mention should be made of METNO that is strikingly different from the other models with respect to the fine-scale spatial variability. Though simulating a reasonable meso-scale distribution of x1d.5, in fact, it exhibits enormous fluctuations from gridbox to gridbox, indicating poor skill to benefit from the relatively high model resolution of 25 km.

**fre** The spatial bias distribution for wet day frequency in autumn bears close resemblance to that observed in wintertime (see Section 5.2.1). Having said that, positive/negative bias patterns in general appear weakened/reinforced in autumn.

**int** Again, striking parallels can be detected between the spatial bias distribution for wet day intensity in autumn and that seen in winter (see Section 5.2.1). In autumn positive/negative biases, in particular those along the northern Alpine rim, tend to be diminished/enhanced.
6 Conclusions

Assessment of how accurately a climate model represents reality provides essential information on its reliability to predict future climate, but also stimulates the continuous process of model development and improvement. Against this background, the present report has undertaken a comprehensive evaluation of average and extreme daily precipitation for the present-day (1961-1990) climate as simulated by fifteen reanalysis driven RCMs from the European Commission’s ENSEMBLES project. The E-OBS gridded dataset, designed within the framework of ENSEMBLES, served as observational reference for this purpose.

All models show -depending on precipitation diagnostic, subregion and season -both convincing performance but also substantial biases. With regard to basic diagnostics (mean precipitation, wet day frequency, precipitation intensity, 90% quantile of wet day precipitation), the evaluation identified KNMI-RACMO as the model that achieves the best agreement with the observations. By contrast, models CNRM-Aladin, DMI-HIRHAM, and OURANOS-CRCM exhibited a considerable number of anomalous deviations (Figs. 5.1.-5.6).

In winter-and springtime, there is a general tendency across models that predominantly positive biases in mean precipitation are larger than the corresponding biases in the other basic diagnostics (Fig. 5.7a,b). This wet model bias essentially comprises the combined overestimation of wet day frequency and precipitation intensity. In the Northern Alpine subregion, performance for mean precipitation varies exceptionally among models in wintertime (from -33% up to +85%). Overall, however, intermodel variability for basic diagnostics tends to be smaller in the Northern Alps than in the Southern Alps. Comparison of biases in the two subregions further revealed a tendency for larger errors in wintertime basic diagnostics in the Northern Alps. More precisely, in winter the prevailing overestimates north of the Alps in general exceed the over- or underestimates south of the Alps. It is somewhat different in summer, when the majority of models have larger biases in basic diagnostics in the Southern Alps. In particular, wet day frequency is overestimated to a larger degree and more consistently, whereas precipitation intensity and the 90% quantile of wet day precipitation is underestimated to a larger degree and more consistently.

A majority of models were found to display a lower degree of performance in extreme diagnostics (5-year return values) than in basic precipitation diagnostics (Fig. 5.7c-f). This, however, does not hold in summer of the Southern Alpine subregion (with regard to all basic diagnostics) as well as in winter and spring of the Northern Alpine subregion (restricted to mean precipitation). The tendency for larger biases in extreme diagnostics is surprising in so far as the evaluation of a previous generation of GCM-driven RCMs, the so-called PRUDENCE models (Frei et al, 2006), did not find evidence for it. Both this report and Frei et al (2006) used the same procedure to estimate return values, so that the performance discrepancy between basic and extreme diagnostics is likely not to be connected to methodical aspects. Thus, all in all, the evaluation in this report indicates model deficiencies that have repercussions specifically on precipitation extremes.

In winter, when the occurrence of precipitation extremes in the Alpine region is primarily associated with persistent large-scale advection of moist air, models produce typical spatial error patterns with regard to return values of 5-day precipitation (Fig. 5.8): positive biases along and over mountain rims (northern and southern Alpine rim, northwestern rim of Jura Mountains), negative biases or realistic values in the inner-alpine zone, the Po Valley and partly in the Swiss Plateau and over the Jura Mountains. For most models, the representation of the precipitation intensity process led to strikingly similar spatial bias patterns. This is therefore reasonable to assume that the following two model mechanisms substantially contribute to the prominent features in the spatial bias distribution for wintertime precipitation extremes: overestimation of topographic precipitation enhancement along and over mountain rims, overestimation of rain shadowing in the lee of mountains (Alps, Apennine, Jura Mountains) and in inner-alpine regions. In other words, the spatial bias patterns signal the models’ inability to correctly reproduce the over spill of precipitation over a topographic barrier. In autumn, the models’ overestimate of precipitation return values becomes also particularly apparent along the southern Alpine rim (Fig. 5.10-5.11). At this time of the year, precipitation extremes south of the Alps are again characterized by pronounced large-scale moisture advection. In this context, mention should be made of the study by Verbunt et al (2007), that demonstrated that incorporating a prognostic precipitation scheme (advection of hydrometeors, falling with typical fall velocities, by the 3D wind field) into a high resolution LAM brings about a remarkable improvement in the spatial precipitation distribution over complex topography.
A previously unresolved general problem of RCMs in summer has again been encountered in this report: most models underestimate precipitation intensity in the Southern Alpine region and in the Mediterranean (Fig. 5.9). This reflects in dry biases for the 90% quantile of wet day precipitation and counteracts the general tendency of models towards moderate to large wet biases for rare extremes. Studies provide evidence that the deficiencies in the summertime intensity process are likely associated with failings in the parameterization for convection (Hagemann et al, 2001) and with an inadequate representation of water and energy balance processes, in particular with the overestimation of the summertime decline in soil moisture (Noguer et al, 1998; Hagemann et al, 2004; Hirschi et al, 2005; Hirschi et al, 2006).

Qualitatively comparing against PRUDENCE RCMs (Frei et al 2006), in winter and summer the performance of the ENSEMBLES RCMs for mean precipitation and for rare extremes (5-year return value of 1-day precipitation) was found to differ, considered in aggregate, in two ways. 1) Northern Alpine region (winter, summer), Southern Alpine region (winter): overestimates are more common among ENSEMBLES RCMs and, especially for rare extremes, tend to be larger. 2) Southern Alpine region (summer): underestimates are less common among ENSEMBLE RCMs and tend to be smaller. The first difference in general stems mainly from an increased intensity process, whereas the second appears to depend primarily on an increased frequency process. Given that the ENSEMBLES RCMs represent a newer generation of models and are driven by reanalysis data, it is somewhat surprising that no general improvement has been found in their representation of average and extreme precipitation by comparison with the GCM driven PRUDENCE RCMs. A possible explanation for this might be the gridded observational dataset used. Indeed, E-OBS – the standard observational reference of the ENSEMBLES project and consequently of the present evaluation – has a lower spatial density of contributing observing stations over the Alpine region as compared to the Frei and Schär (1998) observational dataset used for evaluating the PRUDENCE RCM. If only because of the elevated and complex topography of the Alps, this fact may have clear undesirable effects on our evaluation. Thus to assess the robustness of our results, one could quantitative compare these two observational datasets.
Acknowledgements

We are grateful to Christoph Frei from MeteoSwiss for early discussions and providing R software packages for extreme value analysis and gridded data plotting. Acknowledgements also to Urs Beyerle and DaniLüthi for continuous technical assistance. A special mention also to Olivia Romppainen and Marc Wüest, for advice and assistance with project management aspects.
7 References


Beniston M. (2006), Mountain weather and climate. A general overview and a focus on climatic change in the Alps. Hydrobiologia, 562, 3-16.


Giorgi F. (2008), Regionalization of climate change information for impact assessment and adaption. WMO Bulletin, 57 (2).


Hagemann S. et al (2001), The summer drying problem over south-eastern Europe: Sensitivity of the limited area model HIRHAM4 to improvements in physical parameterization and resolution, Physics and Chemistry of the Earth, Part B, 26, 391-396.


