Danish Meteorological Institute
Ministry of Transport

Danish Climate Centre Report 07-03

Analysis of temporal changes in precipitation intensities using PRUDENCE data

Fredrik Boberg, Peter Berg, Peter Thejll and Jens Hesselbjerg Christensen

Colophone

Serial title:
Danish Climate Centre Report 07-03

Title:
Analysis of temporal changes in precipitation intensities using PRUDENCE data

Subtitle:

Authors:
Fredrik Boberg, Peter Berg, Peter Thejll and Jens Hesselbjerg Christensen

Other Contributers:

Responsible Institution:
Danish Meteorological Institute

Language:
English

Keywords:
Precipitation, climate change, PRUDENCE, regional climate modelling, probability density function, robustness test

Url:
http://prudence.dmi.dk/

ISSN:
1399-1957

ISBN:
978-87-7478-552-1

Version:
1

Website:
www.dmi.dk

Contents

1. Dansk resumé

Vi har sammenlignet et ensemble af regionale klimamodel-simuleringer fra EU rammeprogrammet PRUDENCE med daglige observerede nedbørsmængder fra European Climate Assessment, ved at karakterisere nedbøren med sandsynlighedsfordelingsfunktioner af forskellig art. Vi sammenligner fordeling af nedbørsdage og tørre dage i 8 regioner i Europa mod modellerne. Vi finder den model der er bedst til at beskrive nedbøren i hver region. Vi viser hvordan man estimerer resultaternes robusthed.

2. Abstract

We have compared an ensemble of regional climate modelling simulations from the European framework project PRUDENCE with observed daily precipitation from the European Climate Assessment dataset by characterising precipitation in terms of probability density functions of various kinds. We compare distributions for wet days and consecutive dry days across European sub-regions against the models. We find the models that best describe the observations in given regions as well as across regions. We show how to estimate robustness of the results for wet days and climatological drought days using bootstrapping with replacement on individual station observations time series. When comparing the scenario period (2071-2100) with the control run (1961-1990) for the models we see that the precipitation distribution changes for the scenario period with more days with intense precipitation and fewer days with light precipitation. The amplitude of this distribution change increase with time although the crossover point from less light precipitation to more intense precipitation does not show any significant change.

3. Introduction

As climate changes under human influence, a need has arisen for tools to estimate the nature and amplitude of future changes. Climate models stand as the central tool for this task, but model estimates of future change will not be useful if they are biased or unrealistic. We must therefore find ways to determine which models are best at simulating the atmospheric physical system, from a wide range of possibilities offered. A practical problem for this task is that current model and hardware capabilities are not up to modelling regional climate in a fully global climate model simulation. Rather, AGCMs are used to provide low resolution boundary conditions for a region inside which a detailed regional climate model (RCM) is run (Christensen et al., 1998). This gives much higher resolution in the model output and the possibility of using more complex and detailed model physics. PRUDENCE (Christensen et al., 2007) is a European-scale study aiming at quantifying the uncertainties in predictions of future climate and its impacts, using an array of climate models and impact models and expert judgement on their performance. Furthermore, the study will focus on interpreting these results in relation to European policies for adapting to or mitigating climate change. Here in this study we will use this array of PRUDENCE climate models to characterise precipitation and drought conditions in future climate scenarios.

We must find ways to reliably determine which models perform best based on a comparison with observed climate before using the identified model to predict future climate. We will describe here our work to do so with a method based on statistical metrics evolved specially for the task. The nature of the metrics used to inter-compare models and reality must take into account the properties of precipitation statistics. This varies from region to region based on the type of climate. Some regions may have long periods of drought interspersed with heavy showers while others may be dominated by meso-scale systems that frequently bring frontal precipitation to the area. We should not assume that one metric will help us find the 'best model' in all physical aspects, so we must look to strategies that employ several metrics and see if we can determine which metrics are good generalisers and which are not. It should be kept in mind that it is not without interest to find metrics that excel locally, as long as these metrics are not forced to work outside their area of applicability and can be shown to be robust where they do work.

A suggestion for the type of metric to consider is given by Perkins et al. (2007) in which Australian precipitation data were compared against an ensemble of GCM model outputs. Probability density functions (PDFs) were constructed in that work and used to generate a 'match metric' based on the overlap of normalised PDFs (model and observations, respectively). A perfect overlap results in a skill score SS of one and the score is close to zero for a poor overlap:

$$SS = \sum_{1}^{N} \min(PDF_{\text{Model}}, PDF_{\text{Observation}}). \tag{3.1}$$

Ranking of the obtained scores from the metrics then gives the model rank in the comparison. We will use that basic idea, but extend it, since the overlap between two PDFs may be dominated by values near the median or the mean while statistics for rare extreme events are not weighted highly. Robustness of the optimal metrics must be estimated so that we can state the quality of our answers, and part of the goal of this project is to determine ways to do that.

We also consider the technique of the crossing-point statistic (e.g. Gutowski et al. (in press)) which emerges when subtracting two PDFs extracted for the same region but for different time periods,

Table 4.1: RCMs used in this study.

	Model	Driving GCM	Institute
1	HIRHAM_D	HadAM3H	DMI
2	CHRM	HadAM3H	ETH
3	CLM	HadAM3H	GKSS
4	HadRM	HadAM3P	HC
5	RegCM	HadAM3H	ICTP
6	RACMO	HadAM3H	KNMI
7	HIRHAM_M	HadAM3H	METNO
8	REMO	HadAM3H	MPI
9	RCAO	HadAM3H	SMHI
10	PROMES	HadAM3H	UCM
11	HIRHAM_De	ECHAM4/OPYC	DMI
12	RCAOe	ECHAM4/OPYC	SMHI

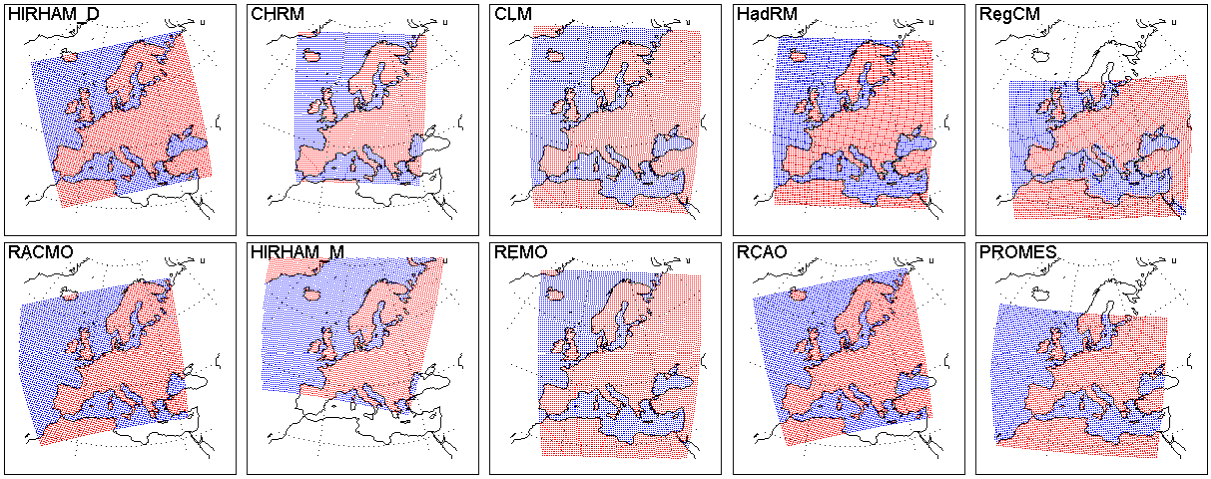

Figure 4.1: Area coverage for ten PRUDENCE model set-ups. Note that coverage is not identical, hence some of the regions defined in Christensen and Christensen (2007) are not defined. Two of the model set-ups (HIRHAM_D, RCAO) include experiments with different driving data giving a total of 12 PRUDENCE model experiments investigated in this study.

indicating which precipitation intensities will make larger or smaller contributions to the total precipitation for a certain time period. The structure of this paper is as follows. Section 4 presents the model data and the observations used in this study. In sections 5 and 6 we present the results when comparing the models with observations for both wet days and for drought conditions. The report ends with a conclusion in section 7.

4. Data

Two different data sets are analysed in this study. The first data set consists of daily precipitation from PRUDENCE RCMs including control (1961-1990) and scenario (2071-2100) runs for each (Christensen and Christensen, 2007). 'Control runs' are forced by observed realistic SSTs, observed greenhouse gases and time-varying realistic volcanic dust/aerosol forcing, while the solar irradiance is held constant. All of the simulations we use here are in a 50 km horizontal resolution, and where future scenarios are discussed we refer to the SRES A2 scenario (Nakićenović et al., 2000). The different models used in this study with their corresponding driving GCM and institute are listed in

Table 4.2: European regions used in this study.

	Region	Acronym
1	British Isles	BI
2	Iberian Peninsula	IP
3	France	FR
4	Mid-Europe	ME
5	Scandinavia	SC
6	Alps	AL
7	Mediterranean	MD
8	Eastern Europe	EA
9	European overlap	EUx

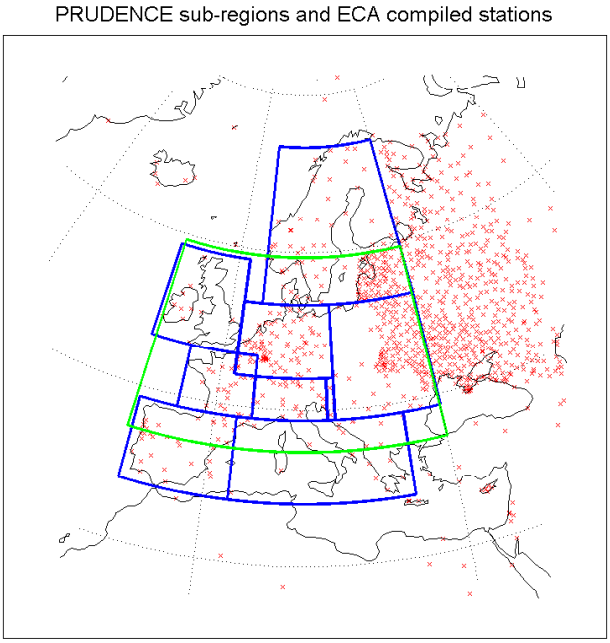

Figure 4.2: Geographical position of ECA compiled precipitation stations (in red) and eight European PRUDENCE sub-areas (in blue). A ninth region (in green), defined as the maximum area covered by the ten models in Figure 4.1, termed 'EUx' will also be used in this work. This region covers latitudes 41.0°N to 59.5°N and longitudes 10.5°W to 30.0°E. Note that region SC is not covered well by models RegCM and PROMES while the IP and MD regions are not well covered by HIRHAM_M.

Table 4.1. Note that the HadRM model is driven by the HadAM3P GCM which main difference to the HadAM3H GCM is improved large-scale cloud and convective anvil parameterisations (Rowell, 2005). The spatial coverage for all these models are shown in Figure 4.1. This study, with an analyse of PRUDENCE model data, is the first part out of two studies. The second study will be undertaken using data from the ENSEMBLES project (Hewitt, 2005) with transient (1951-2100) high resolution, global and regional models validated against quality controlled and high resolution data sets for Europe.

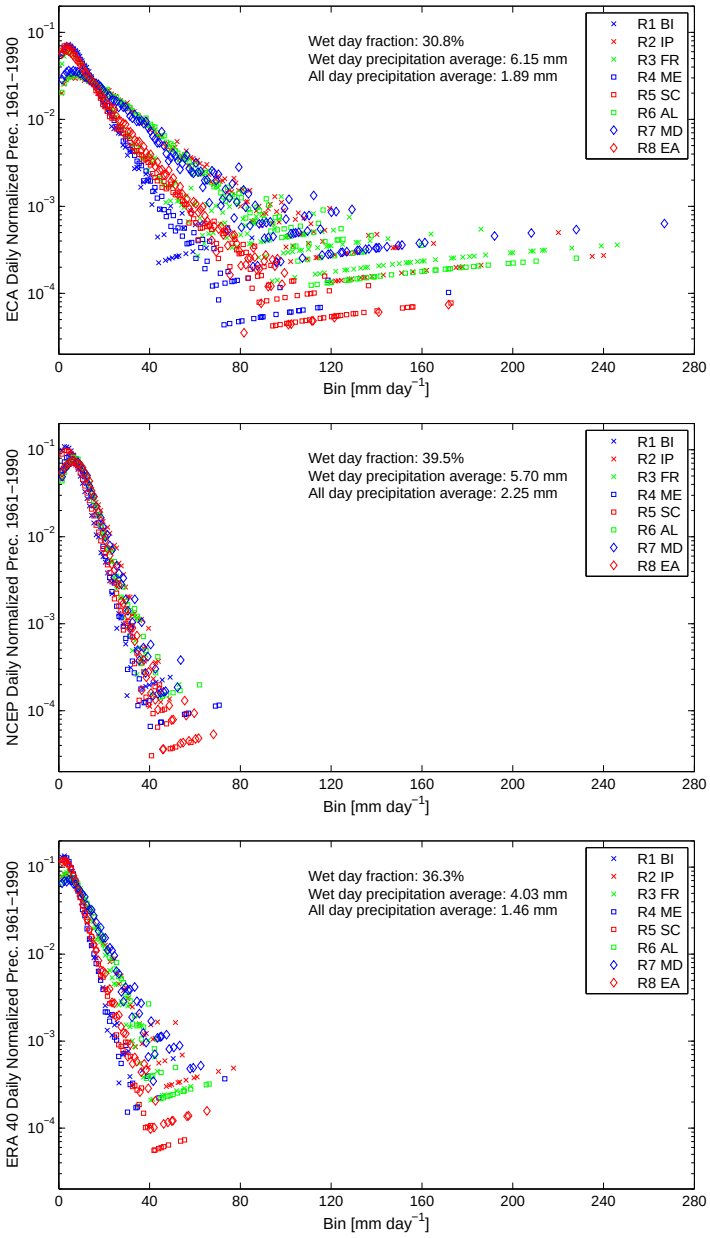

Figure 4.3: Normalised precipitation distribution for ERA40 (bottom panel) and NCEP/NCAR (middle panel) reanalysis data together with ECA observational data (top panel) for the period 1961-1990 for the eight European sub-regions.

To compare the climate change in different climatological areas, nine European sub-regions are defined as given by the boxes in Figure 4.2 (see Figure 4 in Christensen and Christensen (2007) and Table 4.2). After removing data for all ocean grid points, daily precipitation for each model and for each of the European PRUDENCE sub-regions were extracted. Because we want to be able to compare the model data to observations, we define dry days as days with precipitation less than 1 mm, and remove these from the precipitation data. The observations of weak intensity precipitation is not always reliable, and it has earlier been found that low intensity precipitation does not contribute significantly to the total amount of precipitation (Dai, 2001). PDFs were then calculated from the remaining daily data for each region and model through three steps. First the data were binned into bins of 1 mm width starting at 1 mm. Then each bin's frequency of events was multiplied by the bin average to get the precipitation amount. The binned data were finally

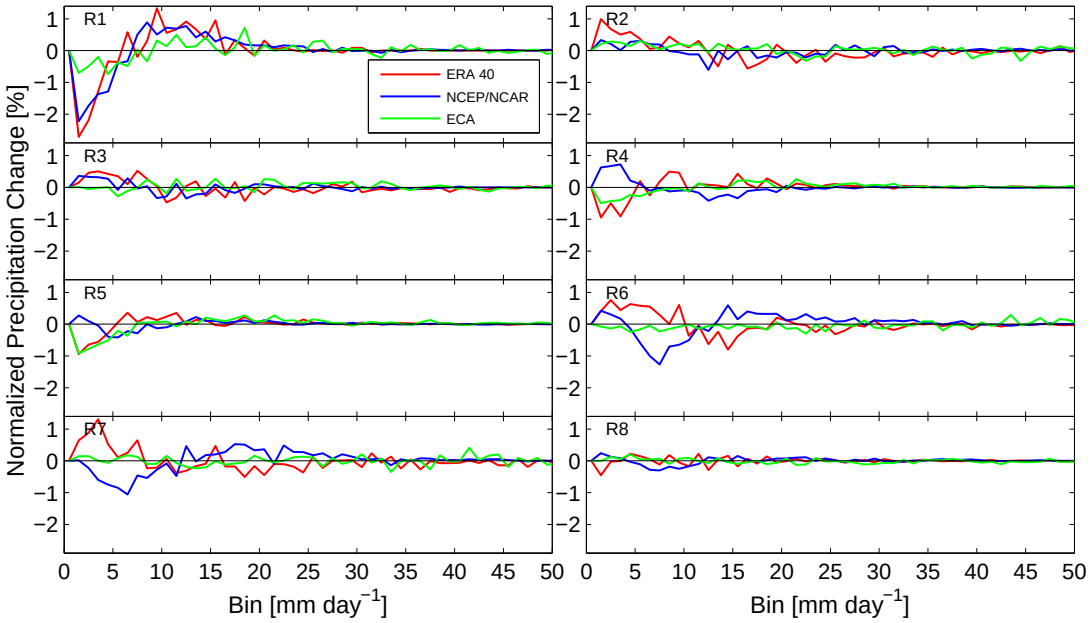

Figure 4.4: Normalised precipitation change for ERA40, NCEP/NCAR and ECA observations. The change is the difference between precipitation distributions for the periods 1980-2002 and 1957-1979. The individual panels show the result for each European sub-region.

normalised by dividing each bin value with the total precipitation amount for the 30 year period for the particular region and model run in question.

The second data set in this study is based on measurements monitored by the European Climate Assessment (ECA) project (http://eca.knmi.nl/). ECA was initiated to combine, present and analyse daily series of meteorological measurements. In this study we analyse series of daily observations of precipitation at ECA stations throughout Europe within the individual PRUDENCE sub-regions (see Figure 4.2). Note the uneven distribution of ECA stations within and between the RCM regions. Daily precipitation with $P \geq 1$ mm day^{-1} for the period 1961-1990 from this data set were extracted, regionally divided, binned and normalised for a direct comparison with the PRUDENCE RCM control runs. Stations with less than 300 wet day measurements were excluded from further analysis.

Observations vs. reanalysis data

The top panel of Figure 4.3 shows the ECA precipitation PDF for 1961-1990 for all eight sub-regions. The other two panels show PDFs for the two major reanalysis precipitation data sets NCEP/NCAR and ERA40 respectively. The overall shapes are similar for the three data sets but the maximum daily precipitation is significantly lower for the reanalysis data compared to observations. Figure 4.4 shows trends in precipitation distributions for the three data sets for eight sub-regions (cf. Figure 6.9). The trends are taken as the difference in precipitation PDFs during the periods 1980–2002 and 1957–1979 and we see large differences between the data sets.

5. Models vs. observations

Figures 5.1 through 5.8 compares binned normalised precipitations for the ten PRUDENCE models, driven by the same GCM and for eight sub-regions, with ECA precipitation data for the period 1961 to 1990. Each figure shows panels with full year precipitation distributions together with

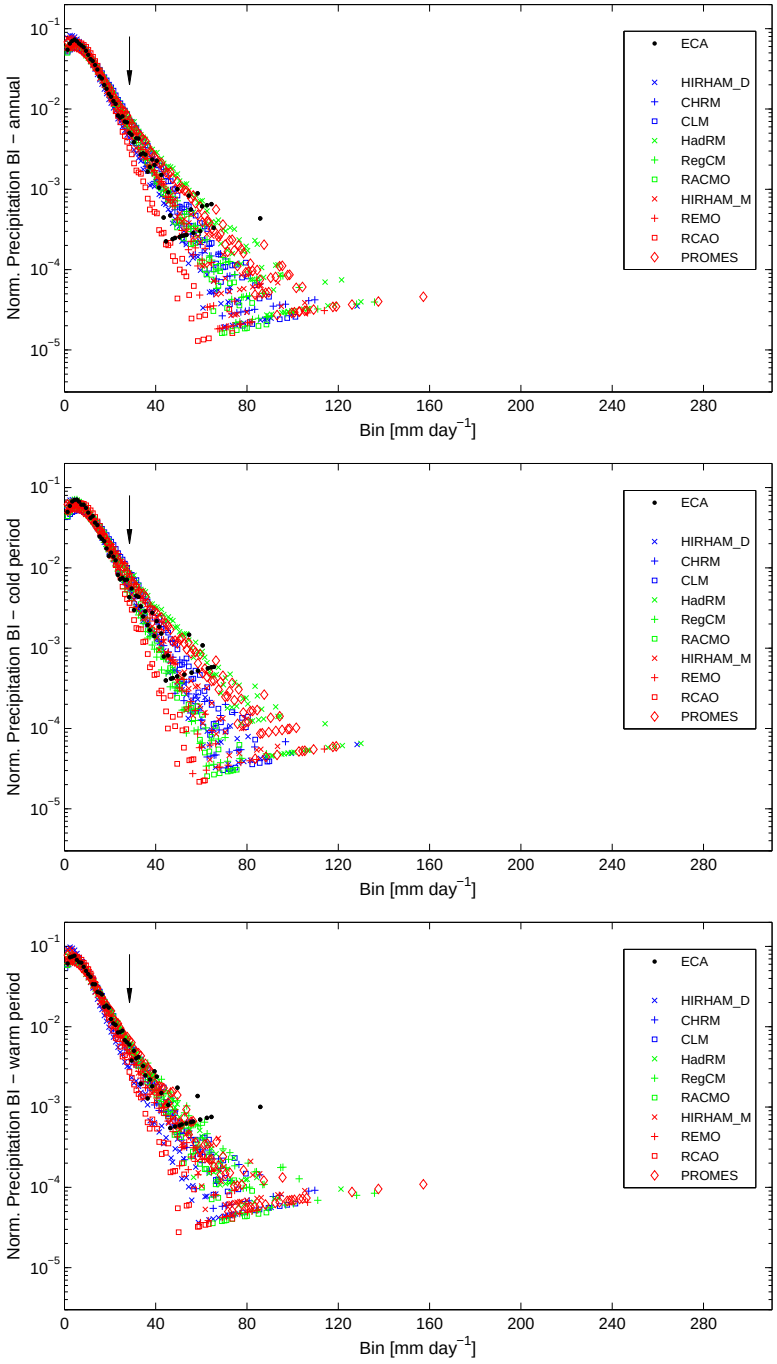

Figure 5.1: Comparison between ECA compiled precipitation observations (black) with precipitation data from ten PRUDENCE RCM models for the period 1961-1990 for PRUDENCE region 1 (BI). Top panel show distributions using the full annual data, the middle panel is for the cold period (October-March) and the bottom panel for the warm period (April-September). Arrows mark the 95$^{\text{th}}$ percentile of observed precipitation.

distributions during the cold and warm periods. Extreme precipitation is, as expected, difficult to model and the majority of the PRUDENCE RCM daily precipitation distributions agrees fairly better with observations during the cold period than for the full annual period for most sub-regions. Note that the distributions can often be divided into three regions in the diagram with a few models above and a few models below the average model distribution, and most models perform similarly in all regions.

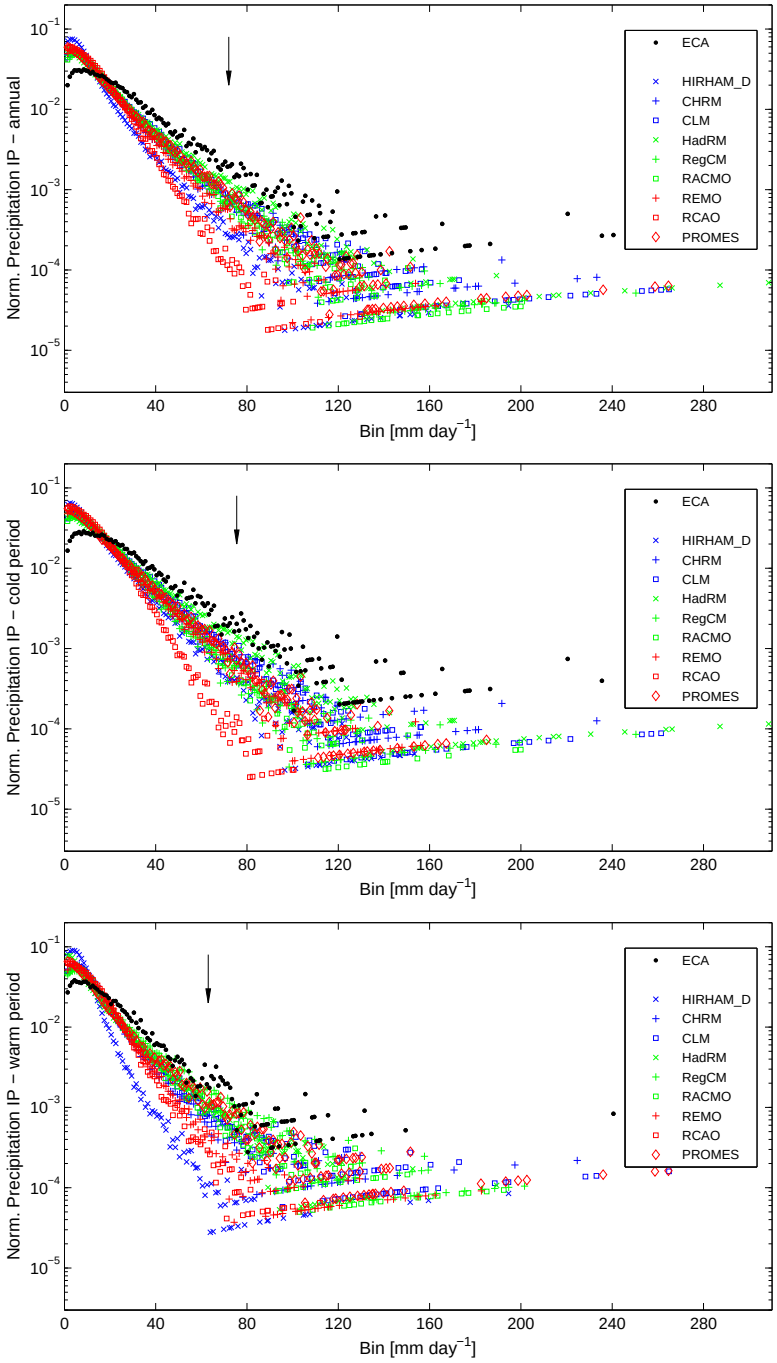

Figure 5.2: The same as Figure 5.1 but for region 2 (IP).

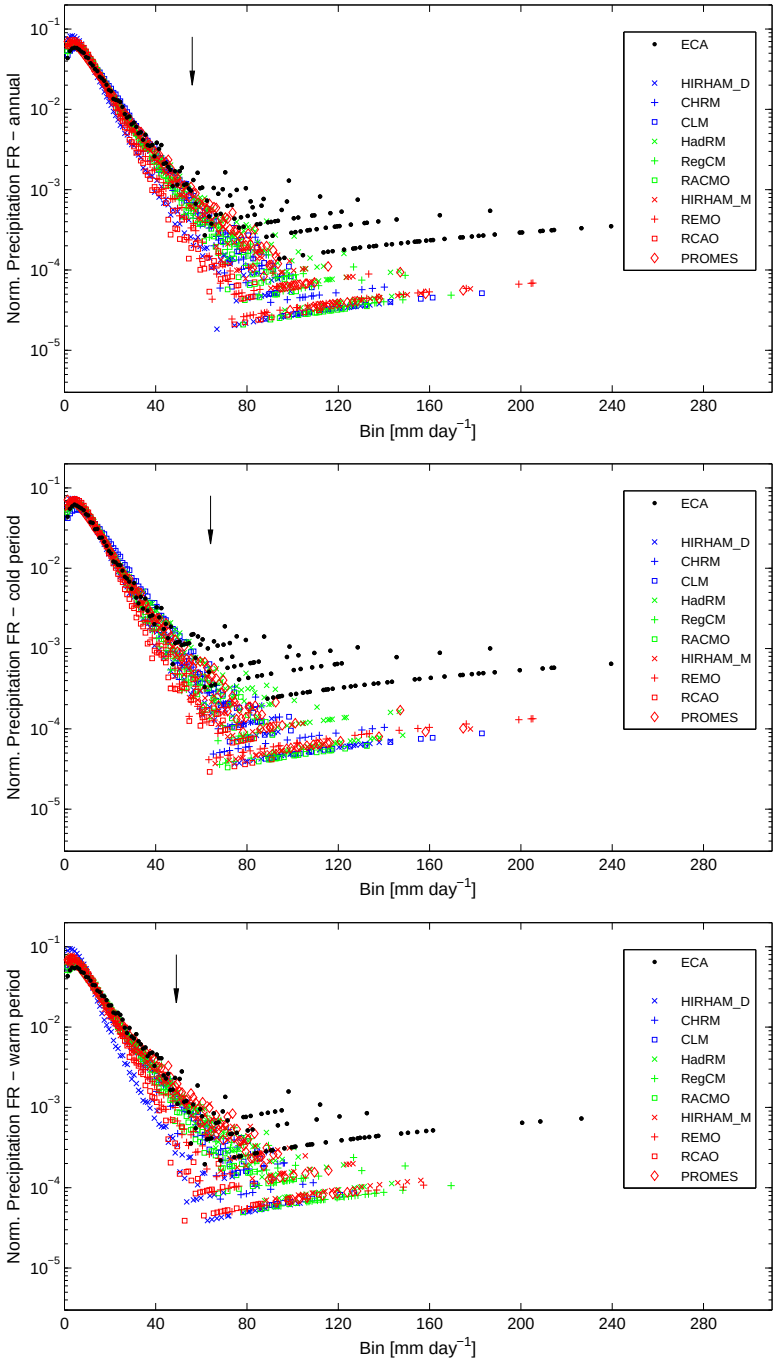

Figure 5.3: The same as Figure 5.1 but for region 3 (FR).

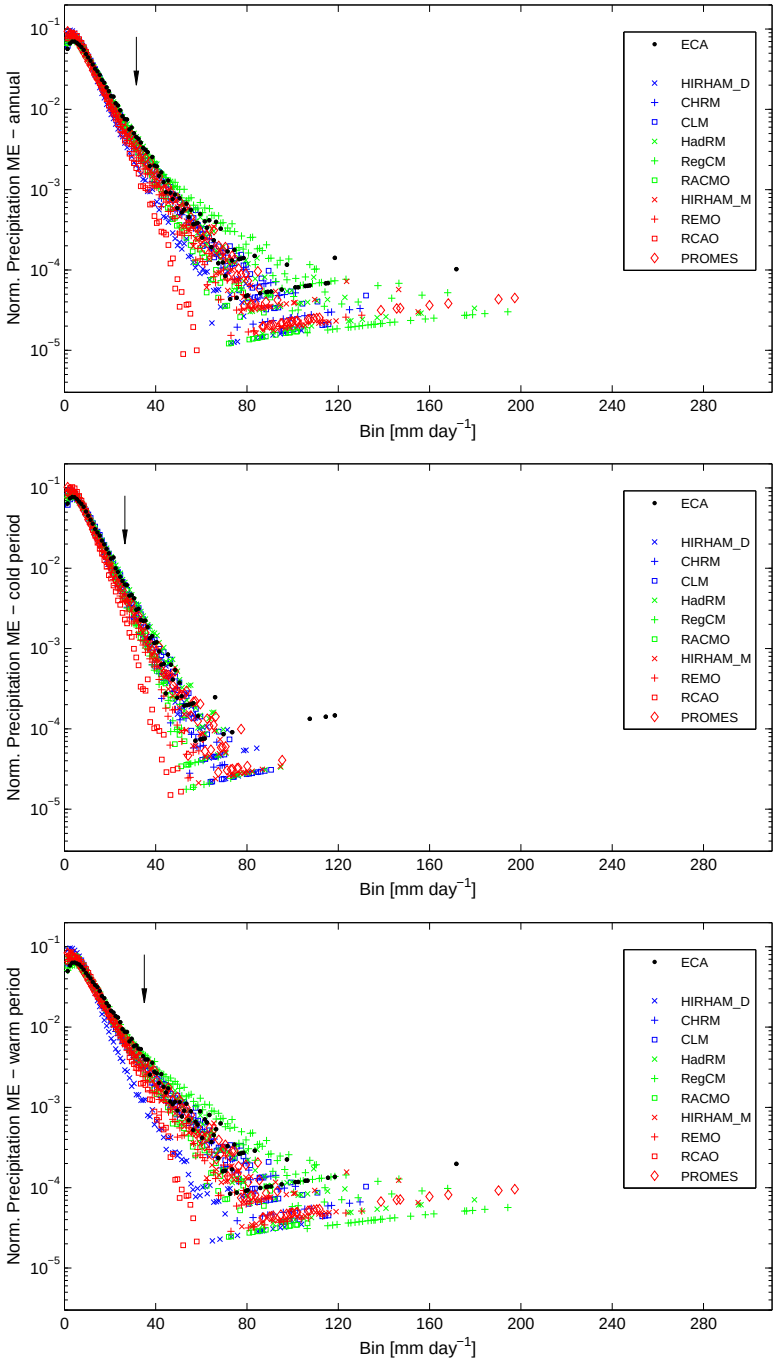

Figure 5.4: The same as Figure 5.1 but for region 4 (ME).

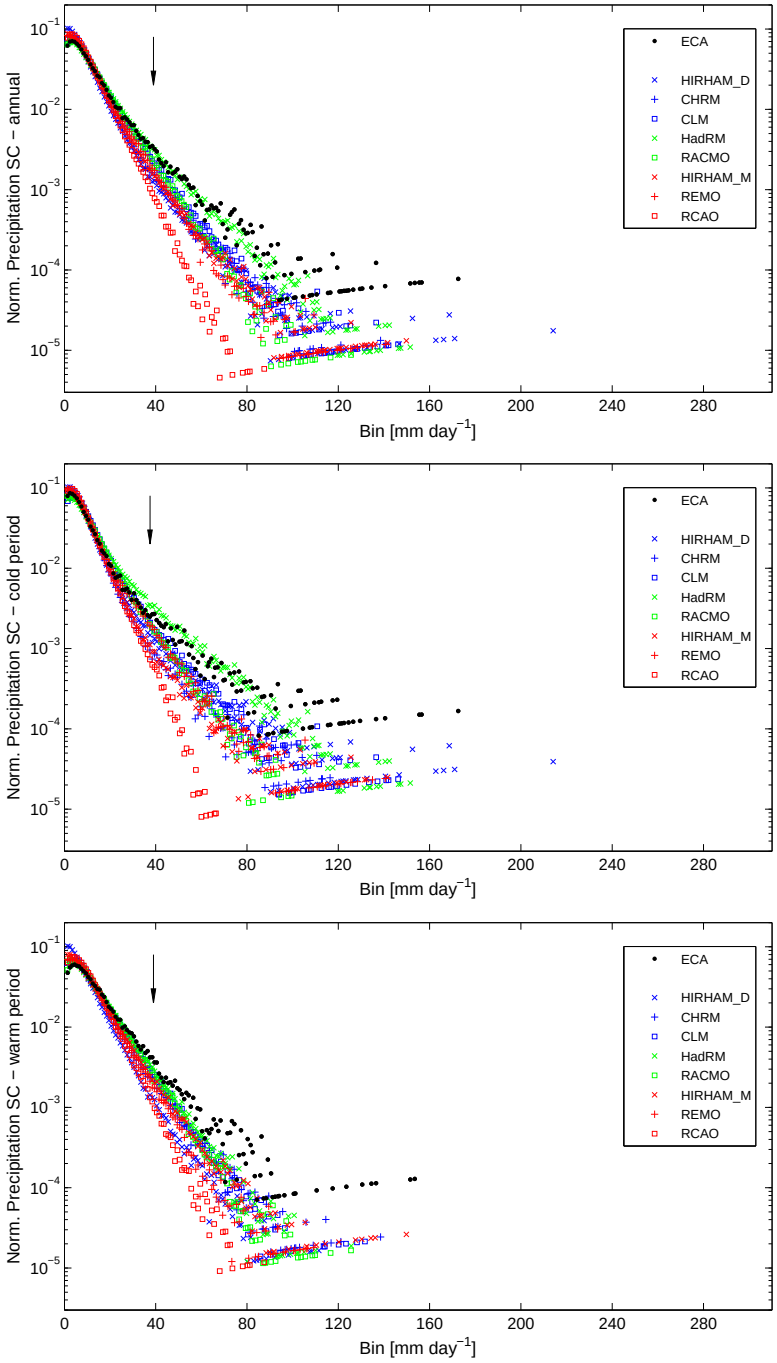

Figure 5.5: The same as Figure 5.1 but for region 5 (SC).

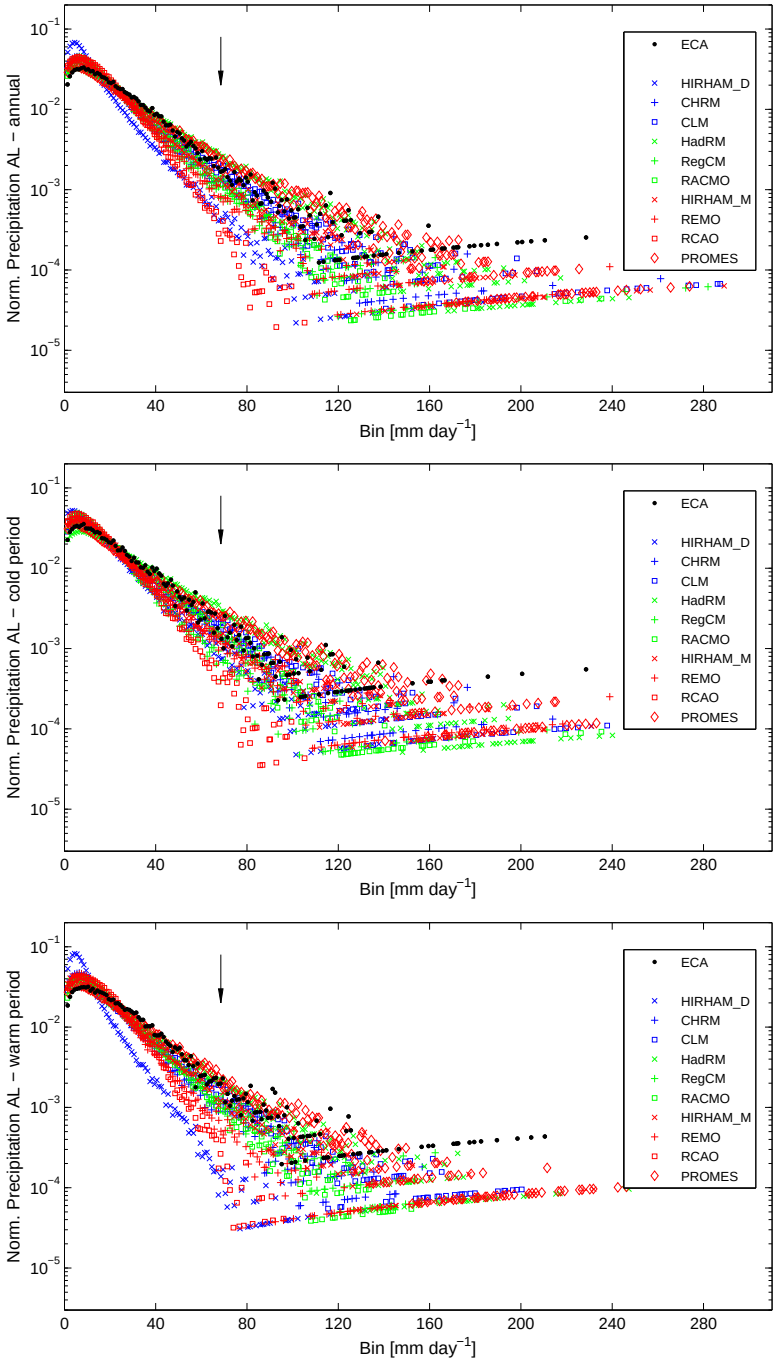

Figure 5.6: The same as Figure 5.1 but for region 6 (AL).

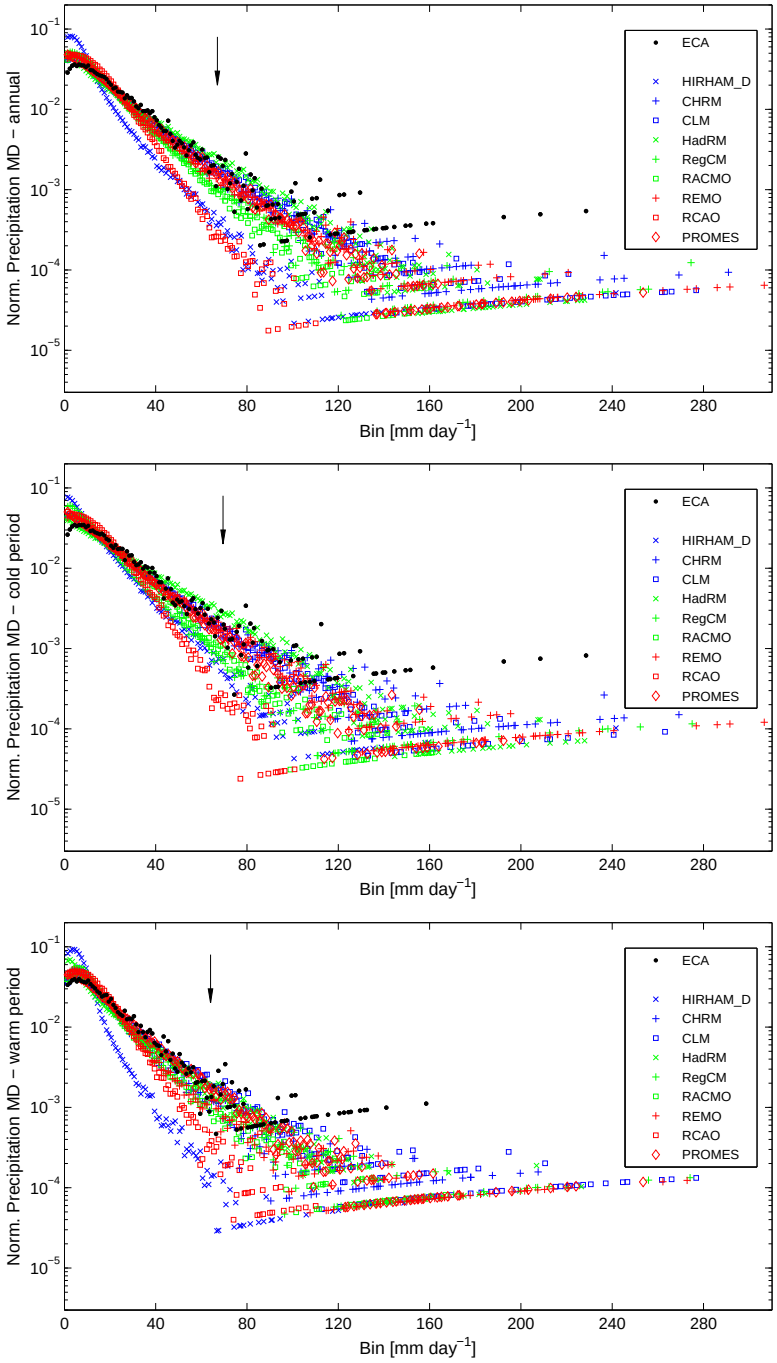

Figure 5.7: The same as Figure 5.1 but for region 7 (MD).

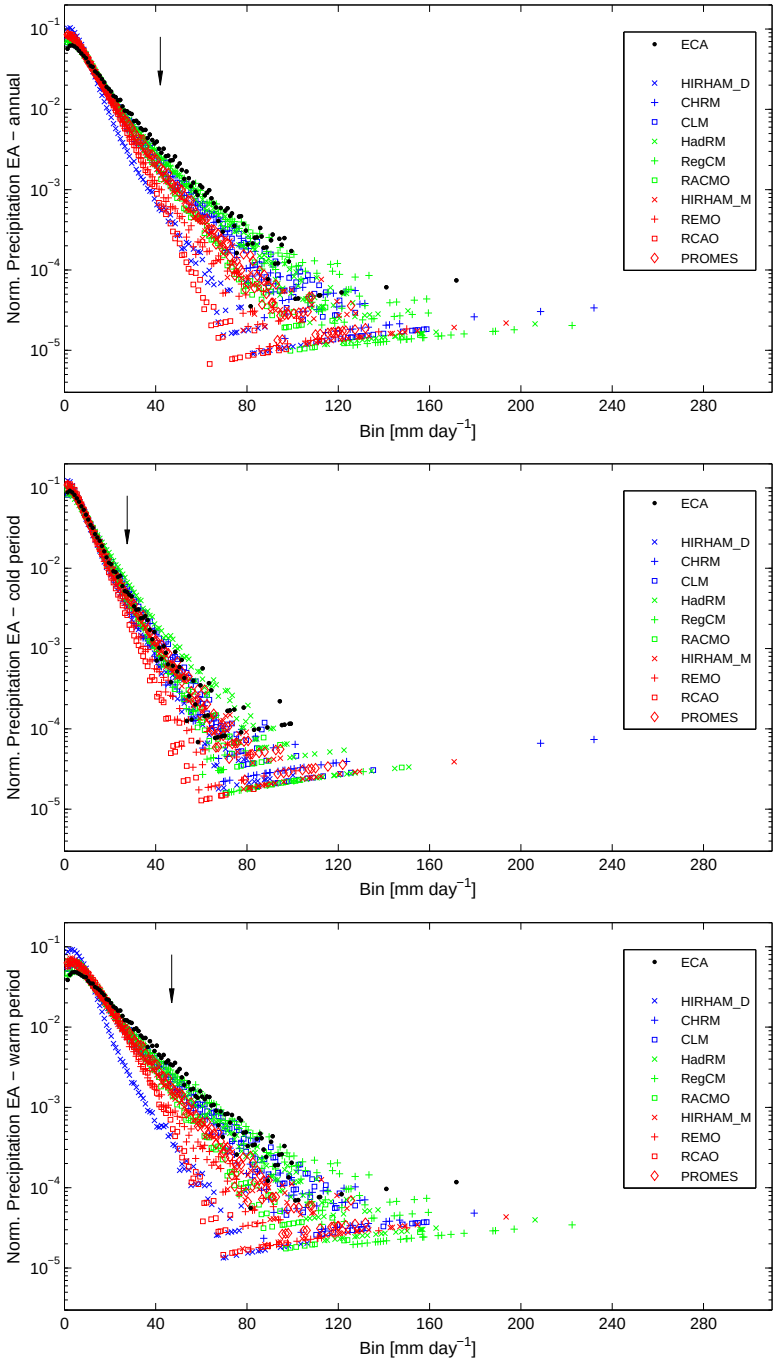

Figure 5.8: The same as Figure 5.1 but for region 8 (EA).

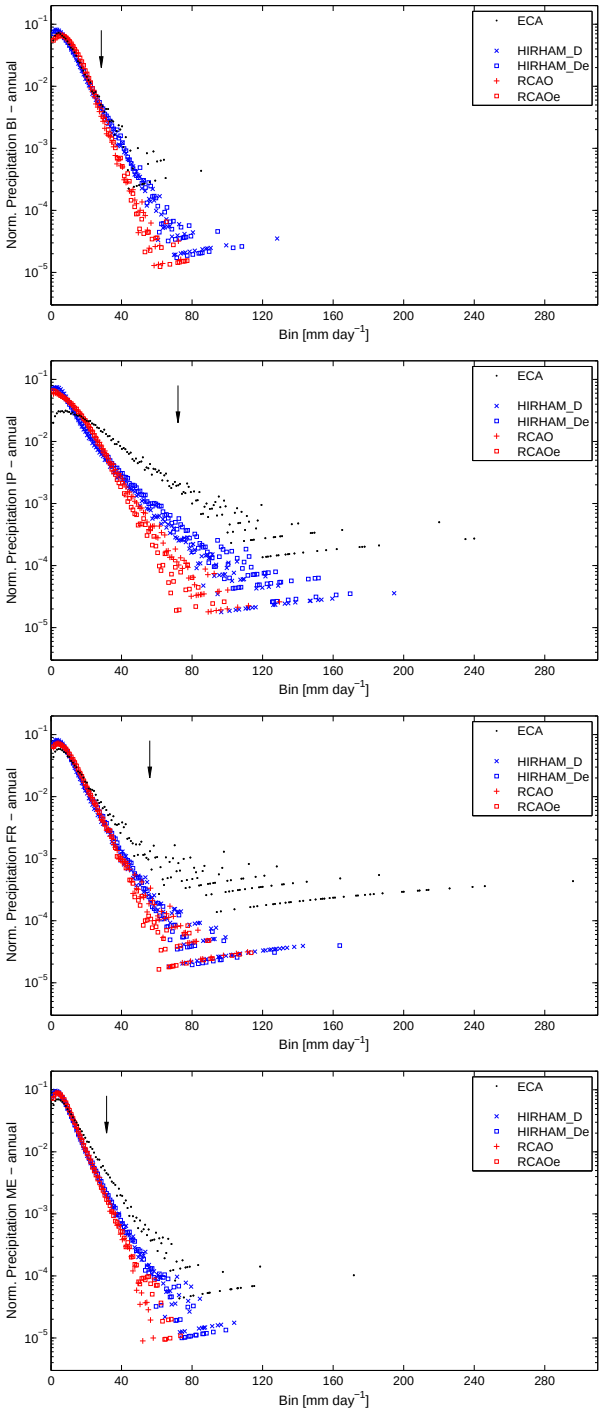

Figure 5.9: Comparison between ECA compiled precipitation observations (black) with precipitation data from PRUDENCE RCM models (HIRHAM_D/e and RCAO/e) individually driven by two different AGCMs for the period 1961-1990 for PRUDENCE regions 1 (top panel) through 4 (bottom panel). Arrows mark the 95th percentile of observed precipitation.

The HIRHAM_D and RCAO models have each been driven by two different GCMs: ECHAM4 (Roeckner et al., 1996) and HadAM3H (Pope et al., 2000). A comparison between precipitation PDFs for these two models together with ECA PDFs are given in Figures 5.9 and 5.10. We find only small deviations in precipitation between the different GCM forcing outputs compared to the different RCM precipitation distributions. We conclude that the different outcomes of the precipitation spectrum from one specific model driven by two different GCMs are negligible in this

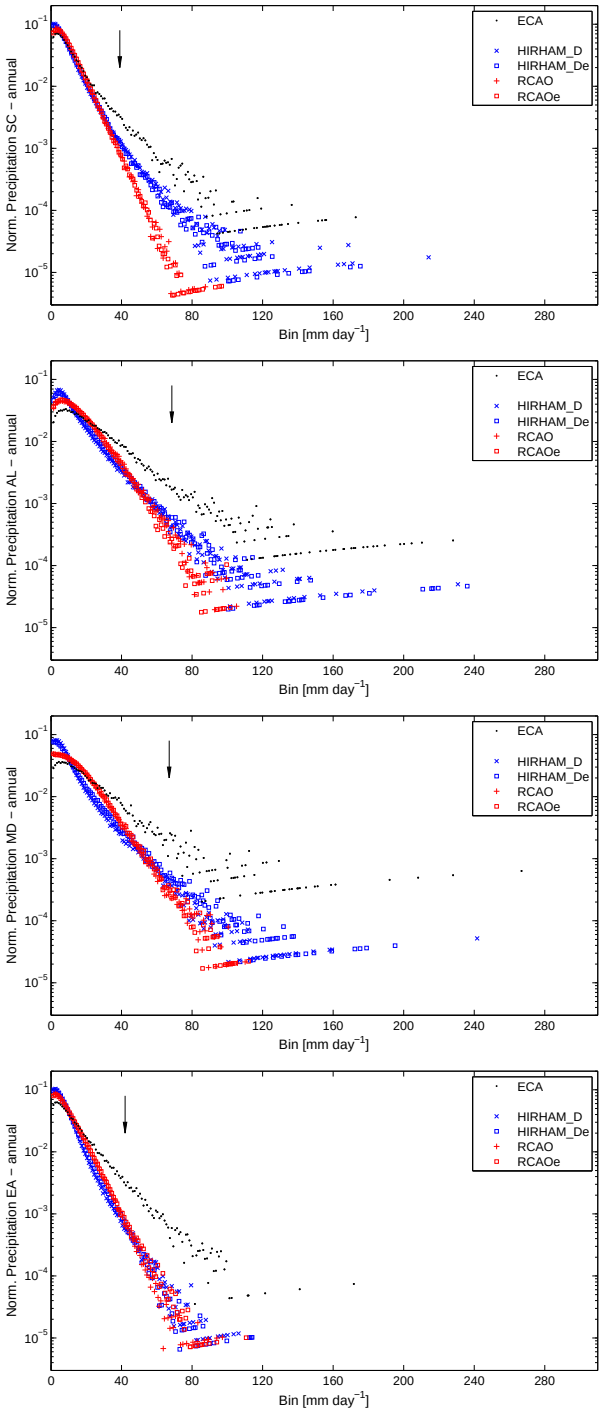

Figure 5.10: The same as Figure 5.9 but for regions 5 (top panel) through 8 (bottom panel).

study compared to different RCMs, at least for the PDFs at moderate precipitation.

The chosen regions and metrics are of course arbitrary. One example where the choice of region might not be physically justified is the SC region, where the precipitation in western Norway is much more extreme than that in the rest of the region. We explored the effect of dividing the region into two smaller regions depending on the skewness of the frequency PDF, i.e. separating the parts of Norway with extreme precipitation, but the ranking of the models was not affected by this. We also divided the European region into a Northern and a Southern region, with the same result. This suggest that the results are not so very dependent on the choice of regions.

Table 5.1: Results of comparing ten PRUDENCE models, all driven by the same GCM, for precipitation against ECA precipitation data, using daily values for the period 1961-1990. Each row of the table gives the region considered (see Table 4.2) and columns describing 4 different metrics with the best model and the winning skill score for each metric. The first part is for all annual values, the second part for the cold period (October–March) and the third part for the warm period (April–September). Note that metrics 1 through 3 are based on fractions of the full PDFs and the values given are fractions of the maximum ECA PDF area for the specific metric interval in question. Note also that the method of common overlap have restrictions when not studying the full PDF. See text for further details.

	Region	Metric 0	Score	Metric 1	Score	Metric 2	Score	Metric 3	Score
Annual	1 BI	RegCM	0.966	HadRM	0.988	HadRM	0.993	RCAO	0.999
	2 IP	HadRM	0.823	HadRM	0.664	HadRM	0.694	RCAO	1.000
	3 FR	CLM	0.953	CLM	0.778	CLM	0.846	RCAO	1.000
	4 ME	RegCM	0.966	HadRM	0.945	HadRM	0.933	RegCM	0.993
	5 SC	HadRM	0.977	HadRM	0.949	HadRM	0.964	RACMO	1.000
	6 AL	HadRM	0.936	HadRM	0.934	HadRM	0.925	RCAO	1.000
	7 MD	PROMES	0.923	HadRM	0.876	CLM	0.863	RCAO	1.000
	8 EA	RegCM	0.949	RegCM	0.801	HadRM	0.838	RegCM	0.996
	9 EUx	CLM	0.975	HadRM	0.985	HadRM	0.977	RCAO	1.000
	Region	Metric 0	Score	Metric 1	Score	Metric 2	Score	Metric 3	Score
Cold period	1 BI	CHRM	0.960	HadRM	0.987	HadRM	0.985	RCAO	0.992
	2 IP	HadRM	0.856	HadRM	0.728	HadRM	0.759	RCAO	1.000
	3 FR	HadRM	0.942	HadRM	0.755	HadRM	0.829	HIRHAM_M	1.000
	4 ME	CLM	0.985	HadRM	0.981	HadRM	0.976	CLM	0.999
	5 SC	REMO	0.959	HadRM	0.963	HadRM	0.975	HIRHAM_M	0.997
	6 AL	CLM	0.939	HadRM	0.962	HadRM	0.957	RCAO	1.000
	7 MD	CLM	0.909	HadRM	0.913	HadRM	0.906	RCAO	1.000
	8 EA	HIRHAM_M	0.982	HadRM	0.989	HadRM	0.993	RACMO	0.995
	9 EUx	PROMES	0.970	HadRM	0.988	HadRM	0.992	RCAO	1.000
	Region	Metric 0	Score	Metric 1	Score	Metric 2	Score	Metric 3	Score
Warm period	1 BI	RACMO	0.963	PROMES	0.903	PROMES	0.935	RCAO	0.998
	2 IP	RegCM	0.870	RegCM	0.709	RegCM	0.736	RCAO	1.000
	3 FR	RegCM	0.945	RegCM	0.772	RegCM	0.831	RegCM	1.000
	4 ME	RegCM	0.965	RegCM	0.953	RegCM	0.953	REMO	0.992
	5 SC	RACMO	0.953	HadRM	0.812	HadRM	0.863	RACMO	0.996
	6 AL	PROMES	0.930	PROMES	0.898	PROMES	0.896	RCAO	1.000
	7 MD	PROMES	0.932	CLM	0.853	PROMES	0.866	RCAO	1.000
	8 EA	RegCM	0.954	RegCM	0.841	RegCM	0.877	RCAO	1.000
	9 EUx	RegCM	0.967	PROMES	0.887	RegCM	0.908	RCAO	1.000

Comparing single RCMs with observations

We need a way to measure which precipitation model is best at regional and larger scales. We therefore compare each PRUDENCE model precipitation against observed precipitation using the ECA data set. Furthermore, we choose only to compare those ten PRUDENCE models driven by the GCM model HadAM3H (Pope et al., 2000; Buonomo et al., 2007). Table 5.1 shows the results of this comparison for the period 1961-1990 when using the full year, using cold period data and using warm period data. All eight regions defined in Christensen and Christensen (2007) are used as well as the model overlap region EUx. Four scoring metrics are devised based on the amount of overlap of the normalised precipitation intensity PDFs. The metrics considered are: (0) - the overlap of the whole normalised precipitation intensity PDF above the wet-day limit, (1) - the far wing of the PDFs

(greater than 20 mm), (2) - the not so far wing of the PDFs (greater than 15 mm), and (3) - the part of the PDF close to the median value (between 2 and 15 mm). Note that metrics 1 through 3 are based on fractions of the full PDFs and the values given are fractions of the maximum ECA PDF area for the specific metric interval in question. Note also that the method of common overlap have restrictions when not studying the full PDF. For instance, the RCAO model is scoring well for metric 3 for all three time periods even though its PDF is far from similar to the observed PDF. This is because of its relatively narrow distribution with an overestimate for low to moderate precipitation resulting in a skill score equal to or close to the area for the observed PDF for that bin interval.

From Table 5.1 we see that no model has scored best in all regions or all time periods, but that HadRM and RegCM do well in metric 0, HadRM often does best in metrics 1 and 2 while RegCM and CHRM do well in metric 3, i.e. around the median value. Overall, the HadRM does best for the cold period while RegCM does best for the warm period.

Skill scores for all models for all regions and time periods are presented in Tables 8.1, 8.2 and 8.3 in section 8.

Robustness of ranking

The robustness of the skill score ranking is evaluated by a Monte Carlo method called bootstrapping with replacement (Efron et al., 1997). From the observations we have a number of time series, one for each region, and by making random selections of these time series we can evaluate how sensitive the skill score is to changes in the sampling of observations. Bootstrapping with replacement allows for the same time series to be selected several times, thus leaving one or more time series out of the calculation of the skill score. By repeating this procedure enough times, e.g. 10,000 times, we get a good statistical description of the variations in the skill scores for the different models, and we can determine the robustness of the ranking of the individual models. Now we can assess whether it is possible to determine if the one model is significantly better than the other at simulating the studied phenomenon.

We look for the distribution of just the highest ranking model, although - as will be done in the case of dry days in a following section - the distribution across all possible ranks can also be investigated. In Table 5.2 we show the percentage distributions of models' probabilities of ranking the highest, across all regions, for 10,000 Monte Carlo bootstrap trials. The values given in red refers to the models having the highest score in Table 5.1. We note that for region 1 (BI), full year and cold period, the best models from Table 5.1 are not ranked first in Table 5.2. In these two cases, the models that were ranked first in Table 5.2 were very close to having the highest score in Table 5.1. The results therefore do not appear to lack robustness. However, all metrics should be tested in this way so that we know whether a model is really the best or shares its rank with other models for a given region.

Comparing composite RCM PDFs with observations

It may be possible to find combinations of model PDFs that score better than single models (Perkins et al., 2007). We tested this in a crude way to gain some first insight. We constructed linear weighted combinations of PDFs and scored these against observations, using the same metrics as in the single-model tests. It was indeed possible to occasionally find linear combinations of model PDFs that scored better. The improvements were modest, but for single models already scoring high it may be worth while to pursue this line of approach. We did not explore nonlinear methods for combining PDFs.

Table 5.2: Results of robustness analysis for wet days. 10,000 Monte Carlo bootstrap trials were performed for all ten models in each region and the percentage of occurrences of first place ranking is recorded in columns 2 through 10. The intensity PDF has been used for the metric containing the whole range of values, i.e. metric 0 in Table 5.1. The top section is for the full year, the second for the cold period and the last section is for the warm period. For comparison, the winning models from Table 5.1 are shown in red font for each region and annual period respectively.

	Full year - 1st rank [%]								
Model	BI	IP	FR	ME	SC	AL	MD	EA	EUx
HIRHAM_D	0.14								
CHRM	19.64								
CLM	4.31	11.01	76.80	4.34	19.78	2.90	27.37		89.94
HadRM	23.18	88.84	16.95	24.07	72.34	69.10	0.75	8.01	10.06
RegCM	17.06		6.06	71.59				91.99	
RACMO	15.79	0.15			7.88	21.45			
HIRHAM_M	1.70		0.05			0.07			
REMO	10.63								
RCAO	6.51								
PROMES	1.04		0.14			6.48	71.88		
	Cold period - 1st rank [%]								
Model	BI	IP	FR	ME	SC	AL	MD	EA	EUx
HIRHAM_D	1.04		0.03		3.24				
CHRM	31.50		5.31		0.05		0.07		7.58
CLM	0.88		2.67	76.23	0.08	48.35	17.22	27.43	10.14
HadRM	11.67	100.00	60.50	23.74	16.08	25.51	31.37		
RegCM	12.15		0.02						
RACMO	16.23		21.27	0.03	20.00	18.40	8.92	22.66	22.69
HIRHAM_M	3.72		9.01		0.24	0.73		49.45	15.59
REMO	8.20				60.31	0.02	0.47		2.83
RCAO	10.13								
PROMES	4.48		1.19			6.99	41.95	0.46	41.17
	Warm period - 1st rank [%]								
Model	BI	IP	FR	ME	SC	AL	MD	EA	EUx
HIRHAM_D									
CHRM	1.17								
CLM	18.58		0.07			0.03	15.70		
HadRM	4.63				12.18				
RegCM	6.08	100.00	99.89	100.00			23.65	100.00	100.00
RACMO	38.40				87.82	9.32			
HIRHAM_M									
REMO	5.62								
RCAO	7.37						0.67		
PROMES	18.15		0.04			90.65	59.98		

Results for drought conditions

A meteorological drought occurs when there is less-than-normal amounts of precipitation in a region. The severity of the drought depends on the temporal and geographical extent of the reduced precipitation. Because some regions are more sensitive than others to the amount of precipitation, it is not possible from meteorological conditions only to generally define a drought. Here we consider only the time interval of less-than-normal precipitation, and make no further analysis of geographical extent or local characteristics. The 'normal' amounts of rain at a station, or grid point, is here defined as the climatological value from the time series, smoothed by a 30 day moving average. A day of drought is defined as when there is less precipitation than the climatological value plus one standard deviation, i.e. typically less than 3–7 mm of precipitation. The number of consecutive days

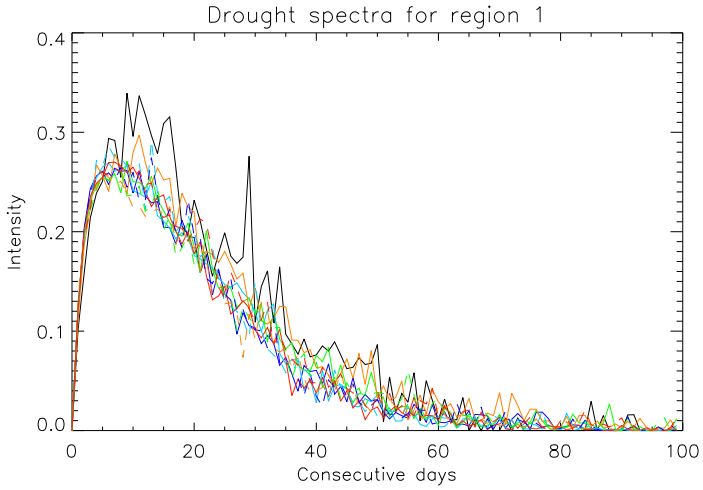

Figure 5.11: Intensity spectrum of the number of consecutive days of drought. One day of drought is defined as when the amount of precipitation on a single day is less than one standard deviation below the climatological value of precipitation for the geographical location. Observations are given as a solid black curve. Model spectra are shown for HIRHAM_D (solid dark blue), CHRM (solid light blue), CLM (solid green), HadRM (solid orange), RegCM (solid red), RACMO (dashed dark blue), HIRHAM_M (dashed light blue), REMO (dashed green), RCAO (dashed orange) and PROMES (dashed red).

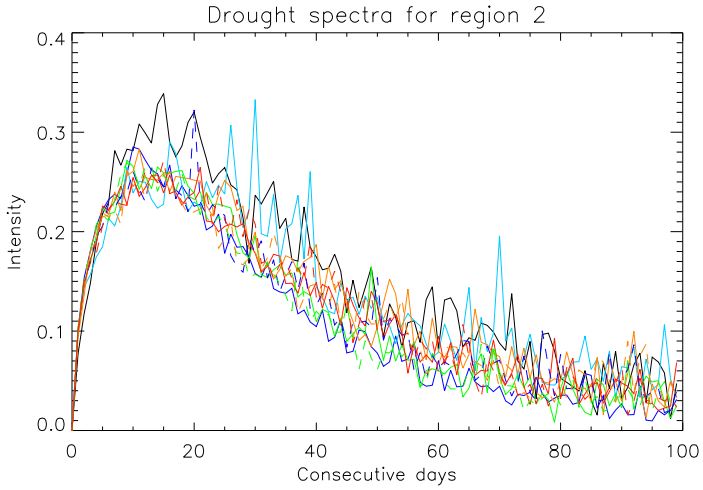

Figure 5.12: Same as Fig. 5.11 but for IP (region 2).

of drought are calculated for each station or grid point, and we study the distribution of the number of consecutive days for the mean value of each of the PRUDENCE regions, as defined above. In general we see that the models overestimate the fraction of droughts with a period of less than one week - and underestimate the fraction of droughts with a period of more than one week (Figures 5.11 through 5.18). A skill score, as defined above, is calculated for each of the models for each region. For most of the regions the models' skill scores are comparable with differences between the top and bottom scores of at most 0.1 percentage units. It is therefore difficult to select one model that performs better than the others, and a robustness analysis is required.

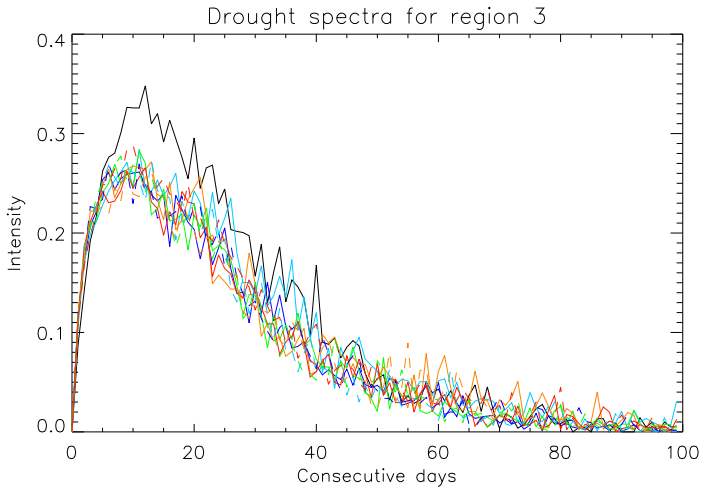

Figure 5.13: Same as Fig. 5.11 but for FR (region 3).

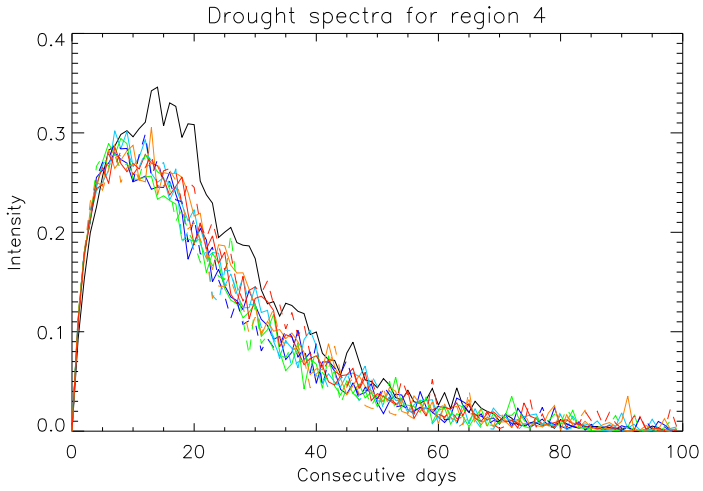

Figure 5.14: Same as Fig. 5.11 but for ME (region 4).

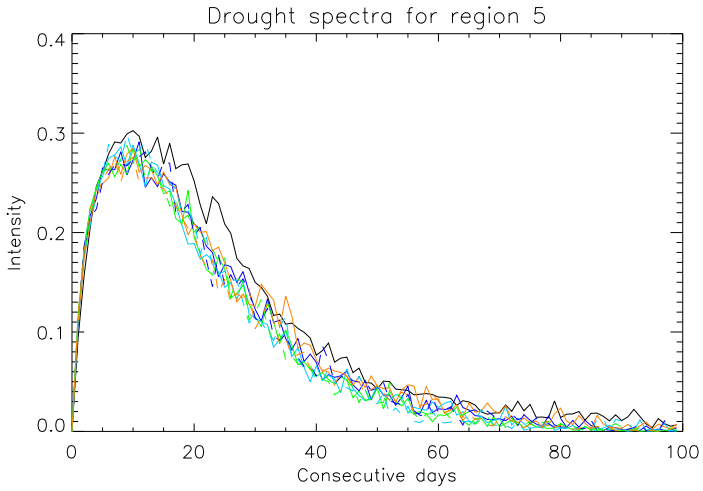

Figure 5.15: Same as Fig. 5.11 but for SC (region 5).

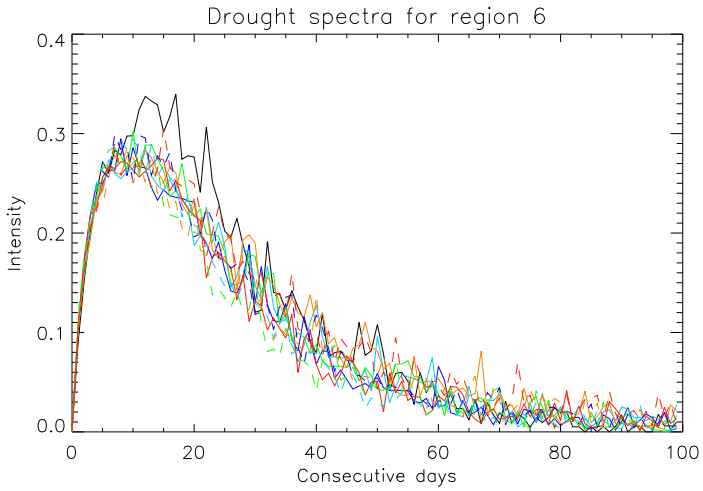

Figure 5.16: Same as Fig. 5.11 but for AL (region 6).

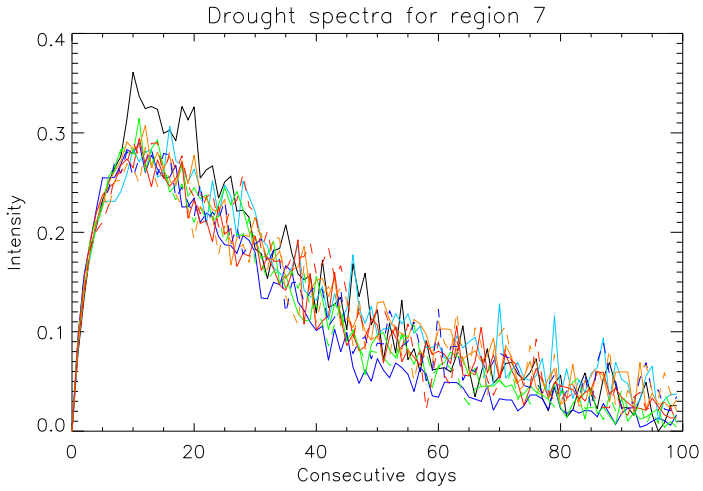

Figure 5.17: Same as Fig. 5.11 but for the MD (region 7).

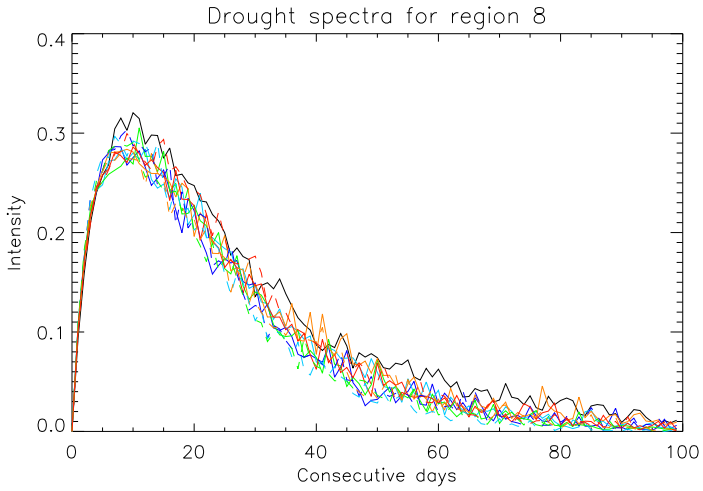

Figure 5.18: Same as Fig. 5.11 but for EA (region 8).

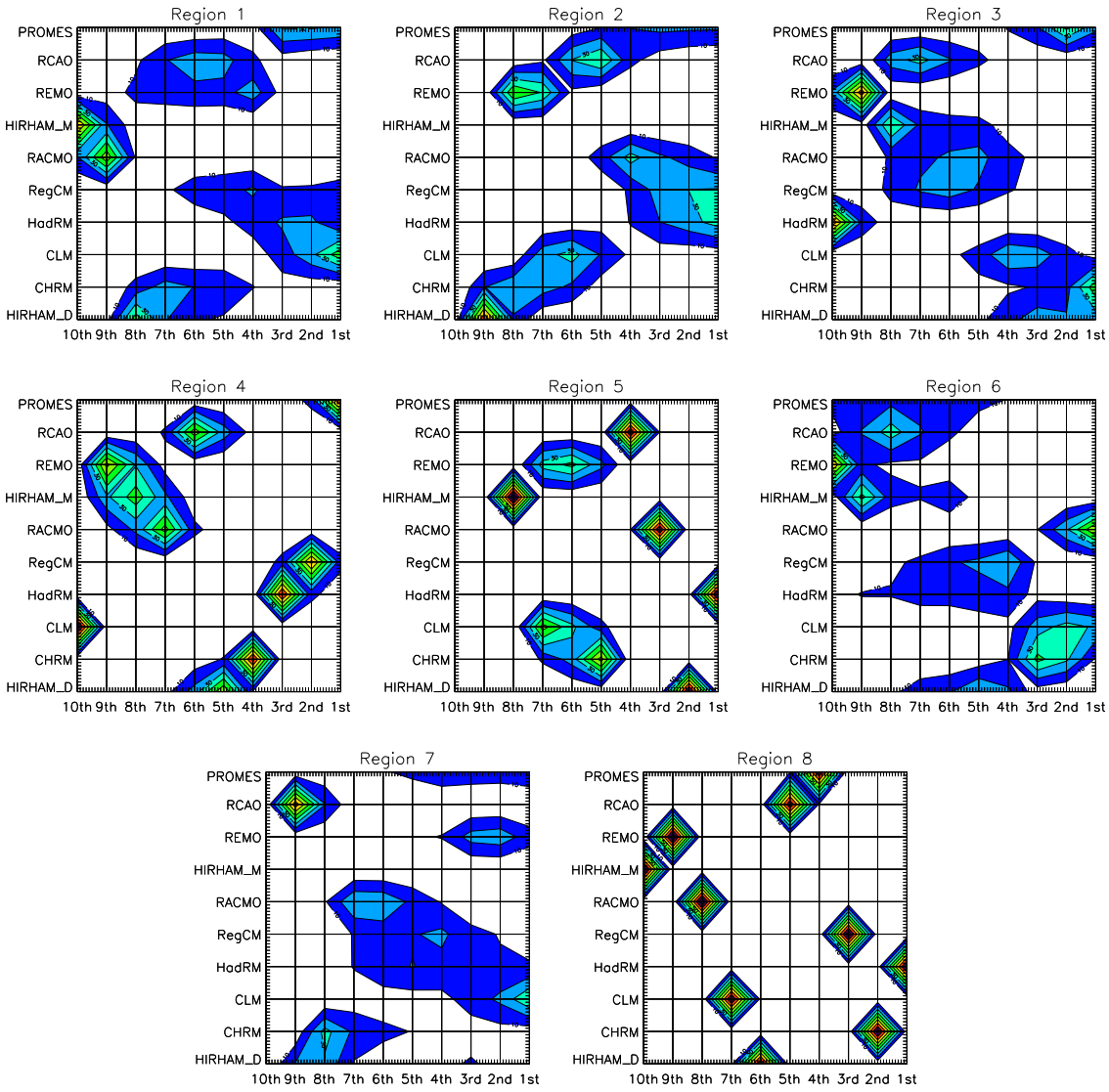

Figure 5.19: The results for the robustness analysis of the skill scores for the eight European sub-regions. Displayed is the percentage of times (out of 10,000) that the model gets a certain ranking, as indicated by the horizontal axis.

To determine the robustness of the skill score ranking of the models we use the method of bootstrapping with replacement, as described above. In 10,000 Monte Carlo simulations of the skill scores we find that whether it is possible to find a unique top scorer depends on the region studied, and the ranking of the models varies a lot between the different regions. In Figure 5.19 the percentage of times that one model ranks at a certain position is presented. The only consistent result for all the regions is that the HIRHAM_M model ranks poorly. For ME, SC and EA it is possible to determine a single top ranking model as the robustness test shows the same model to score the highest more than 80% of the time. For other regions the top scorers are more evenly distributed between two or more models, and it is not possible to determine a top scorer among them. It turns out that for all regions, except for MD, the top scorer in the original skill score test, is also the model with the highest percentage in the robustness test. The robustness analysis shows that it is not possible to single out one model that is a top scorer for all the European regions.

To summarise how the models perform for Europe in general we present the mean of each models

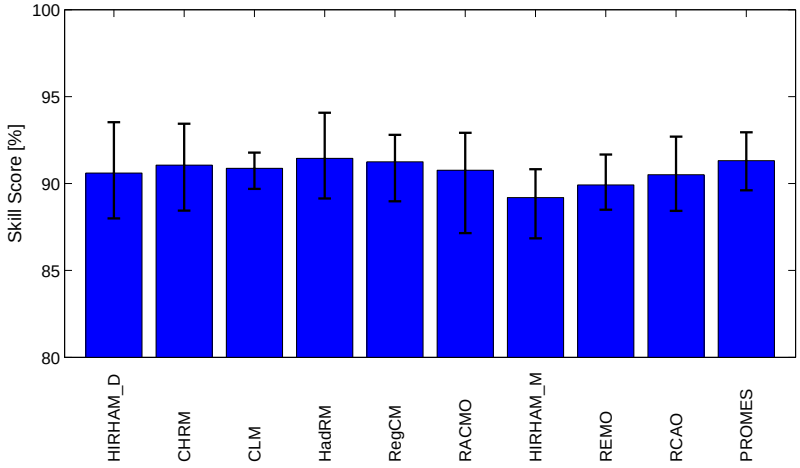

Figure 5.20: The mean skill score (blue bar) together with bars showing the maximum and minimum skill scores for each model and for all sub-regions. Note the vertical axis ranges from 80–100%.

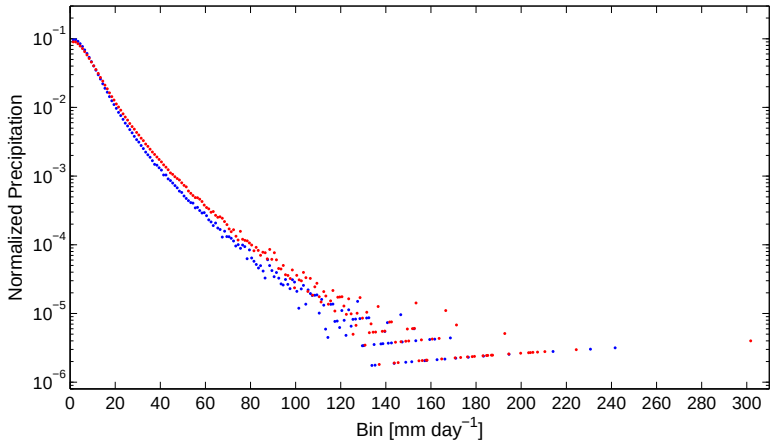

Figure 6.1: Binned normalised precipitations for the full HIRHAM_D land region covering 4150 grid points. Blue symbols refer to the 1961-1990 control run and red symbols represent the 2071-2100 scenario run.

skill score for each sub-region in Figure 5.20. We see that the mean skill scores deviate by about two percentage units, while each model shows variations in skill scores between different regions of one to three percentage units.

6. Crossover point analysis

We next consider a crossover point analysis (Gutowski et al., in press) to study changes in precipitation over long time scales. Daily precipitation PDFs for the first HIRHAM_D experiment for all land grid points are shown in Figure 6.1 for both the control run (blue) and the scenario run (red). This reveals that days with moderate precipitation (less than about 10 mm) will contribute less to the total precipitation in the future scenario while days with higher (more than about 10 mm) precipitation will contribute more to the overall precipitation. The crossover point x_c for this transition is found by subtracting the control precipitation from the scenario precipitation (see Figure 6.2) and determine where the curve goes from negative to positive values. The related transition percentile, P, is calculated as the accumulated precipitation for the control run from 1 to

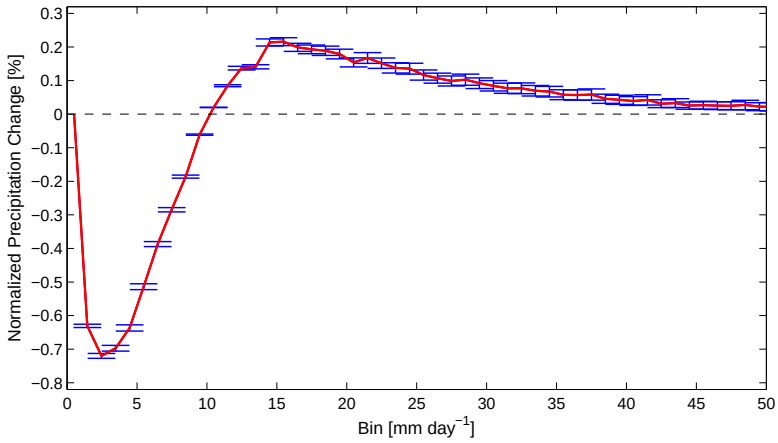

Figure 6.2: Binned normalised precipitation change (scenario minus control run) for the full HIRHAM_D land region. The error bars are calculated using equation 6.1. The curve's crossover point x_c is 10.4 ± 0.2 mm day^{-1} with a corresponding percentile P of $69.2 \pm 0.7\%$.

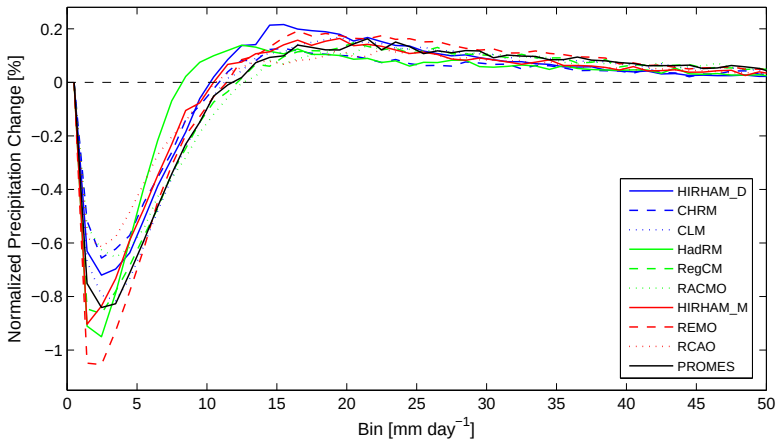

Figure 6.3: Binned normalised precipitation change over land (scenario minus control run) for all ten PRUDENCE experiments driven by the same GCM.

x_c mm day^{-1} divided by the total precipitation.

The error bars in Figure 6.2 are estimated by finding the number of uncorrelated areas U in the full HIRHAM_D model by calculating the autocorrelation between all land grid points. All grid points with a correlation above e^{-1} relative to the grid point in question is defined as belonging to the same correlated area. U is then given by the total number of land points divided by the average size of the correlated area. This is done for the control run and scenario run separately and the errors on the bin means in Figure 6.2 are given, for each bin i assuming a negligible temporal correlation for individual grid points, by

$$\sigma_i = |A_{S,i} - A_{C,i}| \cdot \sqrt{M} \cdot \sqrt{\frac{1}{N_{S,i} \cdot U_S} + \frac{1}{N_{C,i} \cdot U_C}}, \tag{6.1}$$

where the first factor is the normalised precipitation change plotted in Figure 6.2, M is the number of land grid points for the region in question and N_i is the number of values in bin i. For the HIRHAM_D model, U is about 30 for both the control and scenario periods, and we see that the

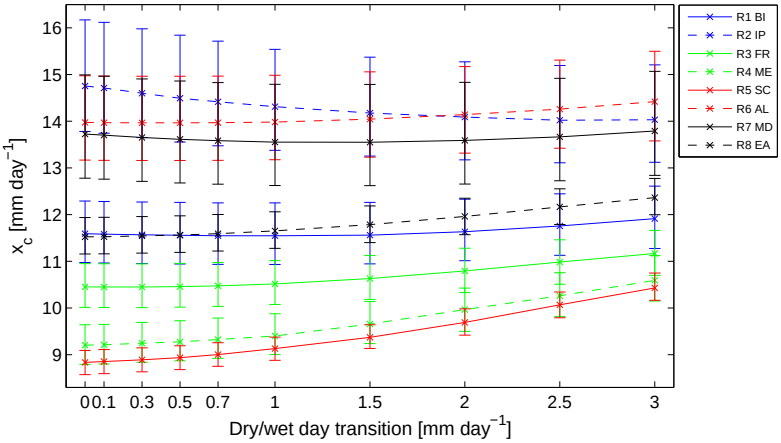

Figure 6.4: Crossover values x_c, with corresponding uncertainties, as a function of definition of wet day for the full HIRHAM_D land region. x_c was calculated for all eight PRUDENCE regions for ten different wet day transition values.

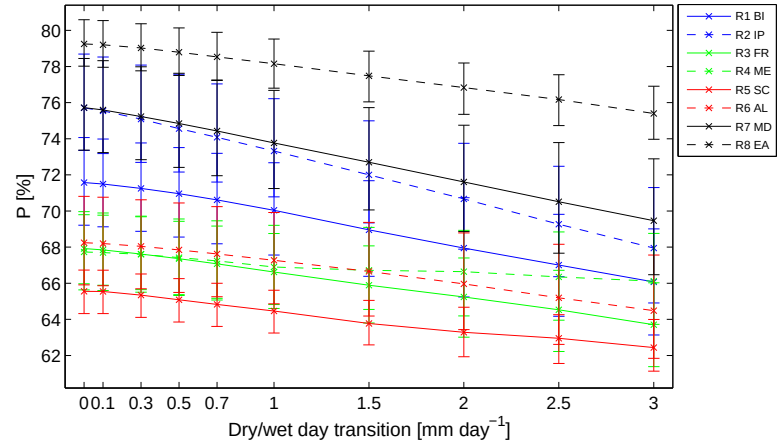

Figure 6.5: Crossover percentiles P, with corresponding uncertainties, as a function of definition of wet day for the full HIRHAM_D land region. P was calculated for all eight PRUDENCE regions for ten different wet day transition values.

amplitude of the precipitation change is significant. Figure 6.3 shows the scenario minus control run precipitation distribution for all PRUDENCE RCMs. The overall shape as well as the amplitude of the curves show strong similarities.

Definition of a 'wet day'

An analysis was undertaken investigating how the choice of dry day/wet day transition will affect the crossover point x_c and adherent crossover percentile P (see Figures 6.4 and 6.5). Transition values of 0, 0.1, 0.3, 0.5, 0.7, 1, 1.5, 2, 2.5 and 3 mm were investigated showing only moderate changes in x_c and P for values up to 1 mm (of the order of 0.2 mm day^{-1} and 1%).

Regional and model differences in x_c and P

Figures 6.6 and 6.7 show crossover values and percentiles for the ten PRUDENCE models respectively. The PRUDENCE regions IP, AL and MD have a relatively large model dependency for

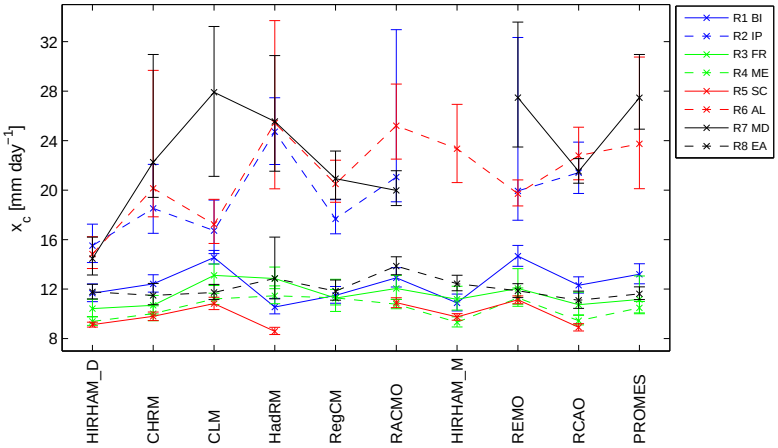

Figure 6.6: Crossover values, as presented in Figure 6.2, for ten models and eight sub-regions. Note the relative small scatter for five regions with x_c values between 9 and 14 mm day^{-1} while the other three regions have a large scatter for the majority of the RCMs.

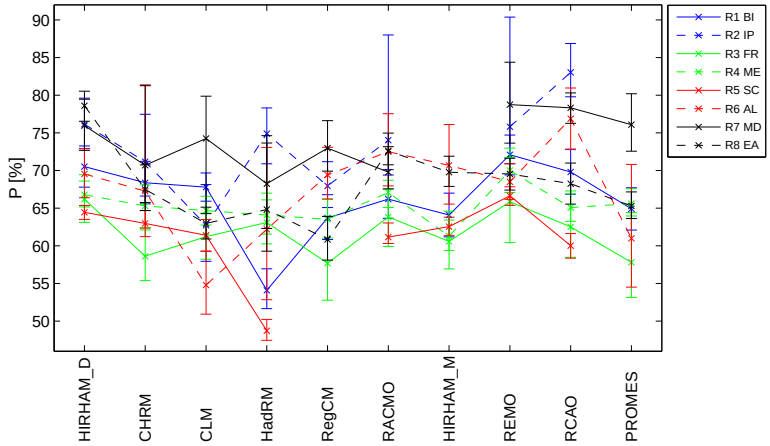

Figure 6.7: Crossover percentiles, as presented in Figure 6.2, for ten models and eight sub-regions. Note the absence of a distinct model dependency between arid and non-arid regions. Note also the large scatter between regions for some models, absent in others.

x_c while the remaining five regions have little dependency. Some models have a large spread in x_c for the different regions (values from 10 to 27 mm day^{-1}) while others have a smaller spread (10 to 15 mm day^{-1}). The result for P shows values in the range 50–83% with no clear distinction between the different models or regions.

Temporal sensitivity of x_c

By dividing the 30-year long scenario period into three 10-year periods we investigate how x_c and P change with time relative to the control run (cf. Figure 6.2). The top panel of Figure 6.8 shows that there is no significant change in x_c or P during the scenario period 2071-2100 for the full HIRHAM_D model. The same holds for all PRUDENCE models (not shown) and is in accordance with scenario precipitation theory where this consistent change can be explained by a gamma distribution having a single transition point between precipitation rates that contribute more/less to the total precipitation (Gutowski et al., in press). Notable is that the amplitude of the difference

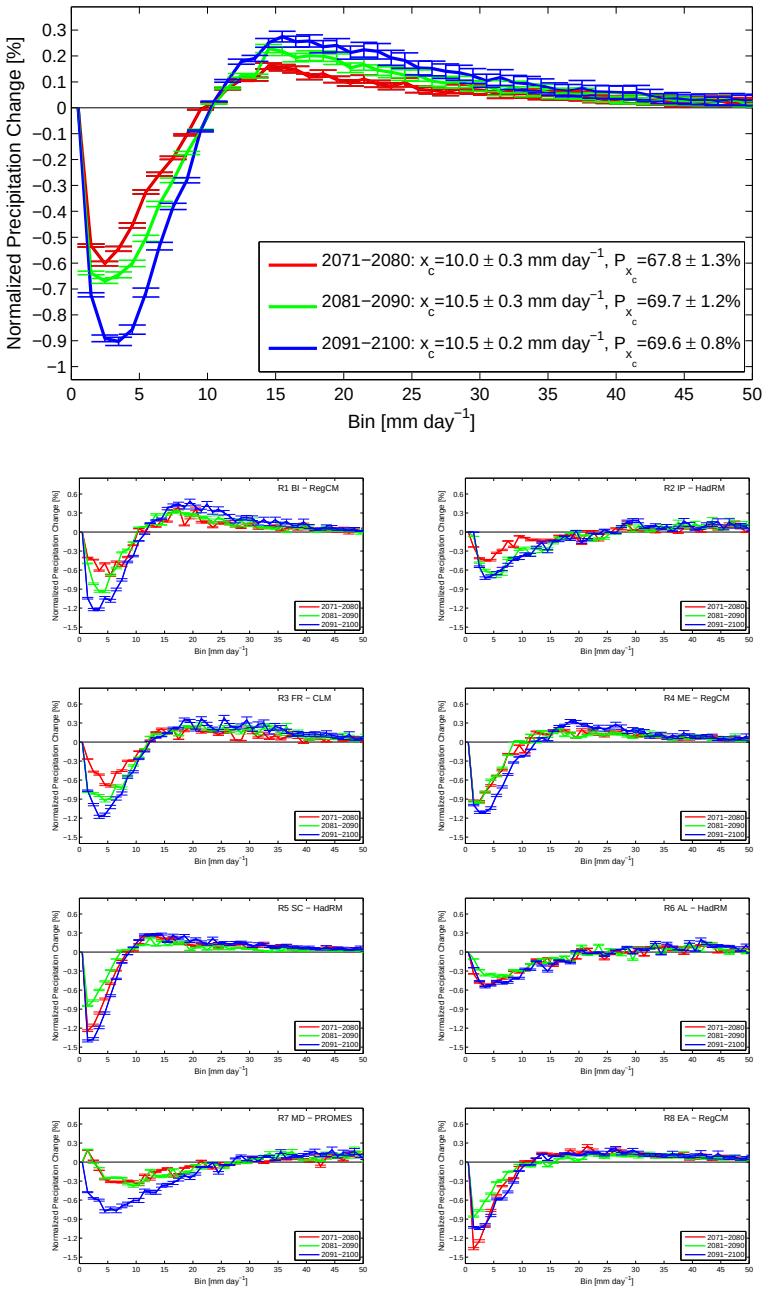

Figure 6.8: Binned normalised precipitation change (scenario minus control run) for three 10-year scenario periods compared to the 30-year control period. The top panel shows the difference for the full HIRHAM_D model land region and the following eight panels represent the individual result for the eight sub-regions, using the models with the best scores in Table 5.2 for the full year, separately. The error bars are calculated using equation 6.1.

between the 10-year scenario periods and the control is increasing with time and the increase is significant. The bottom eight panels of Figure 6.8 show that the same holds when analysing the individual models scoring the best in Table 5.2 for the eight sub-regions individually. A probable exception to this is in the AL sub-region where no noticeable temporal change is seen. The error bars in Figure 6.8 are estimated using Equation 6.1, with U in the range 2-5 for the eight sub-regions.

Figure 6.9 shows trends in ECA compiled precipitation distribution changes for the eight

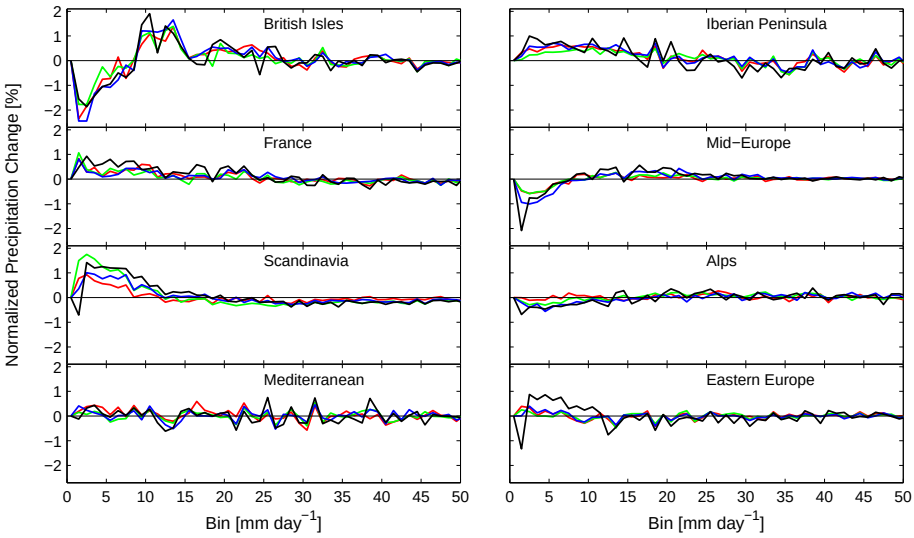

Figure 6.9: Trends in ECA compiled precipitation for the eight PRUDENCE sub-regions. The control period includes observations until 1939. Red, green, blue and black lines represent observations between 1940-1959, 1960-1979, 1980-1999 and 2000-2007 respectively, each subtracted by the control period values.

PRUDENCE sub-regions. The reference period is here taken as all data until 1939 and the test periods are 1940-1959, 1960-1979, 1980-1999 and 2000-2007. Some regions show similar trends as the RCM models shown above whereas other regions show a positive increase in precipitation contribution for moderate precipitation.

Changes in wet day fraction and average precipitation

Figure 6.10 shows the fraction of wet days relative to the total number of days for all RCMs driven by the same GCM together with the average value for the ECA precipitation during the control period. The differences between the control and scenario runs are rather small for the cold period (October–March) while larger differences are visible for the warm period (April–September) giving an overall picture with a larger fraction of wet days during the control period (1961-1990).
Figure 6.11 shows the average precipitation during wet days. We notice large differences between observed and model precipitation for region 2 (IP). The average wet day precipitation is found to be somewhat higher during the scenario period. Figure 6.12 shows the average precipitation during all days of the year. The average precipitation is larger for the scenario run during the cold period while it is smaller during the warm period compared to the control run. These results are consistent with earlier studies of the change in precipitation between the control and scenario simulations (see Fowler et al. (2007) and references therein).

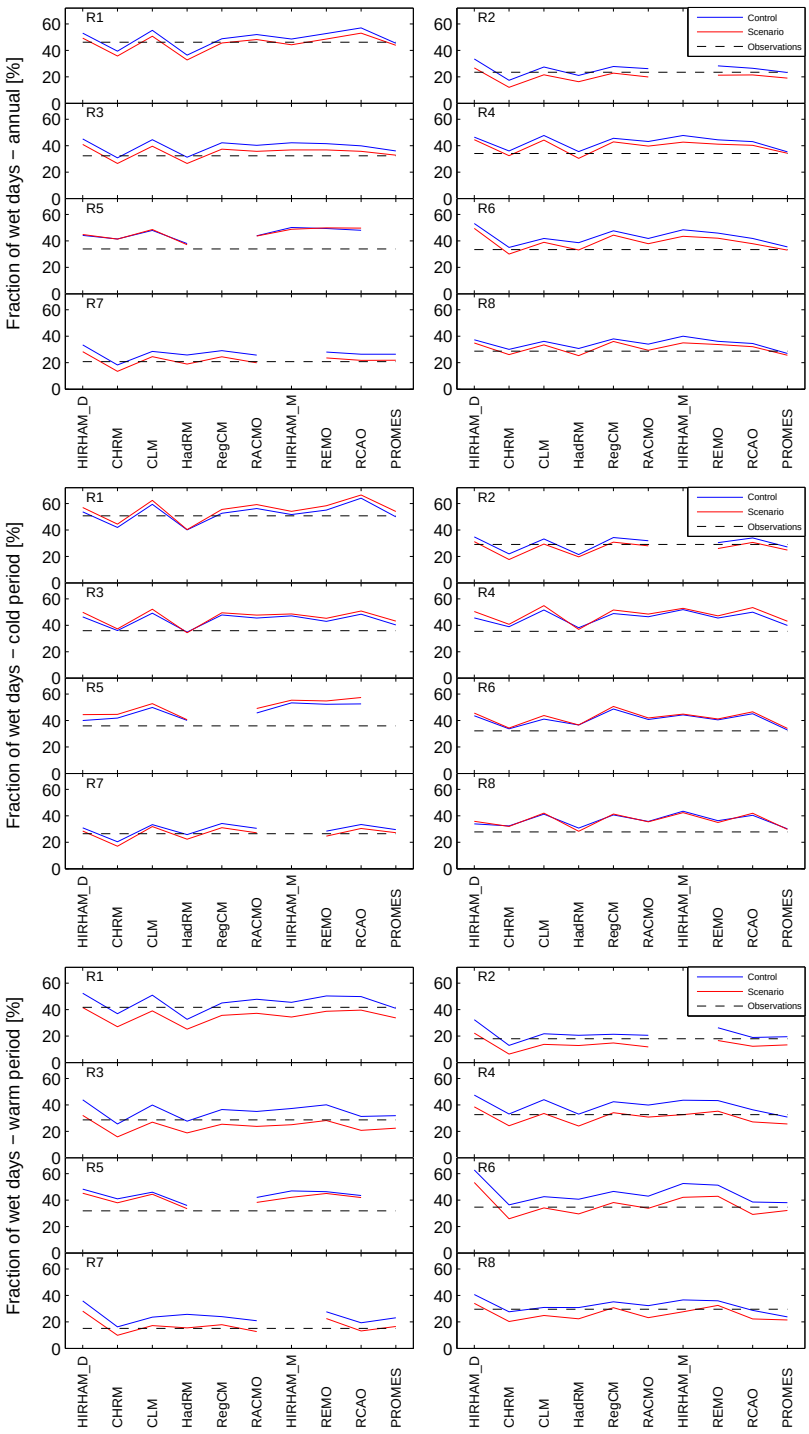

Figure 6.10: Fraction of wet days for ten RCMs and for eight sub-regions. Blue curves represent the control period and red curves are for the scenario period. The horizontal dashed lines are the average values during the control period for ECA observational data for each sub-region. The top panels give values for the full annual period, the middle panels for the cold period and the bottom panels for the warm period.

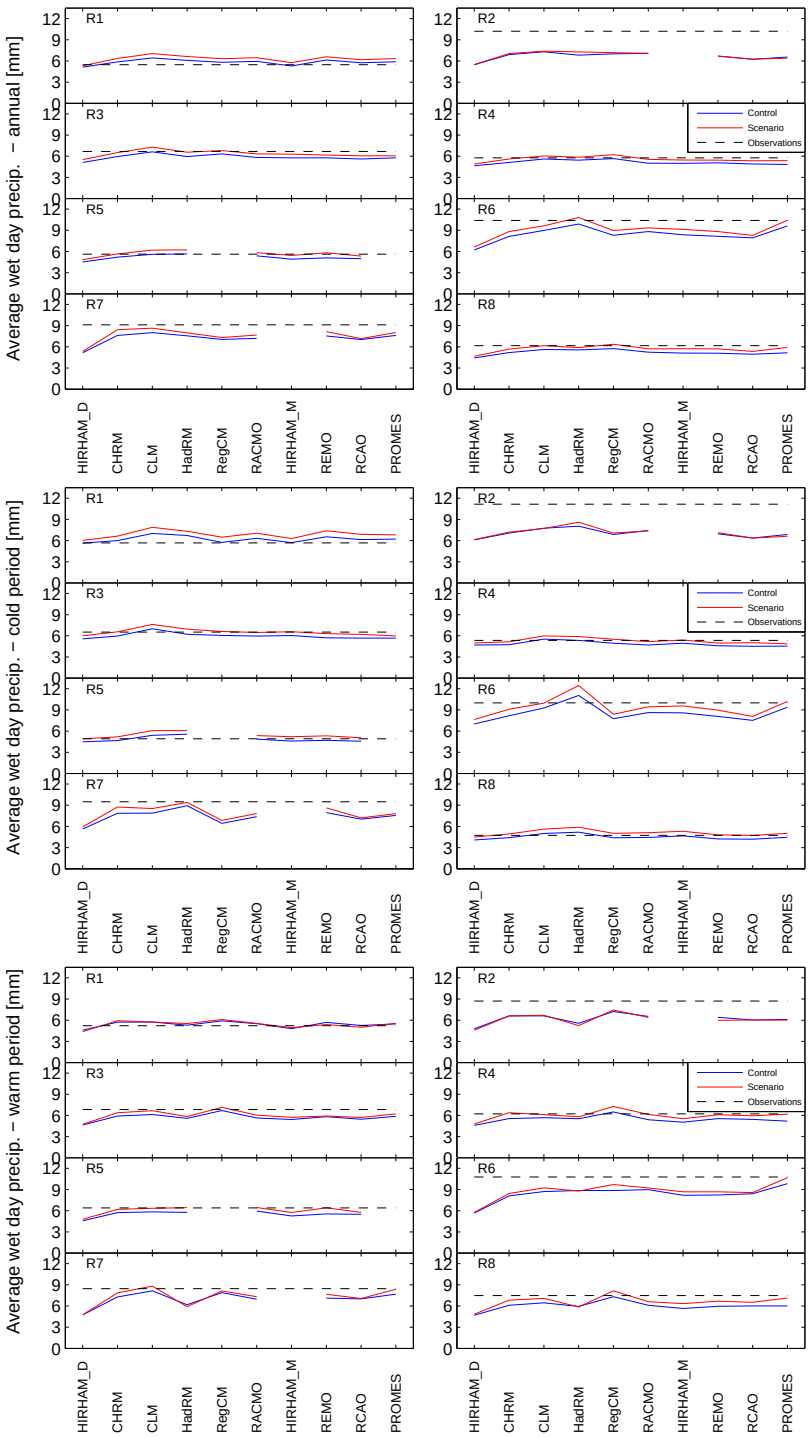

Figure 6.11: Average precipitation for wet days for ten RCMs and for eight sub-regions. Blue curves represent the control period and red curves are for the scenario period. The horizontal dashed lines are the average values during the control period for ECA observational data for each sub-region. The top panels give values for the full annual period, the middle panels for the cold period and the bottom panels for the warm period.

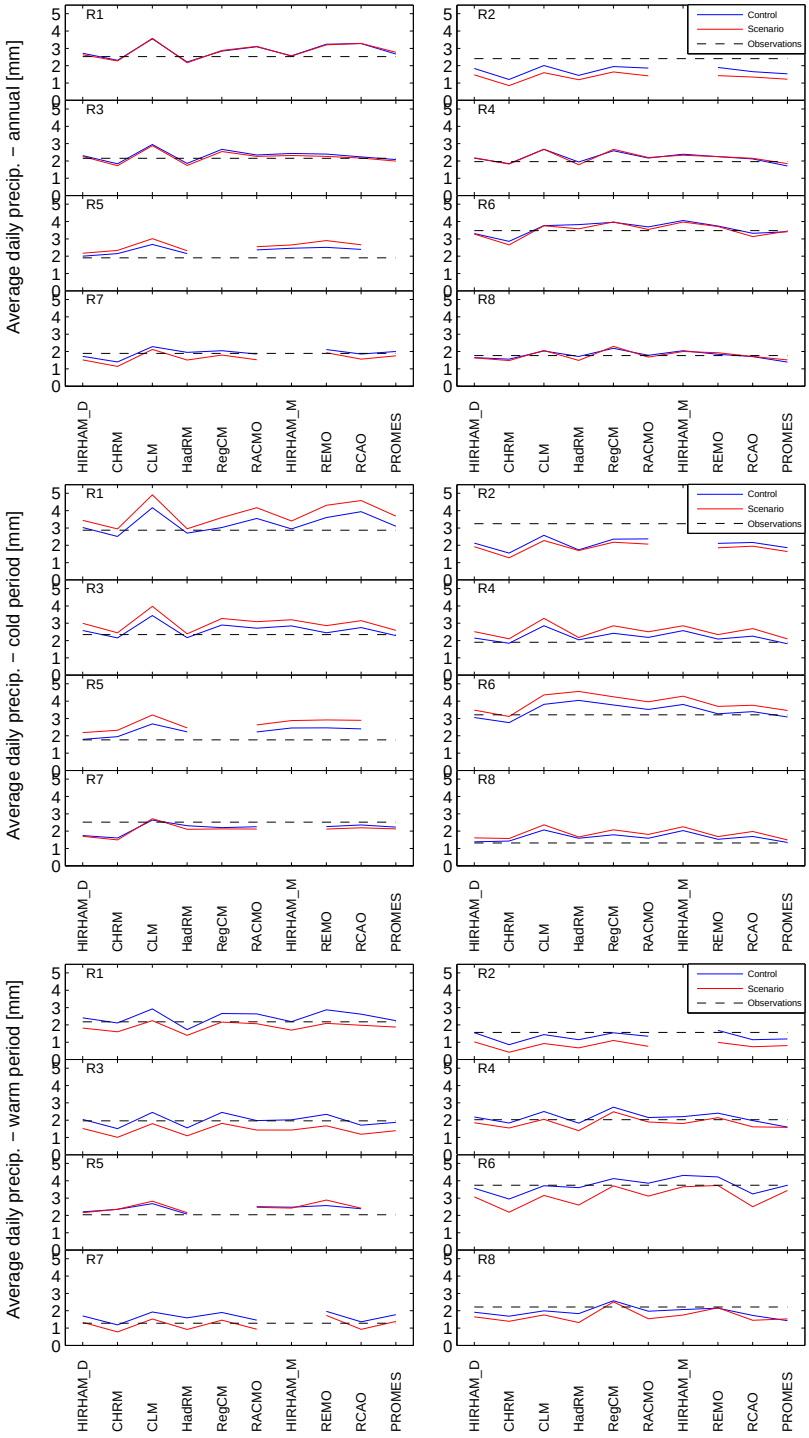

Figure 6.12: Average precipitation for all days for ten RCMs and for eight sub-regions. Blue curves represent the control period and red curves are for the scenario period. The horizontal dashed lines are the average values during the control period for ECA observational data for each sub-region. The top panels give values for the full annual period, the middle panels for the cold period and the bottom panels for the warm period.

7. Conclusions

All of the RCMs analysed in this study show strong similarities with observed precipitation for moderate precipitation during the control period 1961-1990. These similarities have been validated and confirmed using a robustness test. No single model was found to outperform all other models in terms of the metrics we have chosen here. A few models were found to score high in most regions and most time periods whereas a few models were found to almost never score high using our metrics. A test of a linear method of generating an aggregated PDF from several models showed it to perform slightly better than the PDFs of the individual models in some cases.

These results can now be used in two ways. First, we can obtain precipitation PDFs for the scenario period 2071-2100 using models with the best control period scores and then use these PDFs to postulate about future precipitation distribution. Secondly, we can use the analysing and validation methods described in this study on the upcoming ENSEMBLES RCM models (Hewitt, 2005).

Some specific conclusions about using only those RCM models forced by the HadAM3H AGCM: (1) The distribution of consecutive days of drought for the period 1961-1990 is modelled quite well in all of the RCMs, which makes it difficult to determine one single model to perform better than the others in general, even though it is sometimes possible to do so for a specific region. (2) In a study of climate scenarios (2071-2100) relative to the control run (1961-1990), we have found that a significant change in the modelled PDF occurs for model precipitation for the full land region. The change is also evident for specific sub-regions. (3) We have shown that models HadRM, RegCM, and CLM almost always perform better than the others in describing precipitation PDFs. (4) All regions except BI are typically well described by a single top-scoring model while a range of models (CHRM, HadRM, RegCM and RACMO) scores evenly across BI.

In this study we have used the method of common overlap to evaluate the performance of the different RCMs. This method has restrictions when not studying the full PDF. For instance, the RCAO model does score well for metric 3 for all three time periods but it is also found that its PDF is far from similar to the observed PDF. This is because of its relatively narrow distribution with an overestimate for low to moderate precipitation resulting in a skill score equal to or close to the area for the observed PDF for that bin interval.

Acknowledgements Data have been provided through the PRUDENCE data archive, funded by the EU through contract EVK2-CT2001-00132. The present work has been carried out as part of the DMI's contribution to the EU 6th Framewok Programme projects ENSEMBLES (contract number GOCE-CT-2003-505539) and WATCH (contract number GOCE-2006-036946).

References

Buonomo, E., Jones, R.G., Huntingford, C. and Hannaford, J., *On the robustness of changes in extreme precipitation over Europe from two high resolution climate change simulations*, Q. J. R. Meteorol. Soc., in press, 2007.

Christensen, O.B., Christensen, J.H., Machenhauer, B. and Botzet, M., *Very high-resolution regional climate simulations over Scandinavia - Present climate*, Journal of Climate, vol. 11, 3204–3229, 1998.

Christensen, J.H. and Christensen, O., *A summary of the PRUDENCE model projections of changes in European climate by the end of the century*, Climatic Change, vol. 81, 7–30, 2007.

Christensen, J.H., Carter, T.R., Rummukainen, M. and Amanatidis, G., *Evaluating the performance and utility of regional climate models: the PRUDENCE project*, Climatic Change, vol. 81, 1–6, 2007.

Dai, A., *Global precipitation and thunderstorm frequencies. Part I: Seasonal and interannual variations*, Journal of Climate, vol. 14, 1092–1111, 2001.

Efron, B. and Tibshirani, R.J., *An introduction to the bootstrap.* New York: Chapman & Hall, 1993.

Fowler, H.J., Ekström, M., Blenkinsop, S. and Smith, A.P., *Estimating change in extreme European precipitation using a multimodel ensmble*, Journal of Geophys. Res., doi:10.1029/2007JD008619, 2007.

Gutowski Jr., W.J., Kozak, K.A., Arritt, R.W., Christensen, J.H., Patton, J.C. and Takle, E.S., *A possible constraint on regional precipitation intensity changes under global warming*, Journal of Hydrometeorology, in press, 2007.

Hewitt, C.D., *The ENSEMBLES Project: Providing ensemble-based predictions of climate changes and their impacts*, EGGS newsletter, vol. 13, 22–25, 2005.

Nakićenović, N., Alcamo, J., Davis, J., de Vries, B., Fenhann, J., Gaffin, S., Gregory, K., Grübler, A., Jung, T.Y., Kram, T., Lebre La Rovere, E., Michaelis, L., Mori, S., Morita, T., Pepper, W., Pitcher, H., Price, L., Riahi, K., Roehrl, A., Rogner, H.-H., Sankovski, A., Schlesinger, M., Shukla, P., Smith, S., Swart, R., van Rooijen, S., Victor, N. and Dadi, Z., *Special report on emission scenarios*, A special report of Working Group III for the Intergovernmental Panel on Climate Change, Cambridge University Press, 2000.

Perkins, S.E., Pitman, A.J., Holbrook, N.J. and McAneney, J., *Evaluation of the AR4 climate models' simulated daily maximum temperature, minimum temperature, and precipitation over Australia using probability density functions*, Journal of Climate, vol. 20, 4356–4376, 2007.

Pope, V.D., Gallaini, M.L., Rowntree, P.R. and Stratton, R.A., *The impact of new physical parametrizations in the Hadley centre climate model: HadAM3*, Climate Dynamics, vol. 16, 123–146, 2000.

Roeckner, E., Arpe, K., Bengtsson, L., Christoph, M., Claussen, M., Dümenil, L., Esch, M., Giorgetta, M., Schlese, U. and Schulzweida, U., *The atmospheric general circulation model ECHAM-4: model description and simulation of present-day climate*, MPI report 218, p90, Max-Planck-Institut für Meteorologie, 1996.

Rowell, D.P., *A scenario of European climate change for the late twenty-first century; seasonal means and interannual variability*, Climate Dynamics, vol. 25, 837–849, 2005.

8. Appendix

Here we present the complete set of skill scores for all models, all regions, all metrics and all time periods. The result is seen in Tables 8.1, 8.2 and 8.3.

Table 8.1: Results of comparing ten PRUDENCE models, all driven by the same GCM, for precipitation against ECA data, using daily annual values for the period 1961-1990. The table is divided into four parts, one for each metric considered (see Table 5.1). Each part gives the skill score for all models and for all regions. Note that metrics 1 through 3 are based on fractions of the full PDFs and the values given are fractions of the maximum ECA PDF area for the specific metric interval in question. Note also that the method of common overlap have restrictions when not studying the full PDF. See text for further details.

	Skill score for Metric 0								
Model	BI	IP	FR	ME	SC	AL	MD	EA	EUx
HIRHAM_D	0.909	0.646	0.830	0.860	0.867	0.712	0.680	0.775	0.821
CHRM	0.962	0.769	0.911	0.905	0.903	0.862	0.889	0.873	0.918
CLM	0.943	0.813	0.953	0.960	0.964	0.914	0.921	0.927	0.975
HadRM	0.945	0.823	0.945	0.961	0.977	0.936	0.891	0.944	0.965
RegCM	0.966	0.784	0.932	0.966	N/A	0.852	0.857	0.949	0.953
RACMO	0.965	0.802	0.921	0.929	0.962	0.913	0.881	0.904	0.952
HIRHAM_M	0.936	N/A	0.917	0.911	0.913	0.890	N/A	0.884	0.921
REMO	0.966	0.754	0.892	0.914	0.928	0.844	0.890	0.865	0.921
RCAO	0.931	0.717	0.868	0.879	0.888	0.827	0.830	0.852	0.891
PROMES	0.946	0.785	0.935	0.908	N/A	0.930	0.923	0.901	0.960
	Skill score for Metric 1								
Model	BI	IP	FR	ME	SC	AL	MD	EA	EUx
HIRHAM_D	0.619	0.315	0.388	0.447	0.481	0.409	0.335	0.257	0.413
CHRM	0.759	0.518	0.623	0.651	0.542	0.692	0.753	0.565	0.713
CLM	0.960	0.615	0.778	0.837	0.781	0.822	0.856	0.750	0.903
HadRM	0.988	0.664	0.778	0.945	0.949	0.934	0.876	0.800	0.985
RegCM	0.799	0.548	0.673	0.871	N/A	0.656	0.655	0.801	0.803
RACMO	0.869	0.568	0.630	0.696	0.773	0.798	0.698	0.600	0.782
HIRHAM_M	0.750	N/A	0.653	0.682	0.575	0.746	N/A	0.570	0.716
REMO	0.894	0.471	0.510	0.592	0.622	0.633	0.734	0.445	0.659
RCAO	0.483	0.367	0.409	0.422	0.407	0.588	0.548	0.395	0.519
PROMES	0.978	0.567	0.762	0.783	N/A	0.904	0.826	0.649	0.929
	Skill score for Metric 2								
Model	BI	IP	FR	ME	SC	AL	MD	EA	EUx
HIRHAM_D	0.688	0.379	0.485	0.511	0.557	0.477	0.393	0.338	0.486
CHRM	0.826	0.584	0.709	0.676	0.626	0.742	0.780	0.629	0.755
CLM	0.976	0.663	0.846	0.844	0.850	0.843	0.863	0.788	0.923
HadRM	0.993	0.694	0.837	0.933	0.964	0.925	0.862	0.838	0.977
RegCM	0.857	0.610	0.770	0.869	N/A	0.722	0.708	0.837	0.848
RACMO	0.924	0.640	0.736	0.742	0.849	0.837	0.757	0.687	0.842
HIRHAM_M	0.798	N/A	0.732	0.712	0.664	0.794	N/A	0.644	0.765
REMO	0.938	0.554	0.639	0.667	0.711	0.707	0.776	0.558	0.743
RCAO	0.672	0.485	0.557	0.532	0.549	0.675	0.650	0.522	0.643
PROMES	0.987	0.618	0.822	0.778	N/A	0.899	0.852	0.708	0.927
	Skill score for Metric 3								
Model	BI	IP	FR	ME	SC	AL	MD	EA	EUx
HIRHAM_D	0.963	0.983	0.977	0.947	0.956	0.985	0.959	0.952	0.969
CHRM	0.997	1.000	0.998	0.962	0.989	1.000	0.996	0.973	0.991
CLM	0.937	1.000	0.999	0.990	0.999	0.994	0.977	0.984	0.998
HadRM	0.915	0.988	0.989	0.959	0.974	0.941	0.907	0.985	0.947
RegCM	0.994	1.000	1.000	0.993	N/A	1.000	1.000	0.996	1.000
RACMO	0.987	1.000	1.000	0.976	1.000	1.000	1.000	0.994	1.000
HIRHAM_M	0.967	N/A	0.997	0.958	0.989	1.000	N/A	0.982	0.990
REMO	0.977	1.000	1.000	0.977	0.995	1.000	1.000	0.993	1.000
RCAO	0.999	1.000	1.000	0.968	0.995	1.000	1.000	0.989	1.000
PROMES	0.923	0.996	0.979	0.929	N/A	0.962	0.992	0.979	0.969

Table 8.2: Same as Table 8.1 but for cold period (October-March) values.

	Skill score for Metric 0								
Model	BI	IP	FR	ME	SC	AL	MD	EA	EUx
HIRHAM_D	0.943	0.696	0.895	0.914	0.927	0.824	0.762	0.904	0.925
CHRM	0.960	0.765	0.930	0.918	0.910	0.886	0.898	0.940	0.955
CLM	0.909	0.819	0.907	0.985	0.940	0.939	0.909	0.978	0.959
HadRM	0.920	0.856	0.942	0.978	0.934	0.918	0.902	0.937	0.926
RegCM	0.934	0.735	0.913	0.932	N/A	0.834	0.808	0.942	0.936
RACMO	0.948	0.795	0.931	0.947	0.951	0.933	0.892	0.981	0.966
HIRHAM_M	0.951	N/A	0.931	0.944	0.923	0.900	N/A	0.982	0.961
REMO	0.939	0.748	0.902	0.897	0.959	0.873	0.902	0.921	0.959
RCAO	0.927	0.690	0.883	0.865	0.912	0.827	0.822	0.909	0.917
PROMES	0.926	0.776	0.928	0.913	N/A	0.915	0.908	0.960	0.970
	Skill score for Metric 1								
Model	BI	IP	FR	ME	SC	AL	MD	EA	EUx
HIRHAM_D	0.754	0.418	0.520	0.715	0.606	0.590	0.483	0.613	0.661
CHRM	0.796	0.543	0.643	0.633	0.411	0.746	0.799	0.657	0.800
CLM	0.955	0.647	0.728	0.937	0.790	0.867	0.836	0.967	0.955
HadRM	0.987	0.728	0.755	0.981	0.963	0.962	0.913	0.989	0.988
RegCM	0.642	0.475	0.549	0.593	N/A	0.602	0.561	0.749	0.665
RACMO	0.857	0.576	0.642	0.724	0.748	0.847	0.733	0.849	0.891
HIRHAM_M	0.763	N/A	0.642	0.725	0.469	0.761	N/A	0.853	0.786
REMO	0.919	0.506	0.517	0.518	0.712	0.697	0.806	0.553	0.786
RCAO	0.514	0.358	0.411	0.311	0.375	0.583	0.551	0.419	0.550
PROMES	0.972	0.569	0.701	0.759	N/A	0.882	0.796	0.875	0.974
	Skill score for Metric 2								
Model	BI	IP	FR	ME	SC	AL	MD	EA	EUx
HIRHAM_D	0.806	0.488	0.640	0.726	0.679	0.661	0.550	0.675	0.733
CHRM	0.849	0.604	0.749	0.664	0.552	0.784	0.819	0.741	0.839
CLM	0.974	0.694	0.811	0.946	0.858	0.887	0.850	0.981	0.969
HadRM	0.985	0.759	0.829	0.976	0.975	0.957	0.906	0.993	0.992
RegCM	0.738	0.550	0.686	0.667	N/A	0.680	0.628	0.790	0.752
RACMO	0.919	0.651	0.751	0.767	0.830	0.872	0.788	0.911	0.925
HIRHAM_M	0.816	N/A	0.751	0.757	0.594	0.807	N/A	0.909	0.849
REMO	0.954	0.574	0.655	0.581	0.793	0.754	0.825	0.645	0.843
RCAO	0.698	0.472	0.577	0.432	0.540	0.666	0.647	0.557	0.678
PROMES	0.972	0.625	0.792	0.748	N/A	0.880	0.829	0.892	0.976
	Skill score for Metric 3								
Model	BI	IP	FR	ME	SC	AL	MD	EA	EUx
HIRHAM_D	0.975	0.999	0.993	0.940	0.977	1.000	0.984	0.924	0.991
CHRM	0.988	1.000	0.999	0.964	0.990	0.996	0.983	0.969	0.995
CLM	0.897	1.000	0.954	0.999	0.993	0.992	0.970	0.987	0.971
HadRM	0.892	1.000	0.979	0.970	0.945	0.863	0.883	0.934	0.907
RegCM	0.985	1.000	1.000	0.986	N/A	1.000	0.998	0.959	1.000
RACMO	0.967	1.000	1.000	0.981	0.990	0.997	1.000	0.995	0.988
HIRHAM_M	0.984	N/A	1.000	0.980	0.997	1.000	N/A	0.992	1.000
REMO	0.929	1.000	0.996	0.954	0.996	1.000	0.983	0.963	0.998
RCAO	0.992	1.000	1.000	0.945	0.996	1.000	1.000	0.963	1.000
PROMES	0.898	1.000	0.973	0.932	N/A	0.943	0.992	0.963	0.955

Table 8.3: Same as Table 8.1 but for warm period (April-September) values.

	Skill score for Metric 0								
Model	BI	IP	FR	ME	SC	AL	MD	EA	EUx
HIRHAM_D	0.861	0.595	0.742	0.798	0.814	0.612	0.610	0.687	0.728
CHRM	0.938	0.783	0.873	0.894	0.894	0.833	0.866	0.854	0.888
CLM	0.963	0.807	0.914	0.924	0.941	0.876	0.924	0.905	0.933
HadRM	0.942	0.778	0.902	0.934	0.944	0.908	0.839	0.889	0.924
RegCM	0.953	0.870	0.945	0.965	N/A	0.864	0.914	0.954	0.967
RACMO	0.963	0.822	0.889	0.913	0.953	0.886	0.857	0.867	0.915
HIRHAM_M	0.901	N/A	0.864	0.880	0.904	0.866	N/A	0.843	0.882
REMO	0.942	0.791	0.875	0.921	0.901	0.818	0.865	0.834	0.888
RCAO	0.910	0.774	0.851	0.899	0.873	0.825	0.844	0.845	0.877
PROMES	0.952	0.808	0.920	0.909	N/A	0.930	0.932	0.887	0.942
	Skill score for Metric 1								
Model	BI	IP	FR	ME	SC	AL	MD	EA	EUx
HIRHAM_D	0.432	0.149	0.242	0.297	0.406	0.260	0.173	0.184	0.249
CHRM	0.679	0.463	0.582	0.667	0.594	0.633	0.658	0.599	0.662
CLM	0.847	0.544	0.691	0.788	0.775	0.751	0.853	0.755	0.819
HadRM	0.872	0.562	0.707	0.909	0.812	0.864	0.752	0.719	0.873
RegCM	0.883	0.709	0.772	0.953	N/A	0.696	0.780	0.841	0.884
RACMO	0.817	0.559	0.591	0.693	0.778	0.738	0.624	0.583	0.704
HIRHAM_M	0.692	N/A	0.638	0.683	0.656	0.717	N/A	0.563	0.681
REMO	0.678	0.449	0.495	0.620	0.568	0.579	0.625	0.444	0.578
RCAO	0.425	0.396	0.418	0.512	0.437	0.592	0.541	0.452	0.521
PROMES	0.903	0.579	0.772	0.818	N/A	0.898	0.849	0.662	0.887
	Skill score for Metric 2								
Model	BI	IP	FR	ME	SC	AL	MD	EA	EUx
HIRHAM_D	0.515	0.213	0.317	0.368	0.485	0.326	0.224	0.256	0.320
CHRM	0.756	0.546	0.639	0.690	0.662	0.694	0.705	0.650	0.707
CLM	0.898	0.600	0.743	0.780	0.823	0.779	0.853	0.779	0.832
HadRM	0.904	0.592	0.756	0.892	0.863	0.860	0.744	0.747	0.866
RegCM	0.918	0.736	0.831	0.953	N/A	0.751	0.811	0.877	0.908
RACMO	0.881	0.624	0.671	0.735	0.844	0.792	0.683	0.659	0.764
HIRHAM_M	0.735	N/A	0.659	0.700	0.723	0.760	N/A	0.622	0.716
REMO	0.785	0.556	0.615	0.710	0.666	0.667	0.702	0.560	0.680
RCAO	0.580	0.519	0.545	0.627	0.568	0.679	0.654	0.586	0.645
PROMES	0.935	0.620	0.804	0.818	N/A	0.896	0.866	0.723	0.888
	Skill score for Metric 3								
Model	BI	IP	FR	ME	SC	AL	MD	EA	EUx
HIRHAM_D	0.942	0.942	0.945	0.938	0.936	0.969	0.933	0.947	0.948
CHRM	0.995	1.000	0.987	0.957	0.987	1.000	1.000	0.979	0.987
CLM	0.985	0.997	0.995	0.969	0.987	0.995	0.977	0.980	0.988
HadRM	0.940	0.947	0.966	0.934	0.971	0.968	0.903	0.972	0.949
RegCM	0.973	0.995	1.000	0.974	N/A	1.000	1.000	1.000	0.999
RACMO	0.992	1.000	0.995	0.968	0.996	1.000	1.000	0.993	0.996
HIRHAM_M	0.931	N/A	0.957	0.929	0.971	0.997	N/A	0.975	0.969
REMO	0.991	1.000	1.000	0.992	0.994	1.000	1.000	0.999	1.000
RCAO	0.998	1.000	0.998	0.989	0.994	1.000	1.000	1.000	1.000
PROMES	0.951	0.979	0.971	0.925	N/A	0.973	0.983	0.984	0.967

9. Previous reports

Previous reports from the Danish Meteorological Institute can be found on:
http://www.dmi.dk/dmi/dmi-publikationer.htm