
STATISTICAL DOWNSCALING OF FUTURE HOURLY
PRECIPITATION EXTREMES IN THE UK USING REGIONAL
CLIMATE MODELS AND CIRCULATION PATTERNS

By

Markus Tobias Rau

Thesis

Submitted to the University of East Anglia

For the degree of

Doctor of Philosophy

Supervisors: Dr.-Ing. Yi He, Dr. Clare Goodess and Prof. András Bárdossy

Tyndall Centre for Climate Change Research and Climatic Research Unit

School of Environmental Sciences

September 2016

© This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with the author and that no quotation from the thesis, nor any information derived therefrom, may be published without the author's prior, written consent

Abstract

Observational trends, physical reasons and modelling results suggest an increase in extreme precipitation with climate warming. In particular, sub-daily precipitation extremes are expected to increase heavily raising concerns about the future impacts of flash floods in urban environments and for small or steep river catchments. In order to quantify the potential risk of flash floods in the future, impact studies often require site-specific sub-daily estimates of precipitation extremes. But in their current stage, most Regional Climate Models (RCMs) are only able to provide areal averaged projections at ca. 12.5km resolution and simulated sub-daily precipitation extremes tend to be heavily biased. As a result, statistical downscaling methods are needed to provide site-specific more reliable projections of sub-daily precipitation extremes.

In this thesis, a statistical downscaling method was developed to project site-specific future hourly precipitation extremes over the UK. Circulation patterns (CPs) were classified using a fuzzy rules based approach to categorize extreme hourly precipitation events according to their corresponding atmospheric conditions. In a next step, an analogue day method was applied to find the most similar day in the past by comparing the RCM simulated daily precipitation and temperature with the observations for each CP. The daily maximum hourly precipitation record on the most similar day was extracted and perturbed based on precipitation duration-temperature relationships conditioned on CPs. Within the field of statistical downscaling techniques, the applied method is best described as a hybrid of the analogue and the regression-based method. It was shown that the method is capable of reproducing observed extreme hourly precipitation over different validation periods. Projections based on the applied statistical downscaling method indicate increases in UK hourly extremes but with high variations depending on the twelve different stations, the two future time periods, the two emission scenarios and the four different GCM-driven RCMs.

Acknowledgements

I would like to thank my supervisors, Yi He, Clare Goodess and András Bárdossy, for their continuous advice, patience, support and dedication throughout the period of my PhD.

I am very grateful to the Tyndall Centre for Climate Change Research which provided the funding for my PhD and also gave me numerous opportunities of getting to know very inspiring academics in the field of climate change research often in a relaxed and very welcoming environment.

I would also like to thank the Climatic Research Unit for providing me a comfortable and quiet research environment to work in. It was also a place where I could find valuable scientific ideas and advice throughout the period of my PhD. I also want to mention all the coffee breaks I had with David Lister, “Harry” Harris, Clive Wilkinson etc. Their friendly, calm and steady nature was often exactly what I needed.

I would also like to thank my two examiners, Stephen Blenkinsop and Tim Osborn, for their interest in my work, and for their thoughtful and constructive comments.

I should not forget the continuous support of my family. I know that you are always there for me. Thank you!!!

Contents

Abstract	i
Acknowledgements	iii
List of tables	viii
List of figures	xi
Chapter 1 Introduction	1
1.1 Motivation.....	1
1.2 Research aim and research objectives.....	2
1.3 Study approach	3
1.4 Thesis structure.....	6
Chapter 2 Literature review	7
2.1 Changes in extreme daily and sub-daily precipitation events	7
2.1.1 Observed trends.....	7
2.1.2 Causes	8
2.1.3 Future projections.....	10
2.2 Climate modelling	12
2.2.1 Global climate models (GCMs).....	12
2.2.2 Regional climate models (RCMs) - Dynamical downscaling.....	12
2.3 Statistical downscaling.....	16
2.3.1 Introduction	16
2.3.2 Perfect prognosis (PP).....	17
2.3.3 Model output statistics (MOS).....	25
2.3.4 Weather generator (WG).....	27
2.3.5 Summary	28
2.4 The use of atmospheric circulation patterns (CPs)	32
2.4.1 Introduction	32
2.4.2 The input configuration of CP-classifications.....	33
2.4.3 Examples of different CP-classification methods.....	37
2.4.4 Examples of CP-classifications over the UK and their relationships with precipitation.....	41
2.4.5 Limitations of CP-classifications.....	44
2.4.6 Evaluation of CP-classification methods	45

2.4.7	Summary.....	47
Chapter 3 Data description and selection.....		50
3.1	Observed daily and hourly station precipitation data.....	50
3.1.1	Introduction.....	50
3.1.2	Quality control and pre-screening.....	51
3.1.3	Station selection (Statistical Downscaling).....	55
3.1.4	Station selection for the calibration of fuzzy rules based CP-classifications.....	69
3.1.5	Precipitation analysis of the selected stations.....	70
3.2	Historical gridded data sets.....	82
3.2.1	NCAR Sea Level Pressure (SLP) data set.....	82
3.2.2	E-OBS temperature data set.....	83
3.3	Regional Climate Model (RCM) data.....	87
3.4	Summary.....	89
Chapter 4 Methodology.....		91
4.1	Introduction.....	91
4.2	CP-classification based on fuzzy rules.....	93
4.2.1	The spatial domain of the fuzzy rules CP-classification.....	98
4.2.2	The quality assessment of the fuzzy rules CP-classifications.....	99
4.3	The statistical downscaling process.....	101
Chapter 5 The fuzzy rules based CP-classification.....		108
5.1	Input calibration parameters.....	108
5.1.1	Set of precipitation time series.....	108
5.1.2	Calibration time periods.....	111
5.1.3	Number of CPs.....	111
5.2	Assessment of the performances of the different CP-classifications.....	111
5.2.1	Results.....	112
5.2.2	The final input calibration parameters.....	117
5.3	The final CP-classification used for statistical downscaling.....	117
5.3.1	Atmospheric pressure patterns of the CP-classification.....	117
5.3.2	The relationships between CPs and extreme hourly precipitation.....	123
5.4	Summary.....	129
Chapter 6 Statistical downscaling based on observed records.....		131
6.1	Relevance of the predictors in terms of extreme hourly precipitation.....	131
6.2	Relationships between precipitation duration and temperature.....	134

6.3	Cross-validation to assess statistical downscaling performance	141
6.4	Summary	150
Chapter 7 Statistical downscaling based on RCM data		153
7.1	Cross-validation to assess statistical downscaling performance	153
7.2	Baseline projections based on GCM-driven RCM data	164
7.3	Future projections based on GCM-driven RCM data	170
7.3.1	Projections of extreme hourly precipitation under RCP8.5	171
7.3.2	Projections of extreme hourly precipitation under RCP4.5	175
7.4	Summary	188
Chapter 8 Summary and Discussion		190
8.1	Summary	190
8.2	Limitations and potential of future developments.....	196
8.3	Concluding remarks	199
Bibliography.....		202

List of tables

Table 2.1 Summary of the advantages and disadvantages of different statistical downscaling methods to estimate precipitation.....	31
Table 2.2 Examples of different input configurations of objective CP-classifications	36
Table 3.1 The different stages of the Met Office climate quality control.....	53
Table 3.2 Numbers of hourly/daily values that are set to missing or to zero due to certain criteria described in Section 3.1.2.	54
Table 3.3 Results of the comparison of 24hr aggregated hourly (>1mm) and daily values (>1mm) for 34 stations described in Section 3.1.3.2..	59
Table 3.4 Detected suspicious records of daily precipitation values by comparing daily with 24hr aggregated precipitation values.....	60
Table 3.5 Step-wise elimination of hourly precipitation stations (initially 530) depending on the specific criterion.	63
Table 3.6 Details of the twelve final selected hourly precipitation stations following the selection criteria in Section 3.1.3 for the time period 1980-2009	64
Table 3.7 Details of the three hourly precipitation stations used in three different precipitation Sets to define fuzzy rules CP-classification over the calibration period (1960-1989).....	70
Table 3.8 Details of the hourly precipitation stations used in the fourth precipitation Set to define fuzzy rules CP-classification over the calibration period (1960-1989).....	70
Table 3.9 Annual precipitation statistics over the time period 1980-2009 derived from the aggregated hourly records at the twelve selected stations	72
Table 3.10 Annual precipitation statistics over the time period 1960-1989 derived from the aggregated hourly records of the three hourly precipitation stations used for the calibration of the fuzzy-rules CP-classification	72
Table 3.11 Annual temperature statistics for the twelve selected stations between 1980 and 2009. Mean temperature and 99.5 th percentiles are given in Celsius	84
Table 3.12 List of the 12.5km RCM runs available from the EURO-CORDEX data archive (CORDEX, 2014).	88
Table 4.1 Four different approaches of subsampling in order to find the most similar day in the past.	102
Table 5.1 Overview of the input precipitation series used for the calibration of fuzzy rules CP-classifications.....	108
Table 5.2 Overview of the different calibration time periods used for the fuzzy rules CP-classification.	111
Table 5.3 Results of the information measures to assess the CP-classification performances	114
Table 5.4 Results of the information measures to assess the CP-classification performances	114
Table 5.5 Seasonal frequency occurrences of the five CPs (+one unclassified) between 1980 and 2009 in [%].....	118
Table 5.6 Maximum hourly precipitation between 1980-2009 for each station and the corresponding daily precipitation, precipitation duration, date and assigned CPs using the NCAR CP-classification.....	126
Table 6.1 The linear regression slopes for the precipitation duration and temperature relationship in summer conditioned on the set of CPs over the three different calibration periods and for the twelve selected stations.....	139
Table 6.2 The different calibration periods and corresponding validation periods used for the cross-validation of the statistical downscaling performance in Section 6.3 and Section 7.1. ...	142

Table 6.3 The averaged 99.5 th percentile error $PE_{99.5}$ in [%] depending on the four different approaches to find the analogue days over the three different validation periods	145
Table 6.4 Same as Table 6.3, but over the ten year period of warmest summers for each station between 1980 and 2009	146
Table 6.5 The linear regression slopes in [mm/year] (using ordinary least squares fit) of the residual time series of the annual maxima for the twelve stations between 1980 and 2009...	146
Table 7.1 The averaged 99.5 th percentile error $PE_{99.5}$ in [%] depending on the four different approaches to find the analogue days over the three different validation periods	157
Table 7.2 Same as Table 7.1 but RACMO22E RCM data is used as the predictors for the statistical downscaling instead of RCA4 RCM data	158
Table 7.3 Same as Table 7.1 but over the ten year period of warmest summers on average for each station between 1980 and 2009 using RCA4 RCM data as the predictors for the statistical downscaling... ..	158
Table 7.4 Same as Table 7.1 but over the ten year period of warmest summer on average for each station between 1980 and 2009 using RACMO22E RCM data as the predictors for the statistical downscaling	159
Table 7.5 The linear regression slopes in [mm/year] (using ordinary least squares fit) of the residual time series of the annual maxima of hourly precipitation for the twelve stations between 1980 and 2009.. ..	159
Table 7.6 Same as Table 7.5 but RACMO22E RCM data is used as the predictors for the statistical downscaling	159
Table 7.7 The averaged 99.5 th percentile error $PE_{99.5}$ in [%] over the reference period 1980-2005 depending on the GCM-driven RCM.....	166
Table 7.8 Observed and simulated frequency occurrences of the five CPs (plus one unclassified) over the reference period 1980-2005 in summer in [%].....	166
Table 7.9 Same as Table 7.8 but in winter.....	166
Table 7.10 Changes in [°C] in the simulated daily mean temperature in summer for the future time period 2030-2055 under the emission scenario RCP4.5 relative to the reference time period 1980-2005.....	176
Table 7.11 same as Table 7.10 but changes are calculated for the future time period 2075-2100 under emission scenario RCP4.5 relative to the reference time period 1980-2005.	176
Table 7.12 same as Table 7.10 but changes are calculated for the future time period 2075-2100 under emission scenario RCP8.5 relative to the reference time period 1980-2005.	177
Table 7.13 Changes in [%] in the simulated daily mean precipitation in summer for the future time period 2030-2055 under the emission scenario RCP4.5 relative to the reference time period 1980-2005.....	177
Table 7.14 same as Table 7.13 but changes are calculated for the future time period 2075-2100 under emission scenario RCP4.5 relative to the reference time period 1980-2005.	177
Table 7.15 same as Table 7.13 but changes are calculated for the future time period 2075-2100 under emission scenario RCP8.5 relative to the reference time period 1980-2005.	178
Table 7.16 Changes in [%] in the simulated 99.5 th daily precipitation percentile in summer for the future time period 2030-2055 under the emission scenario RCP4.5 relative to the reference time period 1980-2005.	178
Table 7.17 same as Table 7.16 but changes are calculated for the future time period 2075-2100 under emission scenario RCP4.5 relative to the reference time period 1980-2005.	178
Table 7.18 same as Table 7.16 but changes are calculated for the future time period 2075-2100 under emission scenario RCP8.5 relative to the reference time period 1980-2005.	179

Table 7.19 Changes in [%] in the estimated median of the 99.5th hourly precipitation percentile based on the 100 bootstrapping samples for the future time period 2030-2055 under the emission scenario RCP4.5 relative to the reference time period 1980-2005. 179

Table 7.20 same as Table 7.19 but changes are calculated for the future time period 2075-2100 under emission scenario RCP4.5 relative to the reference time period 1980-2005. 179

Table 7.21 same as Table 7.19 but changes are calculated for the future time period 2075-2100 under emission scenario RCP8.5 relative to the reference time period 1980-2005. 180

Table 7.22 same as Table 7.21 but the 99.5th hourly precipitation percentiles are estimated without applying the final perturbation step based on the precipitation duration – temperature relationship (see Equation 4.20 and 4.21)..... 180

List of figures

Figure 3.1 Scatter plot of daily precipitation (horizontal axis) and 24-hour aggregated precipitation values (vertical axis) recorded on the same day for the five stations with $\rho_{X,Y} < 0.9$ or $NRMSD > 0.05$ (failing selection criterion defined in 3.1.3.2).....	60
Figure 3.2 Scatter plot of daily precipitation (horizontal axis) and 24-hour aggregated precipitation values (vertical axis) on the same day for the final selected stations with $\rho_{X,Y} > 0.9$ and $NRMSD < 0.05$ (fulfilling all selection criteria).....	61
Figure 3.3 Step-wise elimination of UK hourly precipitation stations. Left: All available 530 hourly precipitation stations; middle: 34 hourly precipitation stations that provide records between 1980-2009 (criterion defined in Section 3.1.3.1); right: 30 hourly precipitation stations that fulfil criteria defined in 3.1.3.1 and 3.1.3.2.....	65
Figure 3.4 The twelve final selected hourly precipitation stations across the UK following the selection criteria in 3.1.3.....	66
Figure 3.5 The twelve final selected hourly precipitation stations across the UK following the selection criteria in 3.1.3 with the HadUKP regions overlapped on the extreme precipitation regions (Jones et al., 2014).	67
Figure 3.6 The twelve final selected hourly precipitation stations across the UK following the selection criteria in 3.1.3 with the elevation map.	68
Figure 3.7 Monthly precipitation averages of the twelve selected stations (blue solid line with asterisk markers) and of the corresponding HadUKP regions (red dotted line with square markers) averaged over the time period 1980-2009.....	73
Figure 3.8 Monthly precipitation averages of the three stations used for the calibration of the fuzzy rules CP-classification (blue line) and of the corresponding HadUKP regions (red line) averaged over the time period 1960-1989.....	74
Figure 3.9 Annual cycle of the 99.5 th percentile of daily precipitation between 1980 and 2009 for the twelve selected stations.	75
Figure 3.10 The daily maximum hourly precipitation distributions between 1980 and 2009 for the twelve selected stations.	79
Figure 3.11 Annual cycle of the 99.5 th percentiles of the daily maximum hourly precipitation between 1980 and 2009 for the twelve selected stations.	80
Figure 3.12 Annual cycle of the frequency occurrences of daily maximum hourly precipitation (>10mm/hr) between 1980 and 2009 for the twelve selected stations.....	81
Figure 3.13 Annual cycle of the monthly 99.5 th percentiles of the daily mean temperature for the twelve selected stations between 1980 and 2009.....	85
Figure 3.14 Observed trends in temperature between 1980 and 2009 for the twelve selected stations. All trends are significant at the 5% level based on the p-value test.....	86
Figure 3.15 Illustration of the spatial domain (contoured area) of the EURO-CORDEX data set. It shows the daily SLP simulated by the RCA4 RCM driven by the ERA-interim reanalysis data. ...	89
Figure 4.1 Flowchart of the applied method to downscale extreme hourly precipitation. SLP stands for sea level pressure, P for precipitation and T for temperature.	93
Figure 4.2 The membership functions of the fuzzy sets $v = 1, \dots, 4$	96
Figure 4.3 The spatial domains of the NCAR gridded SLP data set (Hurrell and Trenberth, 2013), the RCM data sets the fuzzy rules CP-classification	99
Figure 4.4 Subsample of calibration days tC occurring on days classified as CP3 in winter between 1980 and 1999 at the station of Boulmer.....	103

Figure 4.5 Same as Figure 4.4 but using daily precipitation as an additional subsample criterion.....	105
Figure 5.1 Hourly precipitation stations and HadUKP regions used for the calibration of the fuzzy rules based CP-classification.	110
Figure 5.2 Results of the assessment of the CP-classification performances	113
Figure 5.3 Spatial variability of the CP-classification in terms of the Information Measure 1 for the twelve selected stations using different input precipitation Sets	116
Figure 5.4 SLP patterns in [hPa] of the five CPs (+ one unclassified CP) derived from fuzzy rules using the final input configuration described in Section 5.2.2 for the validation time period 1980-2009 in winter (Nov-April).....	120
Figure 5.5 Same as Figure 5.4 but for the validation time period 1980-2009 in summer (May-Sept).....	121
Figure 5.6 SLP anomalies of the five CPs (+ one unclassified CP) derived from fuzzy rules using the final input configuration described in Section 5.2.2 for the validation time period 1980-2009 over the entire year.....	122
Figure 5.7 Extreme hourly precipitation (>10 mm/hr) probability in winter (Nov-April) conditioned on CPs (frequency occurrences: CP1=5.5%, CP2=14.7%, CP3=29.4%, CP4=23.0%, CP5=20.1% and U=7.4%) of the NCAR CP-classification for the twelve selected stations between 1980 and 2009.....	127
Figure 5.8 Same as Figure 5.7 but for summer (May-Oct) conditioned on CPs (frequency occurrences: CP1=7.6%, CP2=15.5%, CP3=29.0%, CP4=24.9%, CP5=18.2% and U=4.9%).....	127
Figure 5.9 99.5 th percentiles of the daily maximum hourly precipitation in winter (Nov-April) conditioned on CPs (frequency occurrences: CP1=5.5%, CP2=14.7%, CP3=29.4%, CP4=23.0%, CP5=20.1% and U=7.4%) of the NCAR CP-classification for the twelve selected stations between 1980 and 2009.....	128
Figure 5.10 Same as Figure 5.9 but for summer (May-Oct) conditioned on CPs (frequency occurrences: CP1=7.6%, CP2=15.5%, CP3=29.0%, CP4=24.9%, CP5=18.2% and U=4.9%).....	128
Figure 6.1 Scatter plot of daily precipitation events (>=0.2mm/day) recorded in summer between 1980 and 2009 at Northolt.....	132
Figure 6.2 Same as Figure 6.1 but for the station at Marham.	133
Figure 6.3 Same as Figure 6.1 but for the station at Cranwell.....	133
Figure 6.4 The relationship between daily precipitation duration and temperature at Camborne in summer (May-Oct) for the time period 1980-1999 conditioned on the set of CPs.	140
Figure 6.5 The values of the linear regression slopes between precipitation duration and temperature in summer conditioned on the set of CPs for the three different calibration periods at Camborne.....	141
Figure 6.6 The part exceeding 0.99 of the empirical cumulative distribution of the estimated and observed daily maximum hourly precipitation records for the most recent validation period 2000-2009.....	147
Figure 6.7 The empirical cumulative distribution exceeding 0.99 of the estimated and observed daily maximum hourly extreme precipitation records for the most recent validation period 2000-2009.....	148
Figure 6.8 The residual time series of the annual maxima of hourly precipitation for the twelve stations between 1980 and 2009 using the four different statistical downscaling methods...	149
Figure 6.9 Scatter plot of daily precipitation events (>=0.2mm/day) recorded in summer at Northolt between 1980 and 1999 (calibration period C1) conditioned on CP3	150

Figure 7.1 The part exceeding 0.99 of the empirical cumulative distribution of the estimated and observed daily maximum hourly precipitation for the most recent validation period 2000-2009	160
Figure 7.2 Same as Figure 7.1 but RACMO22E RCM data is used as the predictor for the statistical downscaling.	161
Figure 7.3 The residual time series of the annual maxima of hourly precipitation for the twelve stations between 1980 and 2009 using the four different statistical downscaling methods based on RCA4 RCM data.	162
Figure 7.4 Same as Figure 7.3 but RACMO22E RCM data is used instead of RCA4 RCM data.	163
Figure 7.5 Estimated and observed hourly precipitation in summer over the reference period 1980-2005 at Boscombe Down.....	167
Figure 7.6 Distribution (above 0.6) of the simulated and observed daily mean precipitation in summer over the reference period 1980-2005 at Boscombe Down.....	168
Figure 7.7 Distribution of the simulated and observed daily temperature in summer over the reference period 1980-2005 at Boscombe Down.....	169
Figure 7.8 Simulated daily temperature distributions in summer based on four different GCM-driven RCM data sets for the station at Boscombe Down.....	181
Figure 7.9 Simulated daily precipitation distributions in summer based on four different GCM-driven RCM data sets for the station at Boscombe Down.....	182
Figure 7.10 Simulated CP frequencies in summer based on four different GCM-driven RCM data sets.....	183
Figure 7.11 Estimated 99.5 th percentiles of the daily maximum hourly precipitation based on GCM-driven RCM data.	184
Figure 7.12 see Figure 7.11.....	185
Figure 7.13 see Figure 7.11.....	186
Figure 7.14 see Figure 7.11.....	187

Chapter 1

Introduction

1.1 Motivation

Both observed trends (see e.g. Hartmann et al., 2013; Madsen et al., 2014; Seneviratne et al., 2012) and modelling results (see e.g. Orłowsky and Seneviratne, 2012; Arnbjerg-Nielsen, 2012; Willems and Vrac, 2011) suggest that extreme precipitation will become more intense in the future under a warming climate caused by increased concentration of greenhouse gases. This is physically plausible under the assumption that a rise in temperature results in higher moisture content in the atmosphere and increased evapotranspiration (Maraun et al., 2008) and could lead to increased flood risk and vulnerabilities (Willems et al., 2012). Some studies (Pall et al., 2007; Willems and Vrac, 2011) used the thermodynamic constraint based on the Clausius-Clapeyron equation to quantify these changes in precipitation extremes. They suggested that future precipitation extremes increase by 7% per one degree warming based on the assumption that specific humidity also increases by 7% per degree of temperature. But this approach has been questioned. For example, Lenderink and van Meijgaard (2008) found increases of 14% per one degree warming for observed and simulated extreme hourly precipitation and Hardwick Jones et al. (2010) showed a point of inflection for high temperature, where the relationship between extreme precipitation and temperature becomes negative.

Generally, it is likely that future increases in sub-daily precipitation extremes are stronger than those in daily precipitation extremes under a warming climate (Beck, 2013; Lenderink and van Meijgaard, 2008; Westra et al., 2014). Sub-daily precipitation extremes are often associated with flash flooding events especially in urban environments or for small and steep river catchments, which can lead to dramatic consequences. As a result, there is an urgent need to quantify the potentially increasing risk of flash floods in the future. But many impact studies require site-specific sub-daily estimates of precipitation extremes, whereas most Regional Climate Models (RCMs) in their current generation can only provide projections of areal averaged precipitation at ca. 12.5km resolution due to limited computing power. The RCM performances are further restricted by their parameterization schemes to simplify complex mechanism. As a consequence, RCM simulated precipitation intensities may be heavily biased

(Buonomo et al., 2007; Fowler and Ekström, 2009; Hanel and Buishand, 2010; Kendon et al., 2014) and adaptation strategies, which are often heavily based on climate modelled projections, may fall short and leave humans more vulnerable to flood risk (He et al., 2013).

In contrast to climate models, statistical downscaling methods are able to provide site-specific projections of precipitation extremes. Those statistical methods often combine large-scale atmospheric (e.g. sea level pressure) and certain surface variables (e.g. temperature) simulated by climate models to project precipitation extremes. Most studies in the past have focused on the downscaling of daily precipitation events (e.g. Beuchat et al., 2012; Hundecha and Bárdossy, 2007; Maraun et al., 2010; Bárdossy and Pegram, 2011), whereas only few attempted to establish statistical relationships between predictors and sub-daily precipitation. Among those, Haberlandt et al. (2014) applied a statistical model to reproduce hourly precipitation extremes and relied on changes in hourly precipitation simulated by the REMO RCM driven by the ECHAM5 GCM to project site-specific future changes over Northern Germany for the future time period 2071-2100. However, it needs to be questioned whether changes in RCM precipitation derived on a grid-scale are applicable to project site-specific precipitation (Olsson et al., 2012). Mezghani and Hingray (2009) used a Generalized Linear Model (GLM) based on daily atmospheric circulation indices to estimate daily regional precipitation. In a next step, they disaggregated the estimated daily regional precipitation variable to site-specific hourly precipitation. But temporal disaggregation is problematic because the relationships between daily and hourly precipitation are likely to change under a warmer climate (Beck, 2013; Westra et al., 2014). Different analogue day methods were tested in Willems and Vrac (2011) to project sub-daily precipitation extremes for one station in Belgium. The best performances were obtained by comparing the exceedance probabilities of observed and simulated daily precipitation conditioned on circulation patterns (CPs) and perturbing the observed sub-daily precipitation on the analogue day based on the Clausius-Clapeyron (CC) relation. However, as mentioned earlier, changes in daily precipitation may be different to the changes in hourly precipitation and the CC relation may not be sufficient to represent changes in hourly precipitation under a warmer climate.

1.2 Research aim and research objectives

As outlined in the previous section, statistical downscaling methods in the past often relied solely on changes in RCM simulated daily precipitation to project extreme hourly precipitation. Two main problems with this approach can be identified: (1) RCM simulated precipitation should always be interpreted as an areal average and thus is unlikely to realistically represent

changes in site-specific precipitation extremes needed for many impact studies, (2) Changes in daily precipitation extremes are unlikely to realistically represent changes in hourly precipitation extremes.

The aim of this thesis is to provide reliable and robust projections of site-specific UK hourly precipitation extremes under a warmer climate based on a statistical downscaling method which uses a set of different predictors. To achieve the aim of this thesis, the research objectives are defined as follows.

1. Identification of certain months or seasons during which hourly precipitation extremes are more likely to occur.
2. Categorisation of synoptic weather conditions that exhibit distinguishable relationships between atmospheric circulation patterns (CPs) and extreme hourly precipitation.
3. Development of a reliable and robust statistical downscaling method that can assign the “correct” exceedance probability to high precipitation values (reliability) over different time periods (robustness). The statistical downscaling method should also be able to perform under a future climate which is likely to be warmer.
4. Projections of extreme hourly precipitation that are representative of the UK spatial variability, the influences of different time periods and emission scenarios. Assessment of the uncertainties associated with the GCM-driven RCM runs and the statistical downscaling method.

1.3 Study approach

Historical UK daily and hourly precipitation records at a station level are extracted from the MIDAS Land Surface Stations database (NCAS British Atmospheric Data Centre, 2012). A screening procedure is applied to eliminate suspicious or duplicated records and outliers within the precipitation data set. Station selection criteria are defined to ensure that the stations used in the method development contain the best homogeneous precipitation time series with a high level of completeness and represent different UK extreme precipitation regions (Jones et al., 2014). Historical gridded sea level pressure from the NCAR SLP data set (Hurrell and Trenberth, 2013) is used for the definition of the CP-classification. The RCM simulated sea level pressure, temperature and precipitation data sets are extracted from the CORDEX data archive (CORDEX, 2014) for the evaluation and projection of extreme hourly precipitation estimates. Two different RCMs are used, the RCA4 and the RACMO22E.

For the twelve selected UK stations, the annual cycle of extreme hourly precipitation is investigated. The performances of multiple fuzzy rules CP-classifications based on different input configurations are assessed. The automated fuzzy rules CP-classifications have the advantage that information of local precipitation characteristics can be incorporated into the classification process. For the selected fuzzy rules CP-classification, the relationship between CPs and extreme hourly precipitation for each CP and season is established and assessed. The number of CPs needs to be restricted, in order to be able to calibrate the statistical downscaling method on large enough subsample sizes for each season and CP.

For the development of the statistical downscaling method, different aspects need to be considered. The method should be able to represent the physical effects of a warmer climate in terms of site-specific extreme hourly precipitation. Another important aspect is that RCM spatially averaged precipitation cannot realistically simulate localized convective precipitation events (Chan, 2013), which are most important in terms of extreme hourly precipitation, and RCM precipitation intensities tend to be heavily biased (Maraun et al., 2010b). The statistical downscaling method should also not rely only on temporal disaggregation relationships between daily and hourly precipitation because they are likely to change under a warmer climate (Beck, 2013; Westra et al., 2014).

Against this background, an analogue day method is developed. By means of the analogue day method, local precipitation characteristics, including extremes, can be reproduced on every temporal time scale for which observed precipitation data is available. Different approaches of finding the analogue days are evaluated, in order to optimise the quality of the statistical downscaling method. Similarly to (Willems and Vrac, 2011), daily precipitation and temperature are used as predictors conditioned on different seasons and CPs. However, in this thesis, the two predictor variables (daily precipitation and temperature) can be combined and considered as equally important within the process of finding the analogue days. The perturbation of observed hourly precipitation as defined in this thesis is novel to the statistical downscaling of hourly precipitation events. In terms of the selection of the predictor variables, using temperature as a predictor is important to represent changes in extreme hourly precipitation under a warmer climate and is needed to take into account future changes in the disaggregation relationships between daily and hourly precipitation. This is because particularly convective precipitation events are expected to intensify with higher temperature (Berg et al., 2013; Molnar et al., 2014) and can often be linked with extreme hourly precipitation (Beck and Bárdossy, 2013; Gregersen et al., 2013). Daily precipitation, as a predictor of hourly precipitation extremes, enables the statistical downscaling method to create a direct link with the historical weather in terms of precipitation. For example, high

temperature in summer can lead to clear skies and dry conditions, but can also result in an increased risk of extreme precipitation associated with convective events (Beck and Bárdossy, 2013). In this context it is also important to note that moisture availability may decline as temperature increases (Hardwick Jones et al., 2010). Therefore, the use of daily precipitation as a second predictor is very important as it provides additional information about the actual amount of precipitation on each day. It needs to be emphasised that RCM simulated precipitation intensities, which can be heavily biased, are only used to find the analogue days. Conditioning the statistical downscaling method on seasons and CPs is important mainly for two reasons. (1) Extreme hourly precipitation is likely to occur more often for certain seasons and CPs. (2) The increases in extreme hourly precipitation with higher temperature are likely to depend on the season and the CP (Blenkinsop et al., 2015). The main limitation of the analogue day method, to only simulate precipitation events observed in the past, is overcome by perturbing the observed hourly precipitation records based on a precipitation duration–temperature relationship.

The aim of this thesis is to provide estimates of hourly precipitation extremes, which are in good agreement with observed records (reliability) over varying time periods (robustness). For this purpose, the 99.5th percentile errors between observed and estimated hourly precipitation extremes are analysed over different validation periods. By calculating the 99.5th percentile error over 100 different bootstrapping samples, the quality of the estimated extreme precipitation is assessed beyond the 99.5th percentile only. Two different RCMs are used, the RCA4 and the RACMO22E. The stationarity assumption, that the statistical downscaling method remain valid under a different climate, is also assessed. Estimates of hourly precipitation extremes are validated over the ten warmest summers. In a second test to assess the stationarity assumption, the residual time series between observed and estimated annual maximum hourly precipitation are analysed for each station.

The best statistical downscaling method is applied to project hourly precipitation extremes at the twelve selected stations using four different GCM-driven RCMs, namely the CM5A-MR driven RCA4, ESM-LR driven RCA4, EC-EARTH driven RCA4 and EC-EARTH driven RACMO22E, over two time periods, 2030-2055 and 2075-2100, and for two different emission scenarios, RCP4.5 and RCP8.5. The ranges of uncertainty are quantified over 100 different bootstrapping samples.

1.4 Thesis structure

This thesis is divided into eight chapters, including this introduction, and is organised as follows. Chapter 2 provides an overview of the relevant literature, including (1) the observed and projected changes in extreme precipitation and the underlying physical reasons for those changes, (2) the subject of dynamical downscaling to project extreme precipitation at higher resolution, (3) the different types of statistical downscaling to estimate precipitation extremes, and finally (4) the use of atmospheric circulation patterns as an important measure to predict extreme precipitation events. Chapter 3 describes the multiple data sets which are applied in this thesis and the process of selecting twelve stations for which hourly precipitation are estimated in Chapter 6 and 7. In Chapter 4, the fuzzy rules based CP-classification method and the statistical downscaling process are described. Chapter 5 presents the results of the fuzzy rules based CP-classification. It examines the performances of 180 different CP-classifications over the UK using different input parameters, and identifies the most optimal CP-classification for the statistical downscaling process. For this finally selected CP-classification, the atmospheric pressure patterns of each CP are discussed, as well as the relationships between the CPs and extreme hourly precipitation. Chapter 6 provides the statistical downscaling results of hourly precipitation extremes based on observed data. First, it derives the relationships between precipitation duration and temperature for different seasons and CPs. In a next step, hourly precipitation extremes are predicted over different validation time periods for the twelve stations. The stationarity assumption of the statistical downscaling process is also assessed. In Chapter 7, the RCM data is used instead of the observed data to predict hourly precipitation extremes. In the first section, precipitation extremes are predicted by the ERA-interim driven RCM data and validated over the same time periods as in the previous chapter to assess the reliability and robustness of the downscaling results. In the following sections, GCM-driven RCM data is used to estimate hourly precipitation over a reference and two future time periods. Chapter 8 summarises the key findings in this thesis. It also discusses uncertainties and current limitations of the presented statistical downscaling results, as well as the potential for future developments.

Chapter 2

Literature review

This chapter presents an overview of studies in the past, which are relevant to the subject of this thesis. It begins with a short overview of observed trends in extreme precipitation over the last decades and then discusses the causes which could have led to these changes. The first section ends with an outlook into the future by discussing projections of extreme precipitation simulated by climate models. In the following section, the subject of climate models is introduced with a specific focus on extreme sub-daily precipitation. After that, different types of statistical downscaling methods are presented. The chapter ends with a section about the use of atmospheric circulation patterns.

2.1 Changes in extreme daily and sub-daily precipitation events

2.1.1 Observed trends

Given that the globally averaged surface temperature increased about 0.72°C over the period 1951-2012 (Hartmann et al., 2013), multiple studies in the past assessed the potential impact on observed extreme precipitation. On a global scale, Seneviratne et al. (2012) concluded that more regions experienced statistical significant increases in the number of extreme precipitation events than decreases over the second half of the 20th and early 21st century but there are strong regional and seasonal variations. Furthermore, inhomogeneities in the measurement conditions and instrumentation technologies, as well as inconsistent methods for quality assessment and control of precipitation records (Westra et al., 2014) may have contributed to differing results between studies in the past.

Over western Europe, de Lima et al. (2014) found a tendency towards more daily precipitation extremes in Portugal over the 1976-2007 period, but not over the 1941-2007 period.

Hundecha and Bardossy (2005) found increases in daily extremes in winter and decreases in summer between 1958 and 2001 in western Germany. Similarly, Maraun et al. (2009) showed an increase in daily extreme events in winter and a decrease in summer between 1961 and 1995 over the UK. It was suggested that the decrease in summer appeared to be a return to earlier levels before 1960. Fowler and Kilsby (2003) also found a decrease in extreme daily precipitation in summer between 1961 and 2000. For spring and autumn, they showed an increase in daily extremes over the UK. Maraun et al. (2008) updated the trend analysis from

1961 to 2006 and found that the decrease in summer extremes might have reversed over the most recent ten years in all English/Welsh regions except Northeast England. Overall, the increase in winter daily extremes is consistent throughout the UK, while complex changes in daily extremes occurred in spring and autumn with opposite changes in different regions of the UK.

In terms of sub-daily extremes, Beck (2013) found that the precipitation extremes over short time periods intensified more rapidly compared to daily extremes during the second half of the 20th century in South West Germany. Ntegeka and Willems (2008) analysed 10min precipitation extremes from the Uccle station in Belgium between 1898 and 2004 and found that extremes were clustered in the 1910s-1920s, the 1960s and 1990s. Leahy and Kiely (2011) assessed sub-daily precipitation extremes during the second half of the 20th century in Ireland and showed an increase since the late 1970s particularly in the western part of the country. Over the UK, an increase in the number of extreme hourly precipitation occurred between 1980 and 2009 in summer according to Beard (2010). It was also shown that the number of hourly extreme events declined over the same time period for the other three seasons (winter, spring, autumn). Paulson (2010) found an increase in the number of extreme sub-hourly extreme events between 1988 and 2007 in the southern UK. It needs to be mentioned that a 20-year time series is very short for providing a robust trend analysis. On the other hand, sites with a long time series of historical sub-daily precipitation are likely to have used different measurement techniques over their operating lifetime. This can lead to systematic inhomogeneities when analysing precipitation records from a single location (Westra et al., 2014).

More studies assessing trends in extreme precipitation over Europe were recently summarised in Madsen et al. (2014). However, most of them were based on daily precipitation extremes due to the lack of long high-quality records at sub-daily scale (Westra et al., 2014).

2.1.2 Causes

On a global scale there is medium confidence that human induced global warming has contributed to an increase in extreme precipitation as discussed in Seneviratne et al. (2012). They stated that observed changes in extreme precipitation appear to be consistent with the effects of human induced global warming (e.g. an increase in atmospheric moisture content), but a direct cause-and-effect relationship between global warming and extreme precipitation could not be found. Min et al. (2011) compared observed and simulated changes in extreme daily precipitation over the second half of the twentieth century on Northern Hemisphere land areas and concluded that human induced global warming has indeed contributed to an

intensification in extreme precipitation. The main challenge in explaining changes in climate extremes is to separate climate oscillations (natural variability) from the potential impacts of human induced global warming. For example, Ntegeka and Willems (2008) found cyclic patterns of 10min precipitation extremes between 1898 and 2004 for one station (Uccle) in Belgium and concluded that no strong evidence of a link between global warming and an increase in 10min precipitation extremes can be drawn.

However, the hypothesis that global warming leads to an increase in extreme precipitation is physically plausible (Maraun et al., 2008; Westra et al., 2014). The Clausius-Clapeyron (CC) equation states that specific humidity, which represents the atmospheric moisture content, is expected to increase about 7% per degree of warming (Seneviratne et al., 2012). As a result, Pall et al. (2007) suggested the thermodynamic constraint based on the CC equation as a predictor for extreme precipitation at mid-latitudes.

In terms of hourly precipitation extremes, however, other studies (Haerter and Berg, 2009; Lenderink and van Meijgaard, 2008) found increases at a rate close to 14% per degree of warming compared to the 7% increase suggested by the CC relation. Lenderink and van Meijgaard (2008) argued that upward air motions may be intensified by an increase in latent heat release to explain super CC scaling (14% increase) of convective precipitation. They showed that a super CC relation was only found for extreme hourly precipitation in summer, when convective precipitation dominates. For extreme hourly precipitation in winter and for daily extremes, a rate close to the CC relation was still found. Berg et al. (2013) showed that increases in convective precipitation intensities with temperature indeed exceed the CC relation, while large-scale frontal precipitation extremes increase at a rate close to the CC relation. They also found a decrease in convective precipitation intensities for temperature above ca. 22°C, which can be due to a decrease in relative humidity with rising temperature. This indicates that moisture availability becomes the dominant factor in terms of extreme precipitation for high temperatures (Hardwick Jones et al., 2010). Haerter and Berg (2009) proposed a different explanation of the super CC scaling for hourly extremes. They argued that convective precipitation events are naturally more intense than large-scale events, and therefore the fact that convective (large-scale) events occur more often for high (cool) temperature leads to the super CC relation. In this context it is interesting to note that Molnar et al. (2014) showed a statistically significant trend towards more observed convective storm events between April and September over the last 30 years in the Alps under a warmer climate. Over the UK, (Blenkinsop et al., 2015) found super CC relations of hourly precipitation extremes only under anticyclonic conditions in summer.

Apart from the precipitation-temperature relationship, other factors also need to be considered when estimating changes in precipitation extremes. For example, Kendon et al. (2010) noted that changes in soil moisture and potential feedbacks play an important role in terms of extreme precipitation. Molnar et al. (2014) showed that local (orography) and regional effects limit moisture availability and therefore extreme precipitation. (Blenkinsop et al., 2015; de Lima et al., 2014; Fowler and Kilsby, 2003; Leahy and Kiely, 2011) suggested a link between changes in extreme precipitation and atmospheric circulation, although uncertainties in the projections of changes in storm tracks, particularly for the North Atlantic basin, remain high (IPCC, 2013).

2.1.3 Future projections

There is evidence, in terms of projections simulated by climate models, for extreme precipitation to become more intense and frequent in the future in many regions over the globe as stated in Seneviratne et al. (2012). However, it needs to be kept in mind that future projections of precipitation changes are more uncertain than those of temperature and there is less consistency between different climate models, in particular in terms of extreme precipitation (Tebaldi et al., 2006). For projections of temperature extremes regarding the end of the 21st century, the main source of uncertainty is found in the emission scenarios. In contrast, uncertainties in future projections of precipitation extremes are mostly due to uncertainties within the climate models (Seneviratne et al., 2012).

Globally, Kharin et al. (2007) estimated based on an ensemble of GCMs that over land the present day 20 year return period will reduce to 10.7 years for 2046-2065 and to 9 years for 2081-2100 on average corresponding to the same return level of annual daily precipitation extremes. In this particular ensemble, the change in extreme daily precipitation is most pronounced for Central Asia, South Asia and North America. Orłowsky and Seneviratne (2012) extended this study to a larger number of CMIP3 ensemble and used seasonal instead of annual time scales. It was found that projections of extreme precipitation agree most between different models for high latitudes and in the tropics, and in some mid-latitude regions of the Northern Hemisphere in winter. In general, models agree better in terms of extreme precipitation changes in winter than in summer for most regions. Comprising the evidence of a large number of different studies, the IPCC special report on extreme events suggests that “it is likely that the frequency of heavy precipitation or the proportion of total precipitation from heavy precipitation will increase in the 21st century over many areas of the globe” (Seneviratne et al., 2012).

Over Europe, Buonomo et al. (2007) showed a widespread increase in extreme daily and 30-days precipitation based on two RCMs (HadRM2 and HadRM3) even in regions where mean precipitation decreases. The magnitude of the changes increased with shorter duration and longer return periods. Ban et al. (2015) found that high-resolution climate models project an increase in extreme daily (and hourly) precipitation over middle and southern Europe in accordance with the CC relation. Evidence derived from multiple studies suggests an increase in extreme daily precipitation over northern Europe in particular during winter (Seneviratne et al., 2012). Over the UK, Fowler and Ekström (2009) projected an increase in daily precipitation extremes during winter, spring and autumn. In summer, no increase in extreme precipitation was found. However, it was noted that there is low confidence in projections for summer due to poor model performance in this season. It was also shown that the increases in precipitation extremes are more pronounced for short-duration (daily compared to ten days). Within the framework of the UK climate projections (UKCP09), Jones et al. (2009) used a weather generator to project changes in precipitation extremes and found increases in the future annual daily precipitation maximum across the UK.

In terms of sub-daily precipitation extremes, Hanel and Buishand (2010) showed large differences in the projections of annual hourly precipitation maxima between different 25km RCMs (RACMO2.1, REMO5.7, HadRM3, HIRHAM5 and RCA3) over the Netherlands. The changes in precipitation extremes were much higher for shorter duration (hourly compared to daily). Gregersen et al. (2013) assessed two different 25km RCMs (RACMO and HIRHAM) over Denmark and also found significant increases in hourly precipitation extremes. Over the southern UK, Chan et al. (2014) and Kendon et al. (2014) projected an increase in hourly precipitation extremes in winter based on a 1.5km RCM and a 12.5km RCM. During summer, the 1.5km RCM also projected an increase in extreme hourly precipitation, which was not seen in the 12.5km RCM. It can be assumed that the projections of the high-resolution RCM are more reliable as it reproduced the observed hourly precipitation characteristics more realistically as discussed in Section 2.2.2. It was therefore suggested that previous results derived from coarser-resolution RCMs should be revised since changes in convective events could have been misrepresented.

Instead of only focusing on the RCM simulated precipitation variable, other studies applied statistical downscaling methods to project sub-daily precipitation extremes. For example, Beck and Bárdossy (2013) predicted an increased risk of extreme hourly precipitation over southwestern Germany based on changes in daily MSLP and temperature. Willems and Vrac (2011) showed increases in 15min hourly precipitation extremes for one station at Belgium,

but also highlighted the wide range in projections depending on the climate model and the statistical downscaling method.

For a more detailed summary regarding projected precipitation extremes over Europe, the reader is referred to Madsen et al. (2014).

2.2 Climate modelling

2.2.1 Global climate models (GCMs)

Global climate models (GCMs) project the future climate by taking into account the impact of changing concentrations of greenhouse gases in the atmosphere. Projections based on GCMs provide spatially consistent information for all regions, while large-scale modes of variability, such as the El-Niño-Southern Oscillation (ENSO) and the monsoon systems, which affect precipitation extremes, are also attempted to be simulated in GCMs. However, due to the coarse spatial resolution GCMs cannot represent all the physical processes that are important in terms of precipitation extremes (e.g. cyclones, fronts and convective events). The simplified representation of topography and geography in GCMs is another limitation in simulating reliable patterns of precipitation extremes (Goodess, 2012). Therefore, downscaling techniques are required to improve the poor representation of precipitation associated with GCM projections. There exists two different methods to downscaling, namely dynamical downscaling and statistical downscaling.

First, dynamical downscaling is a procedure of nesting a Regional Climate Model (RCM) within a GCM to represent the physical processes in the atmosphere with a higher resolution for a limited area of interest. It will be discussed in more detail in Section 2.2.2. The second method, statistical downscaling, establishes statistical relationships between predictors and the predictand (see Section 2.3).

2.2.2 Regional climate models (RCMs) - Dynamical downscaling

In general, RCMs contain the same representations of atmospheric dynamical and physical processes as in GCMs but on a finer spatial scale (50km or less) over a sub-global domain (e.g. Europe). They provide a consistent physical simulation of large and small scale features, while also maintaining the interrelationships between precipitation and other climate variables and attempting to simulate the transient characteristics of the different processes (Haberlandt et al., 2014). Because of their higher spatial resolutions, RCMs usually require a decreased model time step and provide a better representation of orographic effects, land-sea contrast and the

mechanism of soil-precipitation feedbacks. They also show an improved representation of small-scale physical and dynamical processes (e.g. convective precipitation) and are able to generate mesoscale circulation patterns that are not captured by GCMs (Buonomo et al., 2007).

RCMs are driven by wind, temperature, humidity and sea surface temperature imposed at the boundaries of the RCM domain, all provided by the GCMs, which generally lead to high agreement between the large-scale fields in the RCMs and the GCMs (Buonomo et al., 2007). However, GCM outputs can differ heavily between each other due to uncertainties in model parameters, numerical schemes, resolution, internal variability of the chaotic climate system dependent on initial conditions (e.g. soil moisture), natural forced variability (e.g. solar forcing) and temporal variability caused by large-scale circulation (e.g. blocking effects and El Niño) (Seneviratne et al., 2012). These discrepancies in the GCM outputs can have considerable effects on the RCM outputs (Deidda et al., 2013). Kendon et al. (2012) noted that precipitation projections rely heavily on the driving model to represent the synoptic variability realistically and any changes (e.g. a shift in the storm tracks) could have a significant impact on precipitation over the UK. For example, Fowler and Ekström (2009) used thirteen 50km RCMs nested into different GCMs to assess changes in the seasonal daily precipitation extremes and found that the largest proportion of uncertainty was caused by the lateral boundary conditions from the GCMs.

But differences in the projections of precipitation characteristics were also found between different RCMs for a given boundary forcing. These discrepancies are caused by different RCM model formulations and are also due to small-scale internal variability generated by the RCMs.

In order to assess the performances of a RCM, it can be driven by reanalysis data. In this case the GCM input data is replaced by quasi observed boundary conditions, which allows the isolated assessment of the RCMs predictive skill. For example, the EURO-CORDEX project provides a set of 30 years ERA-interim reanalysis driven RCMs. Reanalysis-driven RCM simulations not only exclude systematic biases inherited from the GCM, in theory they should also be able to reproduce the actual day-to-day weather. This allows a more comprehensive assessment of the RCM downscaling skill (Maraun et al., 2010b). Although (Chan, 2013; Kendon et al., 2012) noted that reanalysis-driven RCMs cannot be expected to exactly reproduce the day-to-day sequences of the observed data due to the large RCM domain size and due to the fact that observed data are only fed into the RCM at the lateral and sea surface boundaries.

Over more than a decade, RCMs have continuously been improved in the generation of climate projections within international coordinated frameworks such as the PRUDENCE (2001-2004), ENSEMBLES (2004-2009) and EURO-CORDEX (2009-present) project. A very important aspect in this context is the increased spatial resolution over which RCMs generate climate projections. Nowadays, the EURO-CORDEX project provides climate projections for certain regions at a spatial resolution of 12.5km.

However, limitations in RCMs still exist. For example, most RCMs are not able to explicitly resolve physical processes such as radiation, convection, cloud microphysics and land atmosphere interaction affecting the precipitable water. Instead, they use parameterization schemes, which simplify the complexity of the real world processes, and leads to inherent model uncertainties in terms of precipitation. One example is the formation of clouds simulated by two separate schemes. Clouds and precipitation caused by atmospheric processes such as frontal systems or cyclones are resolved by a large-scale scheme, whereas clouds and precipitation due to subgrid-scale convective processes are represented by a convection scheme as discussed in Maraun et al. (2010b). Also, local characteristics of the topography, which affect precipitation patterns, cannot be accurately resolved at the spatial resolution commonly used by RCMs. Spatial inhomogeneity of land use, soil moisture and vegetation processes are other examples that can lead to uncertainties (Seneviratne et al., 2012). Some physical processes involving feedbacks such as land-atmosphere interaction, ocean-atmosphere interactions, stratospheric processes and blocking dynamics are still insufficiently understood. They are of great importance in reproducing climate extremes. In terms of stationarity, assumptions need to be made for the parameterization schemes to be valid under a different climate. Therefore, Maraun et al. (2010b) concluded that although some of the bias in RCMs may be inherited from the GCM simulated lateral boundary conditions, RCMs themselves also contribute largely to the uncertainties in the simulated precipitation.

As a result, RCMs generally overestimate the occurrence of wet days (drizzle effect) and thus mean precipitation (Buonomo et al., 2007). In terms of precipitation extremes, Gregersen et al. (2013) analysed two RCMs (RACMO and HIRHAM) driven by ERA40 data and by GCM (ECHAM5) data and found that none of the RCMs runs reproduced the observed seasonal cycle realistically. Buonomo et al. (2007) applied two different 50km RCMs and found that daily precipitation extremes in the western and upland areas of the UK were underestimated in the HadRM2 and overestimated in the HadRM3 due to different parameterization schemes. In the South and East of the UK, daily precipitation extremes were underestimated in both RCMs. Chan (2013) showed that increasing the spatial resolution (from 50km to 12.5km) of RCMs

helps to reduce the biases in orographic precipitation. But RCMs, which rely on convection parameterization, are still not able to simulate local convective precipitation events realistically (Westra et al., 2014). This particularly affects the representation of extreme sub-daily precipitation.

For example, Gregersen et al. (2013) compared temporal and spatial characteristics of hourly precipitation extremes from three different RCM runs at 25km resolution, representing one ERA40-reanalysis driven run (ERA40 driven RACMO) and two GCM-driven runs (ECHAM5 driven RACMO and ECHAM5 driven HIRHAM). They found that neither the temporal changes nor the spatial structure in observed local extreme hourly precipitation were well reproduced. Comparing the two RCMs (RACMO and HIRHAM) driven by GCM (ECHAM5) data, large differences in hourly extreme precipitation intensities were found, as well as significantly different changes in hourly extreme precipitation intensities between the two RCMs over time. It was also shown that the HIRHAM RCM driven by ECHAM5 GCM data exhibited a general bias towards too many precipitation extremes in spring. In a different study, Hanel and Buishand (2010) found large differences in the distribution between annual hourly precipitation maxima in the Netherlands simulated by five RCMs at 25km (RACMO2.1, REMO5.7, HadRM3, HIRHAM5 and RCA3) and high-quality observed radar data. The negative bias in the location parameter, which is found for the majority of the RCM simulations, implies that the RCMs tend to underestimate annual hourly precipitation maxima. They also showed large differences between hourly precipitation extremes for different RCMs but the influence of the driving model (GCM-driven or 'perfect boundary' ERA40 reanalysis driven) on the hourly precipitation extreme distributions was in most cases small. Several studies (Chan et al., 2014a; Chan et al., 2014b; Kendon et al., 2014) found that even 12.5km RCMs significantly misrepresented extreme hourly precipitation particularly in summer, where hourly extremes tend to be most intense.

Over the last few years, RCMs at 1km to 5km resolution have been tested for small domains, such as the Alpine Region (Ban et al., 2014; Ban et al., 2015; Hohenegger et al., 2008) or the southern part of the UK (Chan et al., 2014a; Kendon et al., 2014; Kendon et al., 2012). These models provide a spatial resolution at which clouds and convection can be explicitly resolved, no convection parameterization is required and the diurnal precipitation cycle tends to better reproduced (Seneviratne et al., 2012). In fact, several studies (Ban et al., 2014; Chan et al., 2014b; Kendon et al., 2014) showed that RCMs can reproduce hourly precipitation, including extremes, characteristics more realistically when run at high resolution in comparison with those runs relying on parameterized convection. The former produced more short-lived high peak intensity precipitation as a result of the convection-resolving structure. In the 12.5km

RCMs, precipitation was too persistent due to light rain and this bias was considerably reduced in 1.5km RCMs (Kendon et al., 2012). Apart from that, the differences between convective and large-scale precipitation were much more realistically simulated for 1.5km RCMs compared to 12.5km RCMs.

However, the 1.5km RCMs still tend to overestimate extremes. This is possibly due to the lack of physical understanding of the complex mechanism leading to convective precipitation (Willems et al., 2012). Although RCMs at kilometre-scale are able to simulate convection on the model grid, they cannot resolve convective plumes and smaller showers (Chan et al., 2014a; Kendon et al., 2012). It also needs to be pointed out that even the 1.5km RCM is not able to provide projections of precipitation on a local-scale, which is particularly important in terms of representing convective precipitation events and no clear evidence was found that the 1.5km RCM is superior to the 12.5km RCMs at the daily scale (Chan, 2013).

Due to computational limitations, previous studies have been limited to a single season (Hohenegger et al., 2008), selected events (Attema et al., 2014) or small spatial domains (Ban et al., 2014; Chan et al., 2014a; Chan et al., 2014b; Chan, 2013; Kendon et al., 2012). Longer multi-model ensembles simulations are required to draw broader conclusions on how precipitation extremes will change under a future climate. This will improve as computational power increases and more high-resolution modelling studies become available. Until then, the use of RCMs at coarser resolution with convective parameterization for the projection of precipitation events remains necessary (Chan et al., 2014a).

As a consequence of the current limitations, RCM hourly precipitation projections should always be applied with great caution and alternative methods, such as statistical downscaling techniques, should be considered whenever possible (Hanel and Buishand, 2010).

2.3 Statistical downscaling

2.3.1 Introduction

To study the impacts of climate changes, many geo-physical and bio-geochemical models (e.g. rainfall-runoff models) require precipitation as the driving input in finer spatial and temporal scales than what may be currently simulated by Global Climate Models (GCMs). Two types of downscaling methods exist, namely dynamical and statistical downscaling, to bridge the spatial and temporal gap. Section 2.2.2 already discussed the dynamical downscaling method and its associated limitations. In comparison with the dynamical downscaling method, statistical downscaling methods may be more capable of producing finer spatial and temporal scales

with less computational resources. They can also be used to reduce biases in the precipitation variable simulated by the climate models. The underlying assumption is that large-scale atmospheric variables (e.g. MSLP) and certain surface variables (e.g. temperature) are more reliably represented in climate models than precipitation (Beck and Bárdossy, 2013). In the purest form, the idea of statistical downscaling describes a mapping between a predictor X and the expected values of a predictand Y as defined in Equation 2.1.

$$E(Y/X) = f(X, \beta) \quad (2.1)$$

β : A vector of unknown parameters that needs to be calibrated

In the following, statistical downscaling methods are categorised into three different groups. The first group comprises all the perfect prognosis (PP) methods. Those methods establish statistical relationships between climate variables. In the context of climate change projections, PP methods are based on the assumption that the predictors simulated by climate models represent a physically plausible realization of a future climate and that the observed relationships between predictors and precipitation remain stationary in the future.

Methods in the second group, model output statistics (MOS), make direct use of the simulated predictand (precipitation) by climate models.

The last group comprises the so-called weather generators (WGs). WGs represent statistical methods that generate local-scale weather time series resembling the statistical properties of the observed weather (precipitation). In their most basic form, WGs are calibrated against observations on local-scale only and are therefore technically not a downscaling method (Maraun et al., 2010b). In order to take into account climate changes in the future, change factors are derived from climate models.

2.3.2 Perfect prognosis (PP)

A PP method defines a relationship between a set of predictors and the expected value of a predictand. Simple PP methods ignore any residual noise term, while the state-of-the-art PP methods explicitly incorporate natural variability and extreme events by adding noise into the model (Maraun et al., 2010b).

The main advantage of PP methods is that the predictand is simulated based on statistical relationships that can be observed in reality. The quality of the statistical relationships is expected to be improved in cases where the predicted variable is averaged in space and/or time (Wilby et al., 2003).

The main disadvantage of the method is that PPs often only explain a fraction of the observed natural climate variability, in particular in terms of extreme precipitation (Hundecha and Bardossy, 2008). It is also highly sensitive to the selection of the predictor variables and the statistical transfer functions describing the relationship between the predictor and the predictand. The projection of future precipitation extremes is particularly problematic since these extreme events tend to lie at the margin or beyond the range of the calibration data set (Wilby et al., 2002).

2.3.2.1 Predictor selection

In terms of estimating precipitation, several different predictors can be used including variables such as circulation patterns (see Section 2.4), humidity and temperature often in a combined approach. Pure PP downscaling methods do not in general simulate explicitly physical coherence between predictor and predictand. This can be problematic. For example, high temperature (predictor) in summer may be due to dry conditions (predictand) as a result of clear skies or high temperature may cause convective wet conditions. Therefore, the correlations are difficult to interpret (Maraun et al., 2010b).

In a climate change context, it is also important that predictors are able to capture the effects that a warmer climate has on precipitation. In particular, measures of temperature or humidity are needed to represent increases in the water-holding capacity of the atmosphere under warmer conditions. Downscaling approaches which implicitly assume that changes in precipitation are only due to changes in atmospheric circulation are therefore at least questionable. For example, several studies (Haberlandt et al., 2014; Kendon et al., 2014; Wilby and Wigley, 1997) showed that changes in precipitation cannot be explained by changes in CP frequency only.

Other studies (Beck and Bárdossy, 2013; Berg et al., 2013; Blenkinsop et al., 2015; Hardwick Jones et al., 2010; Lenderink and van Meijgaard, 2008; Molnar et al., 2014) specifically assessed the observed relationships between extreme precipitation and temperature. These studies had the common conclusion that the intensity of extreme precipitation increased with higher temperature. Further, some studies (Hardwick Jones et al., 2010; Lenderink and van Meijgaard, 2008) showed that the rate of increase is greater for hourly compared to daily precipitation. (Berg et al., 2013) found that the rise in the intensity of extreme precipitation with temperature is substantially higher for convective precipitation compared to stratiform precipitation events. They suggested that this is due to physical processes as convective plumes entrain more moisture from the surrounding atmosphere with higher temperature. It was also shown that the relative contribution of convective precipitation events to the total

sum of precipitation increased with temperature. However, the intensity of convective precipitation extremes decreased for temperature above 22°C, which may be partly caused by a decline in relative humidity with rising temperature. It indicates that temperature as the only predictor is also not sufficient to estimate extreme precipitation. Similarly, Blenkinsop et al. (2015) and Molnar et al. (2014) noted that extreme hourly precipitation depends on various other factors, such as atmospheric circulation, local topography and complex physical mechanism. Hardwick Jones et al. (2010) used daily mean temperature and maximum daily temperature to investigate the relationship between temperature and the intensity of precipitation extremes. They found consistent results between the two temperature variables: A decrease in precipitation duration as temperature increases. Similarly, Beck and Bárdossy (2013) found a decline in the frequency of wet hours with increasing temperature. They concluded that higher temperature results in fewer wet hours, but causes a shift in the hourly precipitation distribution towards higher intensities in particular in summer for anticyclonic conditions. This is consistent with physical considerations (Trenberth, 1999) that stronger convective uplift combined with higher moisture content due to an increase in temperature should lead to fewer but more intensive precipitation events.

It is important to note that predictors characterized as informative capturing the impact of climate change might be of little use if they cannot be realistically simulated by climate models. In this context, moisture-related variables are generally considered as problematic (Maraun et al., 2010b).

In general, the relationships between predictors and predictand should be stationary (Maraun et al., 2010b). In some studies additional auto correlated driving variables are required, to simulate the precipitation variable more realistically (Beuchat et al., 2012; Mezghani and Hingray, 2009). Apart from climate variables, geographical variables such as elevation, distance to the coast, advective continentality (represents the degree at which incoming air mass travelled over land versus over the ocean) and topographical slope can also be used to downscale precipitation (Vrac et al., 2007). However, taking too many predictors into account leads to over fitting and can decrease the predictive power of the downscaling model.

In terms of the climate model output, Beuchat et al. (2012) noted that GCM projections are more reliable on monthly than on daily timescales and therefore used monthly predictors to downscale daily precipitation. Wilby and Wigley (2000) found that the grid point nearest to the location of the predictand does not always exhibit the strongest predictor-predictand relationship. Karl et al. (1990) showed that regression-based downscaling methods benefit

from the standardization of the predictor variables, so that the corresponding distributions of the observed and the present-day GCM/RCM predictors are in better agreement.

2.3.2.2 Different PP models

Linear and nonlinear regression based models

For linear models, the relationships between the predictors X_i and the mean μ of the predictand Y , such as local-scale precipitation, is described in Equation 2.2.

$$\mu = \beta_0 + \beta_1 X_1 + .. \quad (2.2)$$

β_i : Strength of the influence of the predictor X_i

The predictand Y is Gaussian distributed, with mean μ and some variance which remains unexplained. However, the Gaussian assumption may only be valid for annual precipitation, whereas on shorter time scale (e.g. daily or even sub-daily) distributions of precipitation intensities are more and more skewed.

In contrast, General Linear and additive Models (GLMs) assume the distribution of the local-scale predictand Y no longer to be Gaussian and instead allow a wide range of distributions. The conditional mean μ of the corresponding distribution is still calculated by a linear function of a set of predictors X_i , but the mean μ of the predictand may be transformed by a link function $g(\cdot)$ as shown in Equation 2.3.

$$g(\mu) = \beta_0 + \beta_1 X_1 + .. \quad (2.3)$$

Similarly to the linear models, the variance of the predictand Y is not explicitly simulated and the unexplained variability is represented by drawing random numbers from the modelled distribution of Y . Linear regression-based models therefore generally underrepresent the variance of the precipitation. Instead, Wilby et al. (2002) combined a linear regression-based method with a stochastic weather generator to artificially inflate the variance of the downscaled daily precipitation time series. The combined approach reproduced the seasonal cycle in terms of the precipitation mean and variance but still underestimated the variance of the daily precipitation variable.

For better representing the variance or the extreme tail of the precipitation variable, Vector Generalized Linear Models (VGLMs) have been developed. Instead of only predicting the conditional mean of a distribution, a vector of distribution parameters is predicted as

explained in Equation 2.4. The vector p_k could for example represent the mean $p_1 = \mu$ and the variance $p_2 = \sigma$ of a precipitation distribution (Maraun et al., 2010b).

$$g_k(p_k) = \beta_{k,0} + \beta_{k,1}X_{k,1} + .. \quad (2.4)$$

In general, regression-based models are able to simulate precipitation intensities not observed in the past, which is particularly important in a climate change context (Mezghani and Hingray, 2009). However, a common drawback of regression-based models is the low inter-annual variability in terms of precipitation patterns and that regression-based methods tend to be highly sensitive to the calibration period (Hundechea and Bardossy, 2008). Another issue is that linear regression models perform poorly in reproducing non-linear shapes or even some shape discontinuity in the precipitation variable and normally perform better at the regional scale than at local-scales (Mezghani and Hingray, 2009).

In order to capture nonlinear and nonadditive relationships between the large-scale variable and local precipitation, nonlinear regression-based methods have been developed. Artificial neural networks (ANNs), as applied in Haylock et al. (2006) and Olsson et al. (2001), is one example of this technique. Haylock et al. (2006) compared six statistical methods, including four regression-based ANN, one correlation analysis and one resampling method, and two dynamical downscaling models in their skill to downscale seven indices of extreme precipitation at the station scale. They found that non-linear ANNs were the most skilful in terms of modelling inter-annual variability of the indices but showed a strong negative bias, which implies a tendency to underestimate extremes.

The popular approach of conditioning local-scale precipitation on circulation patterns can be considered as a special case of a regression model. The difference is that instead of a continuous field of predictors, a set of categorical circulation patterns X_l is used to predict the parameters p_k of a precipitation distribution as shown in Equation 2.5.

$$p_k = p_k(X_l) \quad (2.5)$$

The vector p_k describes the distribution parameters of the local-scale predictand and l stand for the different circulation patterns. Beck and Bárdossy (2013) used temperature as an additional predictor to subdivide each CP into five classes. They predicted future regional distributions of hourly precipitation amounts based solely on daily mean sea level pressure and temperature, which are considered as more reliably simulated than precipitation by climate models (Kendon and Clark, 2008). The observed empirical distribution of hourly precipitation

depending on the CP-temperature class was assumed to be constant under a future climate. The set of CPs are considered as the link between large-scale atmospheric conditions and the local precipitation, while temperature is regarded as the climatic key factor triggering local changes in convective uplift and moisture capacity in the atmosphere (Molnar et al., 2014). As a result, Beck and Bárdossy (2013) projected an almost 20% increase in hourly precipitation extremes (99% quantiles) in summer between 2000 and 2060 mainly due to higher temperatures, although they also showed that summer is on average becoming drier. It needs to be noted that these results are derived from an average of 30 stations in order to guarantee a robust estimation of the empirical precipitation distribution. For generating hourly precipitation extremes on a local-scale, the historical precipitation time series would need to be longer and the number of CPs should be reduced. Another limitation is that the two predictors, SLP and temperature, did not include any information about the precipitable water in the atmosphere and therefore may not be enough to predict changes in precipitation patterns reliably. It also needs to be mentioned that new and unobserved conditions, which are likely to occur under a future (warmer) climate, cannot be simulated by this particular method.

In many other studies, statistical downscaling methods have been conditioned on CPs as well to estimate the parameters of the precipitation variable more accurately (Bárdossy and Pegram, 2011; Bárdossy et al., 2002; Willems and Vrac, 2011; Yang et al., 2010). The main advantages and limitations as well as examples of downscaling methods using CPs are further discussed in Section 2.4.

In terms of extreme precipitation analysis, it needs to be noted that the distributions of extreme precipitation tend to be heavily tailed. Regression-based statistical models which do not consider that are likely to misrepresent high extremes (Maraun et al., 2010b). Also, regression-based models have been typically applied for downscaling daily precipitation (Beuchat et al., 2012; Haylock et al., 2006; Hundsdoerfer and Bárdossy, 2008). However, for many hydrological applications sub-daily precipitation is often required. Therefore, different methods have been proposed to disaggregate the simulated daily precipitation estimates to sub-daily precipitation (Koutsoyiannis and Onof, 2001; Mezghani and Hingray, 2009; Molnar and Burlando, 2005; Nguyen et al., 2007). Mezghani and Hingray (2009) found that the explanatory power of the climate predictors is better at the regional scale than at the local-scale and thus downscaled daily precipitation first to a regional scale, before they then disaggregated regional daily precipitation in time and space to local-scale sub-daily precipitation. As a result, they increased the robustness of the relationship between predictor and predictand. For the temporal disaggregation of daily to sub-daily precipitation, they used

an analogue method which was found to preserve the large spatial differences in precipitation maxima between different regions. A fundamental problem in the temporal disaggregation from daily to sub-daily precipitation, however, is caused by the underlying assumption, that the observed relationship between daily and sub-daily precipitation remain valid under a changing climate. But changes in daily and sub-daily precipitation may be considerably different in the future. For example, Beck (2013) found that the ratio between hourly and daily precipitation intensities has been increased during the second half of the 20th century in terms of average and extreme precipitation.

Analogue method

In a nonparametric analogue downscaling method, the most similar weather situation in the past is identified and the corresponding local-scale observation $Y()$ is selected as the predictand for this particular weather situation. It means that the simulated precipitation intensities by climate models, which tend to be uncertain (Beck and Bárdossy, 2013), are only used for finding the analogue day in the past. Both large-scale atmospheric and surface weather variables can be used to identify historical weather situations (Maraun et al., 2010b). For example, Lee and Jeong (2014) focused on daily precipitation to find similar days in the past by applying a k-nearest neighbour-resampling process. Mezghani and Hingray (2009) used daily regional precipitation and temperature values to identify analogue days in the past. In order to improve the selection of analogous weather conditions, they suggested to introduce synoptic atmospheric variables. Buishand and Brandsma (2001), however, showed that incorporating additional synoptic variables does not necessarily lead to improvements in the selection of analogue days. Nevertheless, Willems and Vrac (2011) classified the historical weather situations into different CPs and months before they used daily temperature and precipitation to find the most similar day in the past and to estimate sub-daily precipitation. Both surface variables, daily temperature and precipitation, can be considered as meaningful predictors in terms of projecting sub-daily precipitation. The temperature variable is expected to be of high predictive power in terms of taking into account the impacts of a changing climate. Daily precipitation enables the downscaling method to directly reflect on the historical weather situation in terms of precipitation. However, there was an increased risk of a limited sampling variability as the number of historical weather situations and their corresponding precipitation records became too restricted due to the high number of different classes (CPs and months). It also needs to be mentioned, that temperature differences between the reference and the future period, which are important in terms of perturbing the analogue day, were only calculated distribution-wise but not event-wise. As a result, the physical coherence

between the predictor (temperature) and the predictand (hourly precipitation) was also restricted.

In general, however, a main advantage of analogue day methods is that they are able to capture physical coherence between the predictor variables (e.g. sea level pressure, temperature or daily precipitation) and the predictand (hourly precipitation), as well as spatial coherence between different sites (Maraun et al., 2010b). For example, Mezghani and Hingray (2009) applied the analogue methods to disaggregate daily regional precipitation in space and time and found that the analogue method was able to reproduce observed spatial differences in sub-daily precipitation maxima between different regions. Analogue methods are also able to simulate infrequent but observed extreme hourly precipitation events (Mezghani and Hingray, 2009), as well as site-specific precipitation characteristics (Willems and Vrac, 2011) by not making restrictive assumption of the precipitation distribution (Gangopadhyay et al., 2005). Another reason for using an analogue day method is its technical simplicity along with the fact that it still provides comparable or better performances compared to other statistical downscaling methods (Zorita and Von Storch, 1999).

On the other hand, the main disadvantage of the standard analogue method is that the observed relationships between predictor and predictand are stationary. It means the standard analogue method is not able to simulate precipitation events that have not been observed in the past, which makes the projection of future precipitation events under a different climate unfeasible. As a consequence, Mezghani and Hingray (2009) estimated daily regional weather variables (temperature and precipitation) from GLMs (see above) to identify the analogue day and then used the temporal patterns (not the historical record) for this day to disaggregate the simulated regional weather variables (daily precipitation) in space and time (hour). In this case, the disaggregation factor between daily precipitation and hourly precipitation is assumed to be stationary in the future which may not be valid under a changing climate (Larsen et al., 2009). Instead, Willems and Vrac (2011) suggested the use of a perturbation factor based on the Clausius-Clapeyron (CC) relationship to allow downscaled precipitation intensities by the analogue day method to be higher than the observed records. The CC rate assumes 7% increase in precipitation per 1°C temperature increase, which is equivalent to the rise in the water vapour holding capacity of the atmosphere. However, Lenderink and van Meijgaard (2008) questioned the applicability of the CC equation in terms of projecting hourly precipitation extremes into the future by showing that for a station in the Netherlands extreme hourly precipitation may in some cases increase 14% instead of 7% per degree of warming. Similar results were found for the United States (Mishra et al., 2012) and under subtropical climate conditions for Hong Kong (Lenderink et al., 2011). On the other

hand, Hardwick Jones et al. (2010) showed that the relationship between extreme precipitation and temperature can also become negative for high temperature.

2.3.3 Model output statistics (MOS)

The second statistical downscaling group, model output statistics (MOS), uses the precipitation variable simulated by climate models directly to estimate precipitation events. The main difference between PP downscaling methods and MOS lies in the fact, that for MOS the statistical relationship is calibrated based on simulated predictors and observed predictand (precipitation) instead of observed predictors and predictands in PP methods. The predictand can be derived on the same spatial scale as the predictor (Bárdossy and Pegram, 2011; Lenderink et al., 2007) or the predictand can be downscaled to site-specific precipitation (Willems and Vrac, 2011). The predictors in a MOS method can be either simulated time series or properties of the simulated intensity distribution. The same holds true for the predictand variable; it can either be a precipitation time series or properties of the precipitation intensity distribution. In cases, where the calibration of the MOS method is done for a climate model run, which is not driven by reanalysis data, the MOS can only establish statistical relationships between the distribution of the simulated predictor variable and the observed predictand variable. However, if the climate model is driven by reanalysis data, which means the simulated predictor and the observed predictand are directly related from day to day, regression techniques can be applied as discussed for the PP method. For all MOS methods it needs to be kept in mind that the statistical relationships are model specific and cannot be readily applied to any other climate model (Maraun et al., 2010b).

2.3.3.1 Delta method

The delta method assumes that the future precipitation y_{i+T} at a time $i + T$ can be calculated by perturbing the observed precipitation $y_{i,obs}$ at a time i in the historical records by the ratio of the mean simulated future precipitation \bar{y}_{sim}^{fut} and the mean simulated control (or reanalysis-driven) run precipitation \bar{y}_{sim}^{con} . In this simplest delta approach only future changes in the mean precipitation are considered as described in Equation 2.6, whereas changes in the variance of the precipitation distribution or in the number of wet days are ignored. However, Shabalova et al. (2003) incorporated future changes in the variance of the precipitation distribution into the delta method and produced significantly different results compared to the simplest delta approach. Another option is to derive and apply the change factors for different quantiles of the precipitation distribution.

$$y_{i+T} = y_{i,obs} \frac{\bar{y}_{sim}^{fut}}{\bar{y}_{sim}^{con}} \quad (2.6)$$

All delta methods rely on the climate model's ability to simulate long-term changes in precipitation. Instead of perturbing an observed precipitation time series, Lenderink et al. (2007) used the corrected RCM control run time series and perturbed it by change factors derived from the RCM runs. In this case the RCM control run time series covered 90 years of data, although only 35 years of observations were available. Nevertheless, all delta methods are subject to the following two limitations in the change factors: (1) when they are derived from daily precipitation they may not be the same for sub-daily precipitation (Beck, 2013; Larsen et al., 2009), and (2) when they are derived on a grid-scale they may not be applicable to local-scale. For example, Olsson et al. (2012) showed that the local change factors were generally larger by a factor of 5% to 10% compared to the gridded change factors.

2.3.3.2 Direct forcing method

In contrast to the delta method, the direct forcing method uses the simulated future precipitation y_{sim}^{fut} 'directly' and corrects the biases inherited from the climate model. In the simplest form, it only considers ratio between the mean observed precipitation \bar{y}_{obs}^{con} and the mean control run (or reanalysis driven) precipitation \bar{y}_{sim}^{con} as described in Equation 2.7.

$$y_i^{fut} = y_{sim}^{fut} \frac{\bar{y}_{obs}^{con}}{\bar{y}_{sim}^{con}} \quad (2.7)$$

More sophisticated methods also correct the biases of the variance or of different quantiles (Sunyer et al., 2015). The direct forcing method aims to capture changes in future precipitation, such as wet day probability or skewness of the precipitation distribution. It also provides precipitation values, which are physically consistent as it directly uses the precipitation simulated by the climate model. However, the bias correction may disrupt the internal physical coherence inherited by the climate models, in particular between precipitation and temperature. The direct forcing method also relies heavily on the climate model's ability to reproduce the precipitation temporal and spatial variability realistically (Lenderink et al., 2007). Where the simulated precipitation has simply no skill in capturing the temporal and spatial structure of precipitation, the application of a direct forcing method should not be justified even if the corrected and observed intensity distribution are brought into perfect agreement. It should also be pointed out that the direct forcing method is not able

to downscale precipitation events to local-scale (Maraun, 2013), since the precipitation estimates are restricted to the spatial scale on which the climate model provides data.

Instead of only considering the ratio of the mean precipitation, Bárdossy and Pegram (2011) applied a more flexible tool, so-called quantile mapping, to project future daily precipitation extremes. In this case, the bias correction of the precipitation intensities is conditioned on the precipitation distribution quantiles. However, the quantile mapping method does not explicitly take into account the tail of the distribution, and as a consequence extreme precipitation may not be accurately corrected (Maraun et al., 2010b).

In terms of projecting future precipitation extremes on an hourly time-scale, the main limitation of the direct forcing method is that climate models are not able to represent the characteristics of extreme hourly precipitation sufficiently (Lenderink and van Meijgaard, 2008). However, this is likely to change in the future as RCMs will use finer spatial resolutions and thus resolve physical processes more realistically (Kendon et al., 2012).

2.3.4 Weather generator (WG)

Weather Generators (WGs) can be described as statistical models that generate random sequences of weather variables, with statistical properties that resemble those of the observed weather. The general motivation behind the usage of WGs is their capacity to simulate synthetic time series of unlimited length, the possibility of filling missing values, and their computational efficiency. WGs are also able to simulate precipitation at very high temporal and spatial resolution. In contrast to pure PP and MOS methods, which simulate the temporal and spatial correlations only by the predictors, WGs explicitly aim to generate time series which capture the observed temporal and spatial structure.

For example, Jones et al. (2009) introduced the UK climate projections (UKCP09) based on a weather generator, which provided future daily and hourly precipitation time series. Molnar and Burlando (2005) applied a WG based on cascade models to disaggregate observed daily precipitation to 10 min precipitation. The advantage of the cascade models lies in the fact that they only require a small number of parameters. Beuchat et al. (2011) and Kilsby et al. (2007) used a WG based on Poisson cluster processes, which incorporates physical interpretable quantities such as storm arrival rate, mean cell intensity and mean number of cells per storm, and have been found to provide meaningful simulation down to hourly timescale. The simulated precipitation cells may be seen, at least conceptually, as loosely representing small-scale precipitation-bearing meteorological structures. Burton et al. (2010) used a variant of this WG based on the Neyman Scott Rectangular Pulses (NSPR) model to generate transient site-

specific daily precipitation time series for the time period 1997-2085. The NSPR stochastic model also provided the basis for the standard UK industrial urban drainage design software and has been shown to realistically represent precipitation extremes for engineering impact studies (Kilsby et al., 2007).

Similarly to the MOS delta method, the main disadvantage of WGs in terms of projecting hourly precipitation extremes lies in the fact that WGs rely on change factors. These change factors are normally derived from daily precipitation simulated by climate models (Onof and Arnbjerg-Nielsen, 2009). However, it is highly questionable that changes in hourly precipitation are the same as changes derived from daily precipitation (Larsen et al., 2009). Change factors derived from climate models are also not able to represent localized convective precipitation events (Chan, 2013), which are particularly important concerning hourly precipitation extremes. Apart from that, WGs often underestimate the inter-annual variability (overdispersion) and the frequency of extremes. This is due to the fact that WGs are not able to take into account the climatic processes of the local weather, which exhibit longer-term variability. A possible solution to simulate low frequency aspects more realistically is to condition specific parameters of the WG on covariates, such as large-scale atmospheric circulation (Maraun et al., 2010b). Another option is the Poisson cluster model, in which the parameters of the WG for each day are conditional on the particular weather state observed on that day.

2.3.5 Summary

Statistical downscaling methods are able to add significant value to climate model projections of precipitation. It enables precipitation projections on a local- scale, which are particularly important for the estimation of flash-floods in urban environments and in small or steep river catchments among others.

In contrast, precipitation projections from GCMs/RCMs should always be considered as spatially averaged values (Chan, 2013). Related to the spatial resolution, there is also a minimum temporal scale on which RCMs can provide meaningful information in terms of precipitation characteristics. As mentioned earlier, even RCMs at 12.5 km considerably underestimate extreme hourly precipitation in summer over the UK (Chan et al., 2014a; Chan et al., 2014b; Kendon et al., 2014).

As a consequence, there is a need for statistical downscaling methods to provide projections of sub-daily precipitation extremes. Various statistical downscaling methods have been applied in the past for this purpose (see Table 2.1). However, remaining gaps still exist. They include

uncertainties due to sparse observed data, representation of summer convective precipitation events, sub-daily precipitation, changes in small-scale processes and their feedback on larger scales (e.g. soil-precipitation feedback and its impact on evaporation) as noted in Maraun et al. (2010b), and the assumption of stationarity between the predictor and predictand. In particular the representation of extreme precipitation remains a challenging task. This explains why only few statistical models are available to simulate sub-daily precipitation extremes. Therefore, this thesis will make a valuable contribution to fill the research gap where downscaling of hourly precipitation is yet to be much undertaken.

Statistical downscaling method	Advantages	Disadvantages
Perfect prognosis		
<p>Linear regression based (Hundecha and Bardossy, 2008; Mezghani and Hingray, 2009)</p>	<ul style="list-style-type: none"> • Based on statistical relationship which can be observed in reality • Able to simulate intensities not observed in the past 	<ul style="list-style-type: none"> • Explains only a fraction of the natural climate variability (Low inter-annual variability) • Highly sensitive to the calibration period • Reproduces poorly non-linear precipitation distributions • Is typically applied to downscale daily precipitation, disaggregation is required to estimate sub-daily precipitation
<p>Artificial neural networks (Haylock et al., 2006; Olsson et al., 2001; Wilby et al., 1998)</p>	<ul style="list-style-type: none"> • Skilful in reproducing inter-annual variability of precipitation indices • Based on statistical relationships which can be observed in reality 	<ul style="list-style-type: none"> • Tendency to underestimate precipitation extremes • Explains only a fraction of the natural climate variability
<p>Analogue method (Buishand and Brandsma, 2001; Lee and Jeong, 2014; Mezghani and Hingray, 2009; Willems and Vrac, 2011)</p>	<ul style="list-style-type: none"> • Based on statistical relationship which can be observed in reality • Captures physically coherent relationships between predictor and predictand • Is able to simulate infrequent but observed precipitation extremes • Reproduces local precipitation characteristics 	<ul style="list-style-type: none"> • Explains only a fraction of the natural climate variability • Is not able to simulate events which have not been observed in the past (Disaggregation or perturbation factor is needed) • For short record time series, important meteorological situations may be missing

Model output statistics		
Delta method (Lenderink et al., 2007; Shabalova et al., 2003; Willems and Vrac, 2011)	<ul style="list-style-type: none"> • Preserves observed spatial and temporal variability 	<ul style="list-style-type: none"> • Change factor derived from daily precipitation may not be applicable to project sub-daily precipitation • Change factor are derived on grid-scale and may thus not be applicable to estimate local-scale precipitation
Direct forcing method (Bárdossy and Pegram, 2011; Lenderink et al., 2007)	<ul style="list-style-type: none"> • Is able to capture changes in wet day probability or in the skewness of the precipitation distribution (not restricted by the observed distribution) • Provides physically consistent precipitation intensities as they are directly simulated by the climate model 	<ul style="list-style-type: none"> • Bias correction may disrupt the physical coherence (e.g. between temperature and precipitation) • Relies heavily on the climate model's ability to reproduce temporal and spatial variability • Biases in extreme precipitation may not be accurately corrected since the tail of the distribution is not explicitly considered • Cannot provide precipitation estimates on local-scale (restricted to the resolution of the climate model)
Weather generator		
Poisson cluster processes (Beuchat et al., 2011; Kilsby et al., 2007) Cascade-based models (Molnar and Burlando, 2005)	<ul style="list-style-type: none"> • Generates precipitation series of unlimited length • Computational efficiency • Is able capture the observed temporal and spatial structure 	<ul style="list-style-type: none"> • Change factor derived from daily precipitation may not be applicable to project sub-daily precipitation • Change factors are derived on grid-scale and may thus not be applicable to estimate local-scale precipitation • Takes not into account natural climate variability (low inter-annual variability) • Underestimates the frequency of extremes

Table 2.1 Summary of the advantages and disadvantages of different statistical downscaling methods to estimate precipitation

2.4 The use of atmospheric circulation patterns (CPs)

2.4.1 Introduction

Categorization of synoptic weather conditions into atmospheric circulation patterns (CPs) is commonly used to describe the climate of a certain region (Beck and Philipp, 2010; Huth et al., 2008; Philipp et al., 2010) and CPs can be considered as a synoptic predictor for forecasting or predicting a range of local variables. However, due to the high non-linearity and influence of local factors present in the atmosphere, it is not an easy task to describe the relationship between atmospheric circulation and precipitation in particular. Nevertheless, several studies have shown that precipitation properties can be linked with large-scale CPs over the UK (Lamb, 1972; Svensson et al., 2002), Germany (Bárdossy et al., 2002; Haberlandt et al., 2014), Iberian Peninsula (Corte-Real et al., 1998; Fernández-González et al., 2012) and the United States (Wilson et al., 1992). Naturally, the interest in CP-classification is highest in regions where the day-to-day synoptic variability is the main driver of local weather conditions, which is in the mid and high latitudes of both hemispheres (Huth et al., 2008). CPs have also been used successfully in previous studies to project extreme precipitation occurrences and intensities for different regions across Europe (Bárdossy and Pegram, 2011; Brigode et al., 2013; Garavaglia et al., 2010; Willems and Vrac, 2011). In order to assess future changes in extreme precipitation, GCM/RCM outputs are downscaled based on statistical models with parameters depending on the CP. As noted in Haberlandt et al. (2014), by using CPs, temporal and spatial persistence of precipitation can be accounted for and the relationship of precipitation with other climate variables can be carefully considered. Another motivation to use CPs is based on the assumption, that the predictive capacity of climate models is greatest at the multiple, rather than the single, grid point level (Grotch and MacCracken, 1991). CPs are also a straightforward way to allow for nonlinear, but sensible and physically plausible, relationships between the large-scale predictor (CP) and the local-scale predictand (precipitation). The nonlinearity between predictor and local predictand is captured by the CPs, as each of them is supposed to exhibit a consistent and distinct relationship with extreme precipitation (Bárdossy et al., 2002; Garavaglia et al., 2011). CPs can also be very useful in distinguishing between frontal and convective precipitation events even though convective activity may also occur in frontal systems (Gregersen et al., 2013; Hand et al., 2004).

So far, most of the CP-classifications have been used to analyse extreme precipitation on a daily time scale, whereas only few studies related CP-classifications to sub-daily precipitation extremes. One example can be found in Willems and Vrac (2011), where objective Lamb weather types were used to predict future hourly precipitation intensities derived from

analogue days in the past. In another study, Svensson et al. (2002) analysed the diurnal characteristics of extreme 3hr and 24hr precipitation events based on the Mayes classification, which is similar to the Lamb weather type catalogue. Beck and Bárdossy (2013) applied a fuzzy rules CP-classification and subdivided each day into different CPs and temperature classes. They found that hourly precipitation extremes are strongly intensified with higher temperature under anticyclonic conditions. The strong reaction to temperature in summer for certain CPs was attributed to convective mechanism in the atmosphere. On the other hand, CPs with westerly airflow causing frontal precipitation showed the lowest reaction to temperature.

2.4.2 The input configuration of CP-classifications

First, it is important to note that in terms of spatial and temporal scale, the concept of circulation patterns is different from that of weather regimes. The latter is defined on large spatial scales (typically continental or hemispheric) and longer time scales than daily (typical 5-days to monthly means). They perform well in describing distinct types of surface climate (e.g. hot or cold spells). In contrast, CPs are defined on daily or even sub-daily time scale, depending on the data availability of atmospheric circulation variable and are able to explain climate features, such as extreme precipitation events (Philipp et al., 2007).

In general, atmospheric circulation should rather be considered as a continuum rather than a system with a certain number of clearly defined, well separated states. Therefore, a set of CPs defined over a given region can vary considerably depending on the different methods of the CP-classification (see Section 2.4.3) and the configuration of the classification explained in the following. Several examples of different input configurations of CP-classifications are shown in Table 2.2. Also, several evaluation measures are presented in Section 2.4.6 to assess the performances of CP-classifications and to find the most appropriate input configuration.

2.4.2.1 Pressure variable

CPs are commonly described by atmospheric pressure variables, such as geopotential heights (Bárdossy et al., 2002; Garavaglia et al., 2010) or mean sea level pressure (Bárdossy, 2010; Corte-Real et al., 1999; Jones et al., 1993). Instead of the absolute pressure values, normalized pressure anomalies can also be used (Bárdossy, 2010; Bárdossy et al., 2002). However, In some cases in addition to the pressure variables other climate variables, such as wind (Wilson et al., 1992), temperature or daily precipitation (Willems and Vrac, 2011), supplement the classification process as illustrated in Table 2.2.

2.4.2.2 Calibration time period

As the climate cannot be assumed to be stationary due to natural variability as well as anthropogenically forced changes, the final set of defined CPs is likely to vary depending on the chosen calibration period of the CP-classification.

2.4.2.3 Number of CPs

Defining the number of different CPs is a crucial part within the CP-classification process. Using not enough CPs leads to a loss of information as the CPs only represent averaged weather conditions. Using too many CPs result in very low frequencies of each CP making statistical analysis impossible (Bárdossy et al., 2002). In terms of simulating hourly precipitation, Haberlandt et al. (2014) concluded that even a relatively small number of CPs (eight) conditioned on two seasons resulted in too many subsamples to provide a robust estimation of the model parameter to simulate hourly precipitation events. In other studies, the number of CPs varies widely between depending on the classification method and purpose as shown in Table 2.2.

Several evaluation measures are described in Section 2.4.6 to assess the performances of classifications for different input variables (e.g. number of CPs) and to decide which input variables to be used. However, Philipp et al. (2007) used four different evaluation measures (explained cluster variance, silhouette index, cluster overlap ratio and Pseudo-F) to assess the quality of a CP-classification over the North-Atlantic-European region in terms of the separability between CPs and the variability within each CP conditioned on different number of CPs and found inconsistent signals for an optimal number of CPs. Therefore, they concluded that there is no clear natural separation between classes of CPs and rather a smooth transition between different individual CPs.

Another aspect in the discussion of the number of CPs is the concept of weather regimes as defined. According to Huth et al. (2008), the idea of a small number of naturally existing weather regimes with higher persistence and higher frequencies than other states, which are merely seen as transitions between the weather regimes, gave reason to the use of only a small number of CPs (two to four) in CP-classification schemes. However, Christiansen (2007) questioned the existence of a certain predetermined number of weather regimes in atmospheric circulation data sets generally.

2.4.2.4 Spatial domain

The spatial domain of the pressure data set used for the CP-classification also needs to be chosen wisely. The domain should be large enough to be able to capture all the relevant high and low pressure zones affecting the local climate over a given region. On the other hand, one

needs to be aware of data restriction in terms of the spatial domain, in particular if Regional Climate Models (RCMs) are used. Also, derived objective classifications are determined by complex and highly demanding computer algorithms and thus the information input, in terms of the pressure data, should not exceed the capacity of the algorithms. Predefined CP-classifications methods, such as the LWT (Lamb Weather types) catalogue (Lamb, 1972) and the objective Jenkinson UK catalogue (Jenkinson and Collison, 1977), are restricted by the fact that in order to predefine a precise and physical meaningful set of CPs, the volume of the information input needs to be narrowed down as much as possible. Concerning the spatial resolution, it is common standard for CP-classifications to be defined on spatial domains with resolutions in the range of 1° to 5° (see Table 2.2). Beck and Bárdossy (2013) found very similar results between the 2.5° gridded NCEP/NCAR reanalysis data set (Kistler et al., 2001) and the 5° gridded NCAR Sea Level Pressure data set and concluded that any spatial resolution is suitable as long as it is fine enough to reproduce cyclonic and anticyclonic conditions.

2.4.2.5 Seasonal variations

A final point to consider is the seasonal variation of certain atmospheric pressure systems. For example over the UK, anticyclonic CPs are more frequent in summer and westerly CPs are more dominant in winter (Lamb, 1972). Similarly, the relationship between CPs and precipitation may vary over the year. Winter precipitation is normally linked to large-scale precipitation events caused by frontal systems or cyclones (Chan et al., 2014a; Lenderink and van Meijgaard, 2008), whereas in summer convective localized showers (thunderstorms) often lead to sub-daily precipitation extremes (Beck and Bárdossy, 2013; Gregersen et al., 2013; Maraun et al., 2008; Maraun et al., 2010a; Westra et al., 2014) as higher solar intake triggers convective events. For example, Bárdossy et al. (2002) found that over Greece certain CPs are wet in winter but dry in summer.

Therefore, defining the CPs for different seasons separately by using temporal subsets of the data set, can improve the overall performance of the CP-classification. Evaluation measures (see Section 2.4.6) can be used to quantify the degree of improvement of the CP-classification (Huth et al., 2008). But future seasonal climates may not exactly correspond to the present ones (Gutiérrez et al., 2013) and thus future projections conditioned on seasonal sets of CP-classification may not be representative of the future climate. Another option is to classify only one set of CPs throughout the year, but to evaluate the CP-precipitation relationship in the classification process for different seasons.

Name of CP-classifications	For which precipitation area	Predefined / derived	Large-scale atmospheric variables	Spatial domain	Spatial resolution	# Seasons	# CPs
Fuzzy rules (Haberlandt et al., 2014)	Germany	Derived	Pressure anomalies of sea level pressure	Not known	5° * 5°	1	8
Fuzzy rules (Bárdossy, 2010)	Germany	Derived	Pressure anomalies of sea level pressure	Not known	5° * 5°	1	10-22
Fuzzy rules (Stehlik and Bárdossy, 2003)	Europe	Derived	Pressure anomalies of 700-mbar geopotential heights	30°N-75°N 30°W-50°E	5° * 5°	1	13
Fuzzy rules (Bárdossy et al., 2002)	Germany	Derived	Pressure anomalies of 500-, 700-mbar geopotential heights	35°N-65°N 15°W-40°E	5° * 5°	1	12
K-means clustering based on precipitation (Garavaglia et al., 2010)	France	Derived	700-, 1000-mbar geopotential heights	38°N-50.3°N 6.2°W-12.9°E	2.5° * 2.5°	1	8
K-means clustering based on pressure data (Corte-Real et al., 1998)	Portugal	Derived	Vorticity and wind (strength and direction) based on sea level pressure	25°N-60°N 30°W-10°E	5° * 5°	1	4
Principal Component Analysis (Wilson et al., 1992)	Washington (USA)	Derived	Sea level pressure, 850- mbar geopotential height, 850-mbar temperature, 850-mbar meridional and zonal winds	40°N -60°N 135°W-105°W	5° * 5°	4	4
Lamb Weather Types (Fernández-González et al., 2012)	Spain	Predefined	Sea level pressure to derive wind direction and vorticity	30°N-50°N 25°W-5°E	5° * 10°	1	26
Lamb-Weather Types (Jones et al., 1993)	UK	Predefined	Vorticity and wind (strength and direction) based on sea level pressure	45°N-65°N 20°W-10°E	5° * 10°	1	7 (basic LWTs)
Multiple Classifications (Philipp et al., 2010)	UK	Predefined & derived	Mean Sea Level Pressure	47°N-62°N 18°W-8°E	1° * 1°	1	9 & 18 & 27

Table 2.2 Examples of different input configurations of objective CP-classifications

2.4.3 Examples of different CP-classification methods

2.4.3.1 *Predefined CP-classifications*

Predefined CP-classifications define the final set of CPs (subjective classifications) or the thresholds to define the CPs (threshold-based classifications) prior to the process of assigning observational atmospheric circulation data set to the different CPs. Predefined CP-classifications often have in common a presumed concept of the relationship between atmospheric circulation and precipitation patterns on which the classification is based.

Most of the subjective predefined CP-classifications were derived before computers became widely available. Subjective classifications have the advantage that the expert knowledge and experience can be fully reflected in the development of the classification. However, subjective classifications cannot be reproduced and as a consequence can only be used by a very limited number of applicants in certain geographical regions. In the following, the focus is on objective CP-classifications as, in practice, all subjective classifications have the common drawback that they cannot be used automatically and so considerable manual effort is required to classify each day. Also, subjective classifications are at-risk to be affected by gradual changes in the methodology. For example, Stehlik and Bárdossy (2003) found that the dependence between objective and subjective classifications decreased over time. But they also showed that subjective classifications performed best in identifying CPs different to the precipitation mean. In general, subjective classifications, such as described in Lamb (1972) or Little et al. (2008), can be a very useful tool to validate automated objective CP-classifications based on the knowledge and experience of meteorologists defining the subjective classes. The most popular subjective classification for the UK are the Lamb Weather types (LWTs) described in Lamb (1972).

Threshold-based methods define the CPs indirectly by using thresholds to distinguish between different CPs. In contrast to the subjective method, threshold-based classification allows automated assignments of a given day to a certain CPs. However, the process of defining the threshold involves subjective decision making. The advantage of threshold-based CP-classification compared to subjective methods is their reproducibility and the fast processing of assigning each day automatically to a certain CPs (Philipp et al., 2010). The most popular example of threshold-based methods for the UK is the objective version of the Lamb Weather types described in Jenkinson and Collison (1977). They used sea level pressure to calculate indices of wind direction and vorticity. Each day was then classified by grouping these indices based on the predefined thresholds. The objective version of the LWT was later used in several studies, for example, over the UK (Conway and Jones, 1998; Jones et al., 1993), the Iberian

Peninsula (Fernández-González et al., 2012; Goodess and Palutikof, 1998) and Belgium (Willems and Vrac, 2011).

2.4.3.2 Derived CP-classifications

In contrast to predefined methods, derived CP-classifications use a certain algorithm for the definition of the final set of CPs. They are based on the concept that the observational atmospheric circulation data set has a certain structure, which needs to be identified to define the final set of CPs (Philipp et al., 2010). Most of the derived classification methods are solely based on a large-scale variable, such as mean sea level pressure (MSLP) or geopotential height, and the relationship of CP and precipitation is determined after the classification. Bottom-up classification methods are different as they are directly derived from the local predictand (e.g. precipitation). As a result, derived CP-classifications based on a bottom-up approach may result in a more contrasted or discretized classification in terms of the studied variable (e.g. precipitation). In general, all derived CP-classifications can be considered as objective CP-classifications, although truly objective CP-classifications do not exist in reality since every classification comprises subjective decision making (Palutikof et al., 1997). Derived CP-classifications can be further distinguished in two different groups, the principal component analysis (PCA)-based methods and the optimization algorithms-based methods.

Principal component analysis (PCA)-based methods

Principal component analysis methods determine principal components (PCs), which explain major fractions of the variance of the input data, and any given day is assigned to CPs based on some relation measure to the PC (Philipp et al., 2010). The potential of PCA to be used as a classification method was first discussed by (Richman, 1981) and then further developed by (Gong and Richman, 1995). PCA-based CP-classification methods have been used in several studies, for example, over the United States (Wilson et al., 1992) and the Pyrenees (Esteban et al., 2005). A main advantage of PCA-based classification methods is their ability to incorporate multiple atmospheric variables (e.g. wind, surface pressure and geopotential heights) from large-scale fields in a very efficient way into the classification process and that they can be used as a data reduction technique (Wilson et al., 1992).

Optimization algorithms-based methods

Optimization methods aim to arrange a set of days attributed to CPs in such a way that a certain function is optimized to achieve a minimal variability within each CP (Philipp et al., 2010). In contrast to PCA-based methods, classification based on optimization algorithm only need a relatively small number of selected input variables (Wilson et al., 1992).

Cluster methods are a general tool for finding groups in data sets and they have been applied in different scientific disciplines including CP-classifications. Cluster methods can be distinguished in two groups, hierarchical clustering methods and non-hierarchical clustering methods. In terms of non-hierarchical clustering methods, the most common one is the k -means method.

The k -means method tends to generate very well separated clusters (CPs) but those are equally-sized, which may be unrealistic (Huth et al., 2008). The number of clusters (CPs) has to be defined a priori. However, it is not a trivial task to pre-determine the “correct” number of clusters and k -means clustering algorithms can be potentially unstable. The ordering of days, which is just a matter of chance, together with the initial clustering partitions leads to random local optima for complex data sets (Philipp et al., 2007). One example can be found in Corte-Real et al. (1998), where they defined four CPs, associated with daily precipitation in Portugal, based on daily sea level pressure over the north-eastern Atlantic and western Europe. For this purpose, they used a k -means clustering algorithm coupled with principal component analysis (PCA). PCA was used to condense information in daily sea level pressure fields and to reduce the number of input variables of the k -means clustering algorithm. The objective of the k -means method was to find a set of clusters (CPs) by minimizing the Euclidean distance between the daily SLP patterns and the means of all SLP patterns within a cluster. The optimal number of CPs (four) was determined based on an information measure evaluating the classification performance in terms of identifying wet and dry CPs.

Hierarchical clustering methods start with each daily pressure patterns in its own cluster, then creating a nested sequence of partitions, and finally form one cluster containing all daily pressure patterns. Hierarchical cluster methods are always at risk of the snowball effect, which means that one large cluster is formed and more and more pressure patterns are attached to it, accompanied by other small clusters and unclassified pressure patterns (Huth et al., 2008). One example of an hierarchical cluster method can be found in Garavaglia et al. (2010). They used a Hierarchical Ascendant Classification (HAC) to cluster the population of rainy days into different classes based on a bottom-up approach to build an objective classifications of eight CPs at the daily scale, describing the different precipitation patterns observed over France. A bottom-up approach directly used the target variable (e.g. precipitation) to define the initial structure of the classification. The relevant precipitation patterns were then linked with atmospheric circulation patterns described by geopotential heights or mean sea level pressure (MSLP). This particular method has been successfully applied over France (Garavaglia et al., 2010) and British Columbia (Brigode et al., 2013) for the estimation of extreme precipitation events.

Another method to derive CP-classifications was developed by Bardossy et al. (1995) based on optimized fuzzy rules to describe atmospheric pressure anomalies. The fuzzy rules CP-classification has been applied in several studies, for example, over Germany (Bárdossy, 2010; Bárdossy and Pegram, 2011; Bárdossy et al., 2002; Haberlandt et al., 2014; Yang et al., 2010) using different numbers of CPs (between 8 and 20) and over Greece (Bárdossy et al., 2002). The classification was carried out using normalized pressure anomalies and every CP was described by a fuzzy rule k represented by a vector $v(k) = (v(1)^k \dots v(n)^k)$, where n is the number of grid points. Five different classes v were applied as described in Bárdossy (2010).

$v = 1$ Very high pressure

$v = 2$ Medium high pressure

$v = 3$ Medium low pressure

$v = 4$ Low pressure

$v = 5$ No influence on the CPs

The fifth index was introduced to describe pressure anomalies at those grid points that had no influence on the fuzzy rule. The location and number of such grid points depend on the fuzzy rule. Usually most grid points were assigned to the fifth index. The membership functions were defined as triangular fuzzy number differently in Bardossy et al. (1995) and Bárdossy et al. (2002) due to different normalization formulas to calculate pressure anomalies. The assignment of a given day to a certain fuzzy rule was based on the membership functions, which were calculated for every grid point and every fuzzy rule on any given day. The closer the membership function was to 1, the stronger was the belief that the pressure value at this grid point on this day belonged to the fuzzy rule. The membership values were then combined over all grid points to calculate the degree of fulfilment (DOF) between the fuzzy rule and the actual pressure conditions on a given day. The fuzzy rule for which DOF was maximal was selected as the CP for a given day. Therefore, the fuzzy rules give a qualitative, rather imprecise description of the atmospheric pressure conditions.

In Bardossy et al. (1995), the fuzzy rules, describing each CP, were initially defined based on verbal description of the subjective European Grosswetterlagen/-typen (Hess and Brezowsky, 1969). This prototype fuzzy rules based CP-classification could be categorized as a hybrid classification, because the CPs were defined subjectively but each day was assigned automatically to one of the fuzzy rules based on the membership functions.

In later studies, information of precipitation was directly used in the CP-classification process to define the optimal set of fuzzy rules objectively (Bárdossy, 2010; Bárdossy et al., 2002). In contrast to the bottom-up approach described in Garavaglia et al. (2010), the identification of

precipitation patterns and the definition of CPs were not two separate steps. Instead, the fuzzy rules were optimized automatically based on objective functions, which measure the performance of the classification in terms of identifying dry and wet CPs. These objective functions gave quantitative assessments of the wet day occurrence and precipitation amounts conditioned on a given CP in comparison to the climatological mean during the classification process. The objective functions need to be robust and fairly smooth so that the optimization process does not stop at a local minimum. It was found that CP-classifications using objective functions based on areal precipitation provide better results compared to those based on point precipitation because the distribution of the latter can be highly skewed (Bárdossy, 2010). Due to the very large number of possible fuzzy rules combinations, it was not feasible to compute the objective functions for every possible fuzzy rule combination. Instead, a simulated annealing algorithm (Aarts and Korst, 1988) was applied to solve the optimization problem of finding the fuzzy rules for which the optimization criterion is maximal. It needs to be pointed out that the relationship between atmospheric circulation and local precipitation was established within the classification process. Therefore, the automated CP-classification based on fuzzy rules has the advantage that it is able to incorporate information of local climate variables (precipitation, temperature) into the classification process to define the CPs and is readily applicable for any given region.

2.4.4 Examples of CP-classifications over the UK and their relationships with precipitation

Over the UK, the most prominent example of CP-classification is the Lamb Weather Type (LWT) catalogue first developed by Lamb as a subjective classification method (Lamb, 1972). He introduced seven pure types (Anticyclonic, Cyclonic, Northerly, Easterly, Southerly, Westerly and North-Westerly), 19 hybrid types that were similar to more than one pure type, and one type of unclassified days. Later, an objective version of the Lamb Weather Type classification was constructed by Jenkinson and Collison (1977) using indices of vorticity, wind flow strength and direction derived from gridded mean sea level pressure (MSLP) to attribute each day to one of the Lamb Weather types. Objective versions of the LWT classification continued to be widely used for many applications over the UK (Conway and Jones, 1998; Jones et al., 1993; Wilby and Quinn, 2013). However, one limitation of the LWTs classification is that it is not able to describe complex synoptic situation with binary airflow over the UK. For example, a day with prevailing southerly airflow over Ireland, southern England and Wales, and north-easterly airflow over Scotland can only be classified as southerly or as unclassified (Sweeney and O'Hare, 1992).

Stehlik and Bárdossy (2003) compared six sub-continental CP-classifications pairwise in respect to their mutual dependence. The strongest dependence among all pairs was found between the objective (Jenkinson and Collison, 1977) and the subjective (Lamb, 1972) derived Lamb Weather Types classification in comparison to the Grosswetterlagen (Hess and Brezowsky, 1969), an hybrid classification over Greece (Maheras, 1989) and the fuzzy rule classification applied over Germany and Greece individually (Bárdossy et al., 2002).

In terms of the frequency occurrences of LWTs, Jones et al. (1993) found that anticyclonic and westerly types are the predominant patterns over the UK. Anticyclonic conditions occur more often in spring (MAM), summer (JJA) and autumn (SON) compared to winter (DJF). In contrast, frequency occurrences of westerly types are lowest in spring, continuously increase over summer and autumn, and are highest in winter. Cyclonic types are most likely to occur in summer and least likely in winter.

Several studies showed that the Lamb Weather Types can also be used to explain differences in precipitation characteristics over the UK. For example, Conway and Jones (1998) used three different versions of CP-classifications to generate daily precipitation series over the UK. Two of them used the objective Jenkinson-Collison LWT classification scheme, while the third method applied different categories of vorticity. All three methods realistically reproduced monthly means, persistence and inter-annual variability of UK daily precipitation. Jones et al. (1993) determined correlation coefficients between an area-averaged England and Wales precipitation time series (Wigley et al., 1984) and the occurrence of pure (objective and subjective) LWTs. They found little difference in performance between the subjective and the objective pure LWTs. For the objective classification they used grid-point mean sea level pressure (MSLP) obtained from a relatively coarse 5° latitude by 10° longitude dataset spanning an area from 45°N to 65°N and 20°W to 10°E without using the corner grid points. It was found for all seasons that the anticyclonic LWT exhibits significant negative correlation with precipitation across the UK, whereas cyclonic LWTs showed significant positive correlation with UK precipitation. For the other five LWTs no relationship was found significant for more than one season. Sweeney and O'Hare (1992) analysed the relationship between daily precipitation and 27 (pure and hybrid LWTs. They showed that cyclonic south westerlies is the LWT which causes the highest daily precipitation on average over the UK as a whole, whereas anticyclonic northerlies and north easterlies are the driest.

In terms of the relationship between regional UK precipitation and LWTs, Conway and Jones (1998) analysed the occurrence of the seven pure objective LWTs and area-averaged precipitation over the North East of England (NEE). It was found that the cyclonic LWTs result

in the highest mean precipitation over NEE, while the anticyclonic LWTs are associated with the lowest mean precipitation over NEE. Similarly to the results shown in Jones et al. (1993), they found for all seasons that anticyclonic LWTs and NEE precipitation are significantly negative correlated, whereas cyclonic LWTs and NEE precipitation are significantly positive correlated. For the other five LWTs no relationship was found significant for more than one season. They also showed a significant positive correlation between vorticity (one of the three parameters used in the objective LWT classification scheme) and NEE precipitation.

Sweeney and O'Hare (1992) found that westerly LWTs have the highest impact in terms of daily precipitation over the western Highlands. These results underlined the importance of orographic influences in terms of the relationship between westerly airflow and precipitation. For days with prevailing northerly LWTs, enhanced precipitation is apparent in North Wales and for the East of England, where moist air is transported from the Irish Sea and the North Sea. In terms of easterly LWTs, the highest precipitation is found in southern England, as fronts may move from East to West with a prevailing low pressure system over France or the Bay of Biscay. In summer, days with easterly airflow may also be associated with increased convective activity as already unstable air is directed over warm land surface. In terms of southerly LWTs, it was shown that precipitation is highest where orographic lifting occurs. The mountains of Cork/Kerry, South Wales and the Highlands are the regions where this orographic influence is most pronounced. It was also shown that southerly and cyclonic LWTs are the two pure LWTs, which cause the highest precipitation across the UK.

In terms of extreme events, several studies assessed the relationship between cyclonic conditions and extreme precipitation over the UK. For example, Kendon et al. (2014) showed that about half of the 50 most extreme hourly precipitation events in the future simulated by a 1.5km RCM over the southern UK could be attributed to larger-scale storms (embedded convection within a front, mesoscale convective systems or squall lines), while the other half were associated with individual smaller storms. Almost all of the extreme precipitation events were associated with cyclonic air flow and hot humid conditions (high 850hPa wet-bulb potential temperature). In contrast, Blenkinsop et al. (2015) found that anticyclonic conditions favoured the occurrence of extreme hourly precipitation over the South East of England in the past. Similarly, Maraun et al. (2010a) showed that vorticity, indicating cyclonic activity, was only an important predictor in terms of observed extreme daily precipitation in northern England and Scotland, but its influence was vanishing or even negative in southern England. Blenkinsop et al. (2015) suggested that limited moisture availability may restrict future precipitation intensities particularly under anticyclonic conditions, in order to explain the

strong relationship between cyclonic conditions and future extreme precipitation as projected in Kendon et al. (2014).

Little et al. (2008) classified extreme precipitation events only based on the spatial layout of UK precipitation depth, instead of using any synoptic variables (e.g. geopotential height or mean sea level pressure). Therefore, it falls technically not into the category of CP-classification since the classification is solely based on precipitation data. However, the classification can be used as a comparison tool with the corresponding precipitation characteristics of “real” CP-classification. Five subjective types of UK extreme daily precipitation were defined: orographic, mesoscale convective complexes (MCC), thunderstorm, east coast and depression. Four of them (orographic, MCC events, east coast and depression) inherited characteristic large-scale atmospheric conditions, which can be also attributed to specific classes in a circulation pattern classification. Orographic precipitation events occur only in mountainous areas and are often associated with westerly winds over the UK. MCC events were associated with a small, shallow depression and intense precipitation from convective precipitation with a larger area of stratiform precipitation. They often originate from the Bay of Biscay, mostly affect the southern and south western part of the UK and the predominant month of occurrence is August. The largest UK daily precipitation have been linked with MCCs as noted in Little et al. (2008). Extreme precipitation attributed to the east coast type mostly affect the eastern part of the UK and is defined by depression travelling in a westerly direction. Those depressions slow down over land and bring moist and easterly winds from the North Sea. All precipitation events which were associated with a depression but not exhibiting the specific characteristics of any of the other types were classified as depression. The only type which was lacking clear characteristic synoptic conditions to attribute them to a specific CP is the thunderstorm class. This class represents spatially localized convective cells which occur often in summer.

2.4.5 Limitations of CP-classifications

In terms of predicting precipitation, several studies found that using a set of CPs as the only predictor is not sufficient. For example, Haberlandt et al. (2014) showed that changes in CP frequencies are relatively minor compared to changes in daily precipitation intensity. Wilby et al. (2002) noted that precipitation changes caused by changes in the CP frequencies are often not consistent with the precipitation simulated by the climate model. Bárdossy and Pegram (2011) found that Regional Climate Models (RCMs) predict a change in the relationships between CPs and precipitation intensity in a future warmer climate. Instead, additional predictors are required. For example, Willems and Vrac (2011) used temperature and daily

precipitation conditioned on CPs to simulate changes in hourly precipitation extremes. Beck and Bárdossy (2013) predicted hourly precipitation based on CPs and temperature. Another possibility is to include additional variables (e.g. temperature or moisture fluxes) directly into the CP-classification to predict precipitation patterns (Yang et al., 2010).

Considering hourly precipitation extremes, the highest intensities are usually linked with convective activity (Beck and Bárdossy, 2013; Gregersen et al., 2013; Lenderink and van Meijgaard, 2008). But localized convective showers are likely to lack a distinguishable atmospheric circulation feature and therefore attributing these localized precipitation events to a specific CP remains a challenging task. Nevertheless, some studies (Beck and Bárdossy, 2013; Blenkinsop et al., 2015) argued that localized convective events may be associated with high pressure (anticyclonic) conditions, which show strong positive correlation with UK temperature in summer (Jones et al., 1993). Similarly, Haberlandt et al. (2014) showed that generally dry CPs can have a high potential of causing extreme precipitation in summer.

2.4.6 Evaluation of CP-classification methods

Huth (1996) described several criteria to evaluate the performance of a CP-classification only based on the large-scale atmospheric variable. Among those are the consistency (small changes in the input configuration should only lead to small changes in the CP-classification) and the reproduction of predefined CPs. One important property of CP-classifications, which can be easily evaluated in a quantitative way, is the separability between classes or in other words the ability of the classification to create classes as distinct as possible. It can be measured by means of the pattern correlation ratio (PCR) as described in Huth (1996) based on the correlation between different pressure maps (Overland and Hiester, 1980). Other measures include the Silhouette index (Rousseeuw, 1987), the explained variation (Philipp et al., 2007) and the Pseudo-F index (Caliński and Harabasz, 1974). Philipp et al. (2010) compared different CP-classification methods (subjective, threshold-based, PCA-based and optimization algorithm-based) using CP frequency based properties, such as the variation of CP sizes within each classification, mean persistence of CP occurrences, mean seasonal variation of CP frequencies, mean inter-annual variation of CP frequencies and the mean trend-noise-ratio. It was found that the different CP-classifications exhibit different CP frequency based properties; at least the subjective and the optimization CP-classification methods can be clearly distinguished from the other CP-classification methods. It was also shown that there are some outliers in terms of CP frequency based properties within most of the CP-classification methods. It means that although the CP-classifications have been standardized to a large degree, differences in the CP-classification procedures such as different similarity metrics,

objective functions, data pre-processing and compilers/platforms have also an important influence on the CP-classification performances. As a result, it was concluded that a comprehensive and robust comparison between different CP-classification methods is very hard to fulfil.

In contrast to the measures described above, information measures do not focus on the large-scale atmospheric variable. Instead, they assess how well the CP-classifications describe the variability of the local climatology for a given region. In terms of precipitation, a good CP-classification needs to distinguish well between wet and dry conditions or the CP dependent precipitation characteristics should differ from the unconditional precipitation mean. Most of the applied information measures focused on the occurrence of precipitation (Bárdossy, 2010; Garavaglia et al., 2010; Wilson et al., 1992). Other information measures as in Bárdossy (2010) and Stehlik and Bárdossy (2003) also take the precipitation intensity into account by using a wetness index, mean wet day precipitation amount or the standard deviation of the wet day precipitation amount. The fuzzy rules CP-classification method presented in Bárdossy (2010) and described in more detail in Bárdossy et al. (2002) is unique by integrating information measures directly into the classification process. This means that the final set of CPs is calibrated based on information measures, so called objective functions, whereas for most other classification methods the calibration of the final set of CPs and the use of information measures are two separate steps.

In the following, different aspects of the evaluation process are discussed.

2.4.6.1 Reliability of the CP-classification / statistical downscaling method

In order to assess the reliability of the CP-classification, the quality of the CP-classification is tested using an independent time period for which it is not calibrated. This is the so called validation, also known in the literature as a split-sampling approach (Bárdossy et al., 2002). For example, Garavaglia et al. (2011) compared return levels conditioned on the same CPs but estimated over different time periods. The results are often used to compare different input configurations of CP-classifications and to identify the 'best' CP-classification.

In terms of assessing the reliability of a certain statistical downscaling method, the method is also tested using an independent time period (Beuchat et al., 2012; Hundecha and Bardossy, 2008; Wilby et al., 2002). One particular interesting way to test the sensitivity (robustness) of the downscaling method is to calibrate the model on data from the coldest (driest) years and then validate it on data from the warmest (wettest) years (Willems et al., 2012).

2.4.6.2 Reliability of the simulated GCM/RCM predictor variables

The validation of CP-classifications can also be used to assess the ability of the climate model to reproduce the observed climate variables. For example, the large-scale pressure variable simulated by the climate model and its ability to reproduce the observed circulation pattern characteristics can be assessed by comparing the CP frequency (Haberlandt et al., 2014) or the correlation coefficients between observed and simulated CPs (Corte-Real et al., 1999). Another possibility would be to calculate the correlation coefficient between observed and simulated daily pressure maps. Goodess (2000) found that pressure anomalies simulated by GCMs tend to be smaller in magnitude in comparison to the observational data set. Therefore, CP-classification using GCM data might be biased towards a higher frequency of CPs with predominant moderate pressure anomalies. Kendon et al. (2012) pointed out the importance of climate models to represent atmospheric circulation patterns accurately in terms of the projections of local precipitation. During the validation process, the representation of the CP-precipitation relationship simulated by the climate model in comparison to that based on observed data may also be tested.

2.4.6.3 Stationarity test

As mentioned earlier, stationarity between the large-scale predictor (CP) and the local predictand (precipitation) cannot be assumed. Therefore, the stationarity test is important to assess if the observed CP-precipitation relationship remain valid in the future under a different climate. For example, Schmith (2008) suggested to test if the residuals of the predicted variable exhibit an overall trend for a high fraction of stations. First, the total precipitation anomalies were predicted by large-scale circulation anomalies over a 50-year time period. It was found that there were indeed instationarities in the predictor-predictand relationship. In a next step, relative humidity was used as an additional predictor and the instationarities in the predictor-predictand relationship were mostly removed as a consequence.

2.4.7 Summary

CPs are normally defined on a daily time scale based on atmospheric circulation variables, such as mean sea level pressure or geopotential heights. Those methods can be differentiated in predefined and derived CP-classifications. Predefined methods can be separated further into subjective and threshold-based CP-classifications. In terms of derived CP-classifications, different types exist, such as the PCA-based or the optimization algorithms-based methods. CP-classifications also vary depending on the input configuration (e.g. the pressure variable, the calibration time period etc.).

The theory behind the application of CPs is to link extreme precipitation events with certain large-scale pressure patterns as several studies (Corte-Real et al., 1998; Huth et al., 2008; Jones et al., 2006; Svensson et al., 2002; Wilson et al., 1992) found strong relationships between CPs and precipitation characteristics. Further, the use of CPs can help to understand the physical mechanism behind the occurrence of extreme precipitation better. For example, the relationships between extreme precipitation events and other climate variables can be more carefully investigated by conditioning them on CPs (Beck and Bárdossy, 2013; Haberlandt et al., 2014). CP-classifications can also be a useful tool to differentiate better between convective and stratiform precipitation events (Maraun et al., 2010b). Those reasons and the fact, that the pressure variable tend to be well reproduced by climate models (Kendon and Clark, 2008), make the use of CP-classifications a common predictor in terms of estimating precipitation extremes (Bárdossy and Pegram, 2011; Brigode et al., 2013; Garavaglia et al., 2010; Willems and Vrac, 2011). However, most of the studies focused on daily precipitation and none of them considered specifically the relationship between CPs and UK extreme hourly precipitation. One reason may be that extreme hourly precipitation are often linked with convective events but attributing those convective precipitation events to certain CPs remains a very challenging task due to the lack of distinguishable atmospheric features associated with those types of events. It was also shown that CP-classifications only are not enough to estimate future changes in precipitation characteristics sufficiently (Bárdossy and Pegram, 2011; Haberlandt et al., 2014; Huth et al., 2008; Wilby et al., 2002). This is due to the fact that the relationship between CPs and precipitation cannot be assumed stationary under a changing climate. Instead, additional predictors, such as temperature, are needed.

Considering the evaluation of the CP-classifications, the focus can be either on the large-scale atmospheric variable defining each CP or on the relationship between the CPs and precipitation. In both cases, it is common practice to evaluate the CP-classification over an independent time period different to the calibration period. (Philipp et al., 2010) found that the final results of the evaluation vary not only depending on the CP-classification method and the input configuration (pressure variable, calibration time period, number of CPs, spatial domain etc.) but also due to differences in the CP-classification process (pre-processing of data, compiler/platform, objective function etc.). Therefore, a comprehensive and robust evaluation between different CP-classification methods is very hard to achieve.

The focus of this thesis is only on the fuzzy rules based CP-classification method. This method is used to define different CPs over the UK and the first attempt is made to specifically address the relationships between different CPs and UK extreme hourly precipitation. As this literature review has shown, the use of CPs alone is not enough to estimate precipitation characteristics

sufficiently well. Therefore, the CPs defined in this thesis are combined with additional predictors in a statistical downscaling method to estimate and project UK extreme hourly precipitation.

Chapter 3

Data description and selection

A number of different data sets are used within this thesis. They are described in this chapter. The historical data set of daily and hourly precipitation at a station level in the UK is described in Section 3.1. This data set is used to provide the precipitation characteristics observed in the past for the projection of hourly precipitation extremes in the future. Besides, hourly station precipitation data is also extracted from the data set to define wet and dry circulation patterns in Chapter 5. The historical gridded sea level pressure (SLP) data used for identifying the actual synoptic condition of the circulation patterns is described in Section 3.2. In addition to CPs and daily precipitation, the temperature variable is used as a third predictor in the statistical downscaling method to estimate hourly precipitation extremes. It is also needed for determining the relationship between precipitation duration and temperature. The temperature data set is also described in Section 3.2. Finally, the Regional Climate Model (RCM) data sets, which are required to simulate the predictors (SLP, temperature, daily precipitation) to downscale future hourly precipitation extremes, are presented in Section 3.3.

3.1 Observed daily and hourly station precipitation data

3.1.1 Introduction

The entire data set of UK historical precipitation records used in this thesis has been collated by the UK Met Office and is available under the MIDAS Land Surface Stations database (NCAS British Atmospheric Data Centre, 2012). It comprises records of daily and hourly precipitation with a high spatial coverage across the UK. However, stations tend to be located in valleys, which can lead to an underestimation of orographic precipitation (Chan, 2013). In total, more than 12,000 stations recording daily precipitation have been installed over the last 160 years (1853-2012). In terms of hourly precipitation, 530 stations in total have been installed over the last 64 years (1949-2012). However, some stations only operated for a few years and are thus of very limited value in this study. Also, a large number of stations have been closed over the last decades. In 2012, daily precipitation has been recorded by 2643 stations and hourly precipitation was still measured by 259 stations. From the earliest days of observing in the UK, the standard period for measuring daily precipitation is 0900 UTC – 0900 UTC. Over time, the instruments to measure precipitation developed from ordinary rain gauges to tipping bucket

rain gauges and tilting siphon rain gauges. Tipping bucket rain gauges are stored directly by the Met Office Solid State Event Recorder (SSER) system, which have been operating since the early 1980s. Today, all three types of instruments are in use across the UK.

Ordinary funnel-type precipitation gauges have been used for manual measurements since the earliest days of observing. With such devices, the amount of precipitation is measured using a rain-measuring glass at regular time intervals. The design of ordinary gauges has changed over the years. But, today the Met Office strongly encourages conformity in order to optimize comparability among different stations. The standard design has a rim of 127mm diameter and stands 30cm above the ground (NCAS British Atmospheric Data Centre, 2012).

Tipping bucket rain gauges are particularly suitable for automated sub-daily and sub-hourly measurements. They are based on a small container which, when filled with rain, tips and empties while recording the event electronically. But they do not perform well in freezing conditions.

Another instrument to measure automatically is the tilting siphon rain gauge which produces an autographic record of precipitation accumulation from a pen attached to a float in the precipitation chamber of the instrument.

The unit of precipitation is given in mm and amounts are measured and recorded to the nearest 0.2mm, and where possible, to the nearest 0.1mm (NCAS British Atmospheric Data Centre, 2012).

It needs to be noted that for most daily and hourly precipitation stations, the precipitation gauges have been replaced over the years or changes in the measurement conditions of the station occurred. Also as mentioned earlier, the lengths of the precipitation time series vary largely across different precipitation stations. As a result, there is a need for controlling the quality and pre-screening of the precipitation data to ensure that they are homogeneous, consistent with surrounding stations, and without extreme outliers. Section 3.1.2 describes these quality control measures. Once the data has undergone the quality control procedure, a number of stations that have relatively long time series and are spatially well representative over different UK extreme precipitation regions (Jones et al., 2014) are selected. These stations are then used for the calibration and validation of the statistical downscaling method.

3.1.2 Quality control and pre-screening

The Met Office routinely performs quality control to the data set to identify suspicious precipitation values by comparing each value with surrounding precipitation values. A screening procedure is then applied within this thesis to the Met Office quality controlled data set. It contains a number of measures and criteria summarised as follows.

- Precipitation values associated with certain Met Office quality control attributes are set to missing (see Section 3.1.2.1)
- When multiple precipitation values are available on the same date recorded by the same precipitation gauge, only one value is selected in this study based on the Met Office quality control attributes (see Section 3.1.2.2)
- Extreme outliers in the data are identified and set to missing by applying a certain threshold criteria (see Section 3.1.2.3)
- In cases where values are provided by more than one precipitation gauge for the same station on the same date, the values from those precipitation gauges that stopped operating earlier are set to missing. This ensures that the more recent sections of the precipitation series are more likely to be homogeneous (see Section 3.1.2.4)

3.1.2.1 Elimination of values with certain Met Office quality controlled attributes

Most of the daily and hourly precipitation values have been quality controlled by the Met Office to check that the data is correct and consistent with surrounding data. The exact quality control processes, as well as corrections, undertaken by the Met Office vary from station to station and thus cannot be further detailed in this thesis. Different Met Office quality controls include areal, sequence and internal consistency checks. Corrections undertaken by the Met Office comprise automated corrections, manual corrections, correction obtained retrospectively from the observer/station, corrections derived from a snow/rain equivalence, correction obtained by changing units of measurement or corrections obtained by applying a systematic adjustment. Values which have not undergone the Met Office quality control are set to missing in this study. These values are identified by the *version number=0* in the Met Office database. Values associated with *version number=1* in the Met Office database have been quality controlled by the Met Office and are used in this study.

The quality control attribute *prcp_amt_q* associated with each quality controlled value gives more information about the different aspects of the value's quality. It includes information about trace/accumulated precipitation, corrected precipitation values, suspected values, the different quality control processes and the quality control stage reached.

Hourly precipitation values as well as daily values, which were flagged by the Met Office as suspect and failed the Met Office quality control are associated with the quality control attribute *prcp_amt_q=129*. These values are set to missing in this study.

It is found that accumulated (e.g. precipitation over more than one hour) and trace hourly precipitation values are often not corrected through the quality control process by the Met

Office and can show high differences in precipitation amount compared to other hourly non-accumulated and non-trace measurements made by a different precipitation gauge at the same station on the same date. Therefore, accumulated hourly precipitation values, indicated as *prcp_amt_q=20009* and *prcp_amt_q=20109* in the Met Office database, often due to snowfall, fog, dew or frost are set to missing in this study. Besides, all hourly precipitation trace values which are indicated as suspect by the Met Office are set to zero in this study. These values are flagged with the quality control code of *prcp_amt_q=10129* and *prcp_amt_q=14129* in the Met Office database. In case of daily precipitation, accumulated and trace precipitation values are corrected through the quality control process by the Met Office and are thus not set to missing or corrected in this study.

3.1.2.2 Selection of one single value out of multiple values for the same precipitation gauge on the same date

For those values which have been quality checked by the Met Office, a *qc_level* has been attributed to each value by the Met Office which indicates the stage of the climate quality control (see Table 3.1). The *qc_level* is represented by the last digit of the quality control code, *prcp_amt_q*, in the Met office database. In this study in cases where more than one precipitation value is available for the same precipitation gauge (*rainfall_id*) on the same date, the values with the lower *qc_level* are set to missing and only the value with the highest *qc_level* is used. As a result 1,171,500 hourly and 2,332,700 daily precipitation values are set to missing.

<i>qc_level</i>	Stage of the climate quality control
0	Climate quality control has not run
1	Climate quality control has run
4	Climate quality control has included internal consistency or sequence checks
6	Climate quality control has a final areal check
9	Climate quality control was complete

Table 3.1 The different stages of the Met Office climate quality control.

3.1.2.3 Extreme outliers

Independently of the Met Office quality control, extreme outliers in the data set are identified by using the historical maximum of observed UK daily/hourly precipitation as a threshold in this study. The highest historical daily precipitation has been recorded in Martinstown, Dorset, on the 18th July 1955, 279mm (Met Office). As a result 41 daily precipitation values, exceeding the threshold of 279mm, are identified as extreme outliers and are set to missing in this study (see Table 3.2). In case of hourly precipitation, the highest intensity has been recorded in

Maidenhead, Berkshire, on the 12th July 1901, 92mm (Met Office). As a consequence, nine hourly precipitation values higher than 92mm are identified as extreme outliers and are not used in this study (see Table 3.2). This criterion is simply to identify extreme outliers generally without taking into account the specific precipitation characteristics of each station. Using a single threshold may not be well justified, especially when this single value is applied throughout the UK disregarding regional precipitation characteristics. There is a remaining risk that some outliers are not eliminated. Other methods that can be used to eliminate outliers can involve the use of arbitrarily selected threshold values. Hence, it is worth noting here that the performance of the downscaling work is always limited by the quality of the underlying data sets.

A summary of the numbers of daily and hourly values that are set to missing or to zero is shown in Table 3.2.

Criterion	Comment	Hourly values (56,623,000 in total)	Daily values (81,370,000 in total)
prcp_amt_q= 129	Suspicious precipitation values which failed the Met Office quality control	492	12
prcp_amt_q= 20009	Accumulated precipitation values	82,862	n/a
prcp_amt_q= 20109	Accumulated precipitation values	316	n/a
prcp_amt_q= 10129	Trace precipitation values	43	n/a
prcp_amt_q= 14129	Trace precipitation values	41	n/a
Multiple values	Multiple values for the same gauge on the same date	1,171,500	2,332,700
Extreme outliers	Values > 279mm (daily), 92mm (hourly)	9	41

Table 3.2 Numbers of hourly/daily values that are set to missing or to zero due to certain criteria described in Section 3.1.2.

3.1.2.4 Selection of one single value out of multiple values for the same station on the same date

A *rainfall-id* number has been attributed to each value by the Met Office. A change in the *rainfall-id* number identifies cases where the precipitation gauges have been replaced by newer ones or changes in the measurement conditions of the station have occurred. Changes in conditions include: relocation of the new station more than 400m away from the previous station; the new station is 15m altitude different from the previous station; the new environment is different (degree of shelter). As a result, for the same precipitation station

more than one precipitation value is often available on the same date, as the “older” (initial) precipitation gauge has not always stopped operating.

In this study, where multiple precipitation values are available at the same station on the same date, the *rainfall-id* number is used to select those values from the precipitation gauge with the most recent values. By using precipitation values continuously from the same precipitation gauge, precipitation series are likely to be homogeneous, with an emphasis on the more recent values.

Note, in theory, precipitation values of each precipitation gauges should be only attributed to one unique station. However, there are cases where precipitation values of the same precipitation gauge have been associated to two different stations in the data set. In this study, these values are listed to be considered in case the affected stations are used for further analysis.

3.1.3 Station selection (Statistical Downscaling)

A number of criteria are used to select the hourly precipitation stations that provide homogeneous precipitation time series with a high level of completeness. They are listed in the order of the steps as follows.

1. Relatively high completeness with no single missing year and less than 5% missing values over the time period 1980 to 2009 (see Section 3.1.3.1). A missing year is defined as a calendar year for which no precipitation records are available.
2. Consistency between hourly and daily precipitation records (see Section 3.1.3.2)
3. Homogenization test by means of the RHtest (see Section 3.1.3.3)
4. Spatial coverage of the UK extreme precipitation regions (see Section 3.1.3.4)

The same criteria are also applied to the hourly precipitation stations over the time period 1960 to 1989, which is used for the calibration of the fuzzy rules based CP-classification (see Section 5.1.2).

3.1.3.1 Incompleteness

In this study, only those stations are considered which provide hourly values over 1980-2009. Stations which started to record after 1980 or were closed before 2009 are not considered in this study. Also stations which exhibit missing years (a missing year is defined as a year for which no precipitation records are available) of data between 1980 and 2009 are eliminated. A large number of hourly precipitation stations started to record in the 1970s and early 1980s and as a result the availability of UK observed hourly precipitation values is much higher

between 1980-2009 compared to the time period 1961-1990 commonly used in previous studies (Lenderink et al., 2007; Willems and Vrac, 2011).

During the evaluation time period 1980-2009, there are 38 out of 530 stations that provide hourly precipitation records with no missing years (a missing year is defined as a year with no precipitation records available). The 38 precipitation stations that have more than 5% missing hourly values during 1980-2009 are eliminated. As a result, 34 precipitation stations remained for further selection.

3.1.3.2 Consistency between hourly and daily precipitation values

It is important that the precipitation stations provide consistent hourly and daily records. In order to assess the consistency, the hourly time series of precipitation were aggregated to daily time series and compared with the corresponding daily time series. In the MIDAS database, not every precipitation station that provides hourly records necessarily provides daily records for the same time period. In fact, none of the 34 remaining precipitation stations provide complete daily precipitation values over the time period 1980-2009. Instead, daily precipitation series range from 8 to 29 years of recorded data during the period 1980-2009 (see Table 3.3). This was unexpected and could be explained by the closure of daily precipitation gauges. It can be seen from the MIDAS database that a high number of them closed in 2008. As a result, hourly precipitation time series are aggregated to daily time series in this thesis, where daily precipitation values are required as predictors in the statistical downscaling process.

The Pearson's correlation coefficient $\rho_{X,Y}$ (see Equation 3.1) and the normalized root-mean-square deviation NRMSD (see Equation 3.3) are calculated to analyse how well the existing daily precipitation value match with the corresponding 24hr aggregated hourly precipitation on the same day. The correlation coefficient measures the strength and direction of the linear relationship between 24hr aggregated hourly and daily precipitation, whereas the NRMSD is a measure of the actual difference in precipitation intensities between daily and 24hr aggregated hourly precipitation. For the calculation of the NRMSD, the daily values are considered as the benchmark values, while the 24hr aggregated values are regarded as values predicted by a model or an estimator. A constant underestimation or overestimation of daily precipitation by 24hr aggregated precipitation can only be detected by NRMSD, not by the correlation coefficient. The correlation coefficient gives a value between +1 and -1, where 1 is total positive correlation, 0 indicates no correlation, and -1 represents total negative correlation. The NRMSD ranges from 0 to 1, where a value of 0 indicates total consistency

between the two variables (daily and 24hr aggregated precipitation) and 1 indicates maximum inconsistency between the two variables.

$$\rho_{X,Y} = \frac{1}{N-1} \sum_{i=1}^N \frac{(X_i - \bar{X})}{s_X} \frac{(Y_i - \bar{Y})}{s_Y} \quad (3.1)$$

N : Number of wet days with daily and 24hr aggregated precipitation values

X_i : Values of daily precipitation

Y_i : Values of 24-hr aggregated hourly precipitation

\bar{X} , \bar{Y} : Sample mean of X_i , Y_i

s_X , s_Y : Standard deviation of X_i , Y_i (see Equation 3.2)

$$s_X = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (X_i - \bar{X})^2} \quad (3.2)$$

$$NRMSD = \frac{\sqrt{\frac{\sum_{i=1}^N (X_i - Y_i)^2}{N}}}{X_{max} - X_{min}} \quad (3.3)$$

Discrepancies between 24hr aggregated and daily precipitation are found for nearly every station. But a large number of stations show generally high correlation and low NRMSD values between hourly and daily values (see Table 3.3). Discrepancies between hourly and daily values can be caused by different types of gauges as manual ordinary gauges are often used for daily values while for hourly values automated rain gauges, such as tipping bucket gauges or tilting siphon gauges, are often used. Another potential reason for differences can be due to human errors or inaccuracies in analysing the measurement (Melber and Oswald, personal communication) .

In this study, precipitation stations are not used which provide more than 15 years of daily precipitation values with $\rho_{X,Y} < 0.9$ or $NRMSD > 0.05$ over the time period 1980-2009. In Figure 3.1, the daily and 24hr aggregated values are plotted pairwise for the five stations which failed to meet this selection criterion. The stations at Kirkwall (North Highlands) and Manston (South East England) are both heavily affected by one daily precipitation record. After setting this daily value (02/11/2006, 61.2mm) to missing at Kirkwall, the correlation coefficient increases from 0.77 to 0.98 and the $NRMSD$ decreased from 0.060 to 0.036. As a result, the station at Kirkwall can be potentially used in this study and only this one specific daily value has to be set to missing in further analysis. At Manston, another value (07/01/1999, 1.8mm) from a different measurement gauge, which stopped operating soon after, is available on the

date of the outlier (07/01/1999, 101.8mm). This value shows a higher stage in the Met Office quality control and is almost identical with the 24hr aggregated value on the same date. After using this value at Manston, the correlation coefficient increases from 0.82 to 0.98 and the *NRMSD* decreased from 0.033 to 0.026. As a result, the station at Manston can be used in this study after replacing this one specific record (1.8mm instead of 101.8mm).

The station at Stornoway airport (North Highlands) exhibits a low correlation coefficient $\rho_{X,Y} = 0.76$ also due to one daily precipitation record (06/06/1999, 103.7mm). The correlation coefficient increases considerably to $\rho_{X,Y} = 0.92$, after setting this record to missing.

However, at the same time, the *NRMSD* also increases from 0.043 to 0.070 and thus failed the *NRMSD* criteria. As a consequence, the station at Stornoway airport is not used in this study.

The stations at Wick airport (East Scotland) and Lyneham (South England) exhibit multiple values with considerable discrepancies between hourly aggregated and daily precipitation. As a consequence, these two stations failed to meet the selection criteria and also cannot be used in this study.

It is interesting to note, that the suspected records (see Table 3.4) are all daily and not 24hr aggregated hourly values. However, considering the entire range of precipitation values (small, medium and high values) across all stations, no systematic overestimation of daily values compared to 24hr aggregated hourly values is found. In total, three (Stornoway airport, Wick airport and Lyneham) of 34 precipitation stations cannot be used due to large discrepancies between 24hr aggregated and daily precipitation.

Figure 3.2 shows the values of daily and 24hr aggregated values plotted pairwise for the final selected stations, which meet all the selection criteria defined in Section 3.1.3. For the station at Boscombe Down, one suspicious value of 24hr aggregated hourly records could be identified. This was caused by one hourly record (19/02/1996 00:00, 40.0mm). Such a high hourly precipitation value can only be caused by increased convective activity which is very unlikely to occur at this time of the year to this extent. As a consequence, this hourly precipitation value is eliminated in this study. Similarly for the station at Boulmer, suspected records of hourly precipitation (01/02/1996 10:00, 20.10mm and 20/02/1994 00:00, 20.0mm) are identified and eliminated.

The scatter plot for the station at Ronaldsway is not shown because it comprises less than 15 years of 24hr aggregated hourly and daily values between 1980 and 2009. As a result, the station at Ronaldsway is not used in this study.

srclد	Station name	Overlapping years	Overlapping days (>1mm)	Pearson's correlation coefficient	NRMSD
1046	RONALDSWAY	8	7	0.2181	0.4539
23	KIRKWALL	16	255	0.9828	0.0356
54	STORNOWAY AIRPORT	16	698	0.9243	0.0696
18974	TIREE	17	1085	0.9825	0.0276
161	DYCE	18	875	0.9837	0.0250
384	WADDINGTON	19	1209	0.9940	0.0090
132	KINLOSS	20	1292	0.9907	0.0163
32	WICK AIRPORT	21	345	0.8459	0.0816
775	MANSTON	22	950	0.9813	0.0261
556	NOTTINGHAM WATNALL	22	841	0.9774	0.0245
1070	CARLISLE	23	1545	0.9745	0.0333
1198	ABERPORTH	24	2554	0.9385	0.0213
708	HEATHROW	24	1591	0.9925	0.0114
643	SHAWBURY	24	1917	0.9841	0.0170
461	BEDFORD	25	970	0.9569	0.0271
9	LERWICK	26	3081	0.9798	0.0158
886	LYNEHAM	26	1543	0.7913	0.0548
409	MARHAM	26	1761	0.9871	0.0102
1395	CAMBORNE	26	2597	0.9869	0.0149
1023	ESKDALEMUIR	27	3267	0.9877	0.0177
842	HURN	27	1455	0.9865	0.0244
583	WITTERING	28	1805	0.9767	0.0159
605	BRIZE NORTON	28	2670	0.9894	0.0083
889	BOSCOMBE DOWN	28	2607	0.9790	0.0259
1346	CHIVENOR	29	2709	0.9673	0.0245
613	BENSON	29	1845	0.9810	0.0179
17314	LEEMING	29	2698	0.9798	0.0123
315	BOULMER	29	2683	0.9692	0.0279
386	CRANWELL	29	2546	0.9946	0.0090
862	ODIHAM	29	2537	0.9748	0.0256
235	LEUCHARS	29	2908	0.9917	0.0142
709	NORTHOLT	29	2862	0.9954	0.0077
1145	VALLEY	29	3357	0.9773	0.0214

Table 3.3 Results of the comparison of 24hr aggregated hourly (>1mm) and daily values (>1mm) for 34 stations described in Section 3.1.3.2. Stations in bold refer to those that have outlier values corrected or set to missing. Stations in italic failed the selection criterion defined in Section 3.1.3.2.

Station (Station number)	Date	Initial daily precipitation value [mm]	Corrected daily precipitation value
Kirkwall (23)	02/11/2006	61.2	Set to missing
Manston (775)	07/01/1999	101.8	1.8
Stornoway airport (54)	06/06/1999	103.7	Set to missing

Table 3.4 Detected suspicious records of daily precipitation values by comparing daily with 24hr aggregated precipitation values.

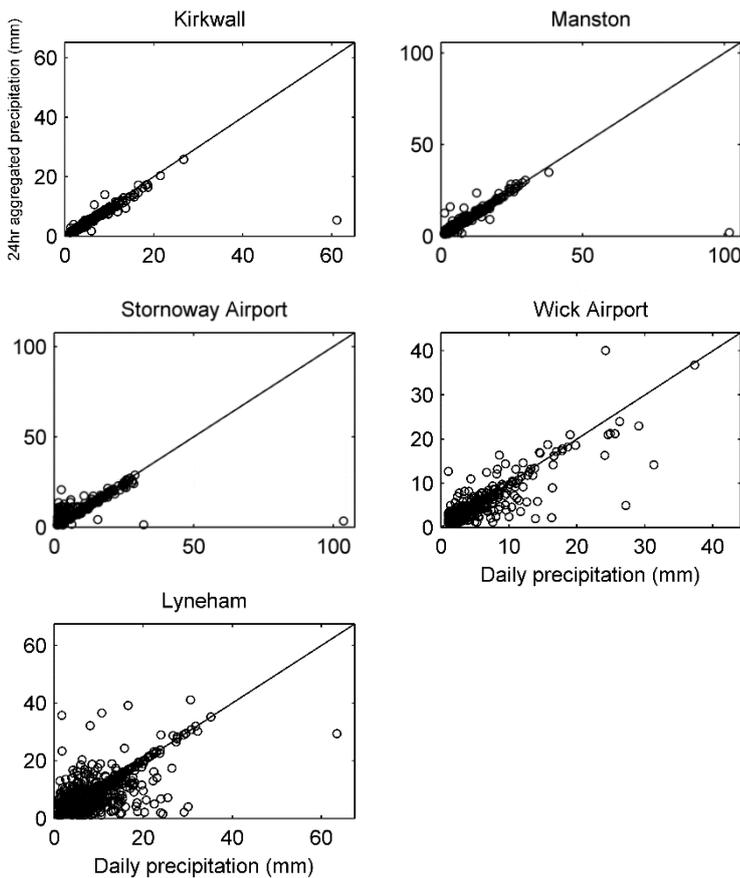


Figure 3.1 Scatter plot of daily precipitation (horizontal axis) and 24-hour aggregated precipitation values (vertical axis) recorded on the same day for the five stations with $\rho_{X,Y} < 0.9$ or $NRMSD > 0.05$ (failing selection criterion defined in 3.1.3.2).

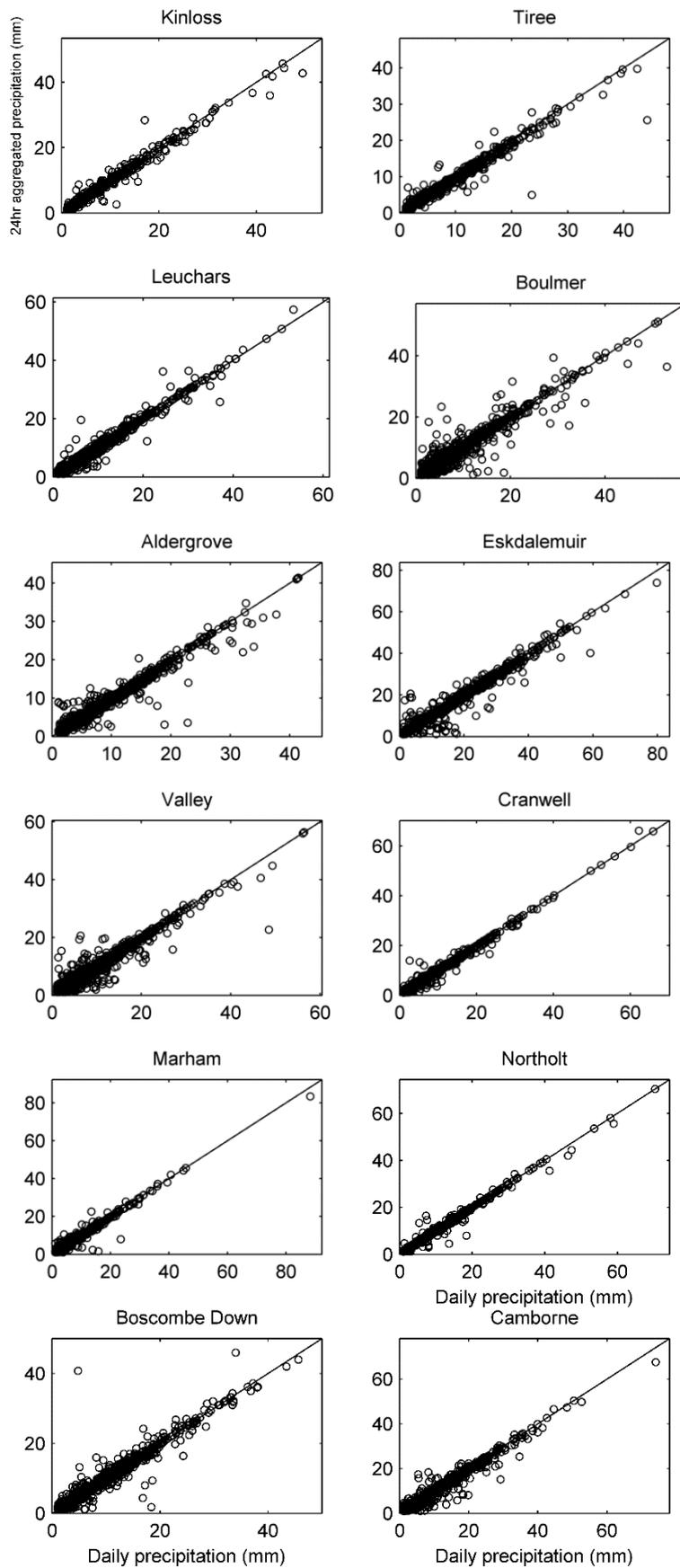


Figure 3.2 Scatter plot of daily precipitation (horizontal axis) and 24-hour aggregated precipitation values (vertical axis) on the same day for the final selected stations with $\rho_{X,Y} > 0.9$ and $NRMSD < 0.05$ (fulfilling all selection criteria).

3.1.3.3 Homogenization test by means of the RHtest

The RHtest V3, developed under the auspices of ETCCDI (Zhang) is a tool to homogenize precipitation time series by detecting artificial changing points and adjusting the biased precipitation amounts. In this study, it was used solely as a selection tool to identify and eliminate stations with an inhomogeneous time series. No adjustment of precipitation amounts was conducted. The RHtest was applied on monthly total precipitation amounts derived from the corresponding hourly precipitation values for each station because hourly/daily precipitation time series are much noisier and thus more difficult to test for changing points compared to monthly precipitation. Also, the monthly total precipitation amounts were logarithm transformed before being used in the RHtest as recommended in Wang and Feng (2010).

In particular those stations with multiple *rainfall-ids* associated to it, indicating a replacement or changes in the measurement conditions, are at risk of exhibiting inhomogeneities (changing points) in their precipitation time series. In this study, all of the 30 stations tested on homogeneity exhibit multiple *rainfall-ids* within the MIDAS database, but none of them showed any changing points based on the RHtest. Hence none of the 30 stations are eliminated in this step.

3.1.3.4 Spatial coverage of UK extreme precipitation regions

In Table 3.5 and Figure 3.3, the step-wise elimination of hourly precipitation stations, according to the criteria described in Section 3.1.3.1 and Section 3.1.3.2, is shown. Figure 3.4 illustrates the 14 different UK extreme precipitation regions defined in Jones et al. (2014) based on spatial differences of daily extreme precipitation. For this regionalisation scheme, precipitation statistics such as magnitude, seasonal variation and variance of yearly maxima are applied to find coherent regions with different characteristics in terms of daily extreme precipitation over the UK. In this study, twelve of the 14 UK extreme precipitation regions are represented by one or more of the final selected hourly precipitation stations. The two regions which are not represented by any hourly precipitation time series in this study are Mid Wales and South Scotland. In cases, where more than one hourly precipitation stations is available for an extreme precipitation region in the UK, the precipitation station with the lowest incompleteness over the time period 1980-2009 is selected. As a consequence, 18 stations are eliminated. The eliminated stations are still used for calibration purposes in terms of establishing relationships between hourly precipitation and other variables, such as daily precipitation, temperature etc. But the projection of hourly extremes is solely done for the one selected station within each region. As a result, twelve final precipitation stations are selected

to represent the characteristic spatial diversity in precipitation extremes over the UK (see Table 3.6 and Figure 3.4 - 3.6). It should be pointed out that the classification of the 14 UK extreme precipitation regions was solely based on daily precipitation pattern and it is possible that the same analysis based on hourly precipitation patterns could result in different UK extreme precipitation regions.

In previous studies, instead of the 14 UK extreme precipitation regions, the Hadley UK Precipitation (HadUKP) regions were used to analyse changes in current and future extreme precipitation (Fowler and Ekström, 2009). However, the nine HadUKP regions were developed from the England and Wales precipitation regions (Wigley et al., 1984) based on mean precipitation characteristics and do not specifically represent the spatial variability of UK extreme precipitation. In Figure 3.5, the HadUKP regions are shown as well as the final selection of hourly precipitation stations used in this study. Although the HadUKP regions are not assumed to reflect on the specific characteristics of precipitation extremes, it is worth mentioning that each region is still covered by at least one of the twelve precipitation stations.

Criterion	Comment	Number of eliminated stations	Number of selected stations
3.1.3.1	Incompleteness	496	34
3.1.3.2	Comparison of 24hr aggregated and daily values	4	30
3.1.3.3	Homogenization test (RHtest)	0	30
3.1.3.4	Spatial coverage of UK extreme precipitation regions	18	12

Table 3.5 Step-wise elimination of hourly precipitation stations (initially 530) depending on the specific criterion.

Station number	Station name	Extreme regions	Elevation [m]	Completeness	Rainfall-ids	Days with hourly and daily precipitation (>1mm)	Pearson's correlation coefficient	<i>NRMSD</i>
132	KINLOSS	ES	5	0.9910	1	1292	0.9907	0.0163
235	LEUCHARS	FOR	10	0.9941	1	2908	0.9917	0.0142
315	BOULMER	NE	23	0.9931	2	2683	0.9692	0.0279
386	CRANWELL	HUM	63	0.9935	2	2546	0.9946	0.0090
409	MARHAM	EA	21	0.9908	3	1761	0.9871	0.0102
709	NORTHOLT	SE	33	0.9945	1	2862	0.9954	0.0077
889	BOSCOMBE DOWN	WC	126	0.9954	2	2607	0.9790	0.0259
1023	ESKDALEMUIR	SOL	236	0.9774	2	3267	0.9877	0.0177
1145	VALLEY	NW	10	0.9976	2	3357	0.9773	0.0214
1395	CAMBORNE	SW	87	0.9914	1	2597	0.9869	0.0149
1450	ALDERGROVE	NI	63	0.9939	4	3022	0.9834	0.0229
18974	TIREE	NHI	9	0.9889	3	1085	0.9825	0.0276

Table 3.6 Details of the twelve final selected hourly precipitation stations following the selection criteria in Section 3.1.3 for the time period 1980-2009. The associated extreme precipitation regions are given as North Highlands and Islands (NHI), East Scotland (ES), South Scotland (SS), Forth (FOR), Northern Ireland (NI), Solway (SOL), North West (NW), North East (NE), Humber (HUM), Mid Wales (MW), East Anglia (EA), South West (SW), West Country (WC) and South East (SE). Completeness is given as the ratio of available precipitation values per station divided by the total number of hours between 1980 and 2009. *Rainfall-ids* indicate the number of changes in the measurement conditions between 1980 and 2009 according to the Met Office.

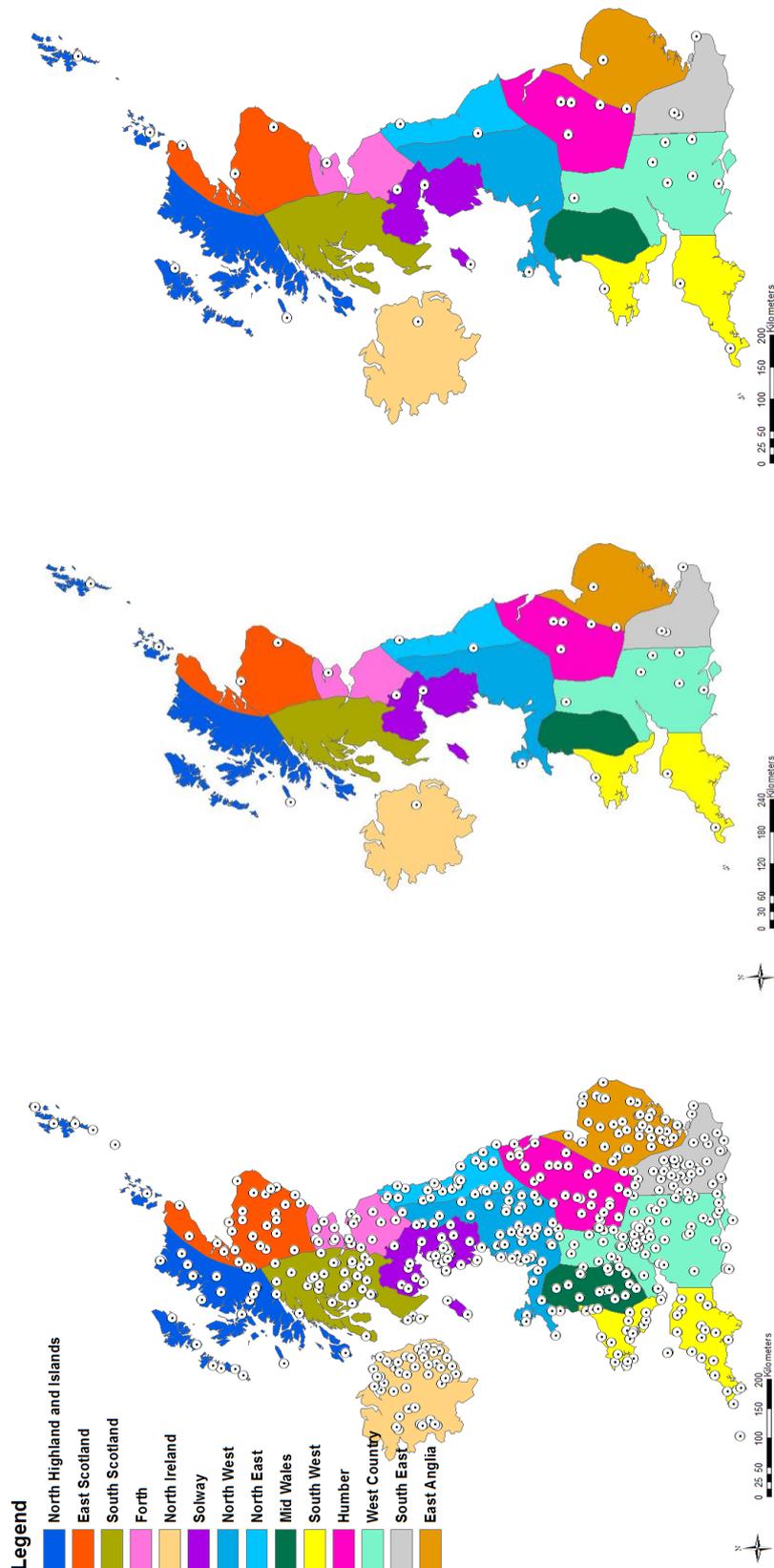


Figure 3.3 Step-wise elimination of UK hourly precipitation stations. Left: All available 530 hourly precipitation stations; middle: 34 hourly precipitation stations that provide records between 1980-2009 (criterion defined in Section 3.1.3.1); right: 30 hourly precipitation stations that fulfil criteria defined in 3.1.3.1 and 3.1.3.2. In colour the different extreme precipitation regions as described in Jones et al. (2014).

Legend

- North Highland and Islands
- East Scotland
- South Scotland
- Forth
- North Ireland
- Solway
- North West
- North East
- Mid Wales
- South West
- Humber
- West Country
- South East
- East Anglia

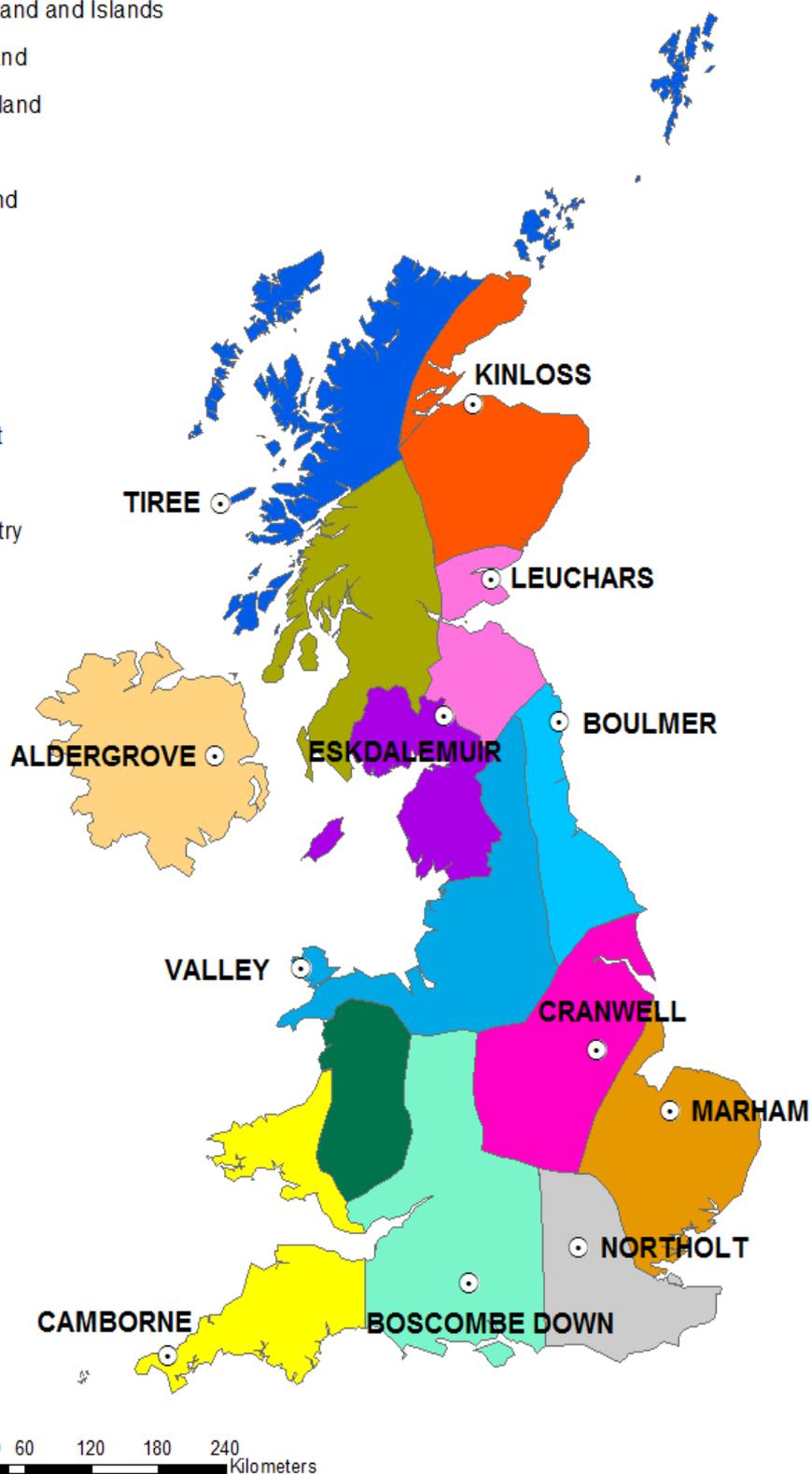


Figure 3.4 The twelve final selected hourly precipitation stations across the UK following the selection criteria in 3.1.3. In colour the different UK extreme precipitation regions as described in Jones et al. (2014).

Legend

- North Highland and Islands
- East Scotland
- South Scotland
- Forth
- North Ireland
- Solway
- North West
- North East
- Mid Wales
- South West
- Humber
- West Country
- South East
- East Anglia

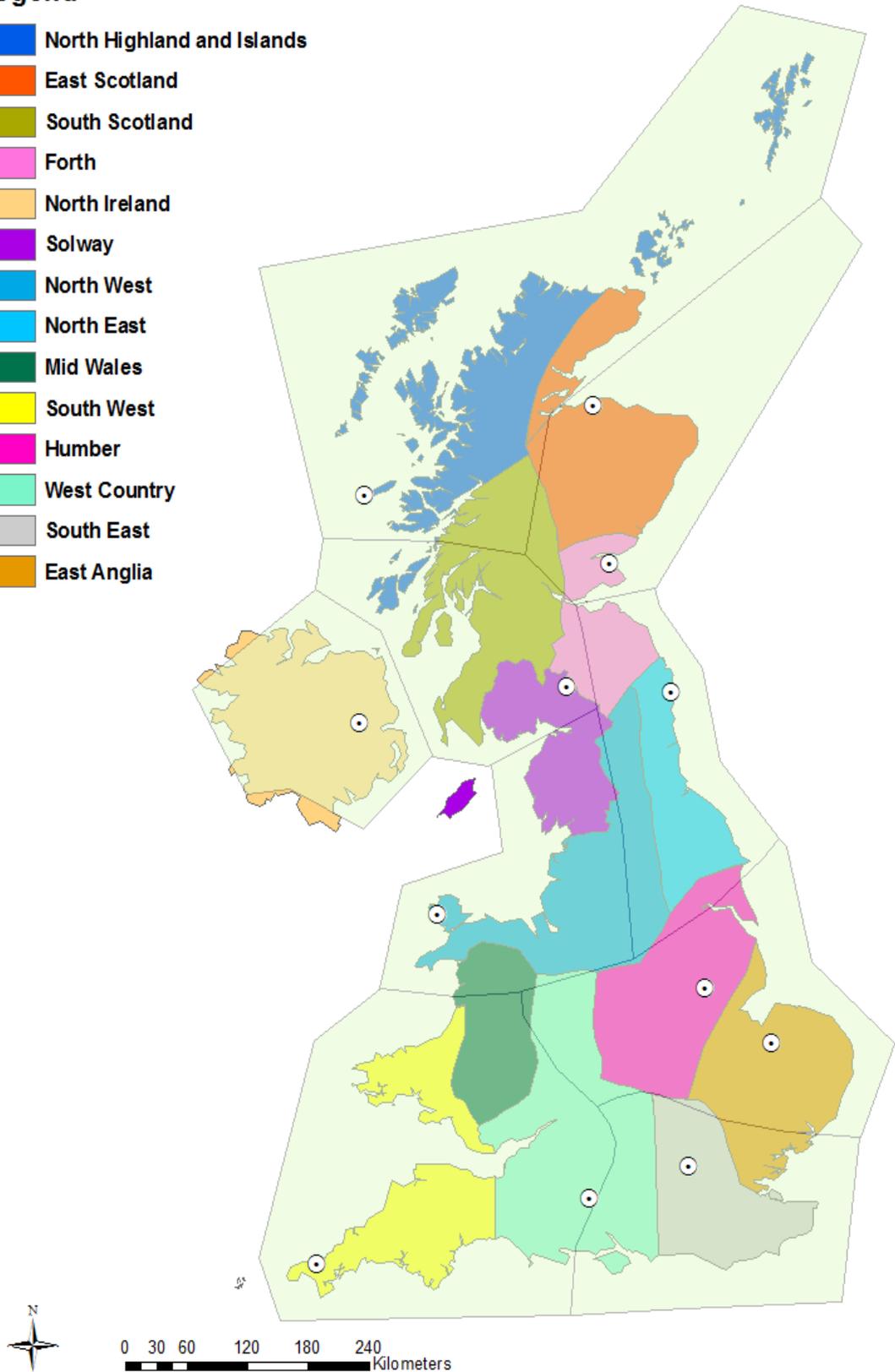


Figure 3.5 The twelve final selected hourly precipitation stations across the UK following the selection criteria in 3.1.3 with the HadUKP regions overlapped on the extreme precipitation regions (Jones et al., 2014).

Elevation [m]
High : 4783
Low : -214

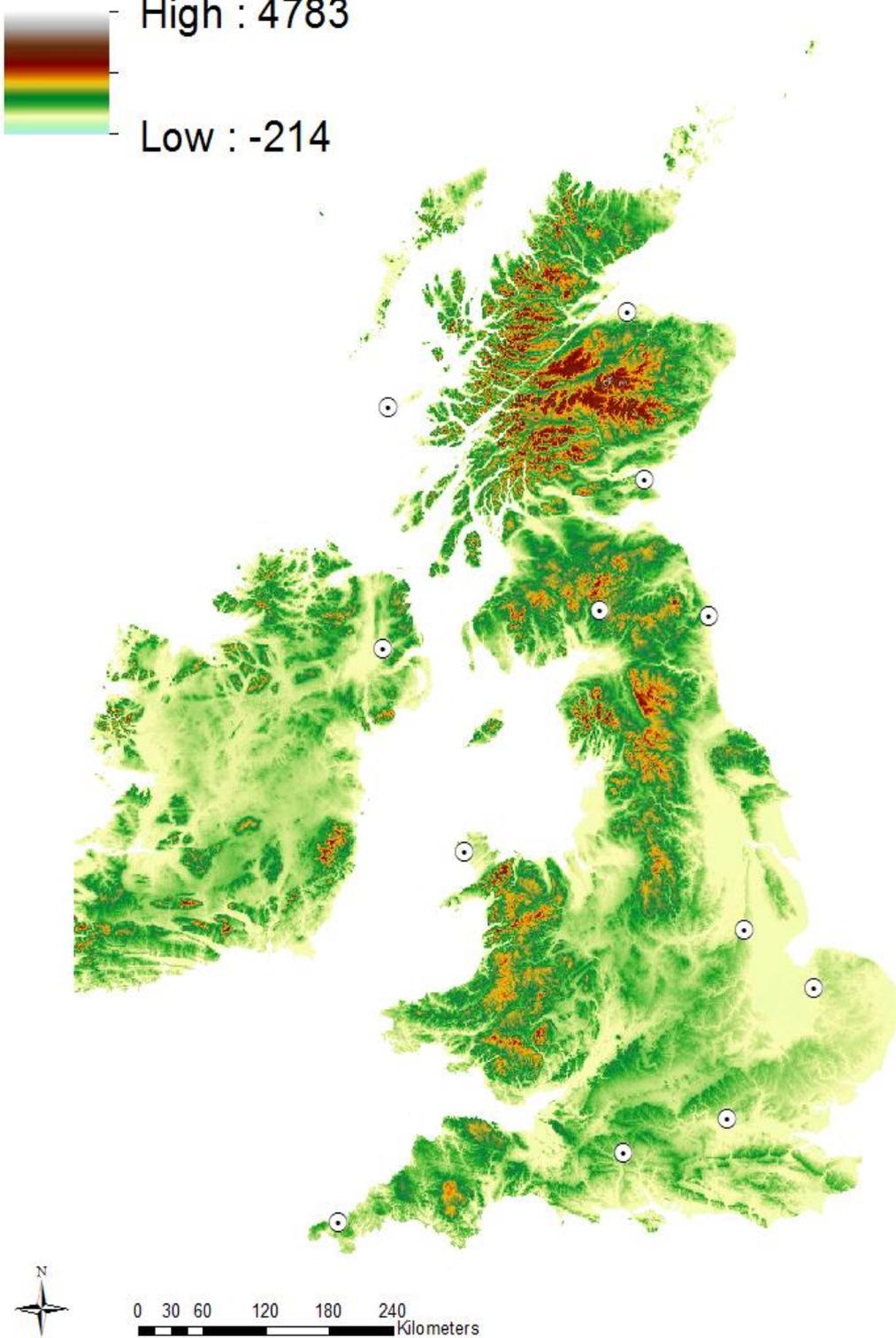


Figure 3.6 The twelve final selected hourly precipitation stations across the UK following the selection criteria in 3.1.3 with the elevation map.

3.1.4 Station selection for the calibration of fuzzy rules based CP-classifications

The same criteria described in Section 3.1.3 are applied to the candidate stations in this section, but during the calibration time period 1960-1989. The selection aims to ensure only the stations with high completeness of hourly precipitation data during the time period 1960-1989 are used for the fuzzy rules based calibration. Table 3.7 shows the three stations that are selected. They do not have missing years (a missing year is defined as a year with no precipitation records available) and have less than 0.5% missing values during 1960-1989. The two stations at Elmdon and Aldergrove provide consistent values between hourly and daily precipitation. The Stornoway airport station could not be tested due to limited daily precipitation availability, but remained in the selection to maximise the number of stations that could be used for the calibration. None of the three stations shows signs of inhomogeneities during 1960-1989 according to the RHtest, despite the fact that four changes occur in the measurement conditions at Elmdon over this time period. The three stations (Elmdon, Aldergrove and Stornoway Airport) also represent three different HadUKP regions respectively (Central England, Northern Ireland and Northern Scotland).

In total, five different precipitation Sets 1-5 are applied and evaluated for the calibration of the different CP-classifications. Precipitation Sets 1-3 comprise different site-specific hourly precipitation series (Table 5.2) derived from these three selected stations (Table 3.7). Precipitation Set 4 uses two series of hourly precipitation, each representing an average over multiple stations (Table 3.8 and Figure 5.1). The two precipitation series represent two different HadUKP regions, namely the South West England (Boscombe Down, Aberporth, Rhoose, Plymouth Mountbatten) and Northern Scotland (Wick Airport, Stornoway Airport, Tiree). For this Set, no selection criteria are applied in order to maximise the number of available stations within the two HadUKP regions. Set 1-4 use the number of wet hours per day as the input in order to distinguish between days with no precipitation, days with short spells of precipitation and days with persistent precipitation. In contrast to the other four precipitation Sets, Set 5 is based on the nine HadUKP regional daily precipitation intensities across the UK.

Station number	Station name	Elevation [m]	Completeness	Rainfall_ids
54	Stornoway Airport	15	0.9987	0
593	Elmdon	96	0.9957	4
1450	Aldergrove	63	0.9994	2

Table 3.7 Details of the three hourly precipitation stations used in three different precipitation Sets to define fuzzy rules CP-classification over the calibration period (1960-1989). Completeness is given as the ratio of available precipitation values per station divided by the total number of hours during 1960-1989. *Rainfall_ids* indicate the number of changes in the measurement conditions during 1960 and 1989 according to the metadata of the MIDAS database.

Station number	Station name	Elevation [m]	Completeness	Rainfall_ids
889	Boscombe Down	126	0.9950	1
1336	Plymouth Mountbatten	50	0.9915	1
1267	Rhoose	65	0.9914	2
1198	Aberporth	133	0.9795	2
18974	Tiree	9	0.9962	1
54	Stornoway Airport	15	0.9987	0
32	Wick Airport	36	0.9973	0

Table 3.8 Details of the hourly precipitation stations used in the fourth precipitation Set to define fuzzy rules CP-classification over the calibration period (1960-1989). Completeness is given as the ratio of available precipitation values per station divided by the total number of hours during 1960-1989. *Rainfall_ids* indicate the number of changes in the measurement conditions during 1960-1989 according to the metadata of the MIDAS database.

3.1.5 Precipitation analysis of the selected stations

3.1.5.1 Monthly/annual precipitation totals

In Table 3.9, the mean annual precipitation totals of the twelve selected stations are shown and the corresponding regional HadUKP precipitation totals between 1980 and 2009. Similarly, Table 3.10 includes the mean annual precipitation totals of the three stations used for the calibration of the fuzzy rules CP-classification and the corresponding regional HadUKP precipitation between 1960 and 1989. For eleven of the twelve selected stations in Table 3.9, the mean annual total precipitation of the respective regional HadUKP precipitation series is higher than the mean annual precipitation totals of the stations. Only the station at Eskdalemuir exhibits a higher mean annual precipitation total compared to its respective regional HadUKP precipitation series. The reason for that lies probably in the elevation of the stations. The station at Eskdalemuir has an elevation of 236m, while the other eleven selected

stations exhibit smaller values of elevation (between 5m and 126m). For the stations used for the calibration of the fuzzy rules CP-classification, lower precipitation totals can be found at Stornoway and Aldergrove compared to the corresponding regional HadUKP precipitation totals (see Table 3.10). This is also likely to be due to the low elevation at Stornoway (15m) and Aldergrove (63m) compared to the regional topography, which are characterized by mountainous areas in Northern Scotland and Northern Ireland. The fact that Aldergrove is located on the leeward side of a mountain range is likely to be another reason that relatively small precipitation totals are measured at Aldergrove compared to the regional average. In contrast, the standard deviation (see Equation 3.2) of the mean annual precipitation totals is not consistently lower than the standard deviation of the respective regional annual HadUKP precipitation total. This is very likely due to the fact that precipitation extremes are smoothed out in the spatially averaged regional HadUKP precipitation series. For all stations, the inter-annual variability of the regional HadUKP precipitation is realistically reproduced indicated by the Pearson's correlation coefficient.

In terms of the spatial variability of the annual precipitation totals, stations in the western part of the UK (Tiree, Eskdaelmuir, Aldergrove, Valley, Camborne and Stornoway airport) show higher annual totals compared to the rest of the stations. The reason for that lies in the fact that stations in the western part of the UK are more exposed to humid airflow coming in from the North Atlantic. Concerning the mean monthly precipitation averages shown in Figure 3.7 and Figure 3.8, it can be seen that the monthly precipitation total of the twelve selected stations and the three stations used in the fuzzy rules CP-classification reproduce the annual cycle of the regional HadUKP precipitation totals realistically.

Stations (HadUKP regions)	Mean [mm]	Standard deviation [mm]	Pearson's correlation coefficient
Kinloss (Eastern Scotland)	651 (793)	96 (91)	0.81
Tiree (Northern Scotland)	1257 (1737)	127 (192)	0.75
Leuchars (Eastern Scotland)	671 (793)	112 (91)	0.84
Eskdalemuir (Southern Scotland)	1663 (1494)	223 (162)	0.89
Aldergrove (Northern Ireland)	854 (1034)	95 (108)	0.86
Boulmer (North East England)	681 (809)	117 (110)	0.75
Valley (North West England)	842 (1059)	112 (135)	0.82
Cranwell (Central England)	605 (670)	88 (84)	0.93
Marham (Central England)	638 (670)	99 (84)	0.85
Northolt (South East England)	633 (745)	113 (106)	0.91
Boscombe Down (South West England)	751 (1067)	121 (120)	0.74
Camborne (South West England)	1044 (1067)	127 (120)	0.61

Table 3.9 Annual precipitation statistics over the time period 1980-2009 derived from the aggregated hourly records at the twelve selected stations. In parentheses are the statistics of the corresponding regional HadUKP precipitation series given. The Pearson's correlation coefficient is calculated for the annual precipitation totals of the station and the corresponding annual HadUKP precipitation totals.

Stations (HadUKP regions)	Mean [mm]	Standard deviation [mm]	Pearson's correlation coefficient
Stornoway Airport (Northern Scotland)	1153 (1580)	151 (201)	0.81
Aldergrove (Northern Ireland)	851 (1026)	90 (99)	0.88
Elmdon (Central England)	676 (654)	85 (80)	0.79

Table 3.10 Annual precipitation statistics over the time period 1960-1989 derived from the aggregated hourly records of the three hourly precipitation stations used for the calibration of the fuzzy-rules CP-classification. In parentheses are the precipitation statistics of the corresponding regional HadUKP precipitation series given. The Pearson's correlation coefficient is calculated for the annual precipitation totals of the station and the corresponding annual HadUKP precipitation totals.

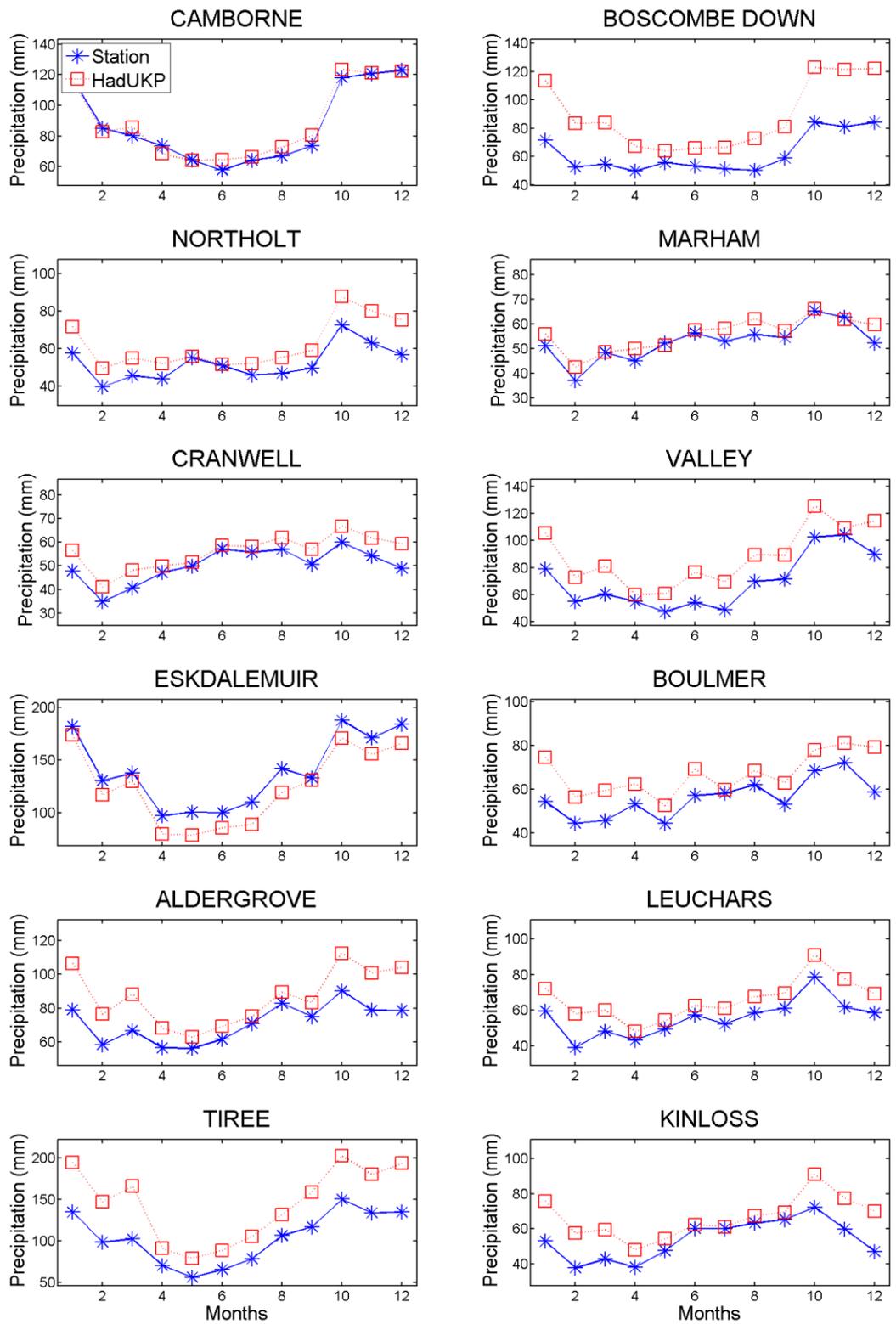


Figure 3.7 Monthly precipitation averages of the twelve selected stations (blue solid line with asterisk markers) and of the corresponding HadUKP regions (red dotted line with square markers) averaged over the time period 1980-2009.

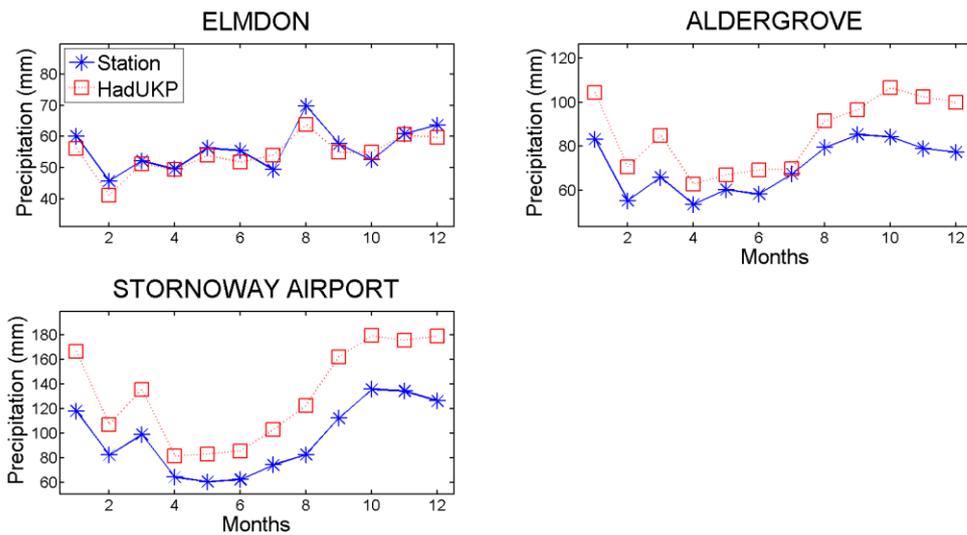


Figure 3.8 Monthly precipitation averages of the three stations used for the calibration of the fuzzy rules CP-classification (blue line) and of the corresponding HadUKP regions (red line) averaged over the time period 1960-1989.

3.1.5.2 Extreme daily precipitation

Figure 3.9 illustrates the monthly 99.5th percentile of daily precipitation for all the twelve selected stations, in order to give a representation of the intensity of daily precipitation extremes through the year. It can be seen that the station at Eskdalemuir exhibit higher intensities of daily precipitation extremes compared to the other stations for most of the months, except May, June and July. The higher intensities are likely due to the effects of orographic precipitation (Chan et al., 2014b; Chan, 2013) as Eskdalemuir is the station with the highest elevation (236m) and is situated in a mountainous area (Svensson et al., 2002). In terms of the annual cycle of the daily precipitation extremes, no clear systematic pattern can be identified over multiple stations. The fact that no pronounced maximum occur in the warmest months indicates the convective mechanism may be less important in driving daily precipitation extremes than in driving hourly precipitation extremes. Instead, stratiform precipitation events caused by cyclonic weather systems are likely to be most important in terms of daily precipitation extremes (Chan et al., 2014b).

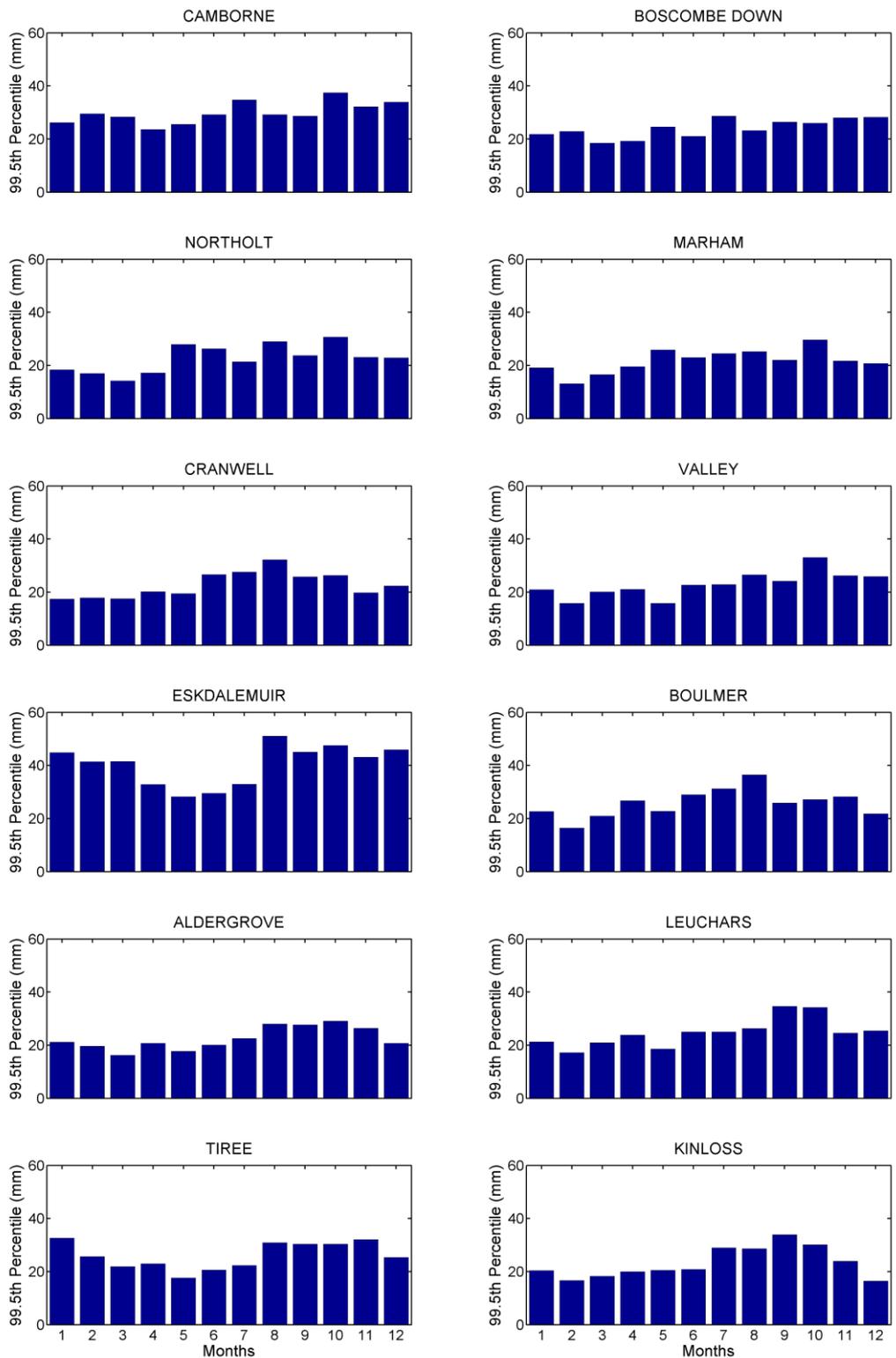


Figure 3.9 Annual cycle of the 99.5th percentile of daily precipitation between 1980 and 2009 for the twelve selected stations.

3.1.5.3 Extreme hourly precipitation

Figure 3.10 shows the statistical distributions of daily maximum hourly precipitation between 1980 and 2009 for the twelve selected stations. The station at Kinloss, which is located furthest North, exhibit a very steep distribution slope resulting in lower high percentile values compared to most of the other stations. For example, the 90th percentile at Kinloss amounts to 1.78mm/hr in comparison to 90th percentile values of 2.93mm/hr at Camborne or 2.81mm/hr at Tiree. However, these existing differences in the distribution for high percentiles are difficult to detect and almost vanish when considering the global distributions only. The reason for this is found in the high discrepancy between relatively high percentile (e.g. the 90th percentile) values and the very most extreme hourly precipitation values. For example, at Kinloss the most extreme hourly precipitation value is more than 30 times higher than the 90th percentiles. This shows the importance of using very high percentiles when conducting extreme hourly precipitation analysis, in order to represent the most extreme events realistically. Besides, for the design of flood prevention measures engineers are often interested in very rare extreme precipitation events (e.g. the 1% annual exceedance probability event, which occurs on average once every 100 years). But the measure of extreme event depends strongly on the available sample size of precipitation records. For example, it would be highly uncertain to estimate the 1% annual exceedance probability event based on a statistical downscaling method which is only calibrated over a 30-year period of precipitation records. As a result, the decision on which measure of extreme to focus on must always be a trade-off between a comprehensible interest in the most extreme precipitation events and the recognition of a limited sample size of precipitation records to calibrate the statistical downscaling method on. It should be also mentioned, that the most extreme events are often estimated by using the extreme value theory (EVT). This approach aims to derive a probability distribution for the extreme events only, which can be used to assess the probability of events that are outside of the observed data range. For example, Jones et al. (2013) applied the EVT to assess changes in the 10-, 20-, 50- and 100-year return levels of UK extreme daily precipitation but acknowledged that estimates of the higher return periods are not necessarily reliable. In this thesis, the focus will be on the 99.5th percentile of extreme hourly precipitation events, which occurs on average at least once every year, and therefore the EVT is not required.

In Figure 3.11, the monthly 99.5th percentiles of the maximum hourly precipitation of the twelve selected stations are shown. Two groups of stations can be identified with distinguishable annual cycle in terms of the 99.5th percentile. The first group comprises seven stations (Kinloss, Leuchars, Aldergrove, Boulmer, Cranwell, Marham, Northolt) and is characterized by higher 99.5th percentiles between May and October compared to the rest of

the year. The stations in this group also exhibit a pronounced maximum in the 99.5th percentiles between June and August. This is the time of the year when convective extreme precipitation events are most likely to occur (Hand et al., 2004). Therefore, the maximum is likely due to increased convective activity in the summer months (Blenkinsop et al., 2015; Chan et al., 2014b; Chan, 2013). It needs to be mentioned that the annual cycle representing the 99.5th percentile of the maximum hourly precipitation differs from the annual cycle of the 99.5th percentile of the daily precipitation. This suggests that summer convective precipitation events have a greater impact on hourly than on daily precipitation extremes (Gregersen et al., 2013).

The second group shows a different annual cycle in terms of hourly precipitation extremes. For the three stations (Tiree, Valley and Camborne) in this group, the second half of the year (July-Dec) is characterized by slightly higher 99.5th percentiles compared to the rest of the year and no pronounced maximum in the 99.5th percentile can be identified for the warmest time of the year (JJA). Therefore, the annual cycle of the hourly precipitation extremes for stations in the second group corresponds with the annual cycle of the daily precipitation extremes shown in Figure 3.9. The differences between the two groups in terms of the annual cycle of hourly precipitation extremes are likely to be due to geographical location, as the three stations in the second group are the stations most exposed to the North Atlantic and as a result more under the influence of westerly synoptic-scale weather systems. Chan (2013) argued that stations in the West are less affected by convective precipitation extremes and more by stratiform precipitation events, which is the dominant feature for winter precipitation (Chan et al., 2014b; Lenderink and van Meijgaard, 2008). As a result, these stations do not exhibit any maximum in hourly precipitation extremes in the summer (JJA) and a stronger relationship may exist between hourly precipitation and daily precipitation intensities. In terms of the convective available potential energy (CAPE), Holley et al. (2014) identified three main CAPE seasons: “land dominated CAPE” between April and September, “sea dominated CAPE” between September and January and “low CAPE” from January to April. As extreme hourly precipitation can often be linked with increased convective activity (Gregersen et al., 2013), this may also explain why the three stations most exposed to the North Atlantic exhibit higher extreme hourly precipitation in the second half of the year (July-Dec).

For the station at Eskdalemuir, orographic effects play an important role in terms of its precipitation pattern (see above). Orographic precipitation events can result in extremely high sub-daily precipitation extremes. For example, the highest 30min UK precipitation record is 80mm at Eskdalemuir on the 26th June 1953 according to the (Met Office).

As shown in Figure 3.11, the monthly maximum of each station of the 99.5th hourly precipitation percentiles is close to the threshold of 10 mm/hr. As a result, hourly precipitation values equal to or higher than the threshold of 10mm/hr are defined as extreme hourly precipitation for the validation of the CP-classifications (see Section 4.2.2 and 5.3.2). Figure 3.12 shows the annual cycle of the frequency of extreme hourly precipitation (\geq for 10mm/hr) for the twelve selected stations. The above defined two distinguishable groups based on the annual cycle of the monthly 99.5th percentile remain not valid in terms of the frequency of extreme hourly precipitation. Over the whole year, the three stations furthest north of the UK (Leuchars, Tiree and Kinloss) exhibit fewer extreme hourly precipitation events compared to the other nine stations. It is likely that this can be attributed to the fact that temperature tend to be lower for higher latitudes (see Section 3.2.2) resulting in reduced convective activity (Chan, 2013) and less intense hourly precipitation events (Berg et al., 2013; Molnar et al., 2014). However, in terms of the annual cycle of the frequency of extreme hourly precipitation all of the twelve stations share one similar pattern that the majority of extreme hourly precipitation events occur between May and October.

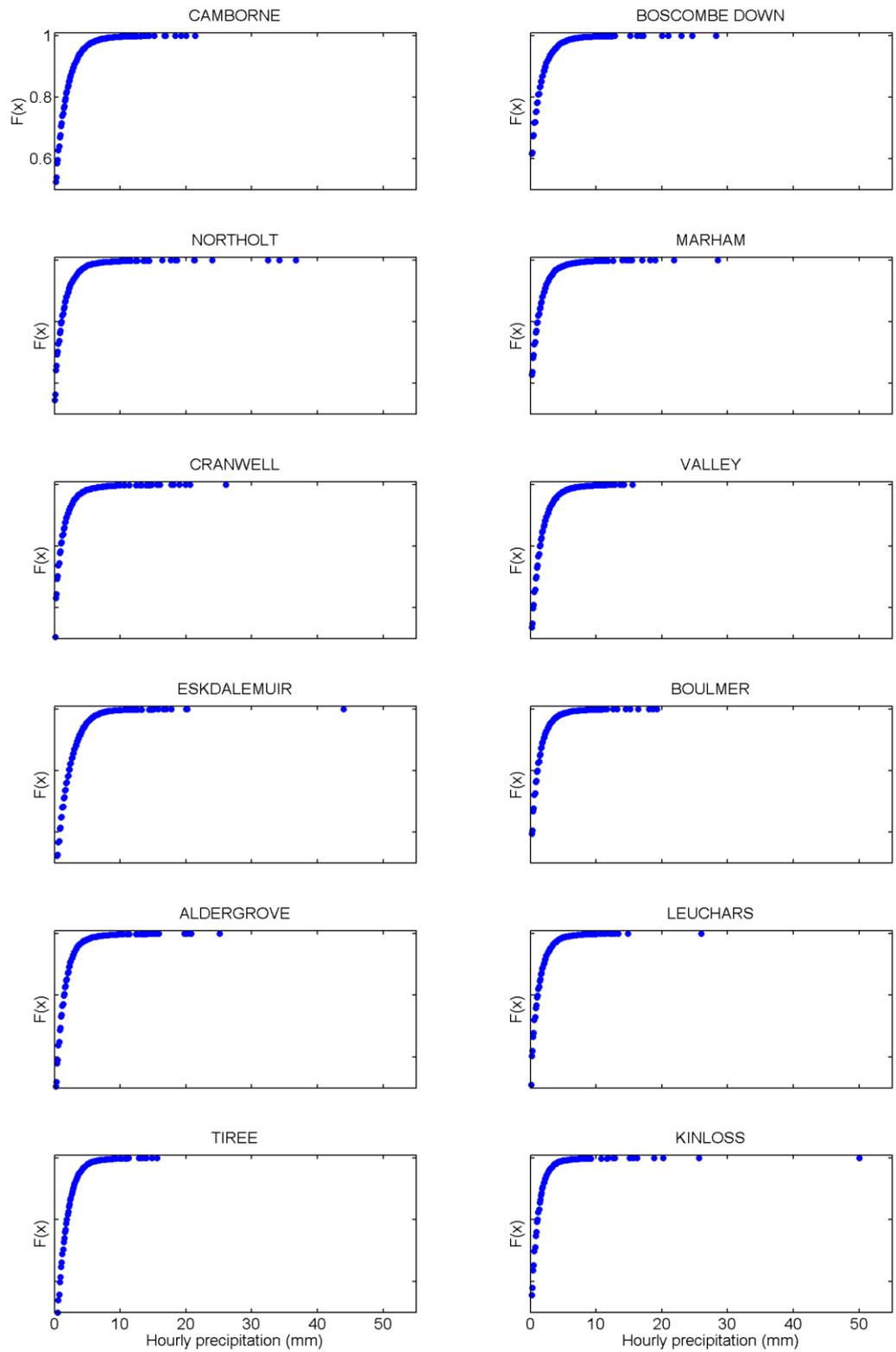


Figure 3.10 The daily maximum hourly precipitation distributions between 1980 and 2009 for the twelve selected stations.

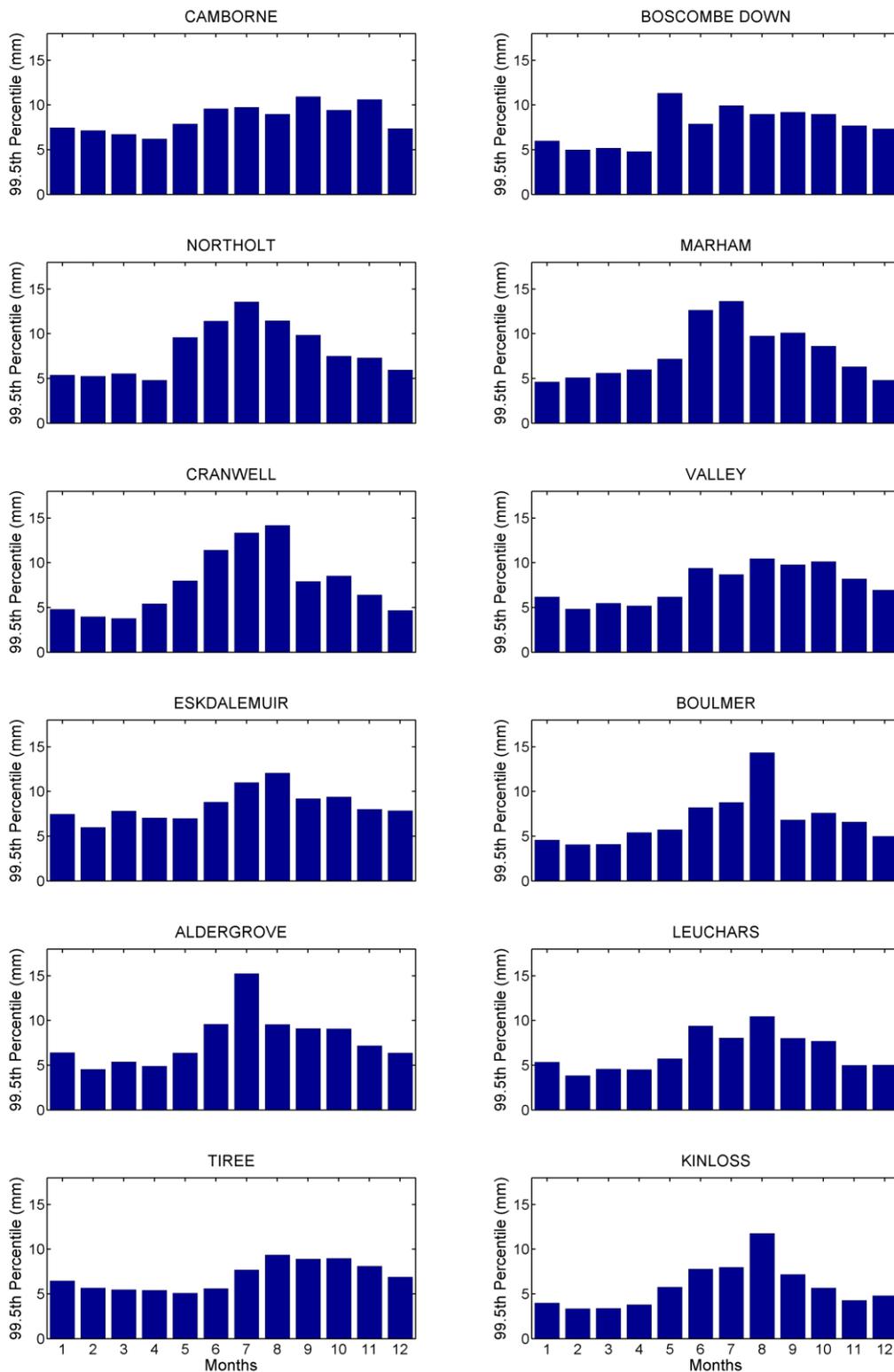


Figure 3.11 Annual cycle of the 99.5th percentiles of the daily maximum hourly precipitation between 1980 and 2009 for the twelve selected stations.

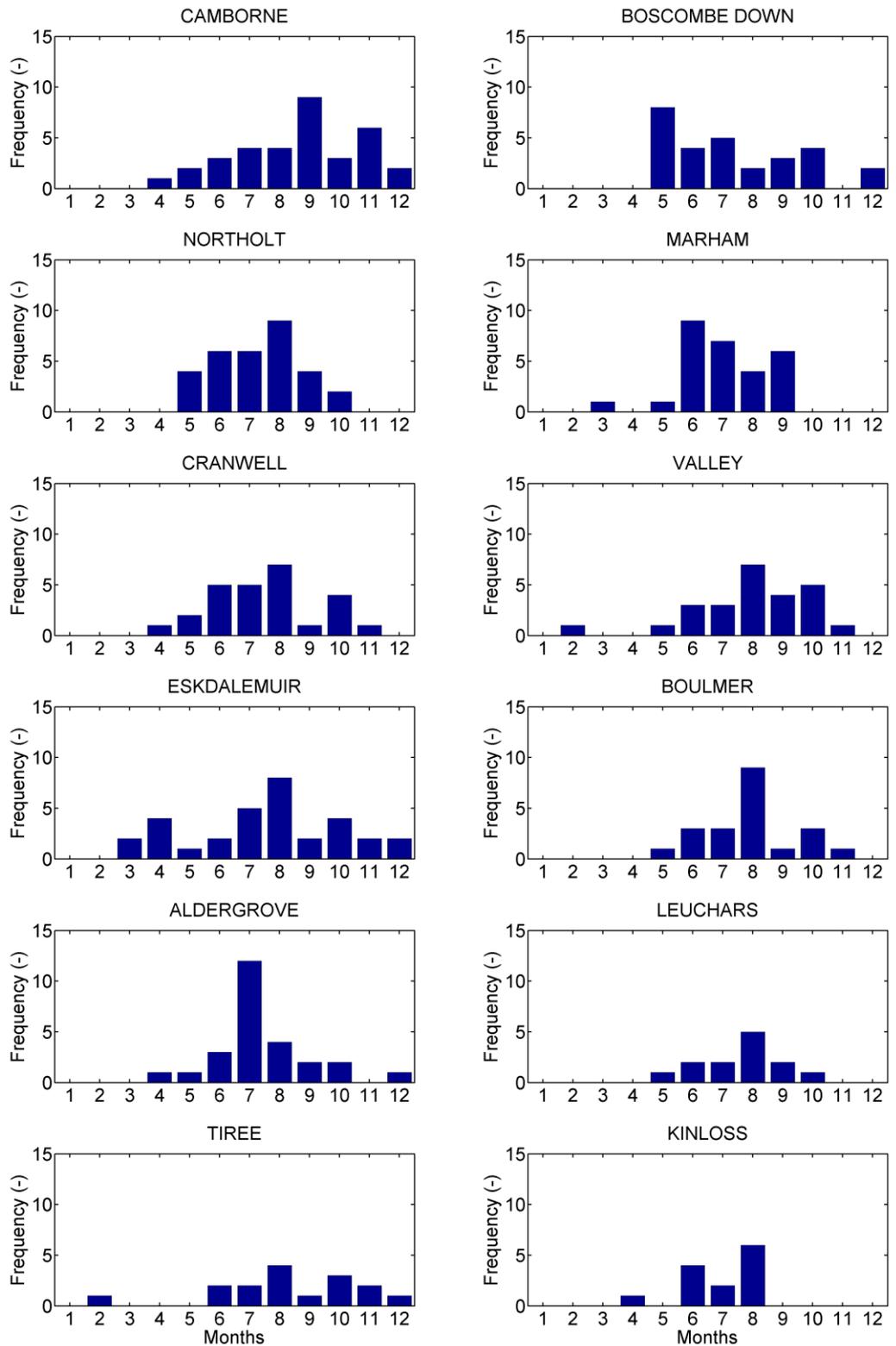


Figure 3.12 Annual cycle of the frequency occurrences of daily maximum hourly precipitation (>10mm/hr) between 1980 and 2009 for the twelve selected stations.

3.2 Historical gridded data sets

3.2.1 NCAR Sea Level Pressure (SLP) data set

The NCAR SLP data set is a 5° gridded analysis based on land station observations, covering 1899 to present for latitudes 30°N - 90°N and longitudes 0°E - 355°E. It provides daily (at 1200 UTC) and monthly data and has been constructed by the National Centre for Atmospheric Research (NCAR). It includes data from several sources and the data is “analysed” which means other variables, in particular surface winds, are taken into account and the grid-scale noise is minimized. The fact that different sources have been used over time along with changes in the corresponding analysis techniques can lead to inconsistencies in the data set. The raw data set also contain numerous erroneous data points. Therefore, the SLP data set has been quality controlled and corrected accounting for changes in instrumentation and station location as explained in Trenberth and Paolino Jr (1980). The quality controlled and corrected data set is available under Hurrell and Trenberth (2013).

The main advantage of the data set is the long time series dating back to 1899, which enables the fuzzy rules based CP-classification used in this thesis to be calibrated from 1960 to 1979. Hence the calibration of the CP-classification is independent from the time periods of the calibration and validation of the statistical downscaling method (1980-2009). In contrast, the ERA-interim reanalysis data set only provides SLP data from 1979-2013. A slight trade-off was made in selecting the NCAR over the ERA-interim reanalysis data in terms of the spatial resolution. The former provide SLP data on a 5° grid, the latter on a 0.75° grid. Given the SLP data set is used to classify the large-scale circulation patterns over the UK, the longer time period available for the calibration of the CP-classification is more important than having a finer spatial resolution. Previous studies used similar spatial resolution to the NCAR resolution for the purpose of CP-classification (Jones et al., 1993; Philipp et al., 2010). It also needs to be noted that SLP, in contrast to precipitation, can be considered as a continuous climate variable with relatively low spatial variability and hence an increase in the spatial resolution only is not likely to result in considerable different results. For example, Bárdossy et al. (2002) found that most of the available grid points providing SLP data have no direct influence on defining the CP-classification. Therefore, it is likely that an increase in the number of grid points, by increasing the spatial resolution, would not benefit the performance of the CP-classification, as the number of characteristic grid points to describe the CPs is not expected to increase simultaneously due to the relatively low spatial variability of the SLP variable. In terms of extreme hourly precipitation, mesoscale features are often considered as an important factor as they have the potential to trigger convective precipitation events (Hand et al., 2004). Therefore, it could be argued that an increase in the spatial resolution of the CP-classification

could benefit the prediction of the occurrence of extreme hourly precipitation. However, due to the small number of CPs used in this thesis it is very likely that the representation of mesoscale features would be smoothed out and would not lead to a better prediction of extreme hourly precipitation independently of the spatial resolution of the SLP data set.

3.2.2 E-OBS temperature data set

The E-OBS data set (Van den Besselaar et al., 2011) provides gridded observed climate variables on a 0.25° resolution between 1950 and 2013. In this study, daily mean temperature is extracted from the E-OBS data set for the statistical downscaling process applied in Chapter 6 and 7. The E-OBS daily mean temperature variable is defined as the mean of the daily maximum and minimum temperature. For each of the twelve stations, the grid point with the smallest distance to the respective station is selected and the temperature at this grid point is used. The gridded E-OBS temperature time series exhibit considerably higher level of completeness compared to the temperature data on a station level available within the MIDAS Land Surface Stations database (NCAS British Atmospheric Data Centre, 2012). Another reason for using the E-OBS data set is the fact the daily mean temperature values are averaged between 0000UTC and 0000UTC (the following day), which is in accordance with the daily mean temperature extracted from the RCM data sets. In contrast, the station temperature data is given each day for the day time between 0900UTC (the previous day) and 0900UTC. Similarly to Mishra et al. (2012), the gridded temperature data set is therefore the preferred choice over station data. Daily mean temperature is used, rather than daily maximum temperature, as it better represents the temperature of the air mass and is less influenced by boundary-layer processes and radiation (Lenderink and van Meijgaard, 2008).

Figure 3.13 shows the monthly 99.5th percentiles of the daily mean temperature for each of the twelve selected stations. This is important as convective precipitation can be linked with high temperature (Maraun et al., 2010b) and often cause extreme hourly precipitation (Gregersen et al., 2013). It has been also shown that precipitation intensifies with higher temperature, in particular for convective events (Berg et al., 2013).

It can be seen that the annual cycles of the monthly 99.5th percentiles are highly similar amongst the twelve stations. The highest temperatures occur in July and August, while the lowest 99.5th percentiles are recorded in December, January and February for all of the twelve stations. In terms of the magnitude of the 99.5th percentiles, it is found that the temperature decreases as the latitude increases (see Table 3.11). The only exception is Camborne, which is the station located furthest south of the twelve stations. In this case, the 99.5th percentiles in

particularly June, July and August are lower compared to the other stations in the southern part of the UK (Boscombe Down, Northolt, Marham and Cranwell). This can be explained by the geographical location of the station. Camborne is in very close proximity to the North Atlantic, which leads to milder summer compared to the stations further inland. In terms of the annual mean, temperatures also tend to decrease with higher latitude. However, the annual mean temperature at Eskdalemuir is the lowest of all twelve stations. This might be due the fact that the stations at Eskdalemuir exhibit the highest elevation (236m) of the twelve stations as given in Table 3.6.

In terms of the inter-annual variability, an increase in the annual temperature is found at all stations between 1980 and 2009. The increases are significant at the 5% level based on the p-value test. The annual increases range from 0.035 °C to 0.057 °C depending on the particular station (see Figure 3.14).

Station (latitude)	Mean temperature	99.5 th percentiles
Camborne (50.22°)	10.95	19.98
Boscombe Down (51.16°)	10.34	21.78
Northolt (51.55°)	11.12	23.39
Marham (52.65°)	10.16	22.03
Cranwell (53.03°)	10.04	21.66
Valley (53.25°)	9.33	19.47
Aldersgrove (54.66°)	9.41	19.44
Boulmer (55.42°)	8.53	18.88
Eskdalemuir (55.31°)	8.05	19.15
Leuchars (56.37°)	8.61	18.77
Tiree (56.5°)	8.73	17.87
Kinloss (57.65°)	8.44	18.83

Table 3.11 Annual temperature statistics for the twelve selected stations between 1980 and 2009. Mean temperature and 99.5th percentiles are given in Celsius. The stations are listed according to their latitude, which is given in parenthesis in the first column.

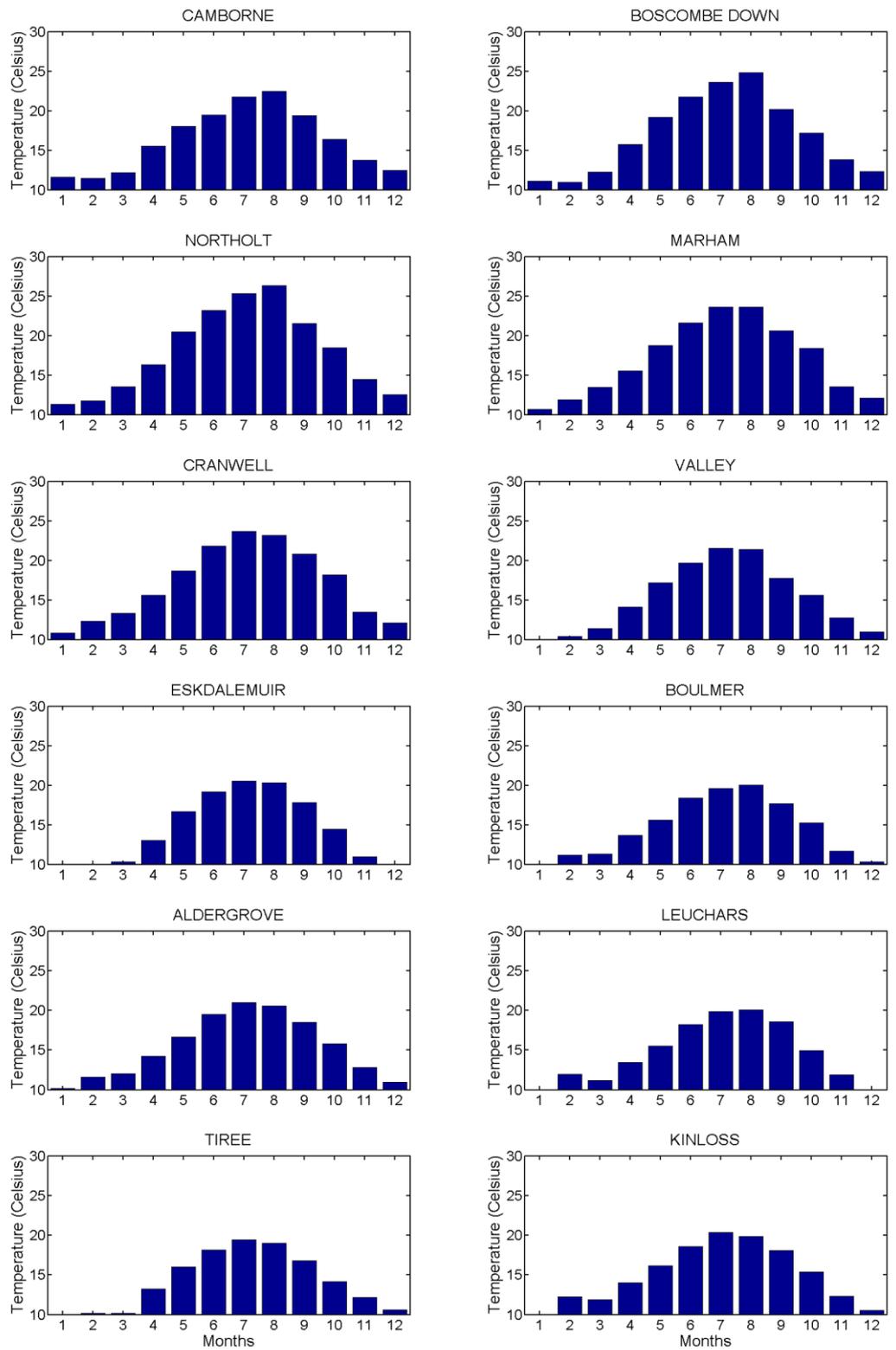


Figure 3.13 Annual cycle of the monthly 99th percentiles of the daily mean temperature for the twelve selected stations between 1980 and 2009.

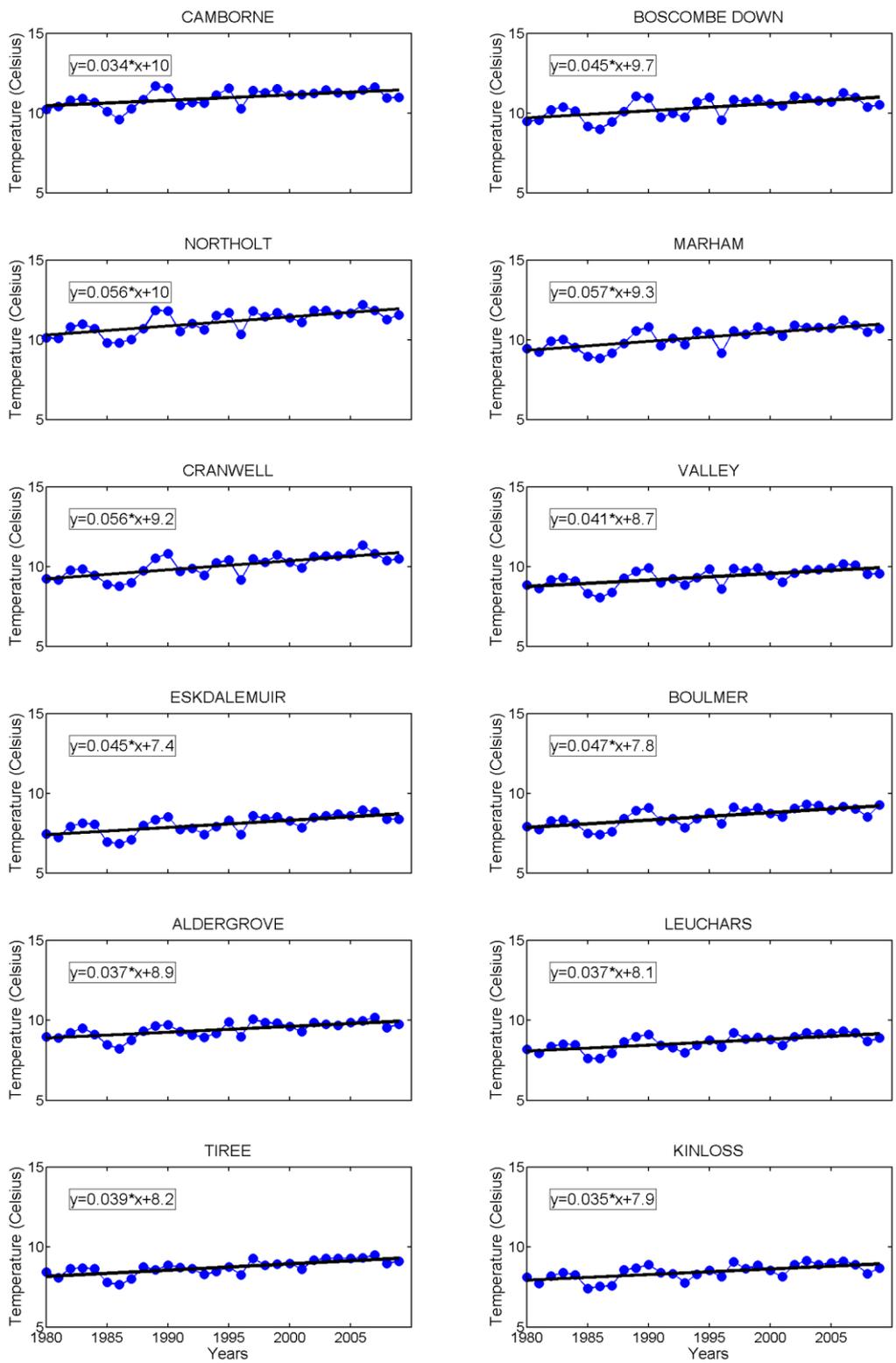


Figure 3.14 Observed trends in temperature between 1980 and 2009 for the twelve selected stations. All trends are significant at the 5% level based on the p-value test.

3.3 Regional Climate Model (RCM) data

Within international coordinated frameworks such as the PRUDENCE (2001-2004), ENSEMBLES (2004-2009) and CORDEX (2009-present) projects, the RCMs have been continuously made available to the wider scientific communities. The CORDEX project represents one of the latest development of RCMs. The EURO-CORDEX is the European branch of the international CORDEX initiative and provides data with spatial resolution of 0.11° (EURO-11, equivalent to ca. 12.5km) and 0.44° (EUR-44, equivalent to ca. 50km).

In Table 3.12, the different 12.5km RCMs are presented for which an ERA-interim driven run is available under the CORDEX data archive (CORDEX, 2014) over Europe. The ERA-interim runs are essential for the statistical downscaling method applied in this study. The time periods of the ERA-interim runs vary depending on the different RCMs. In this study, only the RCA4 RCM and the RACMO22E RCM are used. They both cover the time period from 1980 to 2009, which is the same as the time period used for the calibration and validation of the statistical downscaling methods. The other three available RCMs do not cover the period 1980-2009 when driven by the ERA-interim reanalysis data.

The RCA4 RCM was developed by the Rossby Centre at the Swedish Meteorological and Hydrological Institute (SMHI) in Sweden. It is the fourth version of the Rossby Centre's Regional Climate Models. The Rossby Centre was established in 1997 and has since then developed the [RCA1 (Rummukainen et al., 2001), RCA2 (Jones et al., 2004), RCA3 (Samuelsson et al., 2011)] through to RCA4. The latest RCA4 has undergone numerous physical changes. The focus has been mainly on the land and hydrological processes. For example, the Kain-Fritsch convection scheme has been updated to the Bechtold Kain-Fritsch scheme. It is a mass flux scheme and separates the shallow and deep convection processes. The lake model Flake replaced PROBE and is consistently flux coupled to the atmosphere. The soil hydrology layer in RCA4 is now grouped into forest and open land tiles. The inclusion of soil carbon in the RCA4 decreased the overestimated soil-heat transfer in the RCA3. In addition, the heat transfer in the top soil layer was further decreased for vegetation. These parameterisations regarding the soil characteristics lead to an improved representation of the diurnal temperature range (SMHI). For a more detailed description and a discussion of the performances of the RCMs developed at the Rossby Centre, the reader are referred to Samuelsson et al. (2011).

The RACMO22E was developed by the Royal Netherlands Meteorological Institute (KNMI) in the Netherlands. The version used in this study, referred to as RACMO22E, is an update of the RACMO2 cycle. RACMO22E is based on the ECMWF physics package embedded in the dynamical numerical weather prediction model HIRLAM and few routines to link the physical

and dynamical parts (Sponsor, 2012). The updated convection scheme consists of a reformulation of the triggering of shallow and deep convection over land (Van Meijgaard et al., 2008). It is based on the Tiedtke convection scheme and is parameterized by a mass-flux approach, which is separated in deep, mid-level and shallow convection (DOCUMENTATION–Cy31r1). The main difference between the Tiedtke scheme and the Bechtold Kain-Fritsch scheme, which is used in the RCA4, is the convection closure assumption. In the former, the closure assumption is based on moisture convergence and on Convective Available Potential Energy (CAPE) in the Bechtold Kain-Fritsch scheme (Dierer and Schubiger, 2008). Within the ENSEMBLES project, RACMO2 was found to be the best performing RCM among 14 others in representing climatic features of temperature and precipitation over Europe (Sponsor, 2012).

In addition to RCM runs driven by ERA-interim reanalysis data, CORDEX data sets also archive RCM runs driven by various GCMs. These GCM-driven RCM runs either cover the historical time period 1970-2005 or the future time period 2006-2100. For the future time period, different emission scenarios are used, namely RCP26, RCP45 and RCP85. The availability of GCM-driven RCM runs under the RCP26 emission scenarios is very limited and therefore the RCM runs based on the RCP26 emission scenario are not included in this study.

All daily SLP and near surface temperature values simulated by RCMs are temporally averaged over instantaneous values every three hours from 0000UTC to day+1 0000UTC and assigned to 1200UTC of that day. The extent of the spatial domain of the RCMs run is illustrated in Figure 3.15, where an example of a simulation of Mean Sea Level Pressure (MSLP) over Europe is given.

	ERA-interim	EC-EARTH	CM5A-MR	ESM-LR
RCA4	1980-2010	x	x	x
RACMO22E	1979-2012	x		
HIRHAM5	1989-2011			
CCLM4-8-17	1989-2008			
WRF331F	1989-2008			

Table 3.12 List of the 12.5km RCM runs available from the EURO-CORDEX data archive (CORDEX, 2014). The first column represents the available RCMs and the first row gives the driving ERA-interim reanalysis or GCM data. The red-colour filled cells represent the available runs and the red-colour filled cells with a white x indicate the RCMs runs used in this thesis.

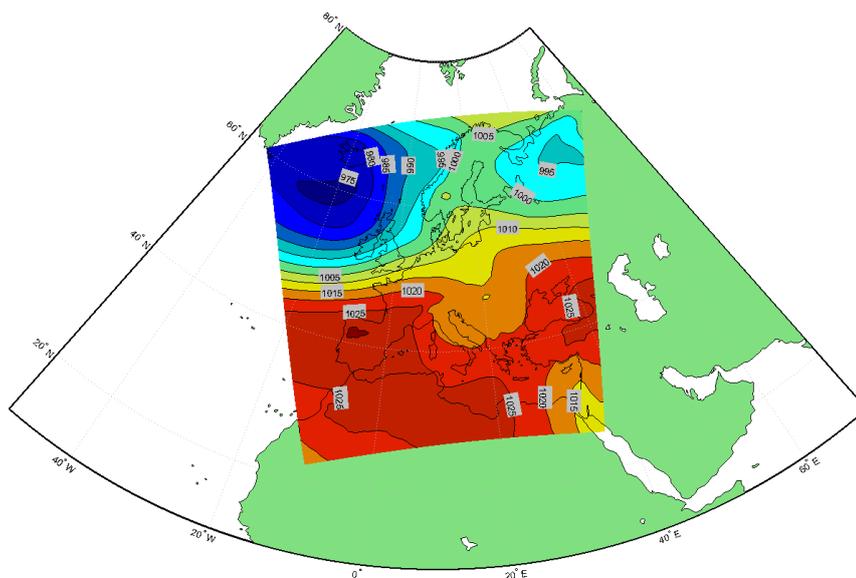


Figure 3.15 Illustration of the spatial domain (contoured area) of the EURO-CORDEX data set. It shows the daily SLP simulated by the RCA4 RCM driven by the ERA-interim reanalysis data.

3.4 Summary

In this chapter, the different data sets (observed site-specific daily/hourly precipitation data, observed gridded SLP/temperature data and multiple RCM runs) applied in this thesis are introduced and discussed. The focus is on the observed site-specific hourly and daily precipitation extracted from the MIDAS Land Surface Stations database (NCAS British Atmospheric Data Centre, 2012). A number of criteria are defined to eliminate extreme outliers, suspicious and multiple available precipitation records. Twelve stations are selected based on the homogeneity and the completeness of the observed hourly precipitation time series. The twelve stations represent twelve of the 14 different UK extreme precipitation regions as defined in Jones et al. (2014). For each of the twelve stations, the annual cycle of extreme hourly precipitation is assessed. Two groups of stations can be distinguished. The majority of the stations show higher extreme hourly precipitation extremes between May and October compared to the rest of the year and a maximum in June, July and August. It is very likely that those stations are strongly affected by convective extreme precipitation events, which tend to occur in summer when temperature are highest (Hand et al., 2004). Similarly, Maraun et al. (2009) assessed the annual cycle of extreme daily precipitation across the UK and also found that extremes occur most often in summer for those stations (regions) that are dominated by convective precipitation extremes. For the second group of stations, extreme hourly precipitation tends to be higher in the second half of the year (Jul-Dec) and shows no pronounced maximum in June, July and August. The three stations (Tiree, Valley and

Camborne) in this group are located in the west of the UK and are most exposed to the North Atlantic. In terms of daily precipitation extremes, (Jones et al., 2013; Maraun et al., 2009) identified a similar annual cycle for stations in the western part of the UK with daily extremes peaking in late autumn or winter. It was argued that those stations are more strongly affected by westerly frontal weather systems. Another explanation for higher hourly extremes in the second half of the year may be given in Holley et al. (2014) where they identified three main convective available potential energy (CAPE) seasons. The “sea dominated CAPE” season, which is highly relevant for those three stations in close proximity to the North Atlantic, was defined from September to January and therefore for months in the second half of the year (with the exception of January).

Chapter 4

Methodology

4.1 Introduction

In this thesis, the categorization of synoptic weather conditions into circulation patterns (CPs) is used to describe characteristic features of different atmospheric patterns and their corresponding relationships to extreme hourly precipitation events. However, the relationship between the atmospheric predictor, CP, and the local predictand, hourly precipitation, may not remain the same in a future climate. As a consequence, using CPs as the only atmospheric predictor may not be sufficient to reflect changes in future extreme hourly precipitation. Thus in this thesis, CPs are first used to subsample extreme hourly precipitation events according to their corresponding atmospheric conditions. In the next step, the effects of a changing climate on the relationship between CPs and extreme hourly precipitation events are analysed for each CP individually. The potential changes are represented by daily temperature and precipitation simulated by Regional Climate Models (RCMs). Daily temperature and precipitation are assumed to be more reliably simulated by RCMs compared to hourly precipitation events (Maraun et al., 2010b). In this thesis, changes in extreme hourly precipitation are therefore derived from the combination of changes in daily temperature and precipitation conditioned on CPs.

The underlying assumptions with respect to using CPs are as follows:

- I. Consistent and distinct relationships between CPs and hourly precipitation can be established to identify CPs at higher risk of extreme hourly precipitation.
- II. The relationships between the predictors, daily precipitation and temperature, and the predictand, extreme hourly precipitation, are different depending on the CPs

Section 4.2 explains the methodology of deriving the CP-classification based on fuzzy rules (Bardossy et al., 1995). This approach has been used in several previous studies over Europe (Bárdossy, 2010; Bárdossy and Pegram, 2011; Bárdossy et al., 2002; Haberlandt et al., 2014; Yang et al., 2010). In Bárdossy et al. (1995), the fuzzy rules were identified subjectively based on the European Grosswetterlagen/types (Hess and Brezowsky, 1969). In later studies (Bárdossy, 2010; Bárdossy et al., 2002), objective functions, which assessed the quality of the CP-classification in distinguishing between dry and wet CPs, were used instead. In order to

evaluate the quality of the CP-classification based on fuzzy rules, information measures were needed. Generally, these objective functions and information measures focused on mean precipitation characteristics. In this thesis, one objective function and one information measure is modified to specifically reflect on the CP-classification's ability to identify days of extreme precipitation events. None of the previous studies have applied the fuzzy CP-classification over the UK. Therefore, the spatial domain of the fuzzy rules CP-classification in this thesis is redefined, in order to be able to capture the characteristic atmospheric circulation patterns over the UK.

In Section 4.3, the focus shifts to the statistical downscaling method. The newly developed method addresses limitations of previous studies in terms of the statistical downscaling process of extreme hourly precipitation as discussed in Section 1.3. It is similar to the analogue day methods presented in Willems and Vrac (2011) where they tested several analogue day methods to project sub-daily precipitation extremes for one site in Belgium. Analogue day methods are able to reproduce local precipitation characteristics and simulate infrequent precipitation extremes on every temporal resolution for which observed data is available. The statistical downscaling process of the analogue method is relatively straightforward to understand and relies on physically coherent relationships between the predictor variables and the predictand. One of the methods in Willems and Vrac (2011) applied a perturbation factor to be able to project precipitation events which have not been observed in the past. In this case, they used daily precipitation, conditioned on different CPs and months, as the only predictor to find the most similar day in the past and applied a perturbation factor based on the Clausius-Clapeyron relationship.

In contrast in this thesis, daily precipitation and temperature, conditioned on different CPs and seasons, are used as combined predictors to find the most similar day (analogue day) over the calibration period. Incorporating temperature as a second predictor is desirable, since sub-daily precipitation extremes are likely to be more strongly intensified with temperature compared to daily precipitation events (Westra et al., 2014). Four different approaches of finding the analogue day are compared. In a next step, the observed maximum hourly precipitation value is extracted from the respective analogue day and perturbed based on the precipitation duration and temperature relationship conditioned on different CPs and seasons. The perturbation of the observed hourly precipitation values is required in order to simulate hourly precipitation values which have not been observed in the past. This is particularly important in terms of taking into account changes in the future climate.

Figure 4.1 summarises the applied method in this thesis to downscale extreme hourly precipitation. On the left side of the flowchart, the calibration of the statistical downscaling method and the derivation of the fuzzy rules CP-classification based on observed data sets are shown. The right side of the flow chart focuses on the estimation of hourly precipitation events using RCM simulated predictors.

Within the taxonomy of statistical downscaling techniques, the applied method is best described as a hybrid of the analogue and the regression-based method.

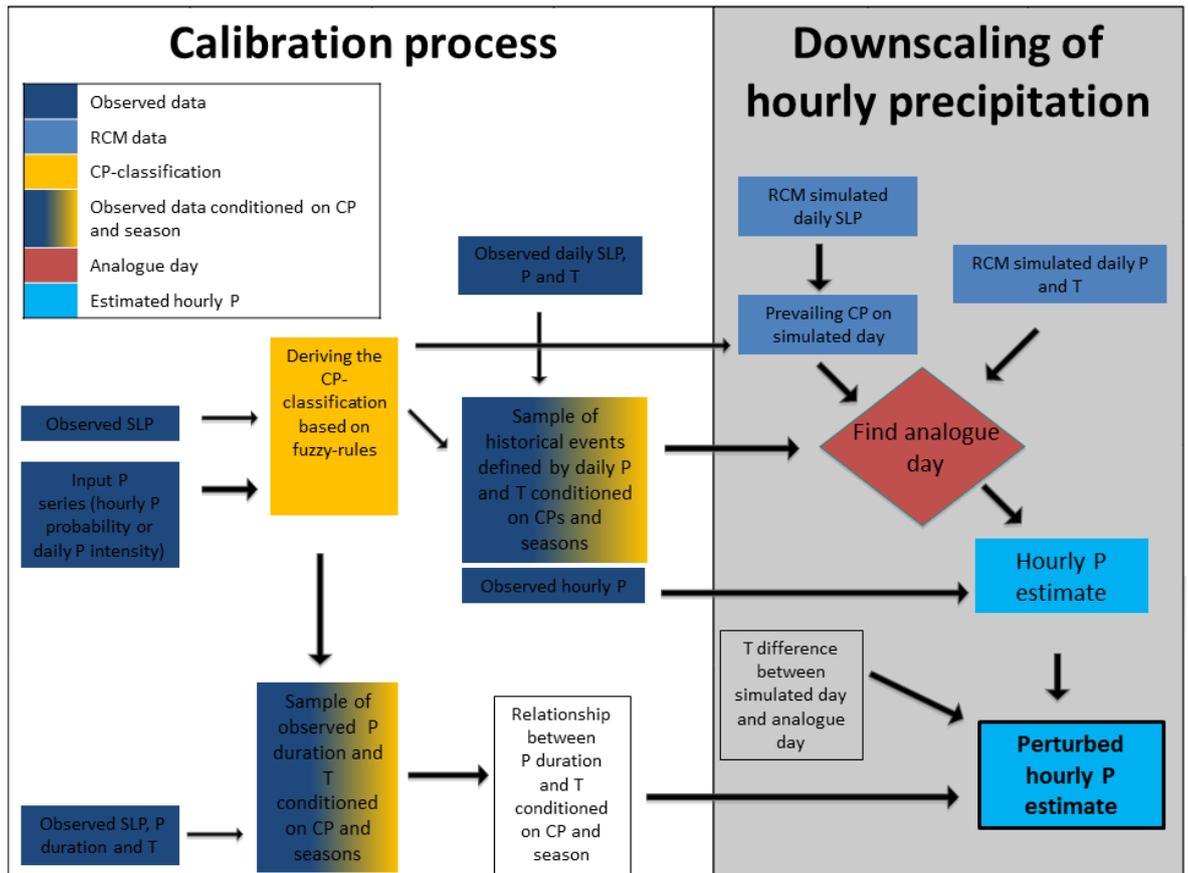


Figure 4.1 Flowchart of the applied method to downscale extreme hourly precipitation. SLP stands for sea level pressure, P for precipitation and T for temperature.

4.2 CP-classification based on fuzzy rules

The fuzzy rules CP-classification applied in this thesis uses the concept of fuzzy sets, in order to represent imprecise statements such as ‘high pressure’ mathematically. For example, the fuzzy set $v = 1$, which represents ‘high pressure’, is described by its membership function μ_1 . The value range of μ_1 is between zero and one. The closer μ_1 is to one, the stronger is the belief that a certain SLP value should be attributed to $v = 1$. The CP-classification based on fuzzy rules was first developed in Bardossy et al. (1995), where the concept of fuzzy sets is

explained in more detail. Initially, the fuzzy rules were defined subjectively (Bardossy et al., 1995), whereas in more recent studies (Bárdossy, 2010; Bárdossy et al., 2002) and in this thesis objective functions are used to derive the fuzzy rules automatically. Each CP is described by a fuzzy rule k and each fuzzy rule is represented by a vector of fuzzy sets $v(k)$ as explained in Equation 4.1.

$$v(k) = (v(1)^k \dots v(n)^k) \text{ as in Bárdossy et al. (2002)} \quad (4.1)$$

n : Number of grid points for which SLP data is available within the selected spatial domain.

Five different fuzzy sets are used:

$v = 1$ (very high pressure)

$v = 2$ (medium high pressure)

$v = 3$ (medium low pressure)

$v = 4$ (very low pressure)

$v = 5$ (is introduced for those grid points that have no influence on the CP)

With the exception of $v = 5$, each fuzzy set uses their own membership functions. The membership function $\mu(i, k)$ shows how strongly the normalized daily SLP anomaly value $g(i, t)$ belongs to a certain fuzzy set $v(i)^k$ for a given day t and grid point i . The normalized daily SLP anomaly values $g(i, t)$ are defined in Equation 4.2. In contrast to Bárdossy et al. (2002), $g(i, t)$ is modified to $g^*(i, t)$ as described in Equation 4.3. As a result, the range of anomaly values $g(i, t)$ is narrowed so that the membership functions $\mu(i, k)$ only need to be defined for values of $g^*(i, t)$ between zero and one (see Equation 4.4). Similarly to Bárdossy et al. (2002), the membership functions are defined as so-called triangular fuzzy numbers (see Figure 4.2).

$$g(i, t) = \frac{p(i, t) - \bar{p}(i, d)}{\sqrt{\frac{1}{N-1} \sum_{j=1}^N (p(i, d, j) - \bar{p}(i, d))^2}} \quad (4.2)$$

$p(i, t)$: SLP pressure for a given day t and grid point i

$\bar{p}(i, d)$: Mean SLP pressure of the d^{th} calendar day of the year at grid point i averaged over all the years N

N : Number of years for which $p(i, d)$ is available

$$\begin{aligned} g^*(i, t) &= 0.3 * g(i, t) + 0.5 & , \text{ if } 0 < g(i, t) < 1 & \quad (4.3) \\ g^*(i, t) &= 0 & , \text{ if } g(i, t) \leq 0 \\ g^*(i, t) &= 1 & , \text{ if } g(i, t) \geq 1 \end{aligned}$$

$$\begin{aligned}
\text{For } v(i)^k = 1 \quad & \mu_1(i, k) = 1 && , \text{ if } g^*(i, t) = 1 && (4.4) \\
& \mu_1(i, k) = 0 && , \text{ if } g^*(i, t) \leq 0.5 \\
& \mu_1(i, k) = 2g^*(i, t) - 1 && , \text{ if } 0.5 < g^*(i, t) < 1.0 \\
\\
\text{For } v(i)^k = 2 \quad & \mu_2(i, k) = \frac{1.3 - g^*(i, t)}{0.5} && , \text{ if } g^*(i, t) \geq 0.8 \\
& \mu_2(i, k) = 0 && , \text{ if } g^*(i, t) \leq 0.4 \\
& \mu_2(i, k) = \frac{g^*(i, t) - 0.4}{0.4} && , \text{ if } 0.4 < g^*(i, t) < 0.8 \\
\\
\text{For } v(i)^k = 3 \quad & \mu_3(i, k) = \frac{g^*(i, t) + 0.3}{0.5} && , \text{ if } g^*(i, t) \leq 0.2 \\
& \mu_3(i, k) = 0 && , \text{ if } g^*(i, t) \geq 0.6 \\
& \mu_3(i, k) = \frac{0.6 - g^*(i, t)}{0.4} && , \text{ if } 0.2 < g^*(i, t) < 0.6 \\
\\
\text{For } v(i)^k = 4 \quad & \mu_4(i, k) = 1 && , \text{ if } g^*(i, t) = 0 \\
& \mu_4(i, k) = 0 && , \text{ if } g^*(i, t) \geq 0.5 \\
& \mu_4(i, k) = 1 - 2g^*(i, t) && , \text{ if } 0 < g^*(i, t) < 0.5
\end{aligned}$$

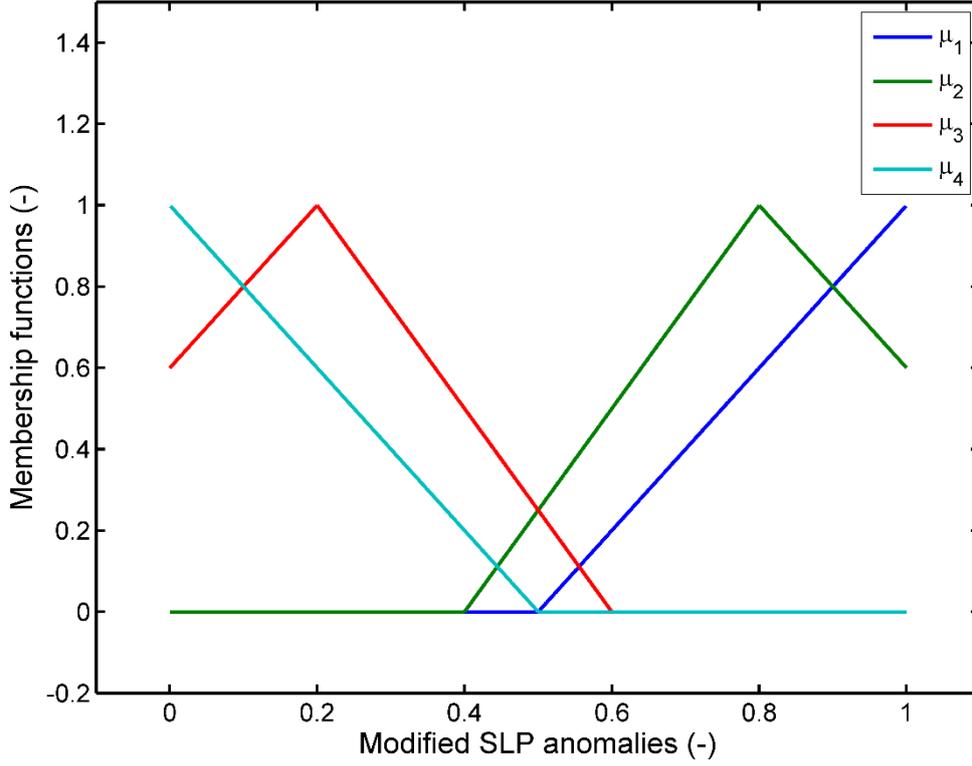


Figure 4.2 The membership functions of the fuzzy sets $v = 1, \dots, 4$

The closer the membership function $\mu_{v(i)^k}(i, k)$ is to one, the stronger is the belief that the SLP anomaly value $g(i, t)$ belongs to the fuzzy set $v(i)^k$. In order to assign each day t to a certain rule k , the degree of fulfilment $DOF(k, t)$ is calculated over all the grid points i :

$$\text{For } v(i)^k = 1 \quad DOF_1(k, t) = 1 - \left(\frac{\sum_i^{D_1} 1 - \mu(i, k)}{D_1} \right)^{1/1.25} \quad (4.5)$$

$$\text{For } v(i)^k = 2 \quad DOF_2(k, t) = 1 - \left(\frac{\sum_i^{D_2} 1 - \mu(i, k)}{D_2} \right)^{1/5}$$

$$\text{For } v(i)^k = 3 \quad DOF_3(k, t) = 1 - \left(\frac{\sum_i^{D_3} 1 - \mu(i, k)}{D_3} \right)^{1/5}$$

$$\text{For } v(i)^k = 4 \quad DOF_4(k, t) = 1 - \left(\frac{\sum_i^{D_4} 1 - \mu(i, k)}{D_4} \right)^{1/1.25}$$

$$DOF(k, t) = DOF_1(k, t) * DOF_2(k, t) * DOF_3(k, t) * DOF_4(k, t)$$

D_v : Number of grid points i classified to class v in rule k

The k for which $DOF(k, t)$ is maximal is chosen as the fuzzy rule (CP) on day t . Before each day can be assigned to a certain fuzzy rule (CP), the fuzzy rules need to be defined first. This is done by an optimization process, which is explained in the following. It is important to keep in

mind that here the definition of the fuzzy rules is done based on observed SLP data (Hurrell and Trenberth, 2013). The subsequent assignment of each day to a certain CP can be done for other SLP data sets as well, including data sets derived from climate models to project extreme hourly precipitation as described in Chapter 7.

Three different objective functions are incorporated into the optimization process to define the fuzzy rules. Each fuzzy rule represents one CP. The aim of the objective functions is to distinguish between wet and dry CPs and thus to define a set of CPs, which explains the variability of precipitation patterns as much as possible. Objective function OF_1 assesses the ability of the CP-classification to identify CPs with a high probability of precipitation and CPs with a low probability of precipitation (see Equation 4.6). The same objective function was also applied in Bárdossy (2010).

$$OF_1 = \sum_{s=1}^S \sqrt{\frac{1}{T} \sum_{t=1}^T (p(CP(t))_s - \bar{p}_s)^2} \quad (4.6)$$

S : Number of precipitation series

T : Number of days

$p(CP(t))_s$: Wet day (≥ 0.2 mm/hr) probability on a day with given CP for precipitation series s . The threshold of 0.2mm is chosen based on the accuracy of the UK observed precipitation records as described in Section 3.1.1.

\bar{p}_s : Wet day probability over all days within time period T for precipitation series s

Objective function OF_2 determines the CP-classification performance in terms of distinguishable wet days associated with different CPs (see Equation 4.7). A similar function was applied as information measure in Bárdossy (2010). In this thesis, an exponent of 1.25 is used to increase the importance of CPs associated with very wet days. Depending on the five different input precipitation Sets (see Table 5.2), a wet day is defined differently. Precipitation Sets 1-4 focus on the number of wet hours per day as the input in order to distinguish between days with no precipitation, days with short spells of precipitation and days with persistent precipitation. In contrast, Set 5 is based on regional daily precipitation intensities across the UK.

$$OF_2 = \sum_{s=1}^S \frac{1}{T} \sum_{t=1}^T \left| \frac{z(CP(t))_s}{\bar{z}_s} - 1 \right|^{1.25} \quad (4.7)$$

$z(CP(t))_s$: Mean maximum 1hr precipitation intensity on wet days with given CP for precipitation series s

\bar{z}_s : Mean maximum 1hr precipitation intensity over all wet days within time period T for precipitation series s

For objective function OF_3 , the focus is on distinguishing different extreme wet day probabilities between the CPs (see Equation 4.8). Again, depending on the input precipitation Set the meaning of the index varies. Using input precipitation Set 1 to 4, an extreme wet day is defined as a day with 24 hours of precipitation. Whereas for input precipitation Set 5, an extreme wet day correspond to daily precipitation equal or higher than 98% of the maximum daily precipitation. This is different to Bárdossy (2010), where only absolute thresholds were used to define the objective functions. The advantage of using relative thresholds is that it takes into account that the definition of an extreme event can vary in terms of precipitation intensity over different stations.

$$OF_3 = \sum_{s=1}^S \sqrt{\frac{1}{T} \sum_{t=1}^T (p_e(CP(t))_s - \bar{p}_{e_s})^2} \quad (4.8)$$

$p_e(CP(t))_s$: Extreme wet day probability on a day with given CP for precipitation series s
 \bar{p}_{e_s} : Extreme wet day probability over all days within time period T for precipitation series s

In this thesis, it was decided to calculate the objective functions for four different seasons separately: winter (Dec-Feb), spring (Mar-May), summer (Jun-Aug) and autumn (Sep-Nov). The final objective of the optimization process is to find fuzzy rules for which the sum of the objective functions is maximal. For this purpose, a simulated annealing algorithm is used (Bárdossy et al., 2002). The optimization process also guarantees that none of the CPs becomes too dominant in terms of their frequency occurrences.

4.2.1 The spatial domain of the fuzzy rules CP-classification

In terms of the spatial domain of the CP-classification, it is important to keep in mind that the classification is also applied to RCM data. Therefore, the chosen spatial domain need to be not only within the spatial extent of the observed NCAR SLP data set (Hurrell and Trenberth, 2013), which is used for the definition of the fuzzy rules, but also within the spatial boundaries of the RCM data sets (Jacob et al., 2014). Figure 4.3 shows the spatial domain of the observed NCAR SLP data sets, the RCM data sets, as well as the chosen spatial domain of the fuzzy rules CP-classification fitted to the two data sets. The chosen spatial domain of the CP-classification was stretched as far as possible to the western boundary of the RCM data sets, in order to represent the large-scale atmospheric features over the North Atlantic. As a result, the chosen spatial domain of the fuzzy rules CP-classification spans from 45°N to 65°N and -25°E to 10°E.

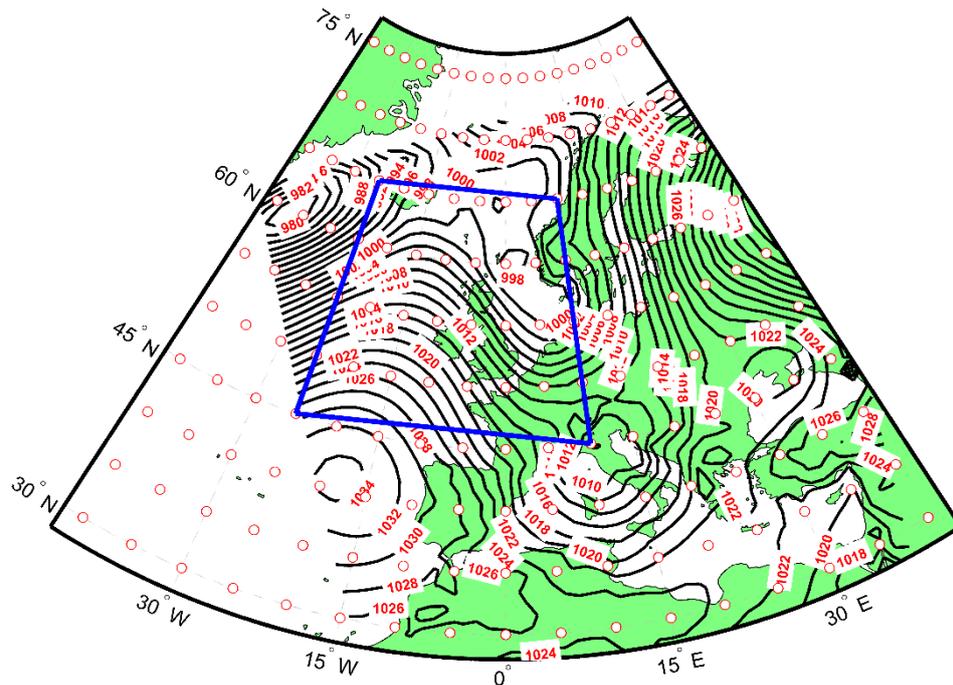


Figure 4.3 The spatial domains of the NCAR gridded SLP data set (Hurrell and Trenberth, 2013), the RCM data sets the fuzzy rules CP-classification. SLP pressure map represents the spatial domain of the RCMs (Jacob et al., 2014). The white points represent the NCAR gridded data set and the blue rectangle represents the spatial domain used for the fuzzy rules CP-classification.

4.2.2 The quality assessment of the fuzzy rules CP-classifications

In total 180 different fuzzy rules have been defined, resulting in 180 CP-classifications based on five different input precipitation series, four different calibration periods and nine different number of CPs as described in Section 5.1. Therefore an objective evaluation of the performance of the CP-classifications is needed, in order to find the “best fitting” CP-classification for the statistical downscaling method. Four different validation measures are applied for this purpose over the validation period 1980-2009, which corresponds to the time period used for the application of the statistical downscaling method (see Chapter 6). Three of them assess different aspects of the relationships between the CPs and precipitation and are referred to as information measure IM_1 , IM_2 and IM_3 in the following. These three information measures are calculated separately for winter and summer and are then averaged. The fourth validation measure focuses on the separability between the atmospheric SLP patterns derived from the observed NCAR gridded SLP data set. A higher value of the validation measure indicates a better performance of the CP-classification.

Information measure IM_1 focuses on the CP-classification's ability to distinguish between CPs with high and low probability of precipitation (see Equation 4.9). It was also used as information measure in Bárdossy (2010).

$$IM_1 = \sum_{s=1}^S \sqrt{\frac{1}{T} \sum_{t=1}^T (p(CP(t))_s - \bar{p}_s)^2} \quad (4.9)$$

S : Number of stations

T : Number of days

$p(CP(t))_s$: Wet day (≥ 0.2 mm/hr) probability on a day with given CP at station s

\bar{p}_s : Wet day probability over all days within time period T at station s

Information measure IM_2 assesses the quality of the CP-classification to identify CPs with different precipitation intensities (see Equation 4.10). A similar information measure was also used in Bárdossy (2010). In this thesis, the exponent of 1.25 is incorporated to emphasize those CPs with intense hourly precipitation.

$$IM_2 = \sum_{s=1}^S \frac{1}{T} \sum_{t=1}^T \left| \frac{z(CP(t))_s}{\bar{z}_s} - 1 \right|^{1.25} \quad (4.10)$$

$z(CP(t))_s$: Mean maximum 1hr precipitation intensity on wet days with given CP at station s

\bar{z}_s : Mean maximum 1hr precipitation intensity over all wet days within time period T at station s

Information measure IM_3 focuses on extreme hourly precipitation (see Equation 4.11). It measures how well the CP-classification distinguishes between CPs with a high probability of extreme events and those CPs with a low extreme event probability. It was also used in the same way as information measure in Bárdossy (2010) with varying thresholds. In this thesis, a threshold of 10mm/hr is selected to define extreme hourly precipitation. A single absolute threshold regardless of the spatial location and seasonality is chosen over a relative threshold, so that the definition of an extreme hourly precipitation event is independent of regional differences and the annual cycle.

$$IM_3 = \sum_{s=1}^S \sqrt{\frac{1}{T} \sum_{t=1}^T (p_e(CP(t))_s - \bar{p}_{e_s})^2} \quad (4.11)$$

$p_e(CP(t))_s$: Extreme precipitation probability (≥ 10 mm/hr) on a day with given CP at station s

\bar{p}_{e_s} : Extreme precipitation probability over all days within time period T at station s

Apart from assessing the quality of CP-classification in terms of its CP-precipitation relationship, it is also important that the different CPs are distinct in terms of their atmospheric circulation patterns. This can be measured by means of the pattern correlation difference (*PCD*) based on the correlation between different pressure maps as discussed in Overland and Hiester (1980). The process of calculating the *PCD* value is described in Equation 4.12 – 4.14. If the daily SLP maps attributed to the same CP (indicated by C_1) are as similar as possible and those attributed to different CPs (indicated by C_o) are as different as possible, the classification is capable of discriminating different circulation patterns. A higher *PCD* value indicates a better performance of the CP-classification.

$$PCD = C_1 - C_o \quad (4.12)$$

C_o : Mean SLP pattern correlation between two days from different CPs

C_1 : Mean SLP pattern correlation between two days attributed to the same CPs

$$\text{For } CP(i) \neq CP(j) \quad C_o = \frac{\sum_i^N \sum_{j \neq i}^N \rho_{ij}}{\sum_i^N \sum_{j \neq i}^N 1} \quad (4.13)$$

$$\text{For } CP(i) = CP(j) \quad C_1 = \frac{\sum_i^N \sum_{j \neq i}^N \rho_{ij}}{\sum_i^N \sum_{j \neq i}^N 1}$$

N : Number of days

$CP(i), CP(j)$: Prevailing CP on a given day i or j

By considering the SLP value at each grid point within the spatial domain, the SLP pattern correlation ρ_{ij} between two days i and j is calculated in the following way:

$$\rho_{ij} = \frac{\sum_m^M P_{im} P_{jm}}{\sqrt{\sum_m^M P_{im}^2 \sum_m^M P_{jm}^2}} \quad (4.14)$$

P_{im}, P_{jm} : SLP deviation from the SLP spatial mean on a given day i or j at grid point m

M : Total number of grid points within the spatial domain

4.3 The statistical downscaling process

The statistical downscaling method applied in this study is based on the analogue day method. Daily precipitation, temperature and the occurring CP on each day are used as predictors to find the most similar day (analogue day) among all the days within the calibration period. Four different approaches (M1, M2, M3 and M4) are compared with each other as shown in Table 4.1. In M1 and M3, each day over the calibration period is subsampled into two seasons,

winter (NDJFMA) and summer (MJJASO), and five (plus one unclassified) CPs. Instead of using two seasons as a subsample criterion, M2 and M4 categorizes every calibration day into one of the twelve calendar months.

Approach	Description
M1	Calibration days are subsampled into seasons and CPs
M2	Calibration days are subsampled into calendar months and CPs
M3	Calibration days are subsampled into seasons, CPs and daily precipitation categories
M4	Calibration days are subsampled into calendar months, CPs and daily precipitation categories

Table 4.1 Four different approaches of subsampling in order to find the most similar day in the past.

As illustrated in Section 3.1.5, extreme hourly precipitation tends to be more intense for certain months. Therefore, the statistical downscaling methods based on M2 or M4 uses the twelve calendar months as a subsample criterion. In M1 and M2, every calibration day t_C within a certain subsample is characterized by its normalized daily precipitation $nP(t_C)$ and temperature value $nT(t_C)$. For the actual day t , on which hourly precipitation is estimated, the relevant subsample is identified and its normalized daily precipitation $nP(t)$ and temperature $nT(t)$ is calculated as described in Equation 4.15 and Equation 4.16. In a next step, the distances $D_{PT}(t, t_C)$ between the actual day t and each calibration day t_C within the relevant subsample are calculated in terms of daily precipitation and temperature as shown in Equation 4.17. The calibration day with the smallest distance is chosen as the analogue day.

$$nP(t) = \frac{P(t) - \min(P_{S,CP})}{\max(P_{S,CP}) - \min(P_{S,CP})} \quad (4.15)$$

$nP(t)$: Normalized daily precipitation for a given day t (or calibration day t_C)

$P(t)$: Daily precipitation for a given day t (or calibration day t_C)

$P_{S,CP}$: Subsample of daily precipitation values for a given season S and CP over the calibration period. In M2 and M4 season S is replaced by month M

$$nT(t) = \frac{T(t) - \min(T_{S,CP})}{\max(T_{S,CP}) - \min(T_{S,CP})} \quad (4.16)$$

$nT(t)$: Normalized daily temperature for a given day t (or calibration day t_C)

$T(t)$: Daily temperature for a given day t (or calibration day t_C)

$T_{S,CP}$: Subsample of daily temperature values for a given season S and CP over the calibration period. In M2 and M4 season S is replaced by month M

$$D_{PT}(t, t_C) = (nP(t) - nP(t_C))^2 + (nT(t) - nT(t_C))^2 \quad (4.17)$$

$D_{PT}(t, t_C)$: Distance between the actual day t and calibration day t_C in terms of daily precipitation and temperature

Figure 4.4 provides an illustrative example of how the analogue day method works. It shows all the calibration days between 1980 and 1999 in winter which are attributed to CP3 at the station of Boulmer. Each day is characterized by its normalized daily precipitation $nP(t_C)$ and temperature $nT(t_C)$. The green triangle represents the actual day t for which hourly precipitation is estimated. The arrow points from the actual day t to the analogue day for which the distance $D(t, t_C)$ is minimal over all the calibration days within this subsample (Boulmer - winter - CP3). In cases where multiple days exist with the same minimal distance to the actual day, the analogue day is randomly selected among those days.

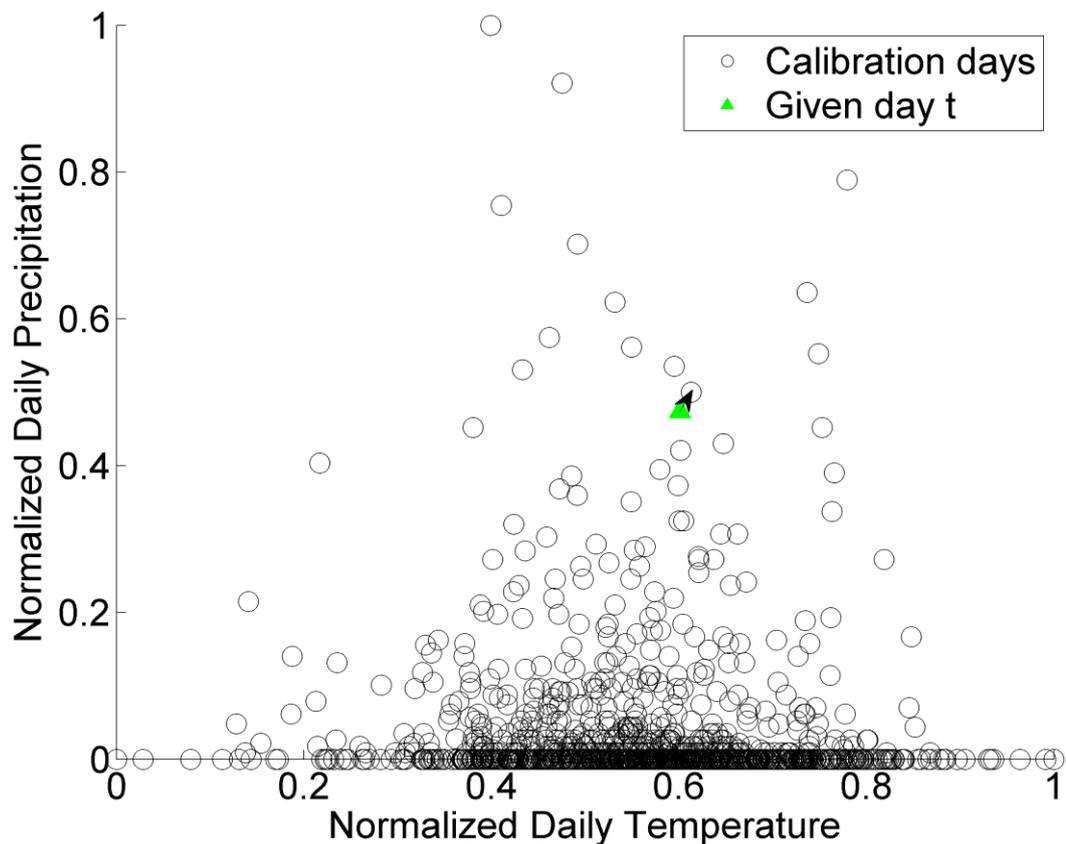


Figure 4.4 Subsample of calibration days t_C occurring on days classified as CP3 in winter between 1980 and 1999 at the station of Boulmer. The green triangle represents the actual day t (11th Jan 2000 - daily precipitation: 10.8mm (normalized: 0.47) - daily mean temperature: 5.53°C (normalized: 0.60)) for which hourly precipitation needs to be estimated. The arrow points to the analogue day for which the distance $D(t, t_C)$ is minimal over all the calibration days within this subsample. The distance is calculated in terms of daily precipitation and temperature. This applies to subsampling approach M1 (and M2 using monthly subsamples).

In the approaches M3 and M4, the daily precipitation value of each calibration day is used as an additional subsample criterion. Similarly to M1 and M2, each calibration day is subsampled into one of the two seasons (M3) or twelve months (M4) and 5 (plus one unclassified) CPs. In a next step, each calibration day is grouped into one of six daily precipitation categories within each subsample as illustrated in Figure 4.5.

- The first category of each subsample comprises all the dry calibration days defined as days with daily precipitation below $1.0mm$.
- The second category includes all the calibration wet days below the 20th daily precipitation percentile of wet days.
- The third, fourth and fifth category consist of calibration days with daily precipitation between the 20th and 40th, 40th and 60th, 60th and 80th daily precipitation percentiles of wet days respectively.
- The sixth category includes all the calibration days with daily precipitation exceeding or equal to the 80th daily precipitation percentile of wet days.

The different percentiles of daily precipitation are calculated for each subsample separately and only consider calibration days with daily precipitation exceeding or equal to $1.0mm$. Each calibration day within a certain daily precipitation category and subsample is characterized by its normalized temperature $nT(t_C)$ as described in Equation 4.16. In order to find the most similar days, the distances $D_T(t, t_C)$ between the actual day and each calibration day, within the relevant subsample and daily precipitation category, are considered only in terms of daily temperature (see Equation 4.18). The calibration day for which the distance is the smallest is chosen as the analogue day. It means that in the approaches M3 and M4, daily precipitation is used to categorize each calibration days and the daily temperature variable is used to find the analogue day in the respective daily precipitation category. The theory behind the application of daily precipitation categories is that RCMs cannot be expected to accurately simulate daily precipitation extremes (Fowler and Ekström, 2009; Maraun et al., 2010b). Instead, RCM simulated precipitation is only used as a qualitative measure to indicate the occurrence of an extreme daily precipitation event.

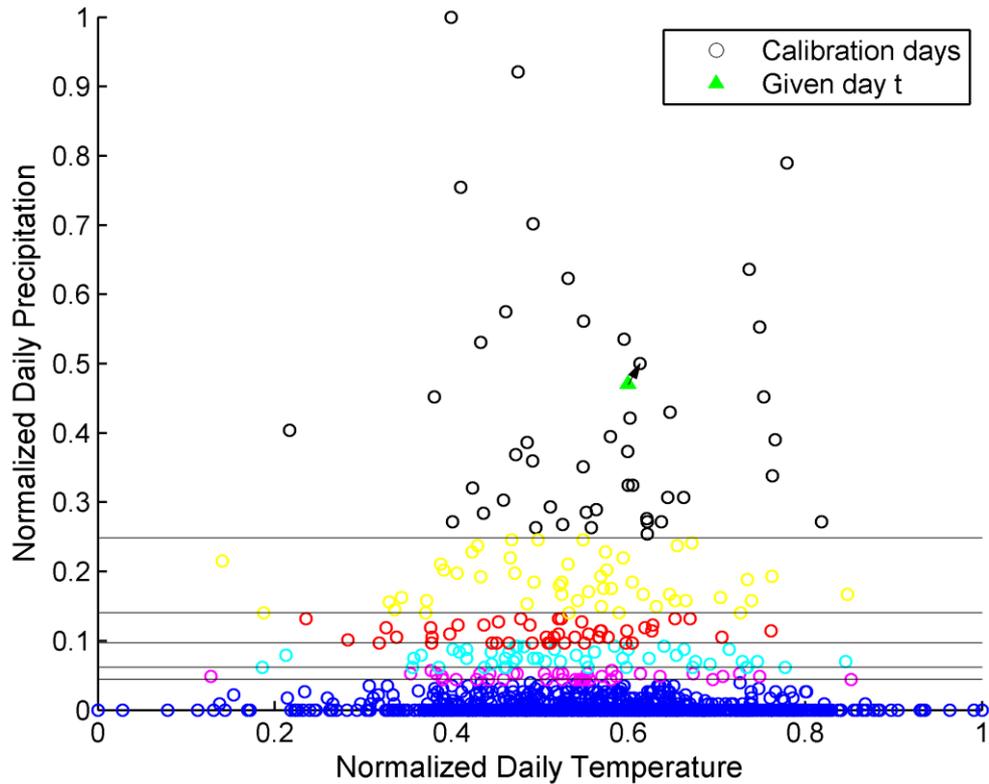


Figure 4.5 Same as Figure 4.4 but using daily precipitation as an additional subsample criterion. The different colours of the calibration days represent the six different daily precipitation categories. The black horizontal lines represent the corresponding thresholds of those categories. The distance is only calculated in terms of temperature. This applies to subsampling approach M3 (and M4 using monthly subsamples).

$$D_T(t, t_C) = (nT(t) - nT(t_C))^2 \quad (4.18)$$

$D_T(t, t_C)$: Distance between the actual day t and calibration day t_C in terms of temperature

In the next step, the observed maximum hourly precipitation value $MHP_{analogue}$ is extracted from the respective analogue day. It is then perturbed based on the precipitation duration and temperature relationships conditioned on the two seasons, winter (NDJFMA) and summer (MJJASO), and five (plus one unclassified) CPs. The perturbation of the observed hourly precipitation values is required in order to simulate hourly precipitation values which have not been observed in the past. This is particularly important in terms of taking into account changes under a warming climate. For each station, season and CP, a linear regression is calculated using an ordinary least squares fit to describe precipitation duration (in hours) as a function of daily temperature. The linear regression is then tested for two separate criteria, the p-value test and the RMSE test. The p-value significance test represents the probability of obtaining the observed sample results (or more extreme results) under the condition that the

null hypothesis (regression slope is equal to zero) is true. The RMSE (Root Mean Square Error) value represents the standard deviation of the residuals and is calculated as described in Equation 4.19 for the mean precipitation duration and mean temperature over five temperature categories:

- cold: 0-20th temperature percentile
- cool: 20th – 40th temperature percentile
- medium: 40th – 60th temperature percentiles
- warm: 60th – 80th temperature percentiles
- hot: 80th – 100th temperature percentiles

If the linear regression is significant at the 5% significance level based on the p-value test and its RMSE value is smaller than 1, the linear regression slope $\beta_{s,S,CP}$ conditioned on a given station, season and CP is used. A RMSE value below 1 means that the precipitation duration values estimated by the linear regression deviate on average not more than one hour from the observed duration records over the five temperature categories. If the two criteria are not fulfilled, the linear regression slope is only conditioned on the station and season without taking into account the CP-classification. In a small number of cases, where the CP-conditional and the CP-unconditional linear regression slope do not fulfil the p-value test or the RMSE test, the perturbation of the hourly precipitation cannot be applied. In most of the cases, the observed maximum hourly precipitation record extracted on the analogue day is perturbed as described in Equation 4.20 and Equation 4.21 for any given day in the past or future. The perturbed maximum hourly precipitation value $MHP(t)$ is then used as the estimated hourly precipitation record on the actual given day t .

$$RMSE = \sqrt{\frac{\sum_{tc=1}^5 (PD(\overline{T}_{tc}) - \overline{PD}_{tc,obs})^2}{5}} \quad (4.19)$$

RMSE: Root Mean Square Error

$PD(\overline{T}_{tc})$: The estimated mean precipitation duration by the linear regression for the temperature mean \overline{T}_{tc} within the temperature category tc

$\overline{PD}_{tc,obs}$: The observed mean precipitation duration within the temperature category tc

$$PD(t) = PD_{analogue} + (T(t) - T_{analogue}) * \beta_{s,S,CP} = 1, 2, 3, \dots, 24 \quad (4.20)$$

$PD(t)$: The estimated precipitation duration on the given day t

$PD_{analogue}$: The observed precipitation duration on the analogue day

$T(t)$: Temperature on the given day t

$T_{analogue}$: Temperature on the analogue day

$\beta_{s,S,CP}$: The statistically significant regression slope conditioned on station s , season S and CP or alternatively station s and season S

$$MHP(t) = MHP_{analogue} * \frac{PD_{analogue}}{PD(t)} \quad (4.21)$$

$MHP(t)$: The perturbed maximum hourly precipitation value on the given day t

$MHP_{analogue}$: The observed maximum hourly precipitation value on the analogue day

The reliability and robustness of the extreme hourly precipitation estimates is assessed over different validation periods (1980-1989, 1990-1999, 2000-2009 and the ten warmest summers of each station). In order to quantify the quality of the estimates, the 99.5th percentile error ($PE_{99.5}$) between the observed and estimated maximum hourly precipitation distribution is calculated as described in Equation 4.22. Similar to the quantile verification score (Maraun et al., 2010a), it assesses the ability of a model to estimate a certain percentile (99.5th) of a distribution. In this thesis, the $PE_{99.5}$ is calculated over 100 bootstrapping samples for each station and validation period. The use of bootstrapping samples helps to assess the quality of the estimated extreme hourly precipitation distribution beyond the 99.5th percentile only. The estimated daily maximum hourly precipitation values are resampled 100 times and for each bootstrapping sample i , the $PE_{99.5,i}$ is determined. In a next step, the average $\overline{PE_{99.5}}$ is calculated for each station and validation period as shown in Equation 4.23.

$$PE_{99.5} = \frac{|OMHP_{99.5} - EMHP_{99.5}|}{OMHP_{99.5}} * 100\% \quad (4.22)$$

$PE_{99.5}$: The 99.5th percentile error in %

$OMHP_{99.5}$: The 99.5th percentile of the observed maximum hourly precipitation distribution

$EMHP_{99.5}$: The 99.5th percentile of the estimated maximum hourly precipitation distribution

$$\overline{PE_{99.5}} = \frac{\sum_{i=1}^{100} PE_{99.5,i}}{100} \quad (4.23)$$

$PE_{99.5,i}$: The 99.5th percentile error for a given bootstrapping sample i

$\overline{PE_{99.5}}$: Averaged 99.5th percentile errors over 100 bootstrapping samples

Chapter 5

The fuzzy rules based CP-classification

Section 4.2 describes the methodology of defining the CP-classification based on fuzzy rules. This chapter discusses the results of the CP-classification. Section 5.1 focuses on the input parameters that are used to derive the sets of fuzzy rules. A total of 180 different sets of fuzzy rules are defined resulting in 180 different CP-classifications based on varying input parameters. In Section 5.2, the quality of each of the 180 different CP-classifications is assessed based on (1) the relationship between large-scale CPs and local precipitation and (2) the separability of the different atmospheric pressure patterns. In Section 5.3, one final CP-classification is selected and the relationships between the different CPs of this selected classification and extreme hourly precipitation is described.

5.1 Input calibration parameters

5.1.1 Set of precipitation time series

Five different sets of precipitation series are tested for the calibration of the CP-classification (see Table 5.1), in order to assess if differences in the performances of the CP-classification occur when the sets are varied. Precipitation Sets 1-3 comprise site-specific hourly precipitation series. Set 4 uses the regional hourly precipitation over two selected UK regions. Set 5 is based on regional daily precipitation across the UK. For the Sets 1–4, the number of wet hours per day is used as the input, distinguishing between days with short spells of precipitation and persistent precipitation over the entire day. Set 5 uses the normalized daily precipitation amounts as the input index.

Set 1	Number of wet hours ($\geq 0.2mm$) per day at one station (Aldergrove–Northern Ireland) resulting in one precipitation series
Set 2	Number of wet hours per day at two stations (Aldergrove–Northern Ireland; Elmdon-Central England) resulting in two precipitation series
Set 3	Number of wet hours per day at three stations (Aldergrove–Northern Ireland; Elmdon-Central England; Stornoway–Northern Scotland) resulting in three precipitation series
Set 4	Spatially averaged number of wet hours per day over two HadUKP regions (South West England, Northern Scotland) resulting in two precipitation series
Set 5	Spatially averaged normalised daily precipitation for the nine HadUKP regions across the UK resulting in nine precipitation series

Table 5.1 Overview of the input precipitation series used for the calibration of fuzzy rules CP-classifications.

Figure 5.1 shows the different precipitation stations and HadUKP regions used as input precipitation series in Set 1-5. Three hourly precipitation stations located in three HadUKP regions (Central England, Northern Ireland and Northern Scotland) are used for Set 1-3. The station selection criteria and the precipitation characteristics are described respectively in the Section 3.1.4 and 3.1.5. The HadUKP regions were specified in Alexander and Jones (2000) and defined initially in Wigley et al. (1984).

Instead of individual stations, two precipitation regions (South West England and Northern Scotland), which are more robust in terms of representing the surrounding climate, are used in Set 4. In this case, the regional precipitation series are calculated as the numbers of wet hours averaged over multiple stations. Figure 5.1 illustrates the stations used to calculate the averages over South West England (Boscombe Down, Aberporth, Rhoose, Plymouth Mountbatten) and Northern Scotland (Wick Airport, Stornoway Airport, Tiree).

In contrast to Sets 1-4, which are based on hourly records only, Set 5 uses nine regional daily normalised precipitation series across the UK corresponding to the nine HadUKP regions (see Figure 5.1)

Legend

- Aldergrove
 - Elmdon
 - Stornoway
 - ◆ Hourly stations in South West England and North Scotland
 - HadUKP_regions
- Elevation [m]
- High : 4783
 - Low : -214

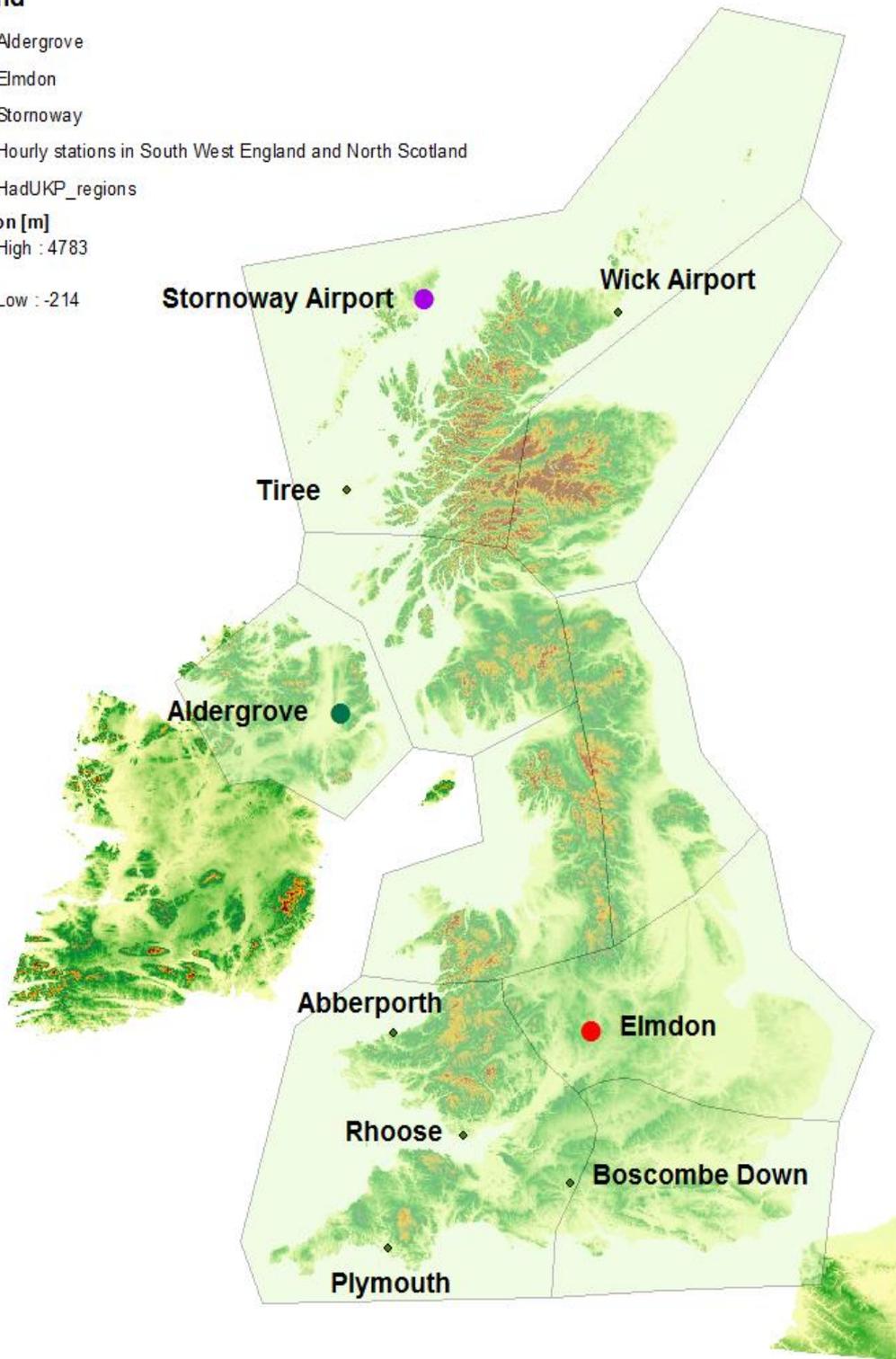


Figure 5.1 Hourly precipitation stations and HadUKP regions used for the calibration of the fuzzy rules based CP-classification.

5.1.2 Calibration time periods

As shown in Table 5.2 four different time periods are tested for the calibration of the fuzzy rules CP-classification, in order to assess differences in the performances of CP-classification based on varying calibration periods. The calibration and validation of the statistical downscaling method is between 1980 and 2009. It does not have any overlap with the time periods 1, 2 and 4, but overlaps with the time period 3 by 10 years.

Time period 1	1960 - 1969
Time period 2	1970 - 1979
Time period 3	1980 - 1989
Time period 4	1960 - 1979

Table 5.2 Overview of the different calibration time periods used for the fuzzy rules CP-classification.

5.1.3 Number of CPs

In this study, different numbers of CPs are tested in terms of their skill to distinguish between dry and wet CPs and in terms of their separability (see Section 5.2). However, as discussed in Section 2.4, the optimal number of CPs does not only depend on the skill of the CP-classification, but also on how robust the statistical downscaling method is for the estimation of precipitation events. For the latter, the number of CPs needs to be limited to a certain extent, in order to provide sufficient precipitation events for each subsample of CP. As a result, the maximum number of CPs is set to ten and the different CP-classifications are validated for a varying number of CPs between four and ten. In order to assess the potential improvement a higher number of CPs could have, the CP-classifications are also tested with a number of 15 and 20 CPs.

5.2 Assessment of the performances of the different CP-classifications

A total number of 180 different fuzzy rule systems are optimized according to the procedure described in Section 4.2 based on five different sets of precipitation input time series, four different calibration time periods and nine different number of CPs (see Section 5.1).

The skill of the CP-classification derived from the optimized fuzzy rules system are assessed over the time period 1980-2009, which is in line with the calibration and validation period used for the statistical downscaling process. The assessment is based on the same twelve precipitation stations (see Section 3.1.3), which are also used for the statistical downscaling process. Two different types of assessment measures are applied. The first type assesses the

relationship between CPs and precipitation and how well the CP-classification distinguishes between wet and dry CPs. In the following, these measures are referred to as information measures IM_1 , IM_2 and IM_3 . Information measure IM_1 assesses the skill of the CP-classification to distinguish between CPs with high and low probability of wet days (see Equation 4.9). IM_2 focuses on the CP-classification's ability to define CPs which result in high intensity of hourly precipitation (see Equation 4.10). IM_3 is based on extreme hourly precipitation (>10 mm/hr) only and assesses how well certain CPs can be linked with these extreme events (see Equation 4.11).

The second type of measures focuses on the atmospheric pressure pattern of each CP. It assesses how different the pressure patterns are between different CPs and is measured by the pattern correlation difference (PCD) as explained in Equation 4.12 to Equation 4.14. The larger the values of the four assessment measures, the better skills the CP-classification has.

5.2.1 Results

The results of the information measures and the PCD value are shown in Figure 5.2 for the time period 1980-2009. Please note that the information measure IM_1 , IM_2 and IM_3 are averaged over winter (Nov-Apr) and summer (May-Oct). IM_1 and IM_3 represent frequency occurrences of (extreme) wet days, whereas IM_2 focuses on precipitation intensities.

It can be seen that the performance improves sharply from 4 CPs to 5 CPs, and then generally increase up to 10 CPs with some fluctuation. From 10 to 15 and finally to 20 CPs, information measures IM_1 , IM_2 and the PCD value do not show much improvement, and decrease in some cases. Whereas, IM_3 shows steady increases up to 20.

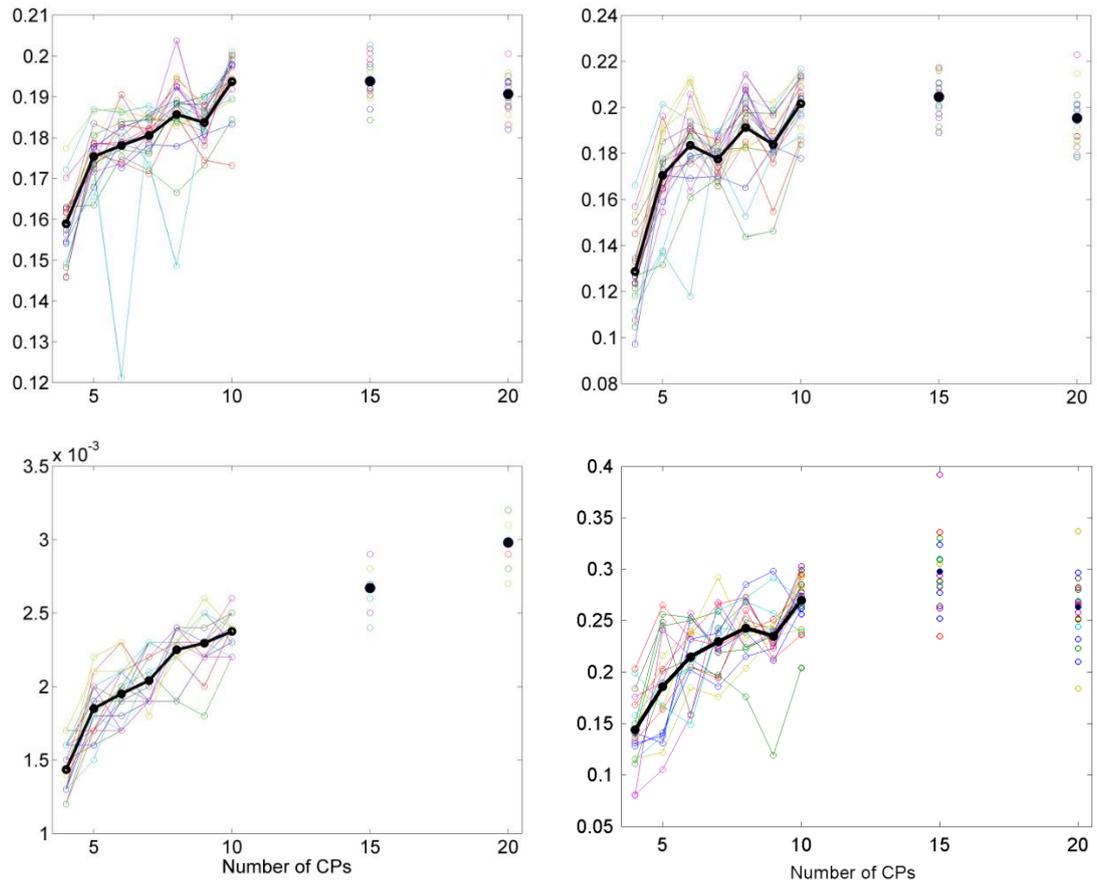


Figure 5.2 Results of the assessment of the CP-classification performances. It shows the performances of all the 180 tested CP-classification runs. Each colour represent a certain combination of the input precipitation Set and the calibration time period (the bold black line represents the average). Top left: Information measure 1 (wet day probability); Top right: Information measure 2 (precipitation intensity); Bottom left: Information measure 3 (extreme precipitation probability); Bottom right: PCD (separability).

Regarding different calibration time periods, calibration period 4 results in slightly better performances of the CP-classifications. This is shown in Table 5.3, where the performances of the CP-classification are averaged over different number of CPs and input precipitation Sets depending on the calibration time period. The calibration period 4 is the only one that contains 20 years, twice as long as the other three periods. The longer time period is likely to have led to a better performance in the classifications.

	Period 1	Period 2	Period 3	Period 4
Information measure 1	0.182	0.182	0.182	0.183
Information measure 2	0.181	0.182	0.182	0.184
Information measure 3	0.0022	0.0022	0.0022	0.0022
PCD (separability)	0.230	0.225	0.232	0.239

Table 5.3 Results of the information measures to assess the CP-classification performances. It shows the performance of the CP-classification averaged over different number of CPs and input precipitation Sets depending on the calibration time period

In Table 5.4, the averaged performances over different number of CPs and calibration periods are summarised depending on the various input precipitation Sets. It can be seen that the performance averages achieved for Set 5 are the highest. In other words, the CP-classifications based on Set 5 outperform those based on Sets 1-4 on average for the twelve selected stations. Set 5 comprises nine HadUKP regional precipitation series covering the entire UK. In contrast, Set 1-3 consist of one to three site-specific precipitation series and Set 4 consists of two pools of multiple site-specific precipitation series to represent two of the nine HadUKP regional precipitation series. It shows that using the nine daily HadUKP regional precipitation series for the calibration of the CP-classification result in the most distinguishable CPs both in terms of precipitation characteristics and atmospheric pressure patterns.

	Set 1	Set 2	Set 3	Set 4	Set 5
Information measure 1	0.176	0.183	0.182	0.183	0.188
Information measure 2	0.167	0.183	0.179	0.185	0.195
Information measure 3	0.0021	0.0022	0.0022	0.0023	0.0023
PCD (separability)	0.210	0.213	0.242	0.244	0.249

Table 5.4 Results of the information measures to assess the CP-classification performances. It shows the performance of the CP-classifications averaged over different number of CPs and calibration time periods depending on the input precipitation Set

In order to assess the spatial variability of the CP-classification performances, the results of IM_1 conditioned on the twelve stations for CP-classifications obtained by using the precipitation Set 1 and the precipitation Set 5 are compared and plotted in Figure 5.3. IM_1 measures the discriminative power of the CP-classification in terms of wet day probabilities. It can be considered as more robust than IM_2 and IM_3 , which are based on precipitation intensities and are thus more affected by extreme values. Regarding the different precipitation Sets, Set 1 (using only one site-specific precipitation series) and Set 5 (using nine regional precipitation series) are most different to each other and thus differences in the spatial variability of the CP-classification performances are likely to be most pronounced between these two Sets.

CP-classifications based on precipitation Set 1 show the best performances for the stations in the western (Aldergrove, Eskdalemuir, Valley) and south western part of the UK (Camborne). This is not a surprising result as Set 1 only consists of hourly records at Aldergrove and as a result the CP-classifications are calibrated only on the specific precipitation patterns at Aldergrove. On the other hand, CP-classifications based on Set 1 show lower skill in distinguishing between wet and dry days for stations in the eastern part of the UK (Marham, Cranwell, Boulmer and Kinloss). Due to their geographical location, these stations are very likely to show different responses to large-scale atmospheric circulation in terms of precipitation compared to Aldergrove. Therefore, it seems reasonable to assume that the poor performances of the CP-classifications for these four stations can be explained by the fact that the CP-classification is only calibrated based on precipitation records at Aldergrove.

A similar spatial patterns can be found using Set 5. The CP-classifications based on Set 5 also show lower skill for the same four stations in the eastern part of the UK (Marham, Cranwell, Boulmer and Kinloss) compared to better performances found for stations in the West (Aldergrove, Eskdalemuir, Valley) and South West (Camborne) of the UK. It may be argued that stations in the western part of the UK generally exhibit a stronger relationship between CPs and precipitation patterns compared to stations in the East. In contrast to CP-classifications based on Set 1, the CP-classifications derived from Set 5 also show high skill for stations in the South (Boscombe Down and Northolt) and South West (Camborne). The IM_1 based on Set 5 in the South West (Camborne), South (Boscombe Down and Northolt), and East (Marham, Cranwell, Boulmer) of the UK are between 10% - 17% higher compared to that based on Set 1. This can be explained by the fact that the CP-classifications based on Set 5 are calibrated by taking into account the precipitation response of the entire nine different HadUKP regions across the UK. In contrast, the IM_1 based on Set 5 for stations in the northern part of the UK (Leuchars, Tiree and Kinloss) is only marginally better, 0% - 5%, than that based on Set 1. Performances for stations in the West of the UK (Valley, Eskdalemuir and Aldergrove) show no improvement using CP-classifications based on Set 5 compared to those based on Set 1. This is due to the fact that CP-classifications based on Set 1 are calibrated using the precipitation records at Aldergrove and should perform well at Aldergrove and the surrounding areas. Nevertheless, the quality of CP-classifications based on Set 5 is still generally good for stations in the western part of the UK. Overall, it can be concluded that CP-classifications based on the nine HadUKP regional precipitation series (Set 5) show better performances than those based on only one local precipitation series (Set 1) across the UK. The differences between CP-classification based on Set 5 and Set 1 are most pronounced in the southern and eastern parts of the UK.

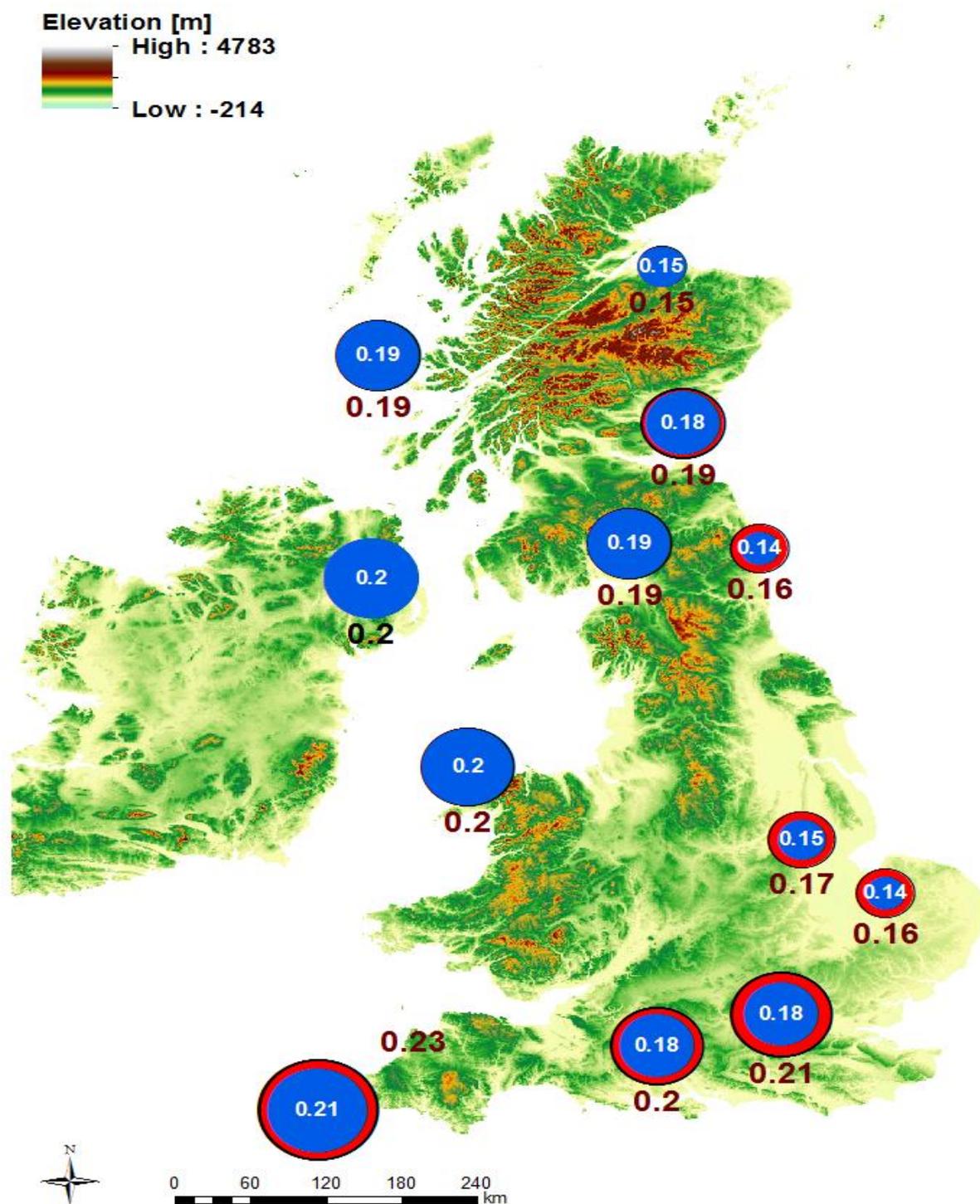


Figure 5.3 Spatial variability of the CP-classification in terms of the Information Measure 1 for the twelve selected stations using different input precipitation Sets. Results are averaged over 36 CP-classifications (four different calibrations time period and nine different numbers of CPs). Blue points with white labelling on top of the points represent the results obtained by using Set 1. Red points with black labelling around the points represent the results obtained by using Set 5.

5.2.2 The final input calibration parameters

Based on the results presented in Section 5.2.1, it appears that the number of CPs and the input precipitation Set are more important factors than the calibration period.

As shown in Table 5.4, CP-classifications based on precipitation Set 5 outperforms the rest of the CP-classification. In terms of the final number of CPs, the performance increases are highest from four to five CPs (plus one unclassified CP). Regarding IM_1 , IM_2 and the PCD value, the skill of the CP-classification improves up to a number of ten CPs. In contrast, IM_3 continues to increase up to the maximum number of 20 CPs. Particularly IM_2 and the PCD value show high fluctuations between the performances of different CP-classification using the same number of CPs. IM_3 is highly important in terms of downscaling hourly precipitation extremes because it assesses the skill of the CP-classification to link hourly precipitation extremes with certain CPs. The decision on the number of CPs is a trade-off between the robustness and skill of the classification. A smaller number of CPs can provide more precipitation events for each CP, whereas a larger number of CPs may lead to a higher performance. Given various factors to consider, the finally selected CP-classification has 5 CPs (plus one unclassified CP), is based on the input precipitation Set 5 and calibration time period 4 (1960-1979). It only compromises IM_3 slightly as the value shows small increase or decrease from 5 CPs at 0.0038, to 6, 7, 8, 9 and 10 CPs at 0.0038, 0.0027, 0.0038, 0.0044, and 0.0040 respectively. It also exhibits the highest PCD value (0.348) among all the CP-classifications based on five CPs. The calibration time period 4 (1960-1979) is independent from the validation period (1980-1989) and also has the best performance (see Table 5.3).

5.3 The final CP-classification used for statistical downscaling

5.3.1 Atmospheric pressure patterns of the CP-classification

Figure 5.4 (in winter) and Figure 5.5 (in summer) shows the selected set of CPs based on the NCAR SLP data (Hurrell and Trenberth, 2013) using the input configuration described in Section 5.2.2 for the time period 1980-2009. In the following, this set of CPs based on the NCAR SLP data set is referred to as the NCAR CP-classification. The corresponding SLP anomalies (see Equation 5.1) of the NCAR CP-classification are illustrated in Figure 5.6 for the time period 1980-2009 over the entire year as the differences between winter and summer are minor.

$$A_{CP,i} = \frac{\overline{P_{CP,i}} - \overline{P}_i}{\sqrt{\frac{1}{N-1} \sum_{j=1}^N (P_{j,i} - \overline{P}_i)^2}} \quad (5.1)$$

$A_{CP,i}$ = SLP anomaly at grid point i for certain CP for the time period 1980-2009

$\overline{P_{CP,i}}$ = Average SLP at grid point i for certain CP for the time period 1980-2009

\overline{P}_i = Average SLP at grid point i over the total number of days N for the time period 1980-2009

$P_{j,i}$ = Daily SLP at grid point i for a certain day j

CP1 and CP4 are the two CPs characterized by low pressure over the UK. CP1 is described by a shallow low-pressure system and CP4 is characterized by a more dominant low-pressure system with a stronger gradient. The location of the low-pressure system is also different between CP1 and CP4. The low in CP1 is centred in the South West (in winter) or South (in summer) of the UK, while the one in CP4 is in the northern part of the UK. In terms of seasonal variations, the low pressure system in CP1 is more extensive trough in winter than it is in summer and the pressure gradient is stronger in winter than in summer for both CPs, CP1 and CP4. Considering the corresponding SLP anomalies, CP1 exhibit negative SLP anomalies in the southern part and positive anomalies in the northern part of the UK. CP4 is characterized by only negative SLP anomalies. The highest negative anomalies for CP4 are found in the centre of the UK. In terms of the annual frequency occurrences, CP4 is much more likely to occur than CP1 (24.1% compared to 6.7%) as shown in Figure 5.6. The CP frequency occurrences are also calculated over four different seasons, in order to compare the results with previous studies based on Lamb Weather Types (LWTs) over the UK (Jones et al., 1993; Lamb, 1972). It is found that CP1 is most likely to occur in summer and CP4 exhibits the highest frequencies in autumn (see Table 5.5). Both CP1 and CP4 are least likely to occur in winter. This is in line with results shown in Jones et al. (1993), where it was found that cyclonic weather types show the lowest frequency occurrences in winter.

	spring (MAM)	summer (JJA)	autumn (SON)	winter (DJF)
CP1	6.98	7.57	7.11	4.97
CP2	13.84	16.35	14.06	15.28
CP3	29.44	30.52	27.68	29.45
CP4	24.93	24.40	26.35	19.83
CP5	18.83	16.75	19.69	21.65
U	5.98	4.41	5.11	8.83

Table 5.5 Seasonal frequency of occurrence of the five CPs (+one unclassified) between 1980 and 2009 in [%]. U stands for unclassified.

In contrast to CP1 and CP4, CP2 and CP3 feature high pressure systems over the UK. Again, the pressure gradient is more pronounced in winter than it is in summer for both of the CPs, CP2

and CP3. The main difference between the two CPs is the location of the high pressure centre. In CP2 it is located East of the UK and in CP3 South West of the UK. Similar features can be found in terms of the SLP anomalies as illustrated. CP2 is characterized by a centre of positive SLP anomalies in the North West of the UK. CP3 exhibit the highest positive SLP anomalies in the South West of the UK. Regarding the CP frequency occurrences, CP3 is almost twice as likely to occur as CP2 (29.4% compared to 15.0%). Both CPs exhibit the highest frequencies in summer as shown in Table 5.5. This is consistent with previous studies regarding LWTs (Lamb, 1972), where it was found that anticyclonic (high pressure) LWTs are more frequent in summer over the UK.

CP5 is characterized by a negative South to North pressure and anomaly gradient resulting in westerly airflow over the UK. This westerly airflow brings humid air from the North Atlantic to the UK. Similarly to the other CPs, the pressure gradient is stronger in winter than in summer. CP5 occurs on 19.2% of all days. It shows the highest frequency occurrences in winter and the lowest in summer. Again this is in line with findings regarding the LWT (Lamb, 1972), where it was shown that westerly CPs are most dominant in winter.

Days which are not attributed to any of the five CPs are unclassified. This category comprises 5.6% of the total number of days.

The main difference between the NCAR CP-classification based on fuzzy rules and the most prominent example of CP-classification over the UK, the Lamb Weather Types (Lamb, 1972; Lamb, 1977), is that the seven pure LWTs (Anticyclonic, Cyclonic, Northerly, Easterly, Southerly, Westerly and North-Westerly) distinguish strictly between different weather types by using predefined threshold. For example, a day with cyclonic conditions in the northern part and south-westerly airflow in the southern part of the UK may be classified as Cyclonic, Southerly or Westerly in the LWT catalogue. In contrast, the fuzzy rules based CP-classification allows CPs with multiple features. For example, CP4 is characterized by cyclonic conditions and south-westerly airflow over the UK. It needs to be mentioned, that an extended version of the LWTs catalogue exists with 19 additional hybrid types (26 LWTs in total). In this case, each hybrid type is defined by a combination of two pure LWT types.

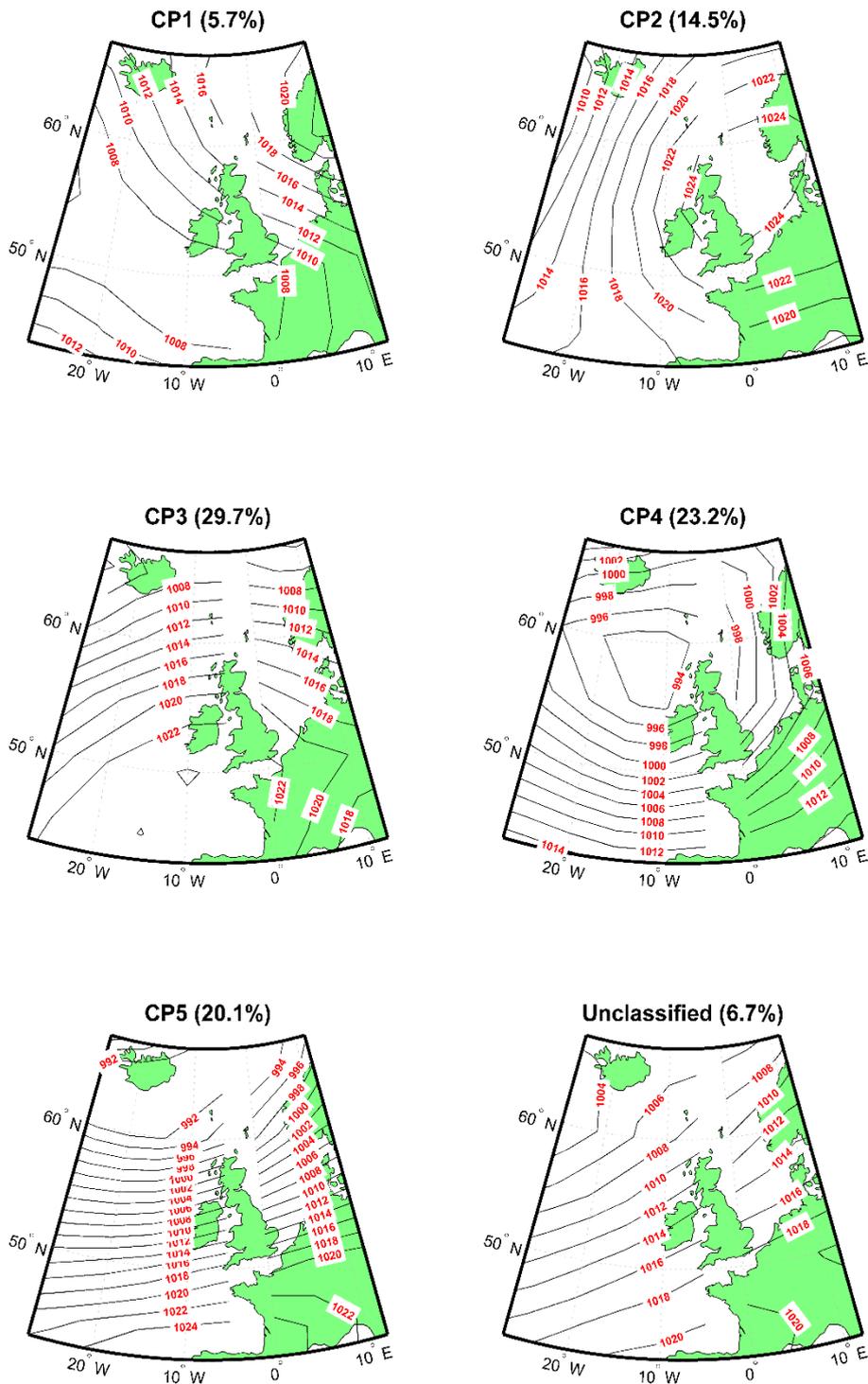


Figure 5.4 SLP patterns in [hPa] of the five CPs (+ one unclassified CP) derived from fuzzy rules using the final input configuration described in Section 5.2.2 for the validation time period 1980-2009 in winter (Nov-April). The frequency occurrences of each CP are given in parentheses. The SLP data comes from the NCAR SLP data set (see Section 3.2.1).

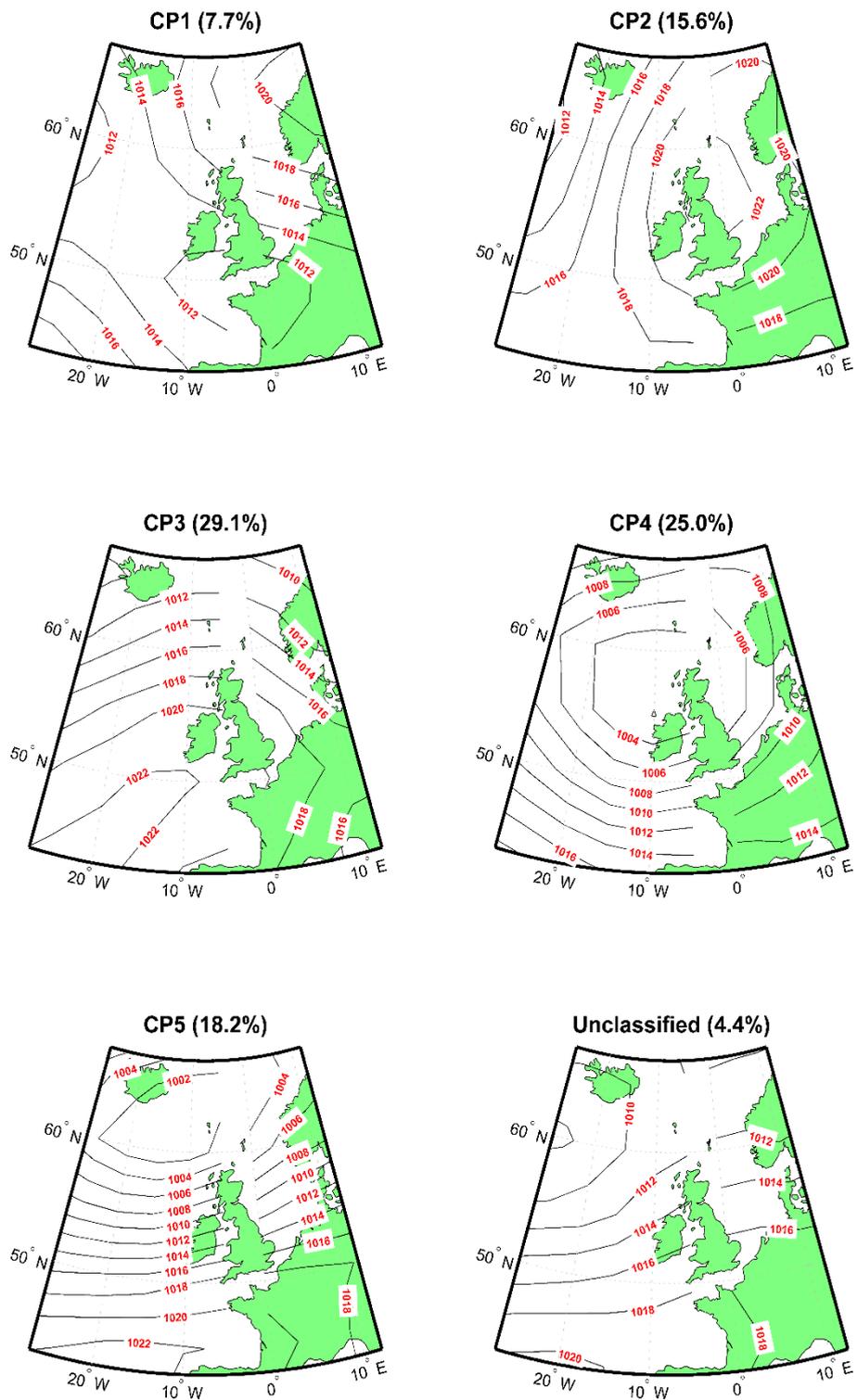


Figure 5.5 Same as Figure 5.4 but for the validation time period 1980-2009 in summer (May-Sept).

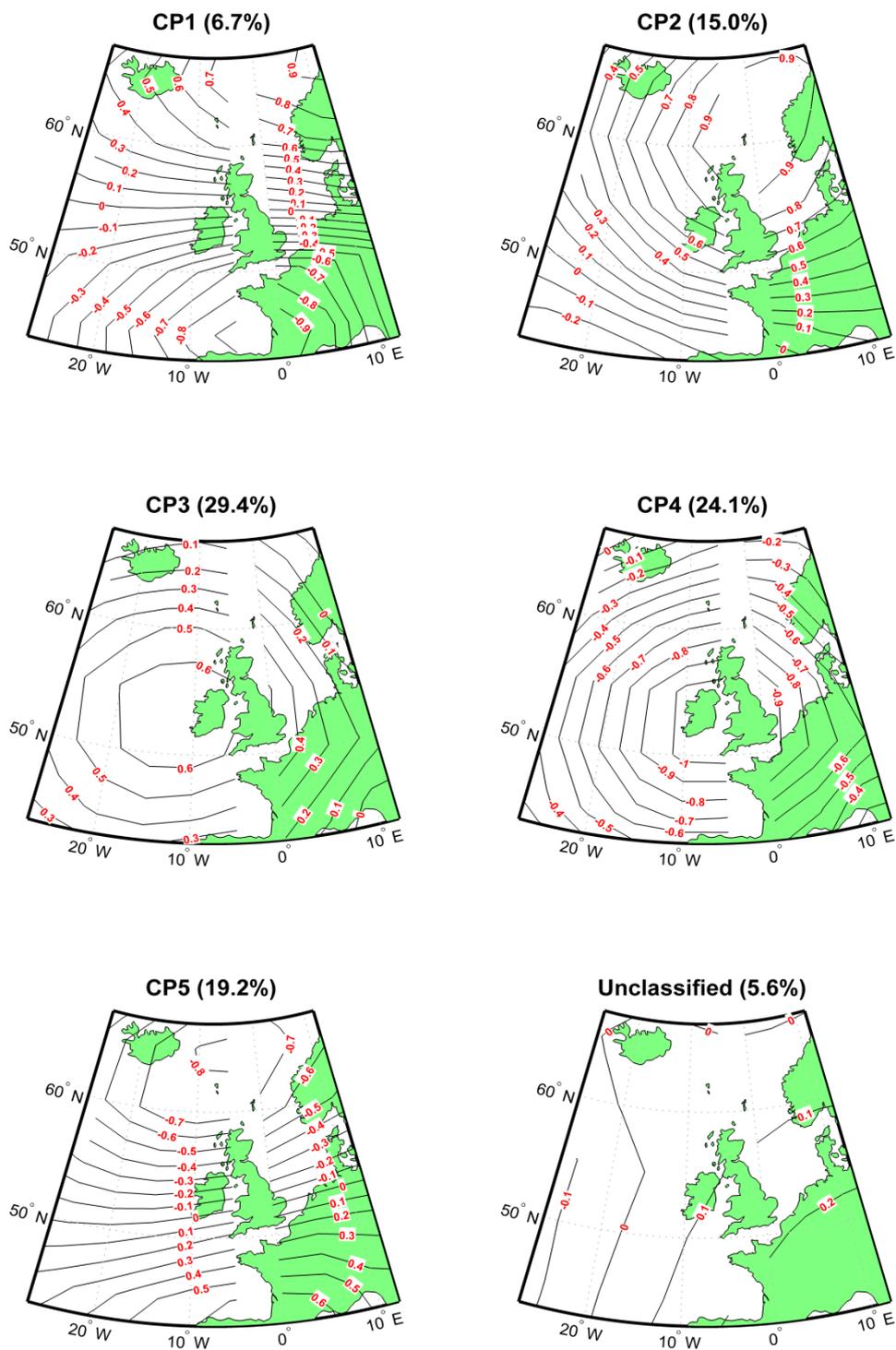


Figure 5.6 SLP anomalies of the five CPs (+ one unclassified CP) derived from fuzzy rules using the final input configuration described in Section 5.2.2 for the validation time period 1980-2009 over the entire year. The frequency occurrences of each CP are given in parentheses. The SLP data comes from the NCAR SLP data set (see Section 3.2.1).

5.3.2 The relationships between CPs and extreme hourly precipitation

In the following, the relationship between the CP-classification based on the input configuration described in Section 5.2.2 and extreme hourly precipitation is investigated. Two different indices of extreme hourly precipitation are applied. The first one is defined as daily maximum hourly precipitation exceeding a threshold of 10mm/hr. This one is used as information measure IM_3 in Section 5.2. The advantage of the first index is its robustness against outliers. This is particularly important for the assessment of the CP-classification, where the subsamples of extreme hourly events can be very small. The second one is defined as the 99.5th percentile of daily maximum hourly precipitation. This is used for the statistical downscaling of extreme events in Chapter 6 & 7. The reason for using the second index is that it helps to analyse differences in extreme hourly precipitation intensities beyond a certain threshold.

Figure 5.7 and Figure 5.8 show the probability $p_{CP,s}$ of hourly precipitation exceeding 10mm/hr (see Equation 5.2). The 99.5th percentiles are displayed in Figure 5.9 and Figure 5.10. Both indices are calculated using the final set of CPs and two seasons over all the precipitation records between 1980 and 2009. This is in line with the statistical downscaling method, where observed hourly precipitation records are also subsampled into five CPs (+one unclassified CPs) and two different seasons.

$$p_{CP,s} = \frac{N_{CP,s}}{n_{CP,s}} \quad (5.2)$$

$p_{CP,s}$ = Probability of extreme hourly precipitation (≥ 10 mm/hr) for a given CP in season s

$N_{CP,s}$ = Number of extreme hourly precipitation (≥ 10 mm/hr) events for a given CP in season s

$n_{CP,s}$ = Number of days for a given CP in season s

In winter, the stations at Eskdalemuir and Camborne record considerable more extreme hourly precipitation exceeding 10mm/hr than any other of the ten stations. Eskdalemuir is situated in a mountainous area (Svensson et al., 2002) and thus orographic effects play an important role in terms of its precipitation patterns over the entire year. Another explanation for the high frequency of extreme hourly precipitation at Eskdalemuir may be that the North Channel acts as a trigger for convective precipitation under cyclonic conditions particularly in autumn and winter (Sweeney and O'Hare, 1992). Camborne is located at the southwestern edge of the UK and under the influence of a maritime climate, which can bring humid and relatively warm air, particularly in autumn and early winter, to the UK. It is also highly exposed to vigorous frontal

and cyclonic weather systems travelling in from the North Atlantic, which are often associated with extreme precipitation in winter (Lenderink and van Meijgaard, 2008; Rodda et al., 2009). In this context it is interesting to note that the 99.5th percentiles of winter extremes tend to be generally higher for the stations Camborne, Valley, Eskdalemuir, Aldergrove and Tiree in the western part of the UK with a higher exposure to the North Atlantic. It also needs to be mentioned that most of the extreme hourly precipitation events in winter are associated with the two cyclonic CPs, namely CP1 and CP4. Cyclonic conditions can lead to so-called atmospheric rivers as the cold front of an extratropical cyclone sweeps up moisture in the warm sector while catching up with the warm front (Dacre et al., 2015). Champion et al. (2015) showed that atmospheric rivers can be linked with extreme precipitation in winter but cannot explain the occurrence of summer precipitation extremes. However, they only focused on daily precipitation extremes and it may be that atmospheric rivers show a stronger relationship with sub-daily precipitation extremes in summer.

In this study, indices of summer extreme hourly precipitation tend to be considerably higher than in winter across all stations, as it was already discussed in Section 3.1.5. It can also be seen that for stations in the southern part of the UK (Camborne, Boscombe Down, Northolt, Marham and Cranwell), indices of extreme precipitation in summer are higher than for stations in the North (Leuchars, Tiree and Kinloss). This can be explained by the fact that convective precipitation, which is particularly important in terms of hourly precipitation extremes (Gregersen et al., 2013), is strongly linked with high temperatures (Berg et al., 2013; Maraun et al., 2010b; Molnar et al., 2014). As a consequence, convective precipitation events are more likely to occur and are more intense in the southern part of the UK (Chan, 2013), where temperatures are higher than in the North of the UK (see Section 3.2.2). Another reason is that the stations in the South, in particular Camborne, Boscombe Down and Northolt, are more heavily affected by mesoscale convective complexes (MCCs). As described in Little et al. (2008), MCCs occur mostly in summer over southern or southwestern parts of the UK. They can be characterized by shallow low pressure zones originating over the Bay of Biscay, travelling from France to the UK and leading to intense thunderstorms. MCCs may be attributed to CP1 as this circulation pattern is also characterized by a shallow low pressure system over the southern part of the UK in summer. As shown in Figure 5.8 and Figure 5.10, CP1 leads to the highest extreme indices for stations (Camborne, Boscombe Down, Northolt, Marham and Cranwell) in the southern part of the UK in summer. For the remaining stations (except Kinloss), CP1 and CP4, which are both characterized by low pressure systems over the UK, are the most important CPs in terms of extreme hourly precipitation in summer. At Kinloss, the station located furthest North in this study, extreme hourly precipitation is most likely to occur for CP1

and CP5. CP5 is also characterized by a low pressure system to the North of the UK. Overall, it can be summarized that for each of the twelve stations, extreme hourly precipitation can be strongly linked with the occurrence of cyclonic conditions. This is in line with the results of other studies in the past. For example, (Kendon et al., 2014; Mishra et al., 2012) found a strong relationship between low pressure systems and extreme hourly precipitation events. (Sweeney and O'Hare, 1992) assessed the UK daily precipitation response with each of the Lamb Weather Types. They found that cyclonic south-westerly conditions lead to the highest precipitation intensities due to the existence both of moist tropical air masses and a low pressure system off the north-west coast. Conway et al. (1996) analysed the relationship between three different airflow indices (vorticity, strength of airflow and direction of airflow) and UK daily precipitation. It was shown that vorticity, a measure of cyclonic activity, exhibited the strongest influence on precipitation intensities.

Table 5.6 lists the maximum hourly precipitation events between 1980 and 2009 for each of the twelve selected stations and the prevailing CPs derived from the NCAR CP-classification on those days. Eight of the twelve maximum hourly precipitation events occur on days attributed to either CP1 or CP4. This means the two CPs characterized by low pressure systems over the UK are closely related to the maximum hourly precipitation events. It can be observed that the maximum hourly precipitation events of the four stations furthest West of the UK (Tiree, Aldergrove, Valley and Camborne) all occur on days attributed to CP4. This CP is characterized by a weather system transporting westerly airmasses from the North Atlantic to the UK under cyclonic conditions. It is also found that the three stations (Tiree, Valley and Camborne) closest to the North Atlantic are the stations with the longest precipitation duration on the day of the maximum hourly precipitation. It may be that those stations are less affected by short and very intense convective showers due to their close proximity to the Sea.

Station	Maximum hourly precipitation	Daily precipitation	Precipitation duration on this day	Date (YYYY-MM-DD HR:SS)	CP on this day
KINLOSS	50.0 mm	50.3 mm	2 hours	1982-07-11 03:00	CP02
TIREE	15.6 mm	51.8 mm	15 hours	2006-10-25 02:00	CP04
LEUCHARS	26.0 mm	45.6 mm	10 hours	2002-07-30 23:00	CP01
ALDERGROVE	25.1 mm	26.9 mm	3 hours	1986-05-07 18:00	CP04
BOULMER	19.2 mm	47.4 mm	8 hours	2008-08-01 07:00	CP04
ESKDALEMUIR	44.0 mm	53.0 mm	7 hours	1982-12-30 20:00	CP03
VALLEY	15.5 mm	50.2 mm	17 hours	1981-10-01 16:00	CP04
CRANWELL	26.1 mm	38.5 mm	5 hours	1982-08-04 21:00	CP01
MARHAM	28.5 mm	36.7 mm	5 hours	1982-06-06 16:00	CP02
NORTHOLT	36.7 mm	55.7 mm	4 hours	1992-05-30 02:00	CP01
BOSCOMBE DOWN	28.2 mm	40.4 mm	7 hours	1989-05-24 15:00	CP03
CAMBORNE	21.4 mm	66.2 mm	12 hours	2004-08-17 20:00	CP04

Table 5.6 Maximum hourly precipitation between 1980-2009 for each station and the corresponding daily precipitation, precipitation duration, date and assigned CPs using the NCAR CP-classification.

Four of the twelve maximum hourly precipitation records cannot be attributed to either CP1 or CP4. This is probably due to the fact that extreme hourly precipitation events are often strongly linked with increased convective activity (Beck and Bárdossy, 2013; Maraun et al., 2008). However, attributing these convective precipitation extremes, in particular localized convective showers, to certain CPs is not an easy task (Little et al., 2008). It also needs to be noted that this table cannot be considered as a representative overview of the most extreme hourly precipitation events between 1980-2009 across the UK. The twelve selected stations only provide precipitation records at a specific point in space. Therefore, it could be argued that it is almost by chance to what extent a specific storm, which leads to an extreme hourly precipitation events somewhere over the UK, is captured by one of the twelve stations. This is particularly the case for convective precipitation events as they are often very localized.

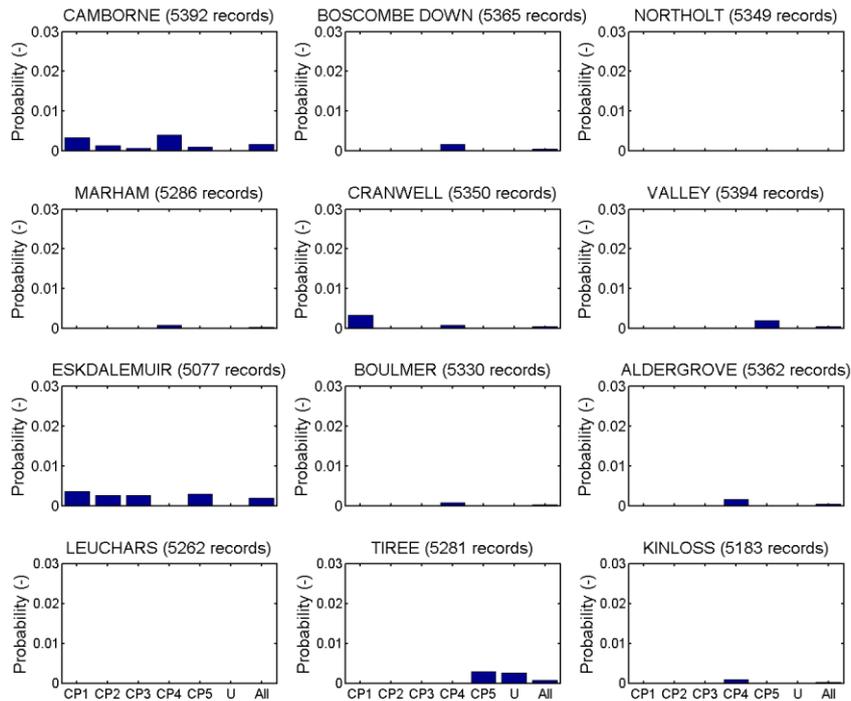


Figure 5.7 Extreme hourly precipitation (>10 mm/hr) probability in winter (Nov-April) conditioned on CPs (frequency occurrences: CP1=5.5%, CP2=14.7%, CP3=29.4%, CP4=23.0%, CP5=20.1% and U=7.4%) of the NCAR CP-classification for the twelve selected stations between 1980 and 2009. 'U' stands for unclassified CP and 'All' stands for all records unconditional of the CPs.

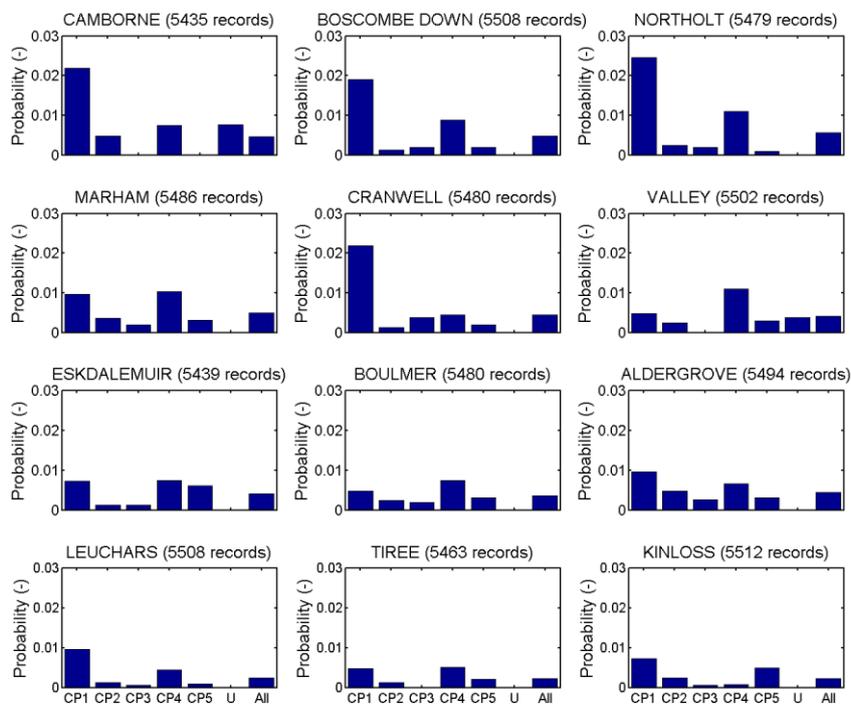


Figure 5.8 Same as Figure 5.7 but for summer (May-Oct) conditioned on CPs (frequency occurrences: CP1=7.6%, CP2=15.5%, CP3=29.0%, CP4=24.9%, CP5=18.2% and U=4.9%).

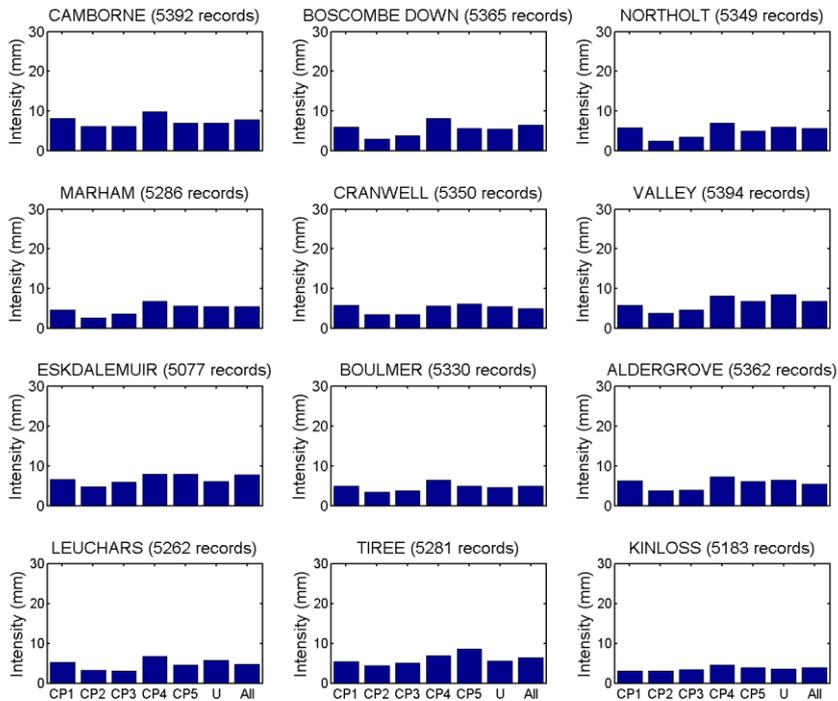


Figure 5.9 99.5th percentiles of the daily maximum hourly precipitation in winter (Nov-April) conditioned on CPs (frequency occurrences: CP1=5.5%, CP2=14.7%, CP3=29.4%, CP4=23.0%, CP5=20.1% and U=7.4%) of the NCAR CP-classification for the twelve selected stations between 1980 and 2009. 'U' stands for unclassified CP and 'All' stands for all records unconditional of the CPs.

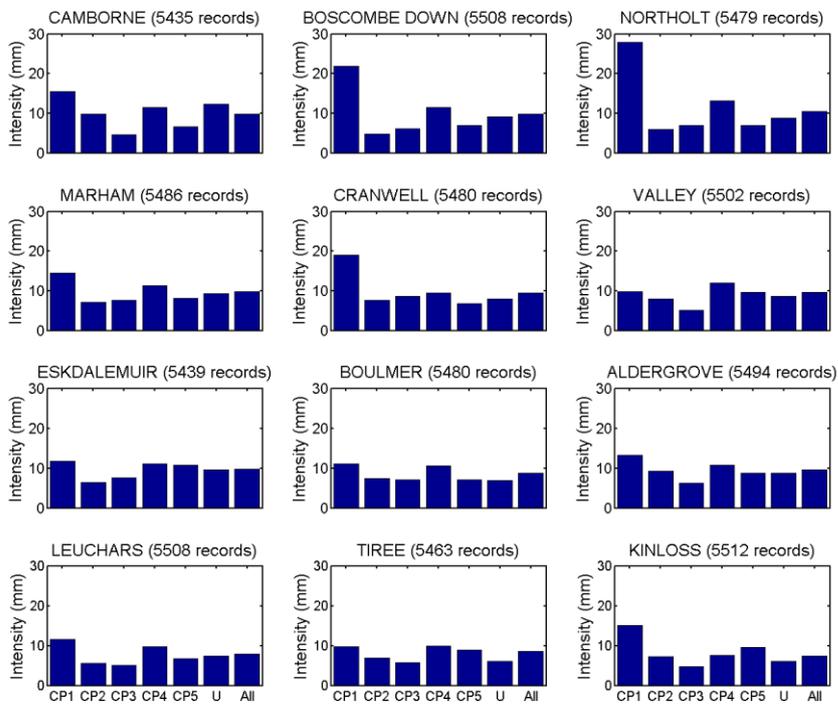


Figure 5.10 Same as Figure 5.9 but for summer (May-Oct) conditioned on CPs (frequency occurrences: CP1=7.6%, CP2=15.5%, CP3=29.0%, CP4=24.9%, CP5=18.2% and U=4.9%).

5.4 Summary

In Chapter 5, the results of fuzzy rules based CP-classifications with 180 different input parameters over the UK are presented and discussed. The input parameters comprise different input precipitation time series, calibration time periods and number of CPs. The best performing CP-classification is further used for the statistical downscaling of hourly precipitation extremes presented in Chapters 6 & 7. In order to determine the best performing CP-classification, the 180 different classifications are assessed based on (1) how well a certain classification represent the relationship between CPs and precipitation and how well the CP-classification distinguishes between wet and dry CPs, and (2) how well a certain classification can separate the pressure patterns represented by different CPs within the classification.

It is found that the CP-classifications exhibit the best performances for input precipitation set 5, which represents spatially averaged daily precipitation for the nine HadUKP regions across the UK. The different time periods used for CP-classifications have no major effect on the skill of the classification. In terms of the number of CPs, the performance of the CP-classification increases with higher number of CPs up to 10 or 20 CPs depending on the assessment measures. However, the highest rate of increase in the performance is reached from 4 to 5 CPs. In addition, certain CP-classifications with a smaller number of CPs show skills comparable to the better performing CP-classifications achieved with higher number of CPs. Another important aspect to consider is that a smaller number of CPs is favourable, in order to provide sufficient data in each subsample for the calibration process of the statistical downscaling method. Otherwise, it could lead to limited sampling variability (Willems and Vrac, 2011) and negatively affect the robustness of the precipitation duration–temperature relationships. Taking into consideration the performance of the classification and the robustness of the statistical downscaling process, the optimum classification has 5 (plus one unclassified) CPs, and is based on the input precipitation set 5 (9 HadUKP regional precipitation series) and the calibration time period 4 (1960-1979).

In the last section of Chapter 5, two important features of this finally selected CP-classification are described: (1) the different atmospheric pressure patterns of the final CP-classification, and (2) the relationships between each CP and extreme hourly precipitation. This is the first study to investigate the relationship between a CP-classification and UK extreme hourly precipitation. It is also the first study to apply a fuzzy-rules based CP-classification over the UK. In contrast to previous applications of the fuzzy rules based CP-classification (Bárdossy and Pegram, 2011; Beck and Bárdossy, 2013; Hundedcha and Bardossy, 2008), the finally selected CP-classification in this thesis only uses a relatively small number of different CPs (five plus one

unclassified). Two of the CPs (CP1 and CP4) are characterized by prevailing low pressure systems, two other CPs (CP2 and CP3) feature high pressure systems and the fifth CP (CP5) is associated with westerly airflow over the UK. The seasonal frequencies of the five CPs are shown to be in line with previous studies using the Lamb Weather Types (Jones et al., 1993; Lamb, 1972). Low pressure dominated CPs (CP1 and CP4) are most likely to occur in summer, whereas the CP featuring westerly airflow over the UK (CP5) show the highest seasonal frequencies in winter compared to any other season. In contrast to the LWTs, the fuzzy rules based CP-classification applied in this thesis is able to distinguish between different types of low (high) pressure systems. For example, CP1 features a relatively weak low-pressure system in the southern part of the UK, whereas CP4 is characterized by a more dominant and widespread low-pressure system with a stronger pressure gradient. In terms of the relationship between CP and extreme hourly precipitation, it is found that the two low pressure dominated CPs exhibit the strongest link with extreme hourly precipitation. Similar results were found in previous studies regarding hourly (Kendon et al., 2014; Mishra et al., 2012) and daily precipitation extremes (Conway et al., 1996; Sweeney and O'Hare, 1992). In particular for stations in the southern part of the UK, CP1 exhibits relatively high probabilities of extreme hourly precipitation. It is suggested that this circulation pattern may be linked with mesoscale convective complexes (MCCs). The two CPs featuring high pressure conditions show a relatively weak link with extreme hourly precipitation but still have the potential to lead occasionally to very extreme events. It is likely that those extreme events are caused by localized convective showers under high pressure conditions (Blenkinsop et al., 2015).

Chapter 6

Statistical downscaling based on observed records

Chapter 6 describes the statistical downscaling process of estimating local hourly precipitation extremes based on observed data. At first, the importance of temperature and daily precipitation as predictors to estimate hourly precipitation extremes is presented in Section 6.1. In the statistical downscaling process, temperature and daily precipitation are used to find the analogue day conditioned on seasons and CPs as described in Chapter 4. The importance of seasonality regarding hourly precipitation extremes is explained in Chapter 3. Only two six-month seasons (winter and summer) are considered, in order to maximise the data available to calibrate the precipitation duration and temperature relationship for each season and CP (see Section 6.2). The relevance of CPs in terms of estimating hourly precipitation extremes is described in Chapter 5. In Section 6.2, the relationship between precipitation duration and temperature is explained. This relationship is used in the statistical downscaling process to perturb the observed hourly precipitation record of the analogue days. Finally in Section 6.3, the results of the statistical downscaling process based on observed records are presented. Estimates of local daily maximum hourly precipitation extremes are compared against observed local daily maximum hourly precipitation records over three different validation periods. In addition, the stationarity assumption is assessed.

6.1 Relevance of the predictors in terms of extreme hourly precipitation

Figure 6.1 - 6.3 show scatter plots of the summer daily temperature and precipitation records between 1980 and 2009 exemplary for three stations in the southern part of the UK (Northolt, Marham and Cranwell). In this part of the UK, temperature are highest (see Section 3.2.2). As a results, those station are under an increased risk of convective precipitation, which often cause extreme hourly precipitation (Gregersen et al., 2013). The relevance of daily precipitation and daily mean temperature as predictors to estimate extreme hourly precipitation is well illustrated in those figures. It can be seen for all three stations, that there is a link between the occurrence of extreme hourly precipitation events ($>10\text{mm/hr}$) in summer and higher temperature. Such events necessarily require daily precipitation to be higher than 10mm/day . Only considering those records (daily precipitation $> 10\text{mm/day}$), extreme hourly precipitation records are more concentrated on the right half of the scatter plot at mean temperatures higher than approximately 15°C . High temperatures are often

considered as a trigger for convective precipitation and therefore the link between warm days and the occurrence of extreme hourly precipitation is physically plausible. In this context, it should be also noted that those three stations exhibit a pronounced maximum of extreme hourly precipitation for the warmest time of the year (JJA) as described in Chapter 3. This is a fact which underlines the importance of high temperature regarding hourly precipitation extremes. The relevance of the other predictor, daily precipitation, in terms of hourly precipitation extremes is self-evident as extreme hourly precipitation can only occur on days with daily precipitation higher than 10mm/day. Also, as seen in Figure 6.1, extremely high daily precipitation (>45mm/day) may also be strongly linked with very extreme hourly precipitation (>20mm/hr) at Northolt. However, this link is not as strong as the ones for Marham and Cranwell as illustrated in Figure 6.2 and Figure 6.3.

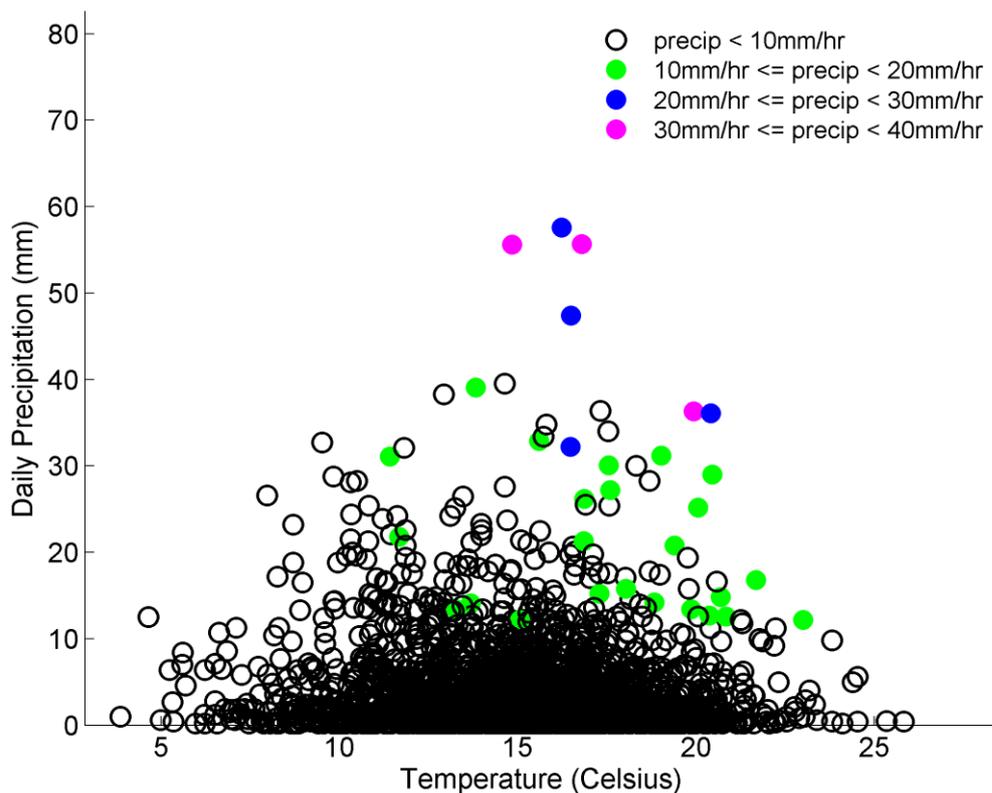


Figure 6.1 Scatter plot of daily precipitation events (≥ 0.2 mm/day) recorded in summer between 1980 and 2009 at Northolt. It shows the E-OBS gridded daily mean temperature in Celsius degrees (x-axis) and site-specific daily precipitation in mm (y-axis), as well as the intensity of daily maximum hourly precipitation represented by different colours.

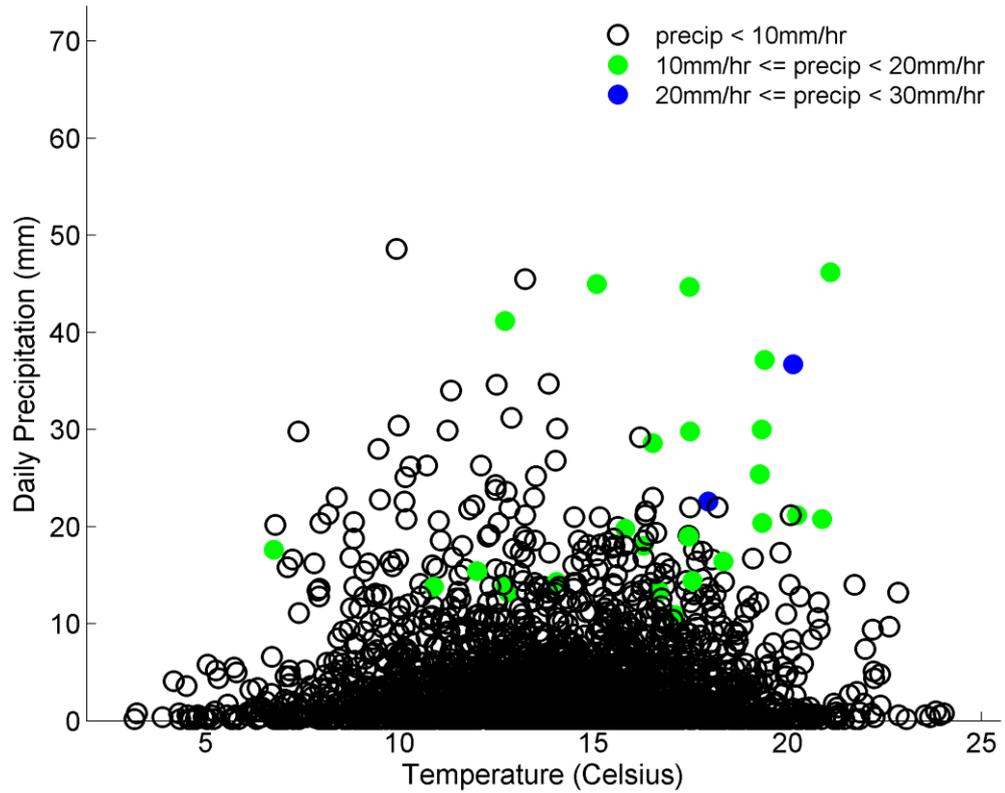


Figure 6.2 Same as Figure 6.1 but for the station at Marham.

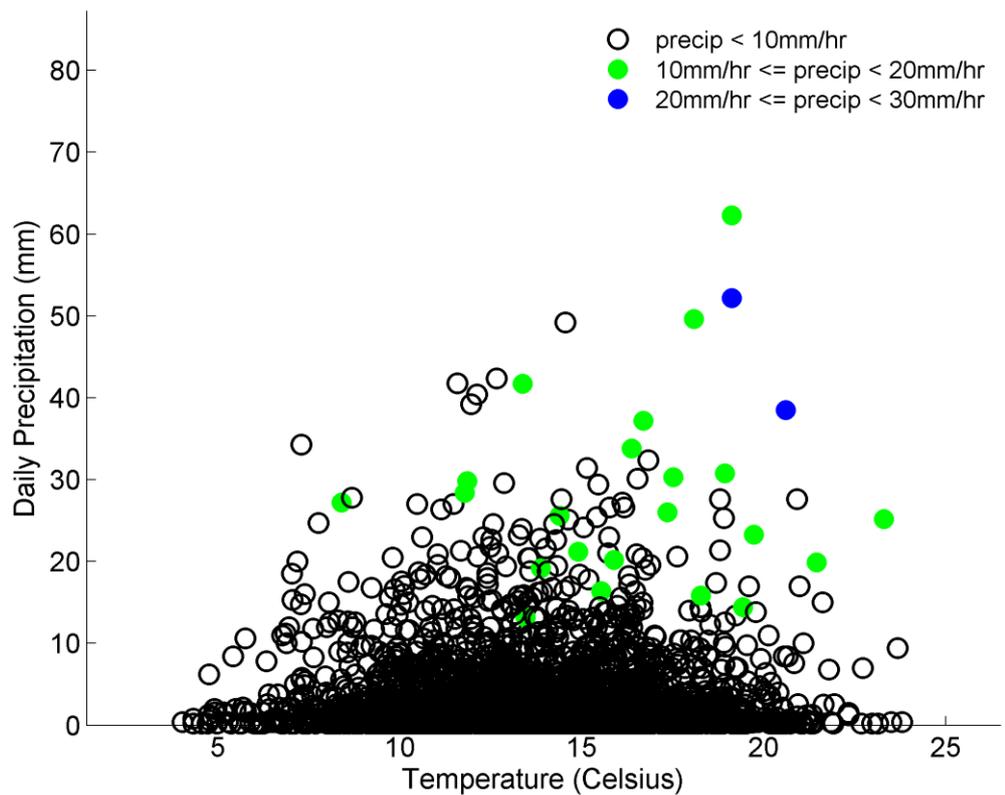


Figure 6.3 Same as Figure 6.1 but for the station at Cranwell.

6.2 Relationships between precipitation duration and temperature

In the following section, the relationships between precipitation duration and temperature conditioned on CPs are analysed for the twelve selected stations. In this thesis, the precipitation duration-temperature relationship is applied to perturb the observed hourly precipitation records in order to overcome the main limitation of the analogue day method, namely that it only simulates precipitation events observed in the past. Previous studies have shown that an increase in temperature often leads to shorter precipitation events. For example, Hardwick Jones et al. (2010) found a decrease in precipitation duration as temperature increases. Similarly, Beck and Bárdossy (2013) showed that higher temperatures lead to a decline in the frequency of wet hours but also causes a shift in the hourly precipitation distribution towards higher intensities. Theoretically, the relationship between hourly precipitation intensity and temperature could also be used to perturb the observed hourly precipitation records. However, it was found that the sample sizes were often too limited to identify reliable hourly precipitation-temperature relationships after subsampling the records into five (plus one) CPs and two seasons. For the same reason, the precipitation duration-temperature relationship applied in this thesis was not calibrated based on extreme events only. In terms of the stationarity assumption, conditioning the observed precipitation duration – temperature relationship on different CPs is important to identify relationships which remain valid under a future warmer climate. Otherwise the relationship between precipitation and temperature may lead to false conclusions. For example, Chan et al. (2016) found a decline in extreme precipitation intensities above 22°C and argued that it may be due to a more frequent occurrence of anticyclonic conditions at high temperatures in the present-day climate.

Figure 6.4 illustrates different relationships between daily precipitation duration and temperature depending on the CPs analysed in this study. It shows the relationships exemplary for the station at Camborne (South West England) for the time period 1980-1999 in summer. Similar results are found for the other stations as discussed in the following. At Camborne, CP1 and CP4 exhibit substantially higher negative regression slopes compared to the rest of the CPs and only the regression slopes of CP1 and CP4 are significant at the 5% level according to the p-value test. In this context, it needs to be mentioned that CP1 and CP4 are both characterized by low pressure areas over the UK. It seems that the effect of higher temperature leading to precipitation events of shorter duration is particularly pronounced for days of low pressure areas over the UK. This holds true for all three different calibration periods at Camborne (see Figure 6.5). A similar picture can be seen for the other eleven stations considered in this study regarding CP4. In total, 32 of the 36 linear regression slopes conditioned on CP4 are

statistically significant and result in RMSE (Root Mean Square Error) values lower than 1 (see Table 6.1). In contrast, for CP1 only the three most southern stations (Camborne, Boscombe Down and Northolt) exhibit a high fraction (9 of 9) of significant regression slopes (with $RMSE < 1$) over the three different calibration time periods. Considering the total number of twelve stations, the fraction of significant regression slopes (with $RMSE < 1$) conditioned on CP1 is only 13 of 36. This can be explained by the fact, that CP1 is characterized by a rather small shallow low pressure area south of the UK affecting mostly stations in the southern part of the UK. Conversely, CP4 is associated with a large low pressure zone over the UK and as a consequence has a more widespread impact. The linear regression slopes associated with CP1 show higher negative values than those for CP4. This means a similar rise in temperature results in a larger drop in the duration of CP1 associated than CP4 associated precipitation events. One possible explanation is that the low pressure area south of the UK, associated with CP1, can be linked with warmer temperature due to its geographical location and hence increased convective activity, compared to the low pressure area centred north of the UK associated with CP4. Considering the other circulation patterns, CP2 and CP3 are characterized by high pressure systems over the UK. The limited moisture availability associated with those weather conditions is often the restricting factor in terms of the actual precipitation intensities. As a result, the frequency occurrences of extreme hourly precipitation events for days attributed to CP2 and CP3 across the twelve stations are relatively low (see Section 5.3.2) and do not affect the duration-temperature relationship substantially. This may explain why no significant relationship can be found between precipitation duration and temperature at Camborne for those two CPs (see Figure 6.4 and Figure 6.5). Similarly, for most of the other eleven stations (see Table 6.1) no significant relationship between precipitation duration and temperature conditioned on CP2 and CP3 can be established. Only the station at Northolt (South East England) and Kinloss (East Scotland) show statistically significant regression relationships (with $RMSE < 1$) in summer over all three calibration periods conditioned on CP2 (Northolt) and CP3 (Kinloss).

Northolt is the station with the highest extreme temperatures as illustrated in Section 3.2.2 and therefore also exhibits an increased risk of convective precipitation events. It is suggested that the statistically significant regression relationships (with $RMSE < 1$) conditioned on CP2 at this station may be caused by the intensification of convective precipitation events with higher temperature on days of anticyclonic conditions. In this context it is interesting to note that statistically significant regression relationships (with $RMSE < 1$) conditioned on CP2 exhibit on average the most negative slopes of all statistically significant regression slopes (with $RMSE < 1$) as shown in Table 6.1. This also indicates increased convective activity as the

strongest reaction to temperature is often found for convective precipitation events (Berg et al., 2013; Molnar et al., 2014).

At Kinloss, located in the northern part of the UK, it is more likely that the strong relationship between precipitation duration and temperature on days attributed to CP3 is due to approaching cold/warm fronts from the North Atlantic under westerly conditions. This is in line with the actual SLP pattern of CP3 characterized by westerly airflow in the northern part of the UK.

Similarly, CP5 can also be associated with westerly airflow but over the entire UK. At Camborne, no statistically significant regression can be identified under CP5 conditions (see Figure 6.4 and Figure 6.5). The same holds true for Valley (North West England) and Boulmer (North East England), which also exhibit no statistically significant regression slope conditioned on CP5 in summer. At Aldergrove (North Ireland) and Leuchars (Forth), only one of the three calibration periods show a statistically significant regression slope conditioned on CP5 in summer. In contrast, at Boscombe Down, Northolt, Marham, Cranwell and Eskdalemuir a different behaviour in terms of the precipitation duration and temperature relationship conditioned on CP5 can be found. In total, 14 of the 15 regression slopes are statistically significant and fulfil the RMSE criterion conditioned on CP5 in summer over all three calibration periods for these five stations (see Table 6.1). It is interesting to note that these five stations are all not in close proximity to the sea. As mentioned above, CP5 is characterized by westerly airflow over the UK and thus precipitation is often caused by cold/warm fronts travelling in from the North Atlantic. It seems that stations further inland of the UK exhibit a pronounced intensification of frontal precipitation events with increased temperature, which is reflected by the statistically significant duration-temperature relationships.

The other two stations with statistically significant regression slopes (and $RMSE < 1$) conditioned on CP5 in summer are Tiree (North Highlands and Islands) and Kinloss (East Scotland). Those two stations, characterized by high latitudes, are under the influence of a low pressure zone north of the UK associated with CP5. As mentioned earlier for CP1 and CP4, the relationship between precipitation duration and temperature appears to be stronger under low pressure conditions than high pressure.

In general, it needs to be noted that the sample size determines the amount of sampling error inherent in a regression analysis. Statistical relationships are harder to detect in smaller samples, while increases in the sample size helps to improve the statistical power of the relationships. For example, 35 of the 36 CP-unconditional relationships between daily precipitation duration and temperature are statistically significant and fulfil the RMSE criterion over the three different time periods (see Table 6.1). Those statistical relationships are based

on at least 975 wet day records, whereas for the regression analysis of some CP-conditional relationships not more than 50 wet day records are available.

The same holds true for the estimation of confidence intervals of the linear regression slopes. The uncertainty of the regression analysis, represented by the range of the confidence intervals, increases with smaller sample sizes. This is exemplarily illustrated for the summer time period 1980-1999 at Camborne in Figure 6.5. The relationship between daily precipitation duration and temperature conditioned on CP1 are estimated based on only 126 wet days, the relationship conditioned on CP4 is based on 603 wet days, and the CP-unconditional relationship between is based on 1275 wet days. As a result, the range of the confidence intervals are largest for the estimated relationship conditioned on CP1 and smallest for the CP-unconditional relationship.

It is also found, that the confidence intervals exhibit a tendency to be skewed towards lower, more negative, linear regression slopes. Therefore, it can be suggested that the estimated regression slopes applied in the statistical downscaling method are rather conservative and may be more extreme in reality. In order to quantify this additional component of uncertainty in the process of estimating extreme hourly precipitation events, the confidence intervals of the linear regression could be incorporated into the statistical downscaling method.

The same analysis was also conducted for the relationships between daily precipitation duration and temperature in winter for all twelve stations over the three different calibration periods (results not shown). The number of significant relationships between duration and temperature (with $RMSE < 1$) in winter (67 out of 252) is considerably smaller compared to summer (131 out of 252) for the twelve selected stations over the three calibration periods, although the results are based on similar sample sizes for summer and winter. It also needs to be mentioned that for stations in close proximity to the North Atlantic (Camborne, Eskdalemuir, Valley and Tiree), the regression slopes in winter exhibit positive values. This means that higher temperatures in winter often result in events with longer precipitation duration for those stations, whereas in summer all stations show a decrease in precipitation duration with increasing temperature. The differences in the duration-temperature relationships between winter and summer can at least partly be explained by the different predominant precipitation types in winter and summer. Winter precipitation is more strongly associated with large-scale precipitation events such as frontal systems or cyclones (Chan et al., 2014a; Lenderink and van Meijgaard, 2008), whereas summer precipitation is often caused by convective events (Beck and Bárdossy, 2013; Westra et al., 2014). For both types of precipitation it holds true that higher temperatures lead to an increase in the water holding capacity of the atmosphere. As a result, not only the intensity but also the duration of a

precipitation event may increase with higher temperatures. In terms of convective precipitation, however, other important aspects also need to be considered. For example, higher temperatures cause a stronger convective uplift, which often leads to more intensive and shorter precipitation events. The same holds true for positive feedback mechanisms with higher temperatures (Berg et al., 2013; Lenderink and van Meijgaard, 2008). It is suggested that the convective factors dominate the duration-temperature relationship in summer and lead to a decrease in the precipitation duration with higher temperatures. In contrast, in winter the convective factors are not as dominant as in summer. Therefore, the surplus of precipitable water with higher temperatures can lead to an increase in the precipitation duration in winter.

		CAMBORNE	BOSCOMBE DOWN	NORTHOLT	MARHAM	CRANWELL	VALLEY	ESKDALEMUIR	BOULMER	ALDERSGROVE	LEUCHARS	TREE	KINLOSS
CP1	C1	-0.24 (126)	-0.35 (108)	-0.28 (116)	-0.49 (105)	-0.31 (101)	(102)	-0.33 (87)	(82)	(95)	(67)	(75)	(64)
	C2	-0.27 (119)	-0.29 (104)	-0.27 (106)	(103)	(99)	(97)	(98)	(85)	-0.38 (84)	(85)	(65)	(75)
	C3	-0.33 (131)	-0.34 (118)	-0.26 (118)	(106)	(108)	(119)	(99)	(95)	(95)	(86)	(76)	(63)
CP2	C1	(84)	-0.47 (51)	-0.25 (39)	(59)	(52)	(76)	(100)	(83)	-0.20 (103)	(72)	(128)	(98)
	C2	(83)	(56)	-0.29 (51)	(62)	(56)	(85)	(107)	(84)	-0.36 (106)	(78)	(142)	(95)
	C3	(71)	-0.46 (53)	-0.34 (52)	-0.33 (57)	(58)	(75)	(103)	(87)	(107)	(76)	(124)	(109)
CP3	C1	(166)	(165)	-0.19 (140)	-0.22 (169)	-0.16 (196)	(184)	(281)	(211)	-0.25 (242)	-0.17 (174)	(334)	-0.15 (272)
	C2	(172)	(150)	(153)	(174)	(192)	(196)	(294)	(225)	(253)	(177)	(358)	-0.13 (279)
	C3	(176)	(167)	(129)	-0.15 (167)	(194)	(188)	(281)	-0.21 (212)	-0.25 (239)	(177)	(350)	-0.26 (265)
CP4	C1	-0.17 (603)	-0.20 (588)	-0.18 (488)	-0.18 (470)	-0.24 (469)	(563)	-0.16 (657)	-0.24 (502)	(589)	-0.26 (510)	-0.16 (540)	-0.21 (470)
	C2	-0.15 (624)	-0.21 (579)	-0.20 (507)	-0.23 (500)	-0.28 (488)	(583)	(713)	-0.25 (503)	-0.13 (637)	-0.23 (531)	-0.18 (598)	-0.32 (519)
	C3	-0.16 (579)	-0.17 (555)	-0.29 (483)	-0.30 (478)	-0.33 (469)	-0.21 (548)	-0.26 (686)	-0.33 (499)	-0.28 (592)	-0.35 (513)	-0.24 (538)	-0.32 (491)
CP5	C1	(238)	-0.18 (178)	-0.25 (147)	-0.18 (160)	(174)	(298)	-0.15 (488)	(187)	(379)	-0.13 (241)	-0.24 (507)	-0.25 (294)
	C2	(259)	-0.18 (199)	-0.23 (178)	-0.18 (200)	-0.17 (217)	(325)	-0.27 (523)	(215)	(415)	(285)	-0.26 (548)	-0.19 (324)
	C3	(253)	-0.18 (191)	-0.36 (159)	-0.24 (186)	-0.19 (207)	(307)	-0.18 (517)	(196)	-0.19 (394)	(278)	-0.35 (521)	-0.20 (314)
U	C1	(58)	(55)	(45)	(55)	(49)	(63)	(83)	(61)	(66)	(48)	-0.55 (80)	-0.48 (67)
	C2	(57)	-0.42 (57)	-0.51 (51)	(48)	(49)	(55)	(74)	(51)	(63)	(57)	-0.46 (71)	-0.61 (52)
	C3	(63)	-0.42 (56)	-0.44 (48)	(65)	(52)	(64)	(83)	(60)	(65)	(51)	(71)	-0.47 (63)
w/o CPs	C1	-0.11 (1275)	-0.21 (1145)	-0.20 (975)	-0.21 (1018)	-0.19 (1041)	-0.07 (1286)	-0.14 (1696)	-0.16 (1126)	-0.14 (1474)	-0.21 (1112)	-0.17 (1664)	-0.20 (1265)
	C2	-0.14 (1314)	-0.21 (1145)	-0.20 (1046)	-0.16 (1087)	-0.17 (1101)	(1341)	-0.10 (1809)	-0.15 (1163)	-0.13 (1558)	-0.15 (1213)	-0.20 (1782)	-0.23 (1344)
	C3	-0.12 (1273)	-0.20 (1140)	-0.27 (989)	-0.22 (1059)	-0.23 (1088)	-0.13 (1301)	-0.14 (1769)	-0.21 (1149)	-0.22 (1492)	-0.18 (1181)	-0.21 (1680)	-0.24 (1305)

Table 6.1 The linear regression slopes for the precipitation duration and temperature relationship in summer conditioned on the set of CPs over the three different calibration periods and for the twelve selected stations. It only shows the linear regression slopes for those relationships that are significant (at the 5% level) and fulfil the RMSE (Root-Mean-Square-Errors) criterion (RMSE<1). The relationships are estimated based on wet days ($\geq 1\text{mm}$) only. The number of wet days are given in parentheses. U stands for unclassified days and w/o CPs stands for all days unconditional on the set of CPs. C1, C2 and C3 represent the three different calibration periods: 1980-1999 (C1), 1980-1989 + 2000-2009 (C2) and 1990-2009 (C3).

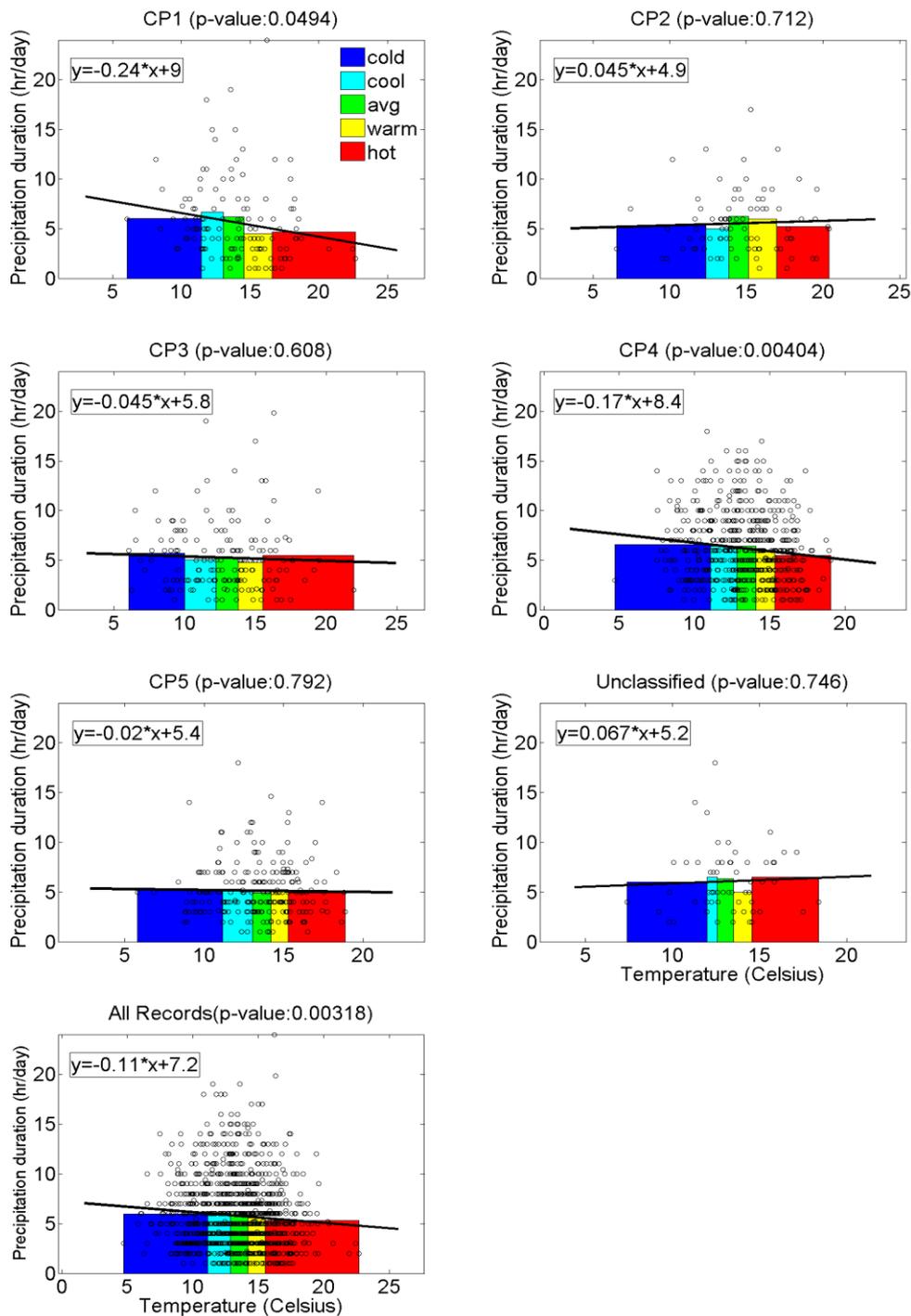


Figure 6.4 The relationship between daily precipitation duration and temperature at Camborne in summer (May-Oct) for the time period 1980-1999 conditioned on the set of CPs. The results of the p-value significance test are given in the title in parentheses for each CPs. The black line indicates the linear regression line and the coloured bins represent mean precipitation duration and mean temperature averaged over five temperature categories for the time period 1980-1999 in summer (cold: 0-20th percentile; cool: 20th – 40th percentile; medium: 40th – 60th percentile; warm: 60th – 80th percentile; hot: 80th -100th percentile).

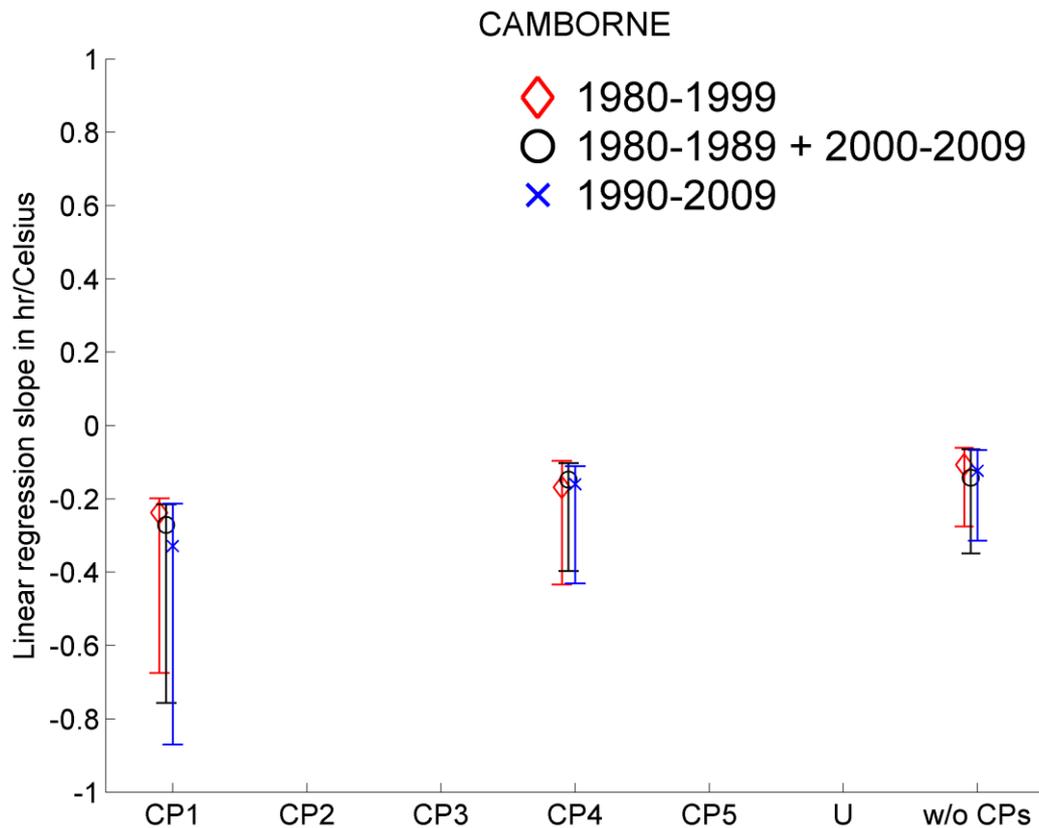


Figure 6.5 The values of the linear regression slopes between precipitation duration and temperature in summer conditioned on the set of CPs for the three different calibration periods at Camborne. It only shows the linear regression slopes for those relationships that are significant (at the 5% level). It also illustrates the confidence intervals at the 10% level of the estimation of the linear regression slopes.

6.3 Cross-validation to assess statistical downscaling performance

In the following section, the results of the estimation of hourly precipitation values based on the statistical downscaling methods described in Chapter 4 are presented. As mentioned before, twelve different stations over the UK are selected as shown in Chapter 3. Three different validation periods are applied as described in Table 6.2 to assess the reliability and robustness of the downscaling method. The three corresponding calibration periods are chosen accordingly over which the historical analogue days are drawn and the precipitation duration-temperature relationship (see Section 6.2) is defined. In order to test whether the downscaling method is able to perform under a warming climate, a fourth validation period is defined comprising the ten warmest summers of each station.

For each validation period and station, the averaged 99.5th percentile error over 100 bootstrapping samples $\overline{PE}_{99.5}$ is calculated. For example, the validation period V1 (1980-1989) comprises 3653 different calendar days. Therefore, each of the 100 bootstrapping sample i

also consists of 3653 validation days. Those 3653 validation days are selected with replacement from the population of 3653 calendar days for each of the 100 bootstrapping samples separately. It means that in each bootstrapping sample, some calendar days occur multiple times while others are not included at all. In a next step, the 99.5th percentile error $PE_{99.5,i}$ is calculated for each of the 100 bootstrapping samples, in order to determine the averaged 99.5th percentile error $\overline{PE}_{99.5}$ (see Equation 4.22 and Equation 4.23). The same procedure is applied in Chapter 7 in order to estimate the averaged 99.5th percentile errors and to project the 99.5th percentiles over 100 bootstrapping samples. The only difference is that Chapter 6 is based on observed predictor variables, whereas Chapter 7 uses RCM data to estimate and project extreme hourly precipitation.

Four different approaches (M1, M2, M3 and M4) of finding the analogue days as defined in Table 4.1 are compared. M1 stands for the analogue day method conditional on summer and winter (but not on calendar months) and no daily precipitation categories are applied. The M2 approach is conditional on the twelve calendar months without any daily precipitation categories. In M3, the analogue day method is conditional on winter and summer and daily precipitation categories are applied. M4 stands for the analogue day method conditional on the twelve calendar months and daily precipitation categories are used. After the analogue day in the calibration period is found, the duration-temperature relationship described in Section 6.2 is applied to perturb the historical daily maximum hourly precipitation record.

Calibration period	Validation period
C1: 1990 – 1999	V1: 1980 – 1989
C2: 1980 – 1989 & 2000 - 2009	V2: 1990 – 1999
C3: 1980 - 1999	V3: 2000 – 2009

Table 6.2 The different calibration periods and corresponding validation periods used for the cross-validation of the statistical downscaling performance in Section 6.3 and Section 7.1.

As shown in Table 6.3, no approach can be identified as the best for all the twelve stations in terms of the estimation of extreme hourly precipitation over different validation periods. On average, the statistical downscaling based on M1 performs best over the three validation periods and the twelve stations. However, the overall differences in the performances between the four approaches are small. The best approach (M1) exhibits a $\overline{PE}_{99.5} = 11.2\%$ compared to the worst performing approach (M4) with a $\overline{PE}_{99.5} = 12.8\%$.

Although averaged differences in the performances between the different approaches are relatively small, M3 and M4 show a tendency to overestimate the most extreme hourly precipitation events. For example, using M3 leads to overestimation of the highest hourly

precipitation extremes at Cranwell and Kinloss and M4 causes overestimation of the highest hourly precipitation extremes at Eskdalemuir and Kinloss for the validation period 2000-2009 as shown in Figure 6.6. It seems that not using daily precipitation as a final predictor increases the risk of overestimating the most extreme hourly precipitation. Instead, the highest hourly precipitation values estimated with M1 and M2, using daily precipitation as a final predictor, show much better agreement with the observed hourly precipitation extremes for those stations.

In terms of the reliability and robustness of the downscaling methods, it can be seen that all four methods show mostly high agreement between the estimated and observed hourly precipitation extremes over the three different validation time periods. However, for the station at Northolt and Marham, using downscaling approach M1 and M2 result in relatively high $\overline{PE}_{99.5}$ (ca. 20%) over validation period 1 (V1) as shown in Table 6.3. Estimates based on M4 show high $\overline{PE}_{99.5}$ over V3 at Northolt. M3 and M4 result in high $\overline{PE}_{99.5}$ at Boscombe Down over V1 (20.9% and 42.6%). At Cranwell, M3 shows low skill in estimating hourly precipitation extremes over V3 ($\overline{PE}_{99.5} = 22.0\%$). It needs to be mentioned that Boscombe Down, Northolt, Marham and Cranwell are all located in the southern and south-eastern part of the UK, where temperatures are higher (see Section 3.2.2). Those stations exhibit an increased risk of convective precipitation events (Chan, 2013), which makes the prediction of extreme hourly precipitation a very challenging task. High differences between estimated and observed extreme hourly precipitation are also found at Leuchars, although this station is less likely to be highly affected by convective precipitation events.

In Figure 6.7, the influence of the perturbation factor on the hourly precipitation estimates is illustrated over the validation period 2000-2009 for all twelve stations. The perturbation factor is derived from the duration-temperature relationship described in Section 6.2. It can be seen, the influence of the perturbation factor on the hourly precipitation estimates is small illustrated by the small differences between non-perturbed and perturbed hourly precipitation estimates based on approach M1. Similar results are found based on the other approaches (not shown). Non-perturbed estimates represent the hourly precipitation recorded on the analogue day, while perturbed estimates are generated by perturbing those observed hourly records on the analogue day using the derived relationship.

The perturbation factor that takes into account changes in the temperature variable is intended to address non-stationarity. In other words, the downscaling method needs to remain valid under a changing climate. In order to test this assumption, Table 6.4 shows the results of the downscaling methods validated over the years with the ten warmest summers

(May-Oct) on average between 1980 and 2009 depending on each station. It thus assesses whether the downscaling methods are able to perform under a climate, which is as different as possible in terms of temperature to the climate it is calibrated on (the 20 coldest summers). On average over the twelve stations all four downscaling methods show even better skill over the ten year period with the warmest summer, indicated by smaller $\overline{PE}_{99.5}$, compared to the averaged results over the three validation periods as shown in Table 6.3. It needs to be noted, however, that the changes in mean temperature are relatively small (0.8°C-1.3°C) compared to the projected temperature increases (up to 3.5°C) in the future scenarios (see Section 7.3). As for the results averaged over the three validation periods, approach M1 performs also best over the ten warmest summer period on average over the twelve stations.

As discussed in Section 3.2.2, the observed temperature increased significantly across all stations over the time period 1980-2009. It is important to assess if the increase in temperature affects the statistical predictor-predictand relationship adversely and leads to a systematic underestimation or overestimation of hourly precipitation extremes. The differences between the observed and the estimated annual maximum of hourly precipitation between 1980 and 2009 are investigated as described in Schmith (2008). A negative value stands for an overestimation of the hourly extremes, while positive represents an underestimation. Table 6.5 shows the linear regression slopes of the differences for each station. For M1 and M2, no significant trends can be found for any of the stations. In other words, the downscaling errors of the annual maximum of hourly precipitation using approach M1 and approach M2 exhibit no signs of significant instationarity. In contrast, applying approach M3 results in significant trends in the residual time series of four stations (Marham, Cranwell, Leuchars and Kinloss), while two stations (Leuchars and Kinloss) show significant trends in the residual time series for downscaling results based on approach M4. The reason for the significant trends using approach M3 and approach M4 can be found in Figure 6.8. At Marham and Cranwell, approach M3 heavily overestimates the annual maxima of hourly precipitation between 2000 and 2009. At Leuchars, the annual maxima tend to be underestimated between 2000 and 2009 by approach M3 and approach M4. At Kinloss, approach M3 and approach M4 tend to heavily underestimate the most extreme hourly precipitation between 1980 and 1989 and heavily overestimate the most extreme hourly precipitation between 2000 and 2009. It shows that using daily precipitation as a final predictor to find the analogue day, which is done for approach M1 and approach M2, reduces not only the risk of heavily overestimating the most extreme hourly precipitation events as mentioned earlier, but also decreases the risk of instationarities in the predictor-predictand relationship.

One limitation of the analogue day method is well illustrated in Figure 6.9. It shows all the precipitation events at Northolt for the summer time period 1980-1999 conditioned on CP3, which represents high pressure conditions in the southern part of the UK. Some studies (Beck and Bárdossy, 2013; Blenkinsop et al., 2015) suggested that high pressure conditions may be associated with localized convective precipitation extremes. Those convective extremes exhibit high return periods (Chan, 2013) and are not always included in the calibration period of the statistical downscaling method. For example, it can be seen in Figure 6.9 that over the calibration period 1980-1999 in summer at Northolt, no hourly precipitation events higher than 10mm/hr occurred on days with prevailing CP3. However, over the validation period 2000-2009, an extreme hourly precipitation event was recorded at Northolt. The problem of the analogue day is that this extreme hourly precipitation event is very unlikely to be reproduced by the statistical downscaling method because no similar precipitation event can be resampled over the calibration period for days of CP3.

		CAMBORNE	BOSCOMBE DOWN	NORTHOLT	MARHAM	CRANWELL	VALLEY	ESKDALEMUIR	BOULMER	ALDERSGROVE	LEUCHARS	TIREE	KINLOSS	MEAN
M1	V1	6.6	9.9	22.0	18.3	8.5	11.1	13.7	14.4	7.4	28.0	13.4	10.3	
	V2	6.0	7.8	10.7	8.2	12.3	12.3	8.9	9.6	10.8	7.3	20.8	11.2	
	V3	7.8	9.6	12.7	4.8	10.0	8.8	6.5	8.7	7.2	19.9	13.2	6.1	
	MEAN	6.8	9.1	15.1	10.4	10.3	10.7	9.7	10.9	8.5	18.4	15.8	9.2	11.2
M2	V1	9.6	7.3	20.5	22.7	16.2	10.8	12.4	10.4	12.1	11.5	8.4	9.4	
	V2	5.4	13.6	10.7	8.4	16.2	8.4	7.6	10.8	18.8	9.8	11.9	12.9	
	V3	9.9	9.2	13.1	13.3	7.4	21.2	6.4	6.7	10.3	21.0	8.8	11.7	
	MEAN	8.3	10.0	14.8	14.8	13.3	13.5	8.8	9.3	13.7	14.1	9.7	11.3	11.8
M3	V1	10.5	20.9	13.1	11.0	5.5	12.5	9.8	10.8	7.7	26.4	11.3	10.6	
	V2	6.6	9.6	11.2	10.8	17.9	9.6	11.3	23.1	9.1	7.8	13.8	10.4	
	V3	8.4	9.2	21.2	11.4	22.0	15.0	14.1	10.8	7.0	17.7	8.8	6.7	
	MEAN	8.5	13.2	15.2	11.1	15.1	12.4	11.7	14.9	7.9	17.3	11.3	9.2	12.3
M4	V1	6.9	42.6	16.8	12.2	11.1	16.9	17.4	12.9	10.4	15.3	7.3	12.5	
	V2	6.0	11.6	8.7	10.6	20.0	16.2	7.5	10.9	15.7	10.2	17.8	9.9	
	V3	6.3	7.6	24.2	8.2	7.0	9.0	7.0	8.5	12.4	25.6	7.8	10.4	
	MEAN	6.4	20.6	16.6	10.3	12.7	14.0	10.6	10.8	12.8	17.0	11.0	10.9	12.8

Table 6.3 The averaged 99.5th percentile error $\overline{PE}_{99.5}$ in [%] depending on the four different approaches to find the analogue days over the three different validation periods. The results are given as an average over 100 bootstrapping samples. Observed data is used as the predictor for statistical downscaling. M1, M2, M3 and M4 represent the different approaches to find the analogue day. V1, V2 and V3 represent the three different validation periods: 1980-1989 (V1), 1990-1999 (V2) and 2000-2009 (V3).

	CAMBORNE (1.1°C)	BOSCOMBE DOWN (1.1°C)	NORTHOLT (1.3°C)	MARHAM (1.2°C)	CRANWELL (1.3°C)	VAELLEY (1.0°C)	ESKDALEMUIR (1.0°C)	BOULMER (1.1°C)	ALDERSGROVE (0.9°C)	LEUCHARS (0.9°C)	TIREE (0.8°C)	KINLOSS (0.8°C)	MEAN
M1	6.8	15.6	10.5	4.8	7.7	8.2	6.7	8.0	11.9	7.7	7.6	11.1	8.8
M2	8.1	17.5	8.6	7.9	6.4	7.4	6.0	10.2	21.5	9.1	7.1	19.3	10.8
M3	9.5	11.3	15.7	7.0	8.9	9.3	18.3	8.7	10.8	6.9	6.2	9.7	10.2
M4	8.1	20.5	15.1	7.4	6.9	7.3	5.7	12.7	10.9	7.2	5.5	16.6	10.3

Table 6.4 Same as Table 6.3, but over the ten year period of warmest summers for each station between 1980 and 2009. The mean temperature difference between the ten warmest summers and 20 coldest summers is given in parenthesis for each station after the station name.

	CAMBORNE	BOSCOMBE DOWN	NORTHOLT	MARHAM	CRANWELL	VALLEY	ESKDALEMUIR	BOULMER	ALDERSGROVE	LEUCHARS	TIREE	KINLOSS
M1	-0.011	-0.019	-0.107	-0.085	-0.128	-0.048	-0.1	0.029	-0.038	0.171	0.068	-0.23
M2	0.02	-0.019	-0.033	-0.104	0.005	0.027	-0.09	-0.016	-0.024	0.063	0.114	-0.281
M3	0.098	0.032	-0.128	-0.388	-0.393	0.043	-0.186	0.083	-0.004	0.196	0.096	-0.538
M4	-0.019	0.081	-0.075	-0.098	-0.13	0.027	-0.196	0.17	-0.077	0.216	-0.008	-0.518

Table 6.5 The linear regression slopes in [mm/year] (using ordinary least squares fit) of the residual time series of the annual maxima for the twelve stations between 1980 and 2009. In bold the regression slopes which are significant at the 5% level according to the p-value test.

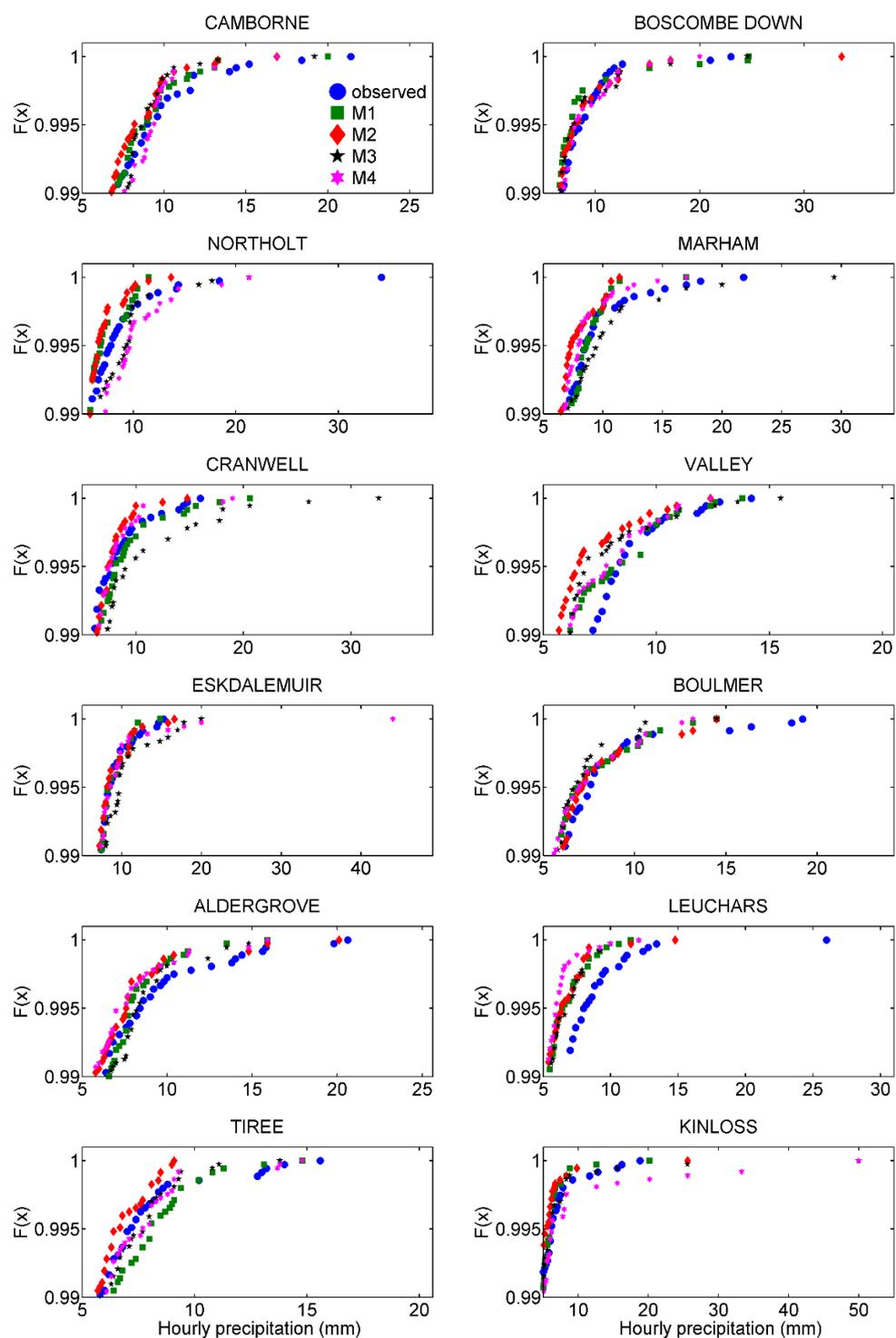


Figure 6.6 The part exceeding 0.99 of the empirical cumulative distribution of the estimated and observed daily maximum hourly precipitation records for the most recent validation period 2000-2009. M1, M2, M3, M4 represent the different approaches to find the analogue day. For all the estimates, a perturbation factor is applied based on the duration and temperature relationship.

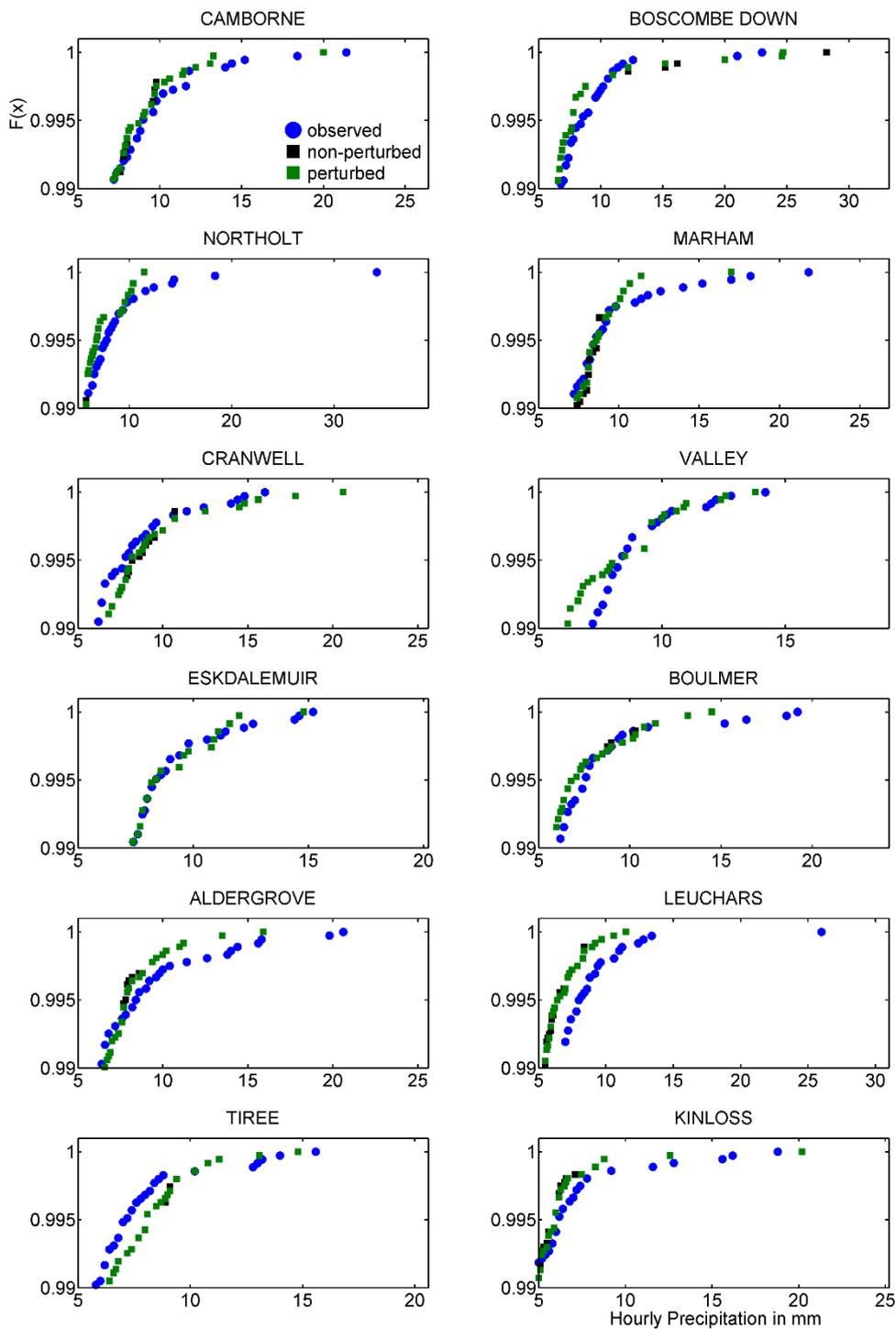


Figure 6.7 The empirical cumulative distribution exceeding 0.99 of the estimated and observed daily maximum hourly extreme precipitation records for the most recent validation period 2000-2009. Estimates are based on approach M1, which showed the best performances on average. Non-perturbed represents records as observed on the analogue days; perturbed represents estimates for which a perturbation factor is applied based on the duration and temperature relationship.

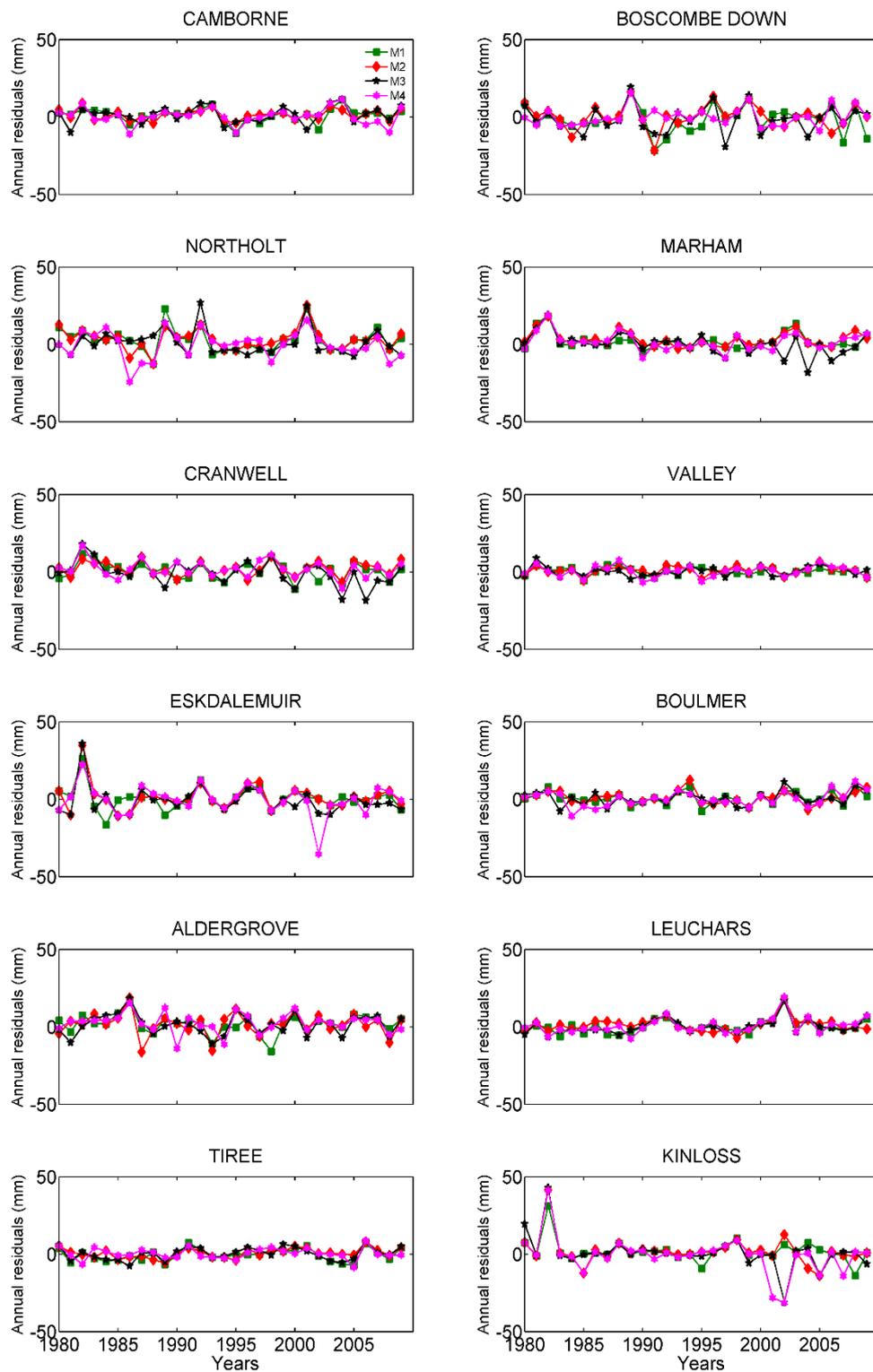


Figure 6.8 The residual time series of the annual maxima of hourly precipitation for the twelve stations between 1980 and 2009 using the four different statistical downscaling methods. M1, M2, M3, M4 represent the different approaches to find the analogue day.

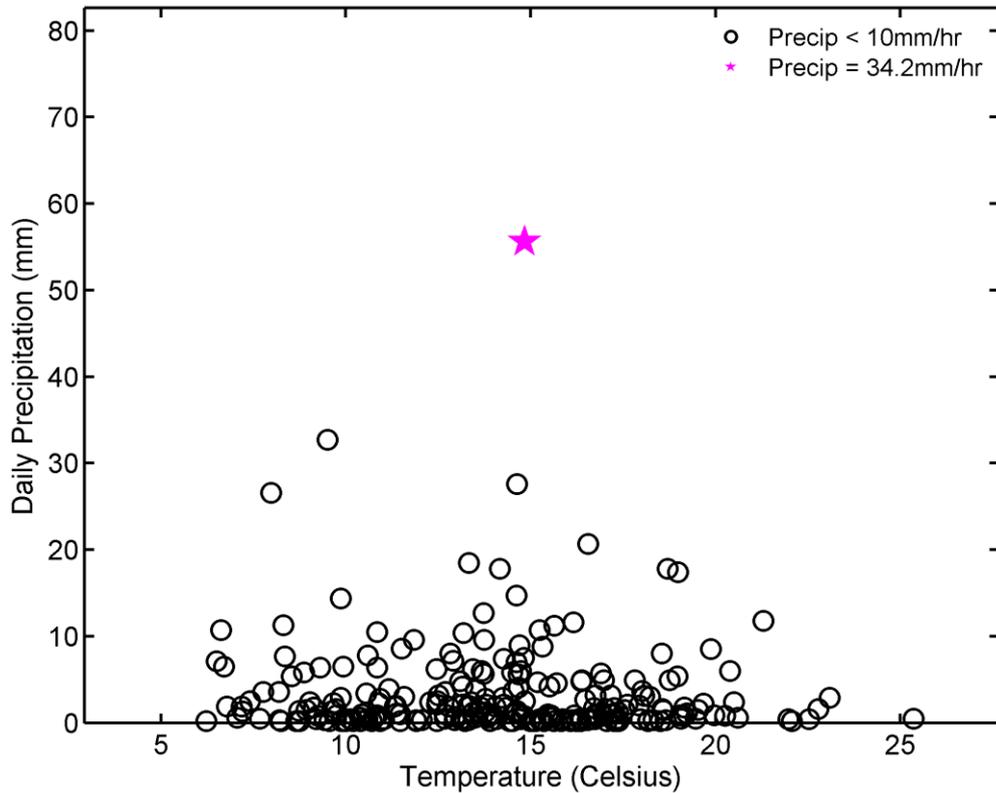


Figure 6.9 Scatter plot of daily precipitation events (≥ 0.2 mm/day) recorded in summer at Northolt between 1980 and 1999 (calibration period C1) conditioned on CP3. It shows the E-OBS gridded daily mean temperature (x-axis) and local daily precipitation (y-axis). The magenta star represents an extreme hourly precipitation event (34.2 mm/hr) on the 9th August 2001 during the validation period V1 (2000–2009).

6.4 Summary

In this chapter, hourly precipitation records are estimated based on the statistical downscaling method described in Section 4.3 using observed daily data. As predictors, daily mean temperature, daily precipitation and the five (plus one unclassified) CPs (see Section 5.3), are used. The importance of the daily temperature and the daily precipitation variable is illustrated in Section 6.1. It can be seen that high daily temperature and/or high daily precipitation often correspond with extreme hourly precipitation.

In the following section, the relationship between precipitation duration and daily temperature conditioned on CPs for each of the twelve stations is investigated. It is found that the duration-temperature relationship in summer is more likely to be statistically significant and to fulfil the Root-Mean-Square-Errors criterion ($RMSE < 1$) for days associated with low pressure systems (CP1 and CP4) over the UK. The relationships in summer conditioned on CP4 are significant and fulfil the RMSE criterion for all twelve stations, while CP1-conditioned

relationships in summer show a tendency to be significant and to fulfil the RMSE criterion for the stations in the southern part of the UK only. It also needs to be noted that statistical relationships are harder to be identified for smaller sample sizes. It means that the probability of the relationships to be significant increases with the size of the record sample it is based on. For example, 35 of the 36 relationships in summer not conditioned on any CP are statistically significant for the twelve stations and over three calibration periods. This is important as the unconditional relationships are used as substitutes in the statistical downscaling process for those CP-conditioned relationships that are either not significant or do not fulfil the RMSE criterion.

The relationships between precipitation duration and temperature are then used to perturb the historical hourly precipitation records on the analogue day following the statistical downscaling process. Results are shown in Section 6.3 for each station and over three different validation time periods (plus the 10 year period with the warmest summers). Four different approaches of the statistical downscaling method are presented depending on how to find the analogue day in the past. It is shown that approach M1, which only subsamples the analogue days into different CPs and different season (winter/summer) without subdividing them further into calendar months and daily precipitation bins, provides the best results on average for the twelve stations and over the three validation periods. However, performances of the four different approaches vary depending on the station. For example, approaches M1 and M2 show better results at Eskdalemuir, which is situated in mountainous area and thus strongly affected by orographic precipitation. In this case, it seems that using daily precipitation as the final predictor only improves the quality of the extreme hourly precipitation estimates. Results based on approach M1 or M2 are also less likely to heavily overestimate the highest percentiles of hourly precipitation. Again this is likely to be due to the fact that M1 and M2 rely stronger on the actual daily precipitation value to find the analogue day compared to M3 and M4. On the other hand, applying approach M1 and M2 lead to high differences between estimated and observed extreme hourly precipitation at Marham and Northolt over validation period V1. Nevertheless, it can be summarized that the statistical downscaling method is able to generate reliable and robust hourly precipitation extremes over different validation time periods at different stations in the UK.

In order to test the performance of the downscaling method under a climate which is warmer than the one it is calibrated on, the four different approaches are also validated over the ten warmest summer period. It is found that the performance of the downscaling method is not negatively affected by the fact that the climates over the validation and calibration period are as different as possible in terms of temperature. In addition, the residual time series between

observed and estimated annual maxima of hourly precipitation are investigated in order to test if the relationships between the combined predictors (CP, daily precipitation and temperature) and predictand (hourly precipitation) remain stationary. For approach M1 and M2, no signs of nonstationarity can be found in the statistical downscaling process, whereas applying approach M3 and M4 results in significant trends in the residual time series for four and two stations respectively.

Chapter 7

Statistical downscaling based on RCM data

Chapter 7 presents the results of the statistical downscaling to estimate local hourly precipitation extremes using 12.5km RCM data. At first, hourly precipitation is estimated based on ERA-interim driven RCM data. The fact that the RCMs are driven under quasi-observed boundary conditions means that the predicted hourly precipitation values can be directly related with the observed records on a day to day basis. The results are compared with the actual observed hourly precipitation extremes in order to find the best performing statistical downscaling approach to find the analogue day (amongst M1, M2, M3 and M4). In Section 7.2 and Section 7.3, the best performing statistical downscaling approach is then used to estimate hourly precipitation extremes over a reference period (1980-2005) and two future periods (2030-2055 and 2075-2100) based on the GCM-driven RCM data.

7.1 Cross-validation to assess statistical downscaling performance

Similar to Section 6.3, the estimated hourly precipitation events using the analogue day method as discussed in Chapter 4 are validated for the same twelve selected stations over the same three different validation periods. Also, the same four approaches (M1, M2, M3 and M4) to find the analogue day are used (see Table 4.1). In contrast to Section 6.3, the predictors are derived from ERA-interim driven RCMs in this section. In order to find the analogue days, daily temperature, precipitation and CPs derived from ERA-interim driven RCMs are compared for each day of the validation period with the observed daily temperature, precipitation and CPs over the calibration period. In a next step, the historical hourly precipitation records on the analogue days are perturbed according to the temperature difference between each given day and the respective analogue day as described in Chapter 4. As the RCM temperature variable may be biased, the RCM simulated temperature value is normalized and the equivalent observed temperature value to the normalized RCM simulated temperature is used to calculate the temperature difference.

The performances of the statistical downscaling method are again validated based on the averaged 99.5th percentile error ($\overline{PE}_{99.5}$) over 100 bootstrapping samples. Table 7.1 and Table 7.2 show the $\overline{PE}_{99.5}$ between observed and estimated hourly precipitation based on the RCA4 and the RACMO22E RCM at 12.5 km resolution. Replacing observed data with the RCM data as

predictors for the downscaling process has a negative effect on the performances in terms of extreme hourly precipitation. Nevertheless, the estimates based on the RCM data still show good agreement with the observed extreme hourly precipitation. It can be seen that using the statistical downscaling approach M1 results in the best performances on average for the two RCMs, the three validation periods and the twelve selected stations. M1 also provides the best results on average over the ten warmest summers (see Table 7.3 and Table 7.4). Similar results are found in Chapter 6, where hourly precipitation are estimated based on observed predictors instead of using RCM simulated predictors. M1 stands for the analogue day method which subsamples the precipitation records only into two different seasons, whereas M2 and M4 subsample the records into twelve calendar months and M3 and M4 also applies six daily precipitation categories. In order to find the most similar day in the past, daily precipitation and temperature are used as equally important predictors in M1. It can be concluded, that additional subsampling of the calibration days into calendar months or daily precipitation categories does not improve the performance of the statistical downscaling method.

Regarding the robustness of the 99.5th percentiles hourly precipitation estimates over different validation periods, it is found that RCM estimates show high agreement with observed precipitation extremes at most stations (Table 7.1 and 7.2). However, at Leuchars relatively high differences occur over V2 for all four approaches using RCA4 RCM data. Large differences between estimates based on RCA4 RCM data and observed hourly precipitation extremes are also found at Kinloss over V1 (for M1, M2 and M3), at Northolt over V2 (for M1 and M3) and over V3 (for M2), at Boscombe Down over V1 (for M3), Cranwell over V3 (M3) and Aldersgrove over V2 (M4). Considering the estimates based on RACMO22E RCM data, large differences between the estimated and observed data in terms of the 99.5th percentiles occur at Leuchars over V1 based on approach M1 and over V2 (M2), at Kinloss over V1 (M1), at Northolt over V3 (M2) and at Boscombe Down over V1 (M3).

It also needs to be noted that the RCM estimates based on the M1 or M2 approach exhibit an increased risk of overestimating the most extreme hourly precipitation events (see Figure 7.1 and Figure 7.2) compared to the estimates based on M1 and M2 predicted by observed variables as illustrated in Figure 6.6. For example, estimates predicted with RCA4 RCM data show overestimation of the most extreme precipitation events at Northolt (for M2), Eskdalemuir (M1 and M2) and Kinloss (M2). The downscaling results using RACMO22E RCM data overestimate the most extreme hourly precipitation also at Eskdalemuir (M1) and Kinloss (M1). The overestimation can be explained by the fact that daily precipitation simulated by ERA-interim RCMs cannot be expected to exactly reproduce the observed day to day sequences of precipitation. This is due to the large size of the RCM domain and also due to the

fact that ERA-interim data is only fed into the RCM at the lateral and sea surface boundaries (Kendon et al., 2012). As a result, the RCM estimates based on M1 and M2 show a similar risk of overestimating the highest precipitation values as the estimates based on M3 and M4, which rely only on temperature as the final predictor.

The influence of the perturbation factor derived from the duration-temperature relationship is relatively small on the estimates based on ERA-interim driven RCMs (not shown). This is similar to the estimates based on the observed predictors shown in Section 6.3.

In order to test the assumption that the statistical downscaling method remains valid under a changing climate, the averaged 99.5th percentile error ($\overline{PE}_{99.5}$) is shown for the ten warmest summers (based on observed data) of each station using RCA4 RCM (see Table 7.3) and RACMO22E RCM data (see Table 7.4) calibrated on the 20 coldest summers (based on observed data) between 1980 and 2009. Very high differences between estimates and observed records are found at Kinloss using the approach M3 for both RCMs. Overall, the estimates show good agreement with observed hourly precipitation extremes over the ten warmest summers. It needs to be noted that the mean temperature increases between the ten years of warm summers and 20 years of cold summers range only from 0.8°C to 1.3°C depending on the stations. In contrast, projections of changes in daily mean temperature can be as high as 4.0°C for the time period 2075-2100 under RCP8.5 (see Table 7.12). Therefore, the ability of the statistical downscaling to perform under a future warmer climate can only be tested to a certain extent based on ERA-interim driven RCM data.

In a second test regarding the stationarity assumption of the statistical downscaling method, it is tested if the temperature increase between 1980 and 2009 has a negative on the predictor-predictand relationship. For this purpose, the residuals between the observed and the estimated annual maximum of hourly precipitation are analysed over time. Estimates based on RCA4 RCM data exhibit signs of significant instationarities in the residuals at Cranwell (using M3), Eskdameluir (M1), Boulmer (M1), Aldergrove (M2) and Kinloss (M2 and M3) as illustrated in Table 7.5. In comparison, estimates based on RACMO22E RCM data are affected by instationarities at Northolt (using M2), Marham (M1), Boulmer (M4) and Leuchars (M2) as shown in Table 7.6. In contrast to the results based on observed predictors, where significant trends only occur for estimates based on M3 and M4 as shown in Section 6.3, significant trends in the residuals are found for all approaches M1 to M4 based on RCM predictors for at least one station. The instationarities in the residual time series may be partly caused by the overestimation of annual maxima at certain years. As mentioned above, daily precipitation simulated by ERA-interim RCMs cannot be expected to exactly reproduce observed

precipitation day by day due to the large size of the RCM domain and due to the fact that ERA-interim data are only fed into the RCM at the lateral and sea surface boundaries (Kendon et al., 2012). As a result, very extreme hourly precipitation estimates are not restricted within the statistical downscaling process to the occurrence of very extreme RCM daily precipitation only. RCM estimates thus derived from M1 and M2 exhibit a similar risk of overestimating the very high precipitation values as estimates based on M3 and M4, which rely only on temperature as the final predictor. For example, applying approach M2 based on RCA4 RCM data results in the overestimation of annual maxima at Kinloss in 2002 and 2009 as indicated by high negative residuals for those years (see Figure 7.3). Similarly, estimates based on RACMO22E RCM data (using M2) overestimate observed annual maxima at Northolt in 2000, 2003, 2004 and 2008 as illustrated in Figure 7.4.

Biases in the three RCM predictors (daily precipitation, temperature and CP frequencies) have also been assessed in terms of statistically significant instationarities between 1980 and 2009 (not shown). For both RCMs, significant positive trends in the biases of daily temperature over time are found. It means, that the RCMs underestimate the increase in observed temperature between 1980 and 2009 (see Section 3.2.2). However, no direct relationship between any significant instationarities in the residual time series of hourly precipitation extremes and in the biases of the three different RCM predictors could be found.

		CAMBORNE	BOSCOMBE DOWN	NORTHOLT	MARHAM	CRANWELL	VALLEY	ESKDALEMUIR	BOULMER	ALDERSGROVE	LEUCHARS	TREE	KINLOSS	MEAN
M1	V1	16.4	21.9	14.6	19.1	4.4	11.0	13.8	22.7	18.3	9.8	7.3	41.2	
	V2	13.3	6.6	44.7	12.1	6.4	7.8	9.6	24.9	12.2	46.4	7.8	9.3	
	V3	10.4	26.3	17.4	9.9	16.1	17.7	34.1	6.1	12.5	15.7	7.7	14.5	
	MEAN	13.4	18.3	26.6	13.7	9.0	12.2	19.2	17.9	14.3	24.0	7.6	21.7	16.4
M2	V1	12.4	35.6	9.1	12.2	7.0	15.3	11.0	11.1	36.6	10.6	8.0	40.3	
	V2	25.9	9.5	14.1	14.5	11.2	24.5	12.4	10.4	19.7	49.2	9.0	15.1	
	V3	6.0	13.7	52.9	9.2	17.5	10.5	13.6	10.5	10.6	21.0	11.0	12.5	
	MEAN	14.8	19.6	25.4	12.0	11.9	16.8	12.3	10.7	22.3	26.9	9.3	22.6	17.1
M3	V1	18.0	42.9	22.6	12.0	5.4	19.9	13.1	9.1	20.4	38.8	8.7	11.5	
	V2	11.4	14.4	51.2	19.1	12.0	21.3	14.8	38.2	10.9	48.1	8.6	9.1	
	V3	12.9	17.7	32.6	29.7	51.8	8.5	25.1	13.7	18.1	10.8	20.3	45.9	
	MEAN	14.1	25.0	35.5	20.3	23.1	16.6	17.7	20.3	16.5	32.6	12.5	22.2	21.4
M4	V1	9.7	31.6	23.9	9.2	15.0	10.3	10.8	9.1	28.4	32.8	6.1	18.3	
	V2	20.2	10.9	16.6	17.4	20.6	30.5	17.4	24.4	54.8	51.3	9.7	15.1	
	V3	5.7	12.4	31.7	8.4	16.9	9.9	31.9	7.0	13.6	19.4	14.9	28.4	
	MEAN	11.9	18.3	24.1	11.7	17.5	16.9	20.0	13.5	32.3	34.5	10.2	20.6	19.3

Table 7.1 The averaged 99.5th percentile error $\overline{PE}_{99.5}$ in [%] depending on the four different approaches to find the analogue days over the three different validation periods. The results are given as an average over 100 bootstrapping samples. RCA4 RCM data is used as the predictors for the statistical downscaling. M1, M2, M3 and M4 represent the different approaches to find the analogue day. V1, V2 and V3 represent the three different validation periods: 1980-1989 (V1), 1990-1999 (V2) and 2000-2009 (V3).

		CAMBORNE	BOSCOMBE DOWN	NORTHOLT	MARHAM	CRANWELL	VALLEY	ESKDALEMUIR	BOULMER	ALDERSGROVE	LEUCHARS	TIREE	KINLOSS	MEAN
M1	V1	11.5	29.6	9.4	18.5	12.0	10.4	8.9	8.7	16.0	44.6	8.0	59.9	
	V2	10.5	16.7	21.0	11.1	7.9	26.4	11.4	17.0	11.6	32.0	16.4	7.0	
	V3	7.2	10.5	10.2	9.2	11.1	16.9	8.7	17.9	14.2	21.6	23.6	13.6	
	MEAN	9.7	18.9	13.5	12.9	10.3	17.9	9.7	14.5	13.9	32.7	16.0	26.8	16.4
M2	V1	12.5	32.0	9.8	13.1	11.7	12.9	8.1	8.4	21.9	31.2	6.9	34.4	
	V2	24.0	9.9	10.8	11.3	9.3	32.3	17.6	25.8	8.3	48.7	11.1	5.7	
	V3	6.6	7.8	41.2	9.0	10.3	16.0	7.2	15.2	11.3	24.0	14.2	15.5	
	MEAN	14.4	16.6	20.6	11.1	10.4	20.4	11.0	16.5	13.8	34.6	10.7	18.5	16.6
M3	V1	8.6	40.6	11.3	12.3	11.7	14.0	7.1	8.7	17.9	38.8	10.4	8.8	
	V2	21.6	8.2	9.8	13.0	7.3	19.5	8.0	32.2	10.1	39.3	9.2	9.2	
	V3	15.6	12.2	17.6	17.4	7.6	12.5	15.6	14.0	10.0	12.0	31.0	23.1	
	MEAN	15.3	20.3	12.9	14.2	8.9	15.3	10.2	18.3	12.7	30.0	16.9	13.7	15.7
M4	V1	7.5	29.1	8.9	11.1	16.5	11.0	8.0	13.3	21.7	22.7	17.4	9.8	
	V2	23.4	8.3	10.2	21.6	11.4	8.7	19.7	24.5	9.2	32.5	7.9	6.5	
	V3	12.4	9.4	17.2	16.7	6.9	14.5	11.6	7.9	10.6	22.2	18.8	22.5	
	MEAN	14.4	15.6	12.1	16.5	11.6	11.4	13.1	15.2	13.8	25.8	14.7	12.9	14.8

Table 7.2 Same as Table 7.1 but RACMO22E RCM data is used as the predictors for the statistical downscaling instead of RCA4 RCM data.

	CAMBORNE (1.1°C)	BOSCOMBE DOWN (1.1°C)	NORTHOLT (1.3°C)	MARHAM (1.2°C)	CRANWELL (1.3°C)	VALLEY (1.0°C)	ESKDALEMUIR (1.0°C)	BOULMER (1.1°C)	ALDERSGROVE (0.9°C)	LEUCHARS (0.9°C)	TIREE (0.8°C)	KINLOSS (0.8°C)	MEAN
M1	8.0	9.2	11.5	10.1	17.8	18.1	27.6	8.8	14.0	18.2	9.5	18.1	14.2
M2	10.2	10.0	15.6	8.4	22.1	27.4	30.7	16.0	10.4	20.4	8.9	17.4	16.5
M3	15.2	14.8	24.3	16.0	24.4	10.2	24.4	14.9	13.5	18.6	7.5	75.5	21.6
M4	17.1	10.4	20.4	7.1	29.7	23.8	28.0	29.0	13.9	35.8	16.2	9.3	20.1

Table 7.3 Same as Table 7.1 but over the ten year period of warmest summers on average for each station between 1980 and 2009 using RCA4 RCM data as the predictors for the statistical downscaling. The mean temperature difference between the ten warmest summers and 20 coldest summers is given in parenthesis for each station after the station name.

	CAMBORNE (1.1°C)	BOSCOMBE DOWN (1.1°C)	NORTHOLT (1.3°C)	MARHAM (1.2°C)	CRANWELL (1.3°C)	VALLEY (1.0°C)	ESKDALEMUIR (1.0°C)	BOULMER (1.1°C)	ALDERSGROVE (0.9°C)	LEUCHARS (0.9°C)	TIREE (0.8°C)	KINLOSS (0.8°C)	MEAN
M1	9.5	19.0	18.0	10.9	13.2	7.1	11.7	11.3	13.6	31.4	17.8	15.6	14.9
M2	6.8	8.2	11.1	14.7	14.6	26.2	4.8	8.7	10.4	29.8	17.4	10.0	13.6
M3	14.8	11.7	10.5	13.5	8.0	5.3	15.1	28.3	12.4	13.2	26.3	64.7	18.7
M4	12.5	13.8	11.1	10.3	7.1	45.4	10.8	32.6	17.6	18.8	14.4	13.8	17.4

Table 7.4 Same as Table 7.1 but over the ten year period of warmest summer on average for each station between 1980 and 2009 using RACMO22E RCM data as the predictors for the statistical downscaling.

	CAMBORNE	BOSCOMBE DOWN	NORTHOLT	MARHAM	CRANWELL	VALLEY	ESKDALEMUIR	BOULMER	ALDERSGROVE	LEUCHARS	TIREE	KINLOSS
M1	0.161	0.226	0.093	0.025	-0.164	0.003	-0.523	0.253	0.042	0.115	0.058	-0.274
M2	0.104	0.136	-0.603	-0.092	-0.270	-0.093	-0.434	0.023	0.252	0.128	-0.062	-0.975
M3	0.018	-0.037	-0.160	-0.286	-0.365	0.016	-0.120	0.040	-0.076	0.114	0.020	-0.916
M4	0.105	-0.070	-0.180	-0.021	-0.170	-0.137	-0.186	0.110	0.034	0.173	0.015	-0.359

Table 7.5 The linear regression slopes in [mm/year] (using ordinary least squares fit) of the residual time series of the annual maxima of hourly precipitation for the twelve stations between 1980 and 2009. RCA4 RCM data is used as the predictors for the statistical downscaling. In bold the regression slopes which are significant at the 5% level according to the p-value test.

	CAMBORNE	BOSCOMBE DOWN	NORTHOLT	MARHAM	CRANWELL	VALLEY	ESKDALEMUIR	BOULMER	ALDERSGROVE	LEUCHARS	TIREE	KINLOSS
M1	0.152	0.055	-0.224	-0.333	-0.057	0.105	-0.489	0.179	-0.124	0.196	0.069	-0.451
M2	0.179	0.199	-0.561	-0.163	0.030	-0.094	-0.132	0.024	0.212	0.224	0.089	-0.223
M3	-0.048	0.121	-0.163	0.109	-0.161	-0.018	-0.190	0.099	-0.041	0.138	0.034	-0.710
M4	0.080	0.098	-0.081	0.020	0.076	-0.008	-0.144	0.233	0.188	0.174	0.049	-0.469

Table 7.6 Same as Table 7.5 but RACMO22E RCM data is used as the predictors for the statistical downscaling.

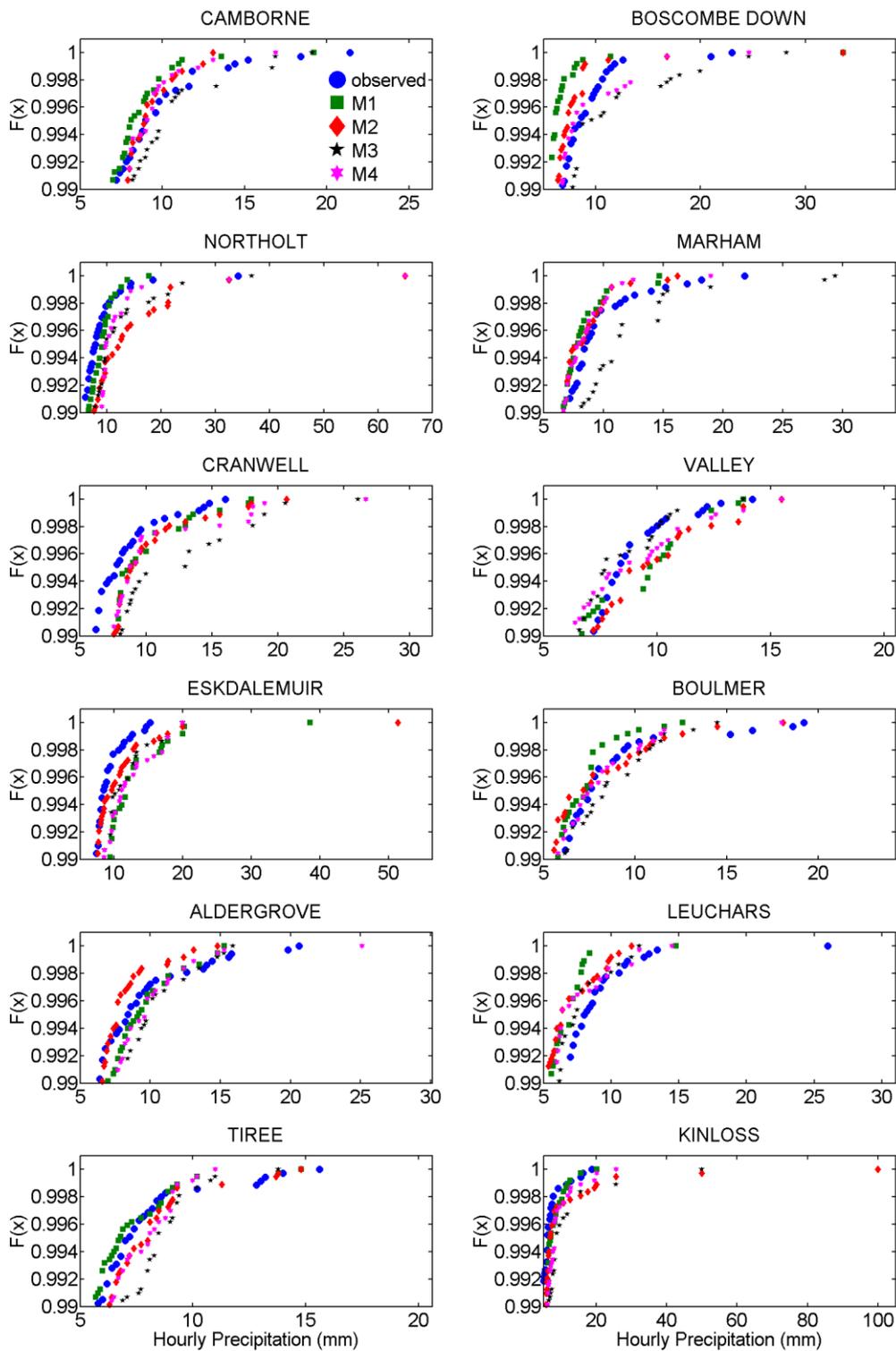


Figure 7.1 The part exceeding 0.99 of the empirical cumulative distribution of the estimated and observed daily maximum hourly precipitation for the most recent validation period 2000-2009. RCA4 RCM data is used as the predictors for the statistical downscaling. M1, M2, M3, M4 represent the different approaches to find the analogue day.

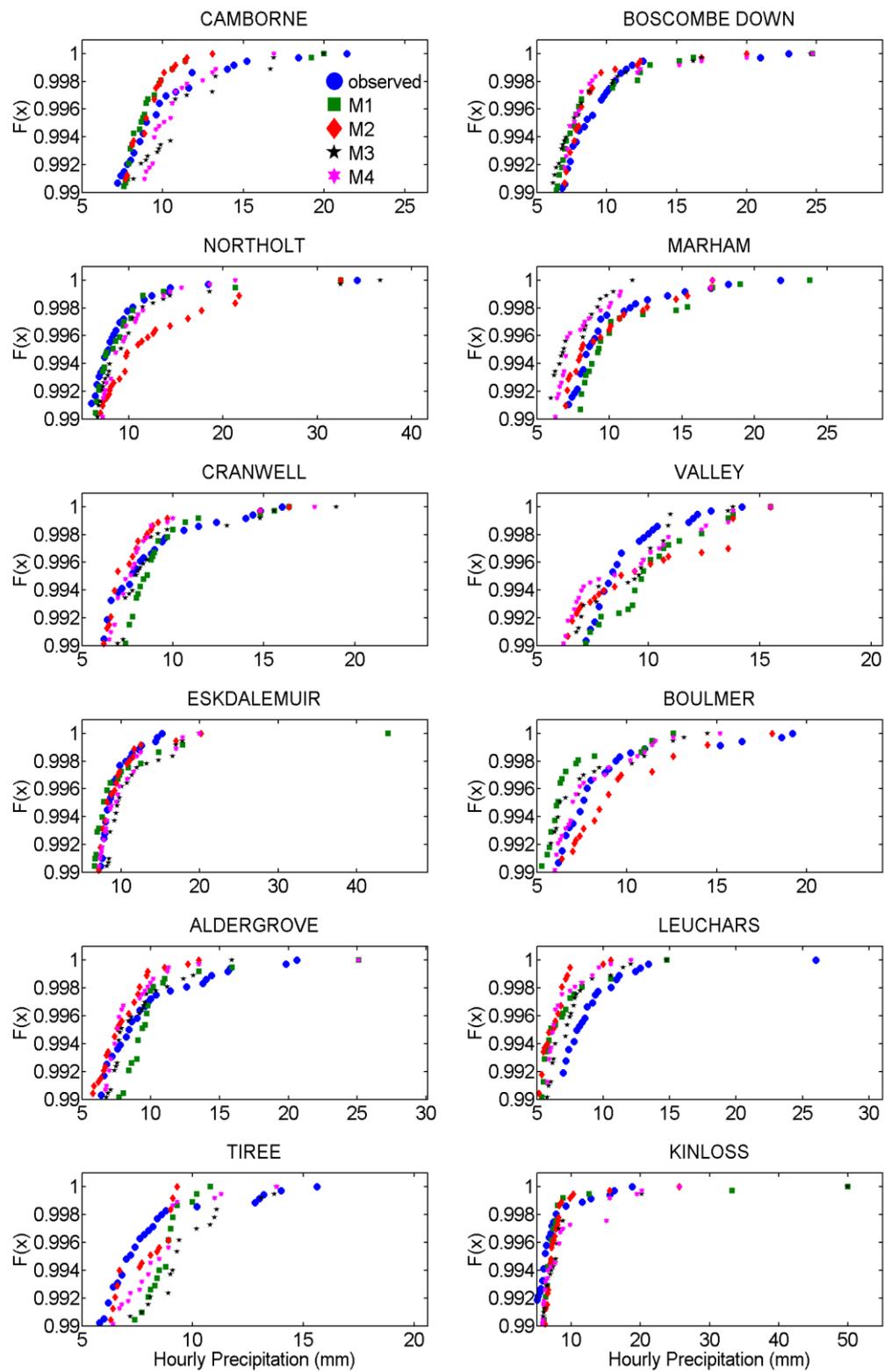


Figure 7.2 Same as Figure 7.1 but RACMO22E RCM data is used as the predictor for the statistical downscaling.

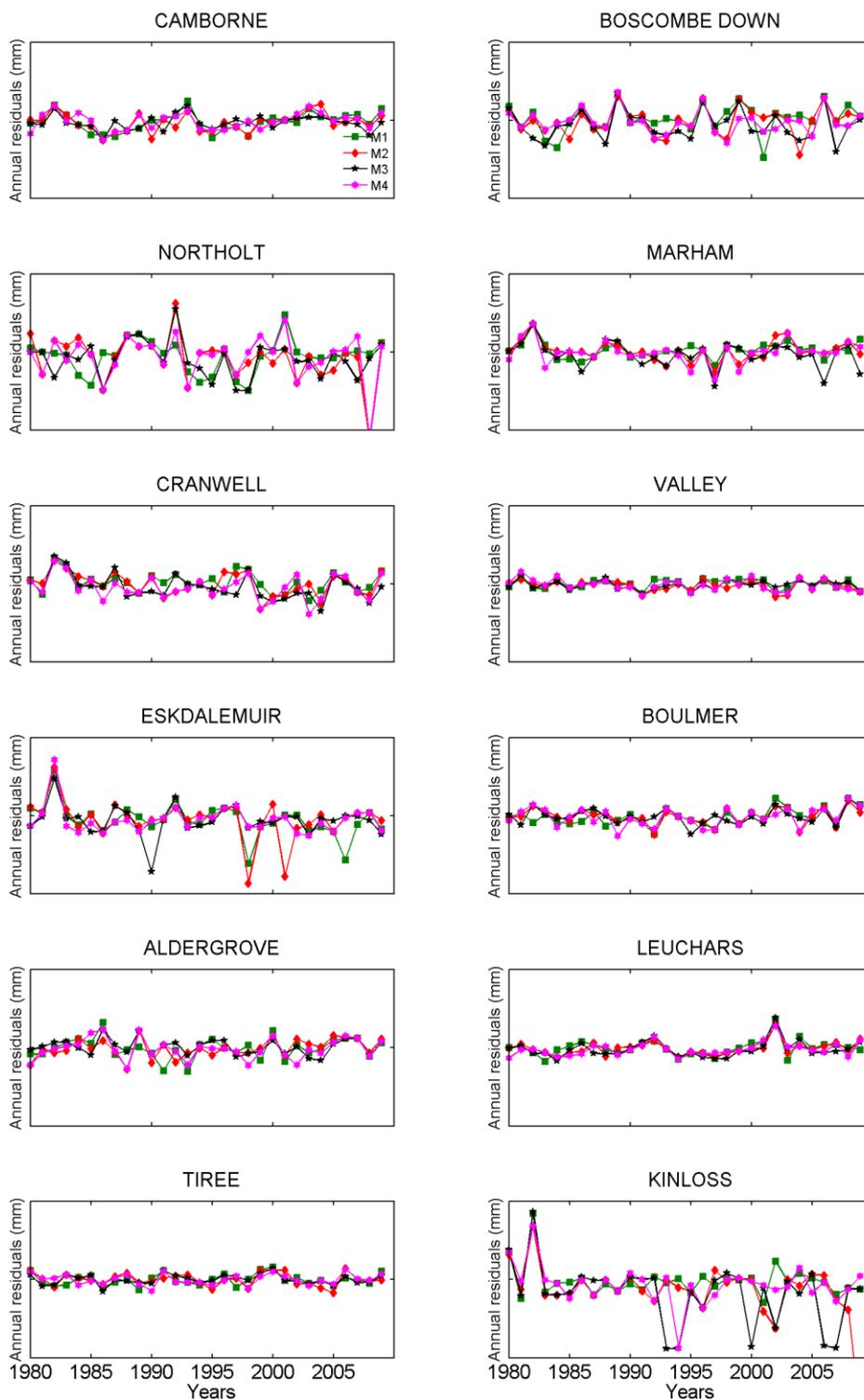


Figure 7.3 The residual time series of the annual maxima of hourly precipitation for the twelve stations between 1980 and 2009 using the four different statistical downscaling methods based on RCA4 RCM data. M1, M2, M3, M4 represent the different approaches to find the analogue day.

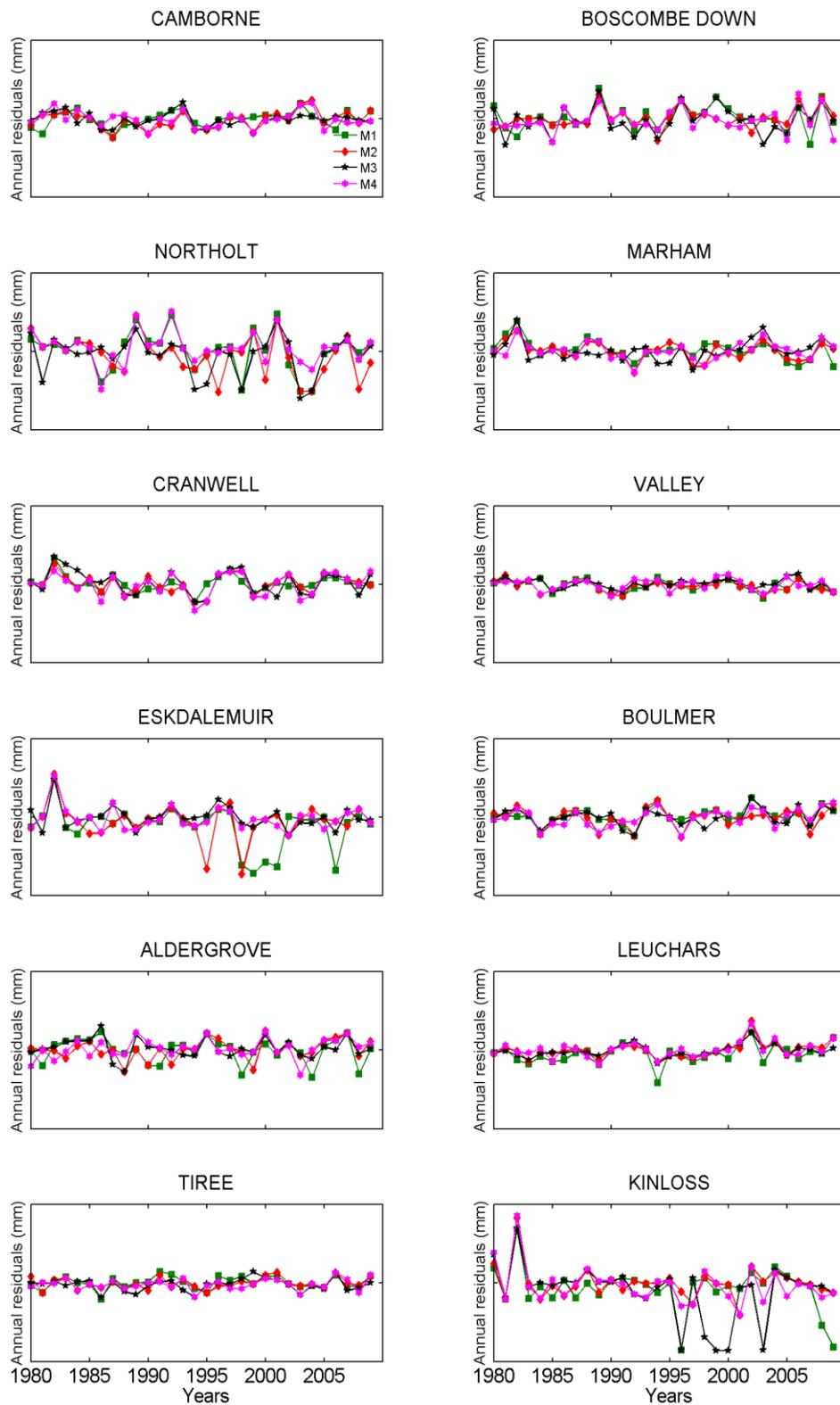


Figure 7.4 Same as Figure 7.3 but RACMO22E RCM data is used instead of RCA4 RCM data.

7.2 Baseline projections based on GCM-driven RCM data

In the following, hourly precipitation extremes are estimated based on GCM-driven RCM data instead of using ERA-interim driven RCM data in Section 7.1. The performances of four different GCM-driven RCMs (CM5A-MR driven RCA4, ESM-LR driven RCA4, EC-Earth driven RCA4 and EC-Earth driven RACMO22E) are compared with each other in terms of their skill to provide reliable extreme hourly precipitation over the reference period 1980-2005. In order to find the analogue days, each day between 1980 and 2005 described by GCM-driven RCM predictors is normalized and compared with all the days over the calibration period 1980-2009 described by normalized observed predictors according to the methodology described in Chapter 4. In this section, the calibration and the reference period are not independent of each other due to the extended length of the reference period. This is to make the length of the reference period equivalent to the length of the two future time periods (2030-2055 and 2075-2100). It means that every hourly precipitation record observed over the reference period can be resampled in the statistical downscaling process. The potential limitation of the statistical downscaling method, that very extreme hourly precipitation only occur over the validation period but cannot be resampled over the calibration period, is not present. In a next step, the precipitation duration and temperature relationship, which is also calibrated over the time period 1980-2009, is applied to perturb each historical hourly precipitation record. As shown in Section 7.1 for estimates derived from the ERA-interim driven RCM data, results based on approach M1 show the best performances. Therefore, in this section, the estimation of hourly precipitation extremes is based only on approach M1 using GCM-driven RCM data from RCA4 and RACMO22E.

Figure 7.5 displays the observed and estimated hourly precipitation values in summer over the reference period 1980-2005 exemplary for the station at Boscombe Down. It can be seen that the estimated hourly precipitation records show very high agreement with the observed hourly precipitation records independently of which RCM driven GCM is used. This holds true for the overall and the extreme precipitation distribution (99th percentiles). Similar results can be found for the other eleven stations in both seasons, summer and winter (not shown). This is reflected in Table 7.7, which shows the averaged 99.5th percentile error $\overline{PE}_{99.5}$ over 100 bootstrapping samples between observed and estimated daily maximum hourly precipitation for the reference period 1980-2005 at the twelve selected stations.

Overall, it needs to be noted that the averaged 99.5th percentile error $\overline{PE}_{99.5}$ based on the GCM-driven RCMs are considerably smaller compared to the averaged 99.5th percentile error $\overline{PE}_{99.5}$ based on ERA-interim driven RCM data (see Table 7.1 and Table 7.2). However,

estimates based on ERA-interim driven RCM data, as shown in the previous section, are validated over a time period which is different to the one it is calibrated on. In other words, the estimates based on the ERA-interim driven RCM data are validated in terms of their ability to generate extreme hourly precipitation events which may have not occurred over the calibration time period. In contrast, the estimates presented in this section based on GCM-driven RCM data are validated over a 26 year reference period, which overlaps with the calibration time period. Therefore, those estimates give an indication of how well the climate predictors are reproduced by the different GCM-driven RCM runs, but do not assess the ability of the statistical downscaling method to perform under a climate different to the one it is calibrated on. It can be concluded that the use of all four GCM-driven RCM data set result in good agreement between estimated and observed hourly precipitation extremes.

In Figure 7.6 and Figure 7.7, the two predictors (daily precipitation and temperature) derived from GCM-driven RCM data and ERA-interim driven RCM data are validated for the station at Boscombe Down in summer. Figure 7.6 shows that the (ERA-interim and GCM-driven) RCA4 RCM has a tendency to overestimate daily precipitation in comparison to the observed distribution. Similar results can be found at most other stations (not shown). Daily precipitation distributions based on ERA-interim driven RCM data are very well reproduced by the GCM-driven RCM runs at all stations.

In terms of the second predictor (daily temperature), Figure 7.7 shows that ERA-interim and GCM-driven RCMs tend to underestimate the daily temperature variable at Boscombe Down. This holds true for most of the other stations in both seasons as well (not shown). It needs to be pointed out that the statistical downscaling process is based on normalized RCM predictor variables, in order to compensate the existing biases between the observed and RCM simulated predictor variables.

Comparing the outputs of the ERA-interim driven RCMs with the GCM-driven RCMs, it is found that the CM5A-MR and EC-EARTH driven RCMs not only underestimate the observed temperature but also the ERA-interim driven RCM simulated temperature at Boscombe Down. Similar results are found for most other stations in both seasons, summer and winter. In theory, the discrepancy in daily temperature between the ERA-interim driven and GCM-driven RCMs could have negative implications on the quality of the estimated hourly precipitation extremes, as the predictors simulated by the GCM-driven RCMs are normalized based on the ERA-interim driven RCMs. However, the good agreement between estimated and observed extreme hourly precipitation (see Table 7.7 and Figure 7.5) shows that the discrepancy in the

temperature variable is not large enough to downgrade the quality of the extreme hourly precipitation estimates considerably.

Considering the large-scale atmospheric variability, it is found that ERA-interim and GCM-driven RCMs are able to reproduce the observed CP frequencies between 1980 and 2005 in summer (Table 7.8) and winter (Table 7.9), as well as the individual pressure patterns of each CP (not shown), realistically. Across all RCMs, CP3 and CP4 are the two CPs with the highest CP frequencies in both seasons and CP1 is least likely to occur.

RCM	GCM	CAMBORNE	BOSCOMBE DOWN	NORTHOLT	MARHAM	CRANWELL	VALLEY	ESKDALEMUIR	BOULMER	ALDERSGROVE	LEUCHARS	TIREE	KINLOSS	MEAN
RCA4	CM5A-MR	4.4	5.2	12.4	4.8	4.5	8.8	9.4	7.2	5.1	8.2	17.1	25.0	9.3
RCA4	ESM-LR	4.9	5.1	8.7	7.8	7.1	7.3	6.1	18.8	15.4	10.4	9.7	51.2	12.7
RCA4	EC-EARTH	3.8	15.3	11.7	5.2	3.9	10.5	6.7	6.8	9.8	6.6	12.2	16.9	9.1
RACMO22E	EC-EARTH	5.5	6.6	19.5	15.8	4.2	6.4	8.7	6.6	12.0	12.7	7.3	14.0	9.9

Table 7.7 The averaged 99.5th percentile error $\overline{PE}_{99.5}$ in [%] over the reference period 1980-2005 depending on the GCM-driven RCM. The results are given as an average over 100 bootstrapping samples. Approach M1 is used to find the analogue day.

RCM	Driven by	CP1	CP2	CP3	CP4	CP5	U
Observed	--	7.8	15.9	29.3	24.8	18.0	4.3
RCA4	ERA-Interim	7.1	15.6	28.5	25.2	18.5	5.1
RCA4	CM5A-MR	7.3	14.7	28.0	27.5	17.3	5.1
RCA4	ESM-LR	6.9	13.1	30.0	25.6	19.0	5.4
RCA4	EC-EARTH	8.6	13.9	29.0	24.9	18.6	5.2
RACMO22E	ERA-Interim	8.0	14.7	29.6	25.2	17.6	4.9
RACMO22E	EC-EARTH	10.0	13.6	27.5	25.7	19.4	3.8

Table 7.8 The observed and simulated frequency occurrences of the five CPs (plus one unclassified) over the reference period 1980-2005 in summer in [%]. U stands for unclassified.

RCM	Driven by	CP1	CP2	CP3	CP4	CP5	U
Observed	--	6.1	14.7	29.8	22.3	20.5	6.6
RCA4	ERA-Interim	6.3	15.6	28.5	24.4	18.8	6.5
RCA4	CM5A-MR	6.5	12.0	28.3	27.7	17.7	7.8
RCA4	ESM-LR	6.6	12.6	28.7	26.8	17.6	7.8
RCA4	EC-EARTH	6.8	14.3	29.3	24.8	18.8	6.0
RACMO22E	ERA-Interim	6.6	14.7	28.4	23.6	19.2	7.5
RACMO22E	EC-EARTH	7.5	13.9	27.9	27.2	18.0	5.6

Table 7.9 Same as Table 7.8 but in winter.

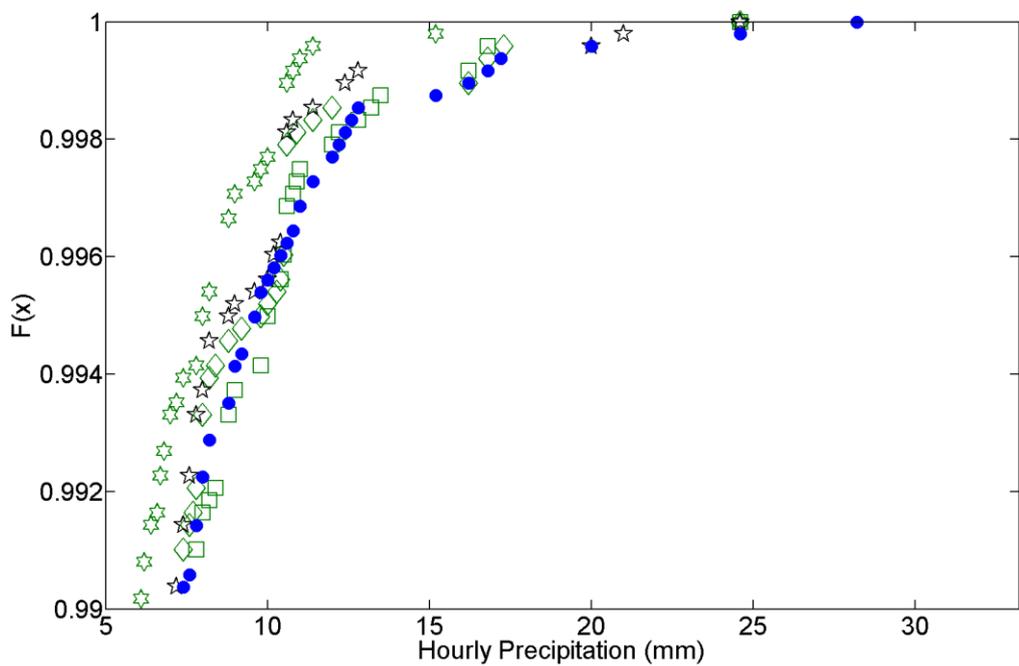
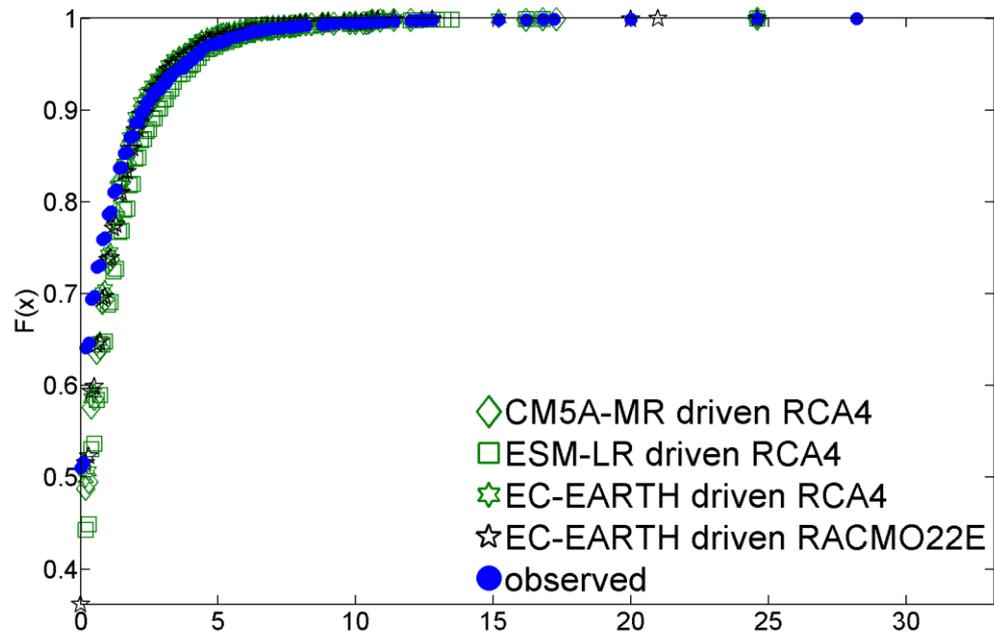


Figure 7.5 Estimated and observed hourly precipitation in summer over the reference period 1980-2005 at Boscombe Down. The statistical downscaling results are shown for the four different GCM-driven RCM runs. Top: The entire distribution of hourly precipitation. Bottom: The part exceeding the 99th percentiles of hourly precipitation.

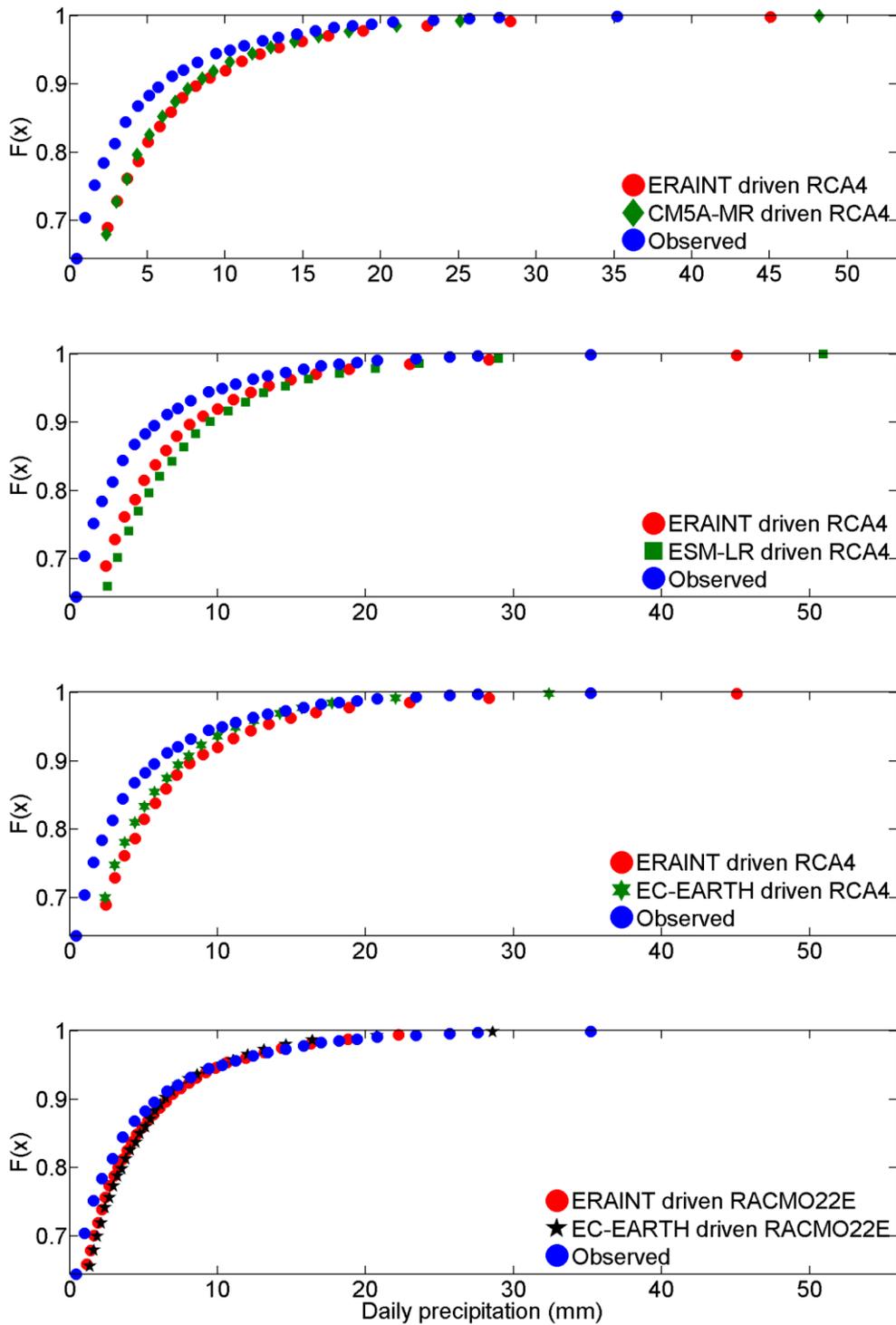


Figure 7.6 Distribution (above 0.6) of the simulated and observed daily mean precipitation in summer over the reference period 1980–2005 at Boscombe Down. The four panels from top to bottom show the four different GCM-driven RCM runs (CM5A-MR driven RCA4, ESM-LR driven RCA4, EC-EARTH driven RCA4 and EC-EARTH driven RACMO22E) and the corresponding ERA-interim (ERAINT) driven RCM runs.

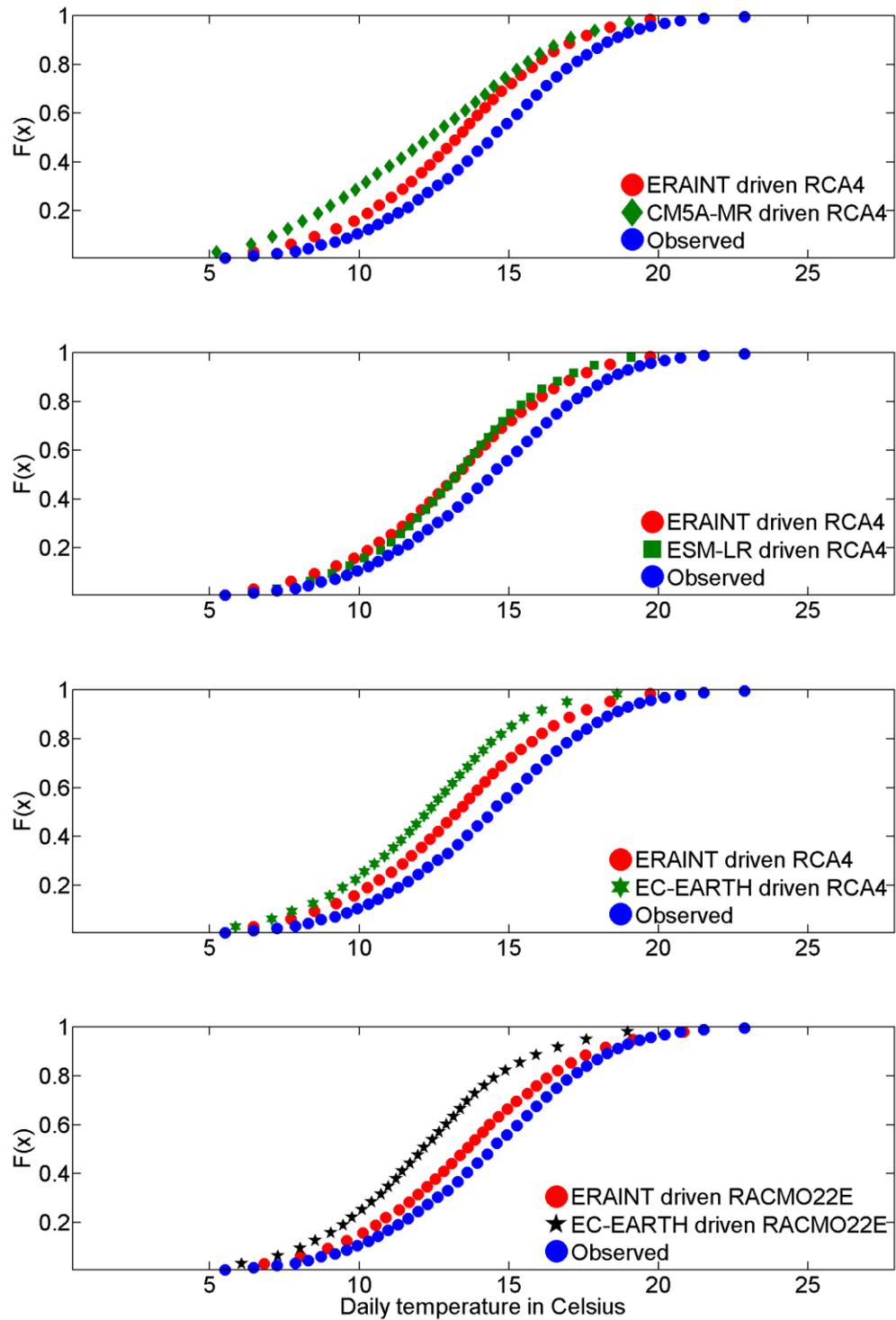


Figure 7.7 Distribution of the simulated and observed daily temperature in summer over the reference period 1980-2005 at Boscombe Down. The four panels from top to bottom show the four different GCM-driven RCM runs (CM5A-MR driven RCA4, ESM-LR driven RCA4, EC-EARTH driven RCA4 and EC-EARTH driven RACMO22E) and the corresponding ERA interim (ERAINT) driven RCM run.

7.3 Future projections based on GCM-driven RCM data

In this section, the estimates of hourly precipitation extremes are presented for two future time periods, 2030-2055 and 2075-2100, based on the GCM-driven RCM data. For the future time period 2075-2100, two different emission scenarios are considered. RCP8.5 corresponds to a high emission scenario resulting in very high temperature increases. RCP4.5 represents a lower emission scenario causing a smaller increase in future temperature compared to the RCP8.5 emission scenario. The distributions of GCM-driven RCM simulated daily temperature in summer are shown exemplary for Boscombe Down in Figure 7.8. The changes in the summer temperature mean over the twelve different stations are given in Table 7.10 – 7.12. It is found that the temperature increases are highest under emission scenario RCP8.5 for the future time period 2075-2100. Considering emission scenario RCP4.5, the increases in temperature are slightly higher for the time period 2075-2100 compared to 2030-2055. The same holds true in winter but the increases in winter temperature are generally smaller (not shown). For the future period 2030-2055, only the emission scenario RCP4.5 is applied because the respective changes in the climate are expected to be similar over the nearer time period between the two emission scenarios (Kirtman et al., 2013).

In terms of daily precipitation, almost all the GCM-driven RCMs simulate future decreases in summer mean precipitation for the six most southern stations, namely Camborne, Boscombe Down, Northolt, Marham, Cranwell and Valley (see Table 7.13 – 7.15). Those decreases are most pronounced for the time period 2075-2100 under emission scenario RCP8.5. An example of the different daily precipitation distributions in summer is shown in Figure 7.9 for the station at Boscombe Down. In winter, almost all the GCM-driven RCMs project increases in daily mean precipitation for all of the twelve stations. The increases in winter are also most pronounced for the future time period 2075-2100 under emission scenario RCP8.5 (not shown).

Considering daily precipitation extremes (represented as the 99.5th percentile), all the GCM-driven RCMs project increases for all of the stations in summer (see Table 7.18) and for almost all the stations in winter (not shown) for the time period 2075-2100 under emission scenario RCP8.5. For 2075-2100 under RCP4.5, the four GCM-driven RCMs project mostly increases in daily precipitation extremes in summer (see Table 7.17) and in winter (not shown) for the twelve stations. Only for 2030-2055 under RCP4.5, the number of projected decreases almost equals the number of increases in daily precipitation extremes in summer (see Table 7.16) and in winter (not shown). The increases are highest in magnitude for 2075-2100 under RCP8.5 in summer and in winter.

In terms of the simulated CP frequencies, it is shown that all the GCM-driven RCMs project a small decrease (up to 3%) in the frequencies of CP4 in summer for the future time period 2075-2100 under both emission scenarios (see Figure 7.10). In terms of the frequencies of the other CPs in summer, no consistent pattern of future changes can be identified. The same holds true in winter (not shown). In general, simulated future changes in CP frequencies are small (below 2%) among all the GCM-driven RCMs. No direct relationship between the changes in CP frequencies and the changes in extreme hourly precipitation can be identified. However, it needs to be noted that other studies suggested future increases in anticyclonic conditions. For example, (Chan et al., 2016) found a shift to more anticyclonic conditions over the UK in the future and argued that this would also lead to a decline in extreme hourly precipitation intensities on days of high temperatures.

Figure 7.11 – 7.14 show the projections of the 99.5th percentile of the estimated daily maximum hourly precipitation for the twelve selected stations over the two different future time periods 2030-2055 and 2075-2100 as well as over the reference time period 1980-2005. The estimates of the 99.5th percentiles are derived from four different GCM-driven RCM runs, two different emission scenarios and 100 different bootstrapping samples for each individual GCM-driven RCM run in order to take into account the different sources of uncertainty in terms of the projections. For the time period 2075-2100 at Eskdalemuir, the overall uncertainty range can be as high as 156% between the smallest projected increase (-30%) of the 100 different bootstrapping samples based on the EC-Earth driven RACMO22E RCM under emission scenario RCP4.5 and the highest projected increase (+126%) of the 100 different bootstrapping samples based on the EC-Earth driven RACMO22E RCM under emission scenario RCP8.5 (see Figure 7.13). In general, increases in hourly precipitation extremes are most pronounced for the future time period 2075-2100 under emission scenario RCP8.5, when the highest increases in temperature also occur. Between 2075 and 2100 under emission scenario RCP8.5, all of the twelve stations show increases in hourly precipitation extremes across all four different GCM-driven RCMs. It is also found that increases in hourly extremes are generally higher in magnitude between 2075 and 2100 under emission scenario RCP8.5 compared to changes for the future time periods 2030-2055 and 2075-2100 under emission scenario RCP4.5 as illustrated in Table 7.19 – Table 7.21.

7.3.1 Projections of extreme hourly precipitation under RCP8.5

Future projected increases in hourly precipitation extremes can be as high as 111.9% (averaged over 100 bootstrapping samples) for the future period 2075-2100 under RCP8.5 (see Table 7.21). As for RCP4.5, the projections of changes vary across the four GCM-driven RCMs.

For example, an average increase in the 99.5th percentile hourly precipitation is estimated to be 36.9% for Kinloss using the CM5A-MR driven RCA4 compared to the maximum increase of 111.9% based on EC-EARTH driven RCA4. The four GCM-driven RCMs simulate more or less similar future responses in daily temperature, precipitation and CP frequencies and therefore the strong discrepancies in future increases of hourly extremes cannot be due to differences in the projection of one particular RCM predictor variable between the four model runs. Instead, it is much more likely that certain types of events, which lead to extreme hourly precipitation, occur more often in certain GCM-driven RCMs than in others. For example, at Kinloss the EC-EARTH driven RCA4 simulate 19 summer days with daily precipitation between 8mm to 15mm, daily temperatures between 14.2°C to 16°C on days associated with CP5. In contrast, only four days with the same predictor characteristics can be found in the CM5A-MR driven RCA4 data set. Those particular events then lead to the estimated extreme hourly precipitation of 16.2mm at Kinloss. This underlines the importance of taking into account multiple GCM-driven RCM runs, in order to estimate the uncertainties induced by the different RCMs in terms of extreme hourly precipitation more accurately.

Future increases in hourly precipitation extremes also exhibit high variations depending on the different stations. For example, at Tiree the average future increase over the four different GCM-driven RCMs (based on the 100 bootstrapping samples) for the future period 2075-2100 under RCP8.5 is only 12.1% compared to an average increase at Kinloss of 75.1%. The other two stations with very high future increases in hourly extremes are Eskdalemuir and Northolt. Hourly extremes at Eskdalemuir increase by an average of 45.8% for the future period 2075-2100 under RCP8.5 over the four GCM-driven RCMs (based on the 100 bootstrapping samples). Eskdalemuir is located in a mountainous area and therefore it is likely that orographic effects play an important role in terms of its extreme hourly precipitation characteristics as discussed in Section 3.1.5. Orographic barriers can also trigger convective precipitation events (Barry and Chorley, 2009). At Northolt, increases in hourly extremes amount to 38.3% on average for the time period 2075-2100 under RCP8.5. Northolt is the station with the highest mean and extreme temperature (see Section 3.2.2). On the other hand, Tiree, the station with the lowest increases in hourly precipitation extremes, is highly exposed to the North Atlantic and exhibits the lowest daily temperature extremes.

It needs to be mentioned that no clear spatial pattern of future changes in extreme hourly precipitation can be identified. This is at least partly due to the restricted number of selected stations in this thesis. As explained in Section 3.1.3, each of the twelve stations represents a different UK extreme precipitation region. But one single station is not sufficient to describe the precipitation characteristics of an entire region particularly in terms of hourly extremes as

those events tend to be very localized. As a consequence, the high increases at Kinloss should not lead to the conclusion that the North Highland and Islands (the region where Kinloss is located) is the UK region for which the highest increases in extreme hourly precipitation can be expected in the future.

It is also important to note that the high spatial variability in the estimated increases of extreme hourly precipitation is not in line with the increases of spatially averaged daily precipitation extremes simulated by the four different GCM-driven RCMs. For example, the highest increases in the 99.5th percentiles of daily precipitation averaged over the four different GCM-driven RCMs (and over 100 bootstrapping samples for each of the model runs) amount to 22.3% at Northolt and the lowest averaged increases are found at Leuchars for which an increase of 13.6% is simulated. In contrast, the difference between the station with the highest and the lowest average increase in terms of the 99.5th percentile of hourly precipitation is more pronounced. As mentioned above, at Kinloss, an average future increase of 75.1% is projected compared to only 12.1% increase at Tiree. Overall, it also needs to be noted that for most of the twelve stations (except Boscombe Down, Marham and Tiree), the increases in hourly precipitation extremes are stronger than the simulated changes in daily extremes.

In terms of the surface warming, the simulated temperature increase for the time period 2075-2100 under emission scenario RCP8.5 is between 2.1°C (at Tiree) and 3.0°C (at Northolt) averaged over the four different GCM-driven RCMs. For the majority of the stations, Valley (8.4 % °C⁻¹), Aldergrove (8.0% °C⁻¹), Camborne (7.6% °C⁻¹), Cranwell (7.4% °C⁻¹), Boulmer (7.4% °C⁻¹), Boscombe Down (5.9% °C⁻¹) and Tiree (5.6% °C⁻¹), this leads to increases (averaged over the four GCM-driven RCMs and 100 bootstrapping samples each) in extreme hourly precipitation that are consistent or only slightly higher (lower) than what the thermodynamic Clausius-Clapeyron (CC) relation would suggest (a rate of ca. 7% °C⁻¹). Four of the twelve stations, namely Kinloss (28.7% °C⁻¹), Eskdalemuir (16.6% °C⁻¹), Northolt (12.8% °C⁻¹) and Leuchars (9.7% °C⁻¹), show increases considerably higher than the CC scaling and only at Marham (4.5% °C⁻¹) is the average increase clearly below the CC scaling.

It is suggested that increases in extreme precipitation exceeding the CC scaling are mostly due to convective precipitation events. Extreme hourly precipitation can often be linked with convective precipitation events (Beck and Bárdossy, 2013; Gregersen et al., 2013) and those convective events are expected to strongly intensify with higher temperatures. Lenderink and van Meijgaard (2008) argued that the increase in latent heat release with higher temperatures may intensify the upward air motions to explain super CC scaling (ca. 14% increases) of

convective precipitation. Berg et al. (2013) found that the convective precipitation response to warming exceeds the CC rate, whereas stratiform precipitation extremes (such as large-scale frontal precipitation) increase according to the CC scaling. However, it is not always an easy task to distinguish between convective and stratiform precipitation events. For example, large-scale frontal systems are normally linked with stratiform precipitation events but approaching cold fronts can also lead to convective precipitation events (Hand et al., 2004). Besides, Sweeney and O'Hare (1992) suggested that cyclonic conditions, normally associated with the occurrence of stratiform precipitation events, can cause increased convective activity. In this thesis, the two CPs (CP1 and CP4) representing cyclonic conditions over the UK show the strongest link with the occurrence of extreme hourly precipitation. Overall, the highest probabilities of extreme hourly precipitation can be found for stations in the South of the UK on days of CP1. It is suggested that CP1 may be linked with mesoscale convective complexes (MCCs). As explained in Little et al. (2008), MCC events are associated with a small, shallow low pressure system and can lead to intense convective precipitation within a larger area of stratiform precipitation. They often originate from the Bay of Biscay and mostly affect the South and South West of the UK.

In this thesis, the observed hourly precipitation records are perturbed based on precipitation duration – temperature relationships (see Equation 4.20 and 4.21) to project extreme hourly precipitation that have not been observed in the past. In order to illustrate the contribution of the final perturbation process to the overall changes, Table 7.22 shows the projected changes of the 99.5th hourly precipitation percentile without perturbing the observed hourly records exemplary for the time period 2075-2100 under RCP8.5. As expected, not applying the final perturbation step leads to smaller overall increases in future hourly precipitation extremes. However, for the station at Eskdalemuir, the increases in hourly extremes are considerably stronger when the observed records are not perturbed. This can be explained by the fact that Eskdalemuir is one of the few stations with positive regression slopes between winter precipitation duration and temperature as described in Section 6.2. It means that higher temperatures lead to precipitation events of longer duration. Therefore, the application of the perturbation process results in a decrease in extreme hourly precipitation in those cases as the perturbed duration of a precipitation event increases with higher temperatures.

Other studies also estimated changes in extreme hourly precipitation over the UK. For example, Kendon et al. (2014) applied a high-resolution 1.5km RCM over the southern part of the UK. They found a spatially averaged increase of 36% in hourly extremes in summer between a 13-year future period (~2100) and a 13-year reference period (1996-2009) under emission scenario RCP8.5. In comparison, the projected increase for the six stations

(Camborne, Boscombe Down, Northolt, Marham, Cranwell and Valley) in the southern UK selected in this thesis amounts only to 21.6 % on average for 2075-2100 under RCP8.5 (see Table 7.21). The differences may be at least partly due to the fact that the average increase of 36% in the study of Kendon et al. (2014) only refers to summer extremes with a high proportion of convective events, whereas in this thesis the projected increases are given for the entire year. It should be also noted that the definition of extreme precipitation varies between the two studies. Kendon et al. (2014) focused on the mean of the upper 5% of wet hourly values, while in this thesis extreme hourly precipitation is defined as the 99.5th percentiles of the daily maximum hourly precipitation.

Dale et al. (2015) combined the results of the high-resolution 1.5km RCM and a climate analogue approach. They found increases in extreme hourly precipitation of 50% at Newcastle and 48% at London between the 2080s and the reference period 1981-2010 using a high emission scenario broadly comparable with the RCP8.5 scenario. In this thesis, the projected increases at Boulmer (geographically closest station to Newcastle) are only 19.6% on average for the future time period 2075-2100 under RCP8.5 relative to the reference period 1980-2005. At Northolt (geographically closest station to London), increases amount to 38.3% on average for 2075-2100 under RCP8.5 (see Table 7.21). However, results are also not directly comparable as Dale et al. (2015) focused on return levels to estimate changes in extreme precipitation, whereas in this thesis changes in the 99.5th percentiles are considered.

It needs to be mentioned that the projected changes in this thesis and in other studies (Kendon et al., 2014; Dale et al., 2015) are greater than the current guidelines produced by DEFRA (2006) and the Environment Agency (2011). According to those guidelines, increases in extreme precipitation are only between 5% and 20% for comparable future time periods (2030s to 2080s). But it needs to be kept in mind that those guidelines are not designed to be used at sub-daily time scales (Dale et al., 2015).

7.3.2 Projections of extreme hourly precipitation under RCP4.5

Over the two future time periods 2030-2055 and 2075-2100 under emission scenario RCP4.5, most stations exhibit an increase in hourly precipitation extremes across all four different GCM-driven RCMs. The increases in magnitude of the hourly extremes tend to be stronger for the later time period 2075-2100, for which also the increase in temperature is higher compared to the future time period 2030-2055. Similarly to the future time period 2075-2100 under RCP8.5, the highest increases in extremes hourly precipitation can be found at Kinloss for both time periods under RCP4.5. Increases in extreme hourly precipitation are smaller for both time periods under RCP4.5 compared to 2075-2100 under RCP8.5. For example, at

Kinloss, the average increase is 19.3% for 2030-2055 under RCP4.5 (see Table 7.19) and 43.7% for 2075-2100 under RCP4.5 (see Table 7.20), whereas for 2075-2100 under RCP8.5 an average increase of 75% is estimated (see Table 7.21). It also needs to be mentioned that the uncertainty range in the projections of hourly precipitation extremes (based on the 100 bootstrapping samples) can be very high for certain stations (Eskdalemuir and Kinloss).

RCM	GCM	CAMBORNE	BOSCOMBE DOWN	NORTHOLT	MARHAM	CRANWELL	VALLEY	ESKDALEMUIR	BOULMER	ALDERSGROVE	LEUCHARS	TREE	KINLOSS
RCA4	CM5A-MR	1.1	1.4	1.4	1.3	1.3	0.9	1.1	1.2	1.0	1.2	0.8	1.2
RCA4	ESM-LR	0.5	0.6	0.6	0.6	0.6	0.5	0.6	0.6	0.6	0.7	0.5	0.7
RCA4	EC-EARTH	0.9	1.2	1.2	1.2	1.1	0.9	0.9	1.0	0.9	0.9	0.8	1.0
RACMO22E	EC-EARTH	1.0	1.1	1.1	1.1	1.0	1.0	1.0	0.9	1.0	1.0	0.8	0.9
MEAN		0.9	1.1	1.1	1.1	1.0	0.8	0.9	0.9	0.9	1.0	0.7	1.0

Table 7.10 Changes in [°C] in the simulated daily mean temperature in summer for the future time period 2030-2055 under the emission scenario RCP4.5 relative to the reference time period 1980-2005.

RCM	GCM	CAMBORNE	BOSCOMBE DOWN	NORTHOLT	MARHAM	CRANWELL	VALLEY	ESKDALEMUIR	BOULMER	ALDERSGROVE	LEUCHARS	TREE	KINLOSS
RCA4	CM5A-MR	1.6	1.9	2.0	1.9	1.9	1.4	1.8	1.8	1.6	1.9	1.5	1.9
RCA4	ESM-LR	0.9	1.1	1.2	1.2	1.2	0.9	1.1	1.1	1.0	1.1	0.8	1.1
RCA4	EC-EARTH	1.8	2.2	2.2	2.1	2.1	1.8	1.8	1.9	1.8	1.8	1.7	1.9
RACMO22E	EC-EARTH	1.5	1.6	1.5	1.5	1.5	1.5	1.6	1.5	1.5	1.6	1.3	1.6
MEAN		1.5	1.7	1.7	1.7	1.7	1.4	1.6	1.6	1.5	1.6	1.3	1.6

Table 7.11 same as Table 7.10 but changes are calculated for the future time period 2075-2100 under emission scenario RCP4.5 relative to the reference time period 1980-2005.

RCM	GCM	CAMBORNE	BOSCOMBE DOWN	NORTHOLT	MARHAM	CRANWELL	VALLEY	ESKDALEMUIR	BOULMER	ALDERSGROVE	LEUCHARS	TREE	KINLOSS
RCA4	CM5A-MR	2.9	3.4	3.4	3.3	3.2	2.5	2.9	2.9	2.7	3.0	2.3	2.9
RCA4	ESM-LR	2.2	2.7	2.8	2.6	2.6	2.0	2.2	2.2	2.1	2.2	1.7	2.1
RCA4	EC-EARTH	3.2	4.0	4.0	3.8	3.7	3.0	3.1	3.2	3.0	3.1	2.6	3.1
RACMO22E	EC-EARTH	2.9	3.4	3.4	3.1	3.1	2.8	2.9	2.7	2.8	2.9	2.2	2.6
MEAN		2.8	3.4	3.4	3.2	3.2	2.6	2.8	2.8	2.7	2.8	2.2	2.7

Table 7.12 same as Table 7.10 but changes are calculated for the future time period 2075-2100 under emission scenario RCP8.5 relative to the reference time period 1980-2005.

RCM	GCM	CAMBORNE	BOSCOMBE DOWN	NORTHOLT	MARHAM	CRANWELL	VALLEY	ESKDALEMUIR	BOULMER	ALDERSGROVE	LEUCHARS	TREE	KINLOSS
RCA4	CM5A-MR	-5.3	-12.2	-5.3	-10.7	-12.2	-7.2	-1.0	1.9	-0.9	1.3	6.6	-1.5
RCA4	ESM-LR	-0.6	-3.2	2.1	-2.8	0.4	-5.1	-1.9	-6.0	-5.0	-3.8	-4.7	-2.8
RCA4	EC-EARTH	-3.9	-3.2	-5.4	-11.4	-9.4	-1.4	-1.6	-9.7	0.2	-4.1	-2.6	-10.6
RACMO22E	EC-EARTH	-5.1	-0.4	-5.0	-3.9	-5.0	-6.5	-5.3	2.3	-6.7	-2.8	-2.3	2.3
MEAN		-3.7	-4.8	-3.4	-7.2	-6.6	-5.1	-2.5	-2.9	-3.1	-2.4	-0.8	-3.2

Table 7.13 Changes in [%] in the simulated daily mean precipitation in summer for the future time period 2030-2055 under the emission scenario RCP4.5 relative to the reference time period 1980-2005.

RCM	GCM	CAMBORNE	BOSCOMBE DOWN	NORTHOLT	MARHAM	CRANWELL	VALLEY	ESKDALEMUIR	BOULMER	ALDERSGROVE	LEUCHARS	TREE	KINLOSS
RCA4	CM5A-MR	-5.0	-7.4	-5.8	-4.7	-5.0	-5.8	6.5	4.2	-4.3	5.0	4.4	-1.0
RCA4	ESM-LR	-9.1	-6.7	-8.3	-8.7	-4.6	-5.0	-1.3	-3.2	-4.8	0.6	-2.1	-2.4
RCA4	EC-EARTH	-5.3	-2.2	-8.9	-13.8	-10.8	-7.3	-9.9	-9.6	-9.6	-5.7	-0.2	-6.6
RACMO22E	EC-EARTH	-4.1	0.1	-1.5	2.0	4.0	4.8	4.5	8.1	5.5	7.5	1.1	5.7
MEAN		-5.9	-4.1	-6.1	-6.3	-4.1	-3.3	-0.1	-0.1	-3.3	1.9	0.8	-1.1

Table 7.14 same as Table 7.13 but changes are calculated for the future time period 2075-2100 under emission scenario RCP4.5 relative to the reference time period 1980-2005.

RCM	GCM	CAMBORNE	BOSCOMBE DOWN	NORTHOLT	MARHAM	CRANWELL	VALLEY	ESKDALEMUIR	BOULMER	ALDERSGROVE	LEUCHARS	TIREE	KINLOSS
RCA4	CM5A-MR	-20.2	-16.6	-12.5	-12.6	-12.0	-3.9	3.4	9.2	4.6	5.3	7.6	7.0
RCA4	ESM-LR	-13.0	-11.9	-7.8	-12.6	-12.4	-8.2	-6.6	-1.0	-7.2	-2.5	1.6	-0.9
RCA4	EC-EARTH	-14.6	-9.8	-12.5	-17.3	-12.2	-3.5	-6.0	-8.2	-6.7	0.9	-1.9	-4.2
RACMO22E	EC-EARTH	-14.3	-9.5	-7.3	-4.4	-3.3	-5.0	-0.5	5.8	-2.3	1.3	3.5	5.2
MEAN		-15.5	-12.0	-10.0	-11.7	-10.0	-5.2	-2.4	1.5	-2.9	1.3	2.7	1.8

Table 7.15 same as Table 7.13 but changes are calculated for the future time period 2075-2100 under emission scenario RCP8.5 relative to the reference time period 1980-2005.

RCM	GCM	CAMBORNE	BOSCOMBE DOWN	NORTHOLT	MARHAM	CRANWELL	VALLEY	ESKDALEMUIR	BOULMER	ALDERSGROVE	LEUCHARS	TIREE	KINLOSS
RCA4	CM5A-MR	16.3	-10.3	9.8	-9.6	-1.4	-5.9	2.5	16.3	5.1	10.0	6.3	-13.5
RCA4	ESM-LR	-5.5	-3.2	7.0	-8.5	3.4	1.4	2.6	-7.3	-8.5	4.0	-9.2	4.0
RCA4	EC-EARTH	8.1	11.1	4.5	-4.2	-5.6	11.2	11.5	-3.8	13.1	-4.0	-1.7	-6.0
RACMO22E	EC-EARTH	7.4	9.1	-1.3	12.2	7.1	-2.2	-1.0	17.0	-5.9	13.7	10.0	20.6
MEAN		8.8	1.7	5.0	-2.5	0.9	1.1	3.9	5.6	1.0	5.9	1.4	1.3

Table 7.16 Changes in [%] in the simulated 99.5th daily precipitation percentile in summer for the future time period 2030-2055 under the emission scenario RCP4.5 relative to the reference time period 1980-2005.

RCM	GCM	CAMBORNE	BOSCOMBE DOWN	NORTHOLT	MARHAM	CRANWELL	VALLEY	ESKDALEMUIR	BOULMER	ALDERSGROVE	LEUCHARS	TIREE	KINLOSS
RCA4	CM5A-MR	28.0	15.9	9.1	6.3	13.8	4.3	10.1	28.7	-3.4	13.4	10.7	13.9
RCA4	ESM-LR	-4.3	2.2	-1.5	-0.2	2.7	-2.0	-0.7	-0.3	-7.5	14.4	7.0	-3.9
RCA4	EC-EARTH	11.2	25.5	2.5	4.3	16.9	7.5	14.5	8.4	4.5	9.2	6.9	-7.3
RACMO22E	EC-EARTH	1.3	3.6	-0.2	21.4	22.2	29.7	1.5	12.1	3.2	17.9	9.3	4.1
MEAN		9.1	11.8	2.5	8.0	13.9	9.9	6.4	12.2	-0.8	13.7	8.5	1.7

Table 7.17 same as Table 7.16 but changes are calculated for the future time period 2075-2100 under emission scenario RCP4.5 relative to the reference time period 1980-2005.

RCM	GCM	CAMBORNE	BOSCOMBE DOWN	NORTHOLT	MARHAM	CRANWELL	VALLEY	ESKDALEMUIR	BOULMER	ALDERSGROVE	LEUCHARS	TIREE	KINLOSS
RCA4	CM5A-MR	5.3	8.9	10.0	1.6	17.4	12.4	24.8	28.7	22.7	9.2	21.7	26.8
RCA4	ESM-LR	16.2	8.5	20.6	10.4	3.8	17.8	9.6	11.9	11.3	14.1	12.9	9.5
RCA4	EC-EARTH	24.6	27.6	17.1	13.6	14.9	24.3	22.7	8.6	18.5	19.9	5.7	3.0
RACMO22E	EC-EARTH	11.0	13.4	22.9	26.6	24.2	19.6	13.7	18.9	11.4	18.0	27.6	8.2
MEAN		14.3	14.6	17.7	13.1	15.1	18.5	17.7	17.0	16.0	15.3	17.0	11.9

Table 7.18 same as Table 7.16 but changes are calculated for the future time period 2075-2100 under emission scenario RCP8.5 relative to the reference time period 1980-2005.

RCM	GCM	CAMBORNE	BOSCOMBE DOWN	NORTHOLT	MARHAM	CRANWELL	VALLEY	ESKDALEMUIR	BOULMER	ALDERSGROVE	LEUCHARS	TIREE	KINLOSS
RCA4	CM5A-MR	6.3	-5.9	25.5	2.3	0.0	4.3	38.4	1.9	16.7	23.9	4.1	43.3
RCA4	ESM-LR	12.0	12.1	27.6	3.7	6.2	3.3	27.1	-4.6	-3.4	8.6	1.6	6.8
RCA4	EC-EARTH	-6.9	15.8	27.1	1.4	1.7	2.2	12.7	3.7	22.7	12.6	-0.9	7.2
RACMO22E	EC-EARTH	5.9	-0.1	9.8	2.0	-1.1	-1.6	51.5	9.6	2.4	11.9	-7.3	19.8
MEAN		4.3	5.5	22.5	2.4	1.7	2.1	32.4	2.7	9.6	14.3	-0.6	19.3

Table 7.19 Changes in [%] in the estimated median of the 99.5th hourly precipitation percentile based on the 100 bootstrapping samples for the future time period 2030-2055 under the emission scenario RCP4.5 relative to the reference time period 1980-2005.

RCM	GCM	CAMBORNE	BOSCOMBE DOWN	NORTHOLT	MARHAM	CRANWELL	VALLEY	ESKDALEMUIR	BOULMER	ALDERSGROVE	LEUCHARS	TIREE	KINLOSS
RCA4	CM5A-MR	21.0	2.9	36.7	3.1	16.6	4.6	62.7	20.6	-2.4	26.0	10.1	14.2
RCA4	ESM-LR	8.9	13.4	22.4	7.9	7.3	-0.5	19.9	3.7	9.6	12.9	2.0	75.9
RCA4	EC-EARTH	8.4	26.4	32.8	0.7	3.5	0.0	5.2	3.6	17.1	22.6	10.2	59.2
RACMO22E	EC-EARTH	2.1	-5.3	0.4	4.2	13.3	4.9	2.3	17.3	6.3	22.4	-3.0	25.4
MEAN		10.1	9.4	23.1	4.0	10.2	2.3	22.5	11.3	7.7	21.0	4.8	43.7

Table 7.20 same as Table 7.19 but changes are calculated for the future time period 2075-2100 under emission scenario RCP4.5 relative to the reference time period 1980-2005.

RCM	GCM	CAMBORNE	BOSCOMBE DOWN	NORTHOLT	MARHAM	CRANWELL	VALLEY	ESKDALEMUIR	BOULMER	ALDERSGROVE	LEUCHARS	TREE	KINLOSS
RCA4	CM5A-MR	22.1	13.1	41.9	14.2	29.4	28.0	37.1	31.8	25.2	39.5	24.8	36.9
RCA4	ESM-LR	21.9	15.0	43.9	10.5	27.7	25.7	36.2	9.8	11.0	18.8	13.1	85.8
RCA4	EC-EARTH	10.8	28.9	46.2	15.6	18.2	9.3	31.7	15.9	29.8	23.2	9.2	111.9
RACMO22E	EC-EARTH	22.0	12.2	21.3	12.2	10.2	18.6	78.0	20.9	13.0	23.8	1.2	65.9
MEAN		19.2	17.3	38.3	13.1	21.4	20.4	45.8	19.6	19.8	26.3	12.1	75.0

Table 7.21 same as Table 7.19 but changes are calculated for the future time period 2075-2100 under emission scenario RCP8.5 relative to the reference time period 1980-2005.

RCM	GCM	CAMBORNE	BOSCOMBE DOWN	NORTHOLT	MARHAM	CRANWELL	VALLEY	ESKDALEMUIR	BOULMER	ALDERSGROVE	LEUCHARS	TREE	KINLOSS
RCA4	CM5A-MR	18.0	10.3	34.8	6.8	19.7	25.1	75.1	30.4	24.7	39.1	25.2	38.1
RCA4	ESM-LR	20.7	14.0	37.4	6.3	19.1	23.2	62.0	9.6	11.4	15.6	13.7	85.6
RCA4	EC-EARTH	10.6	22.8	37.7	12.1	9.6	9.3	34.3	10.6	31.6	22.0	6.1	113.0
RACMO22E	EC-EARTH	23.4	10.6	18.0	12.3	8.3	19.3	90.3	26.1	13.9	26.9	4.0	60.6
MEAN		18.2	14.4	31.2	9.4	14.2	19.2	65.4	19.2	20.4	25.9	12.3	74.3

Table 7.22 same as Table 7.21 but the 99.5th hourly precipitation percentiles are estimated without applying the final perturbation step based on the precipitation duration – temperature relationship (see Equation 4.20 and 4.21).

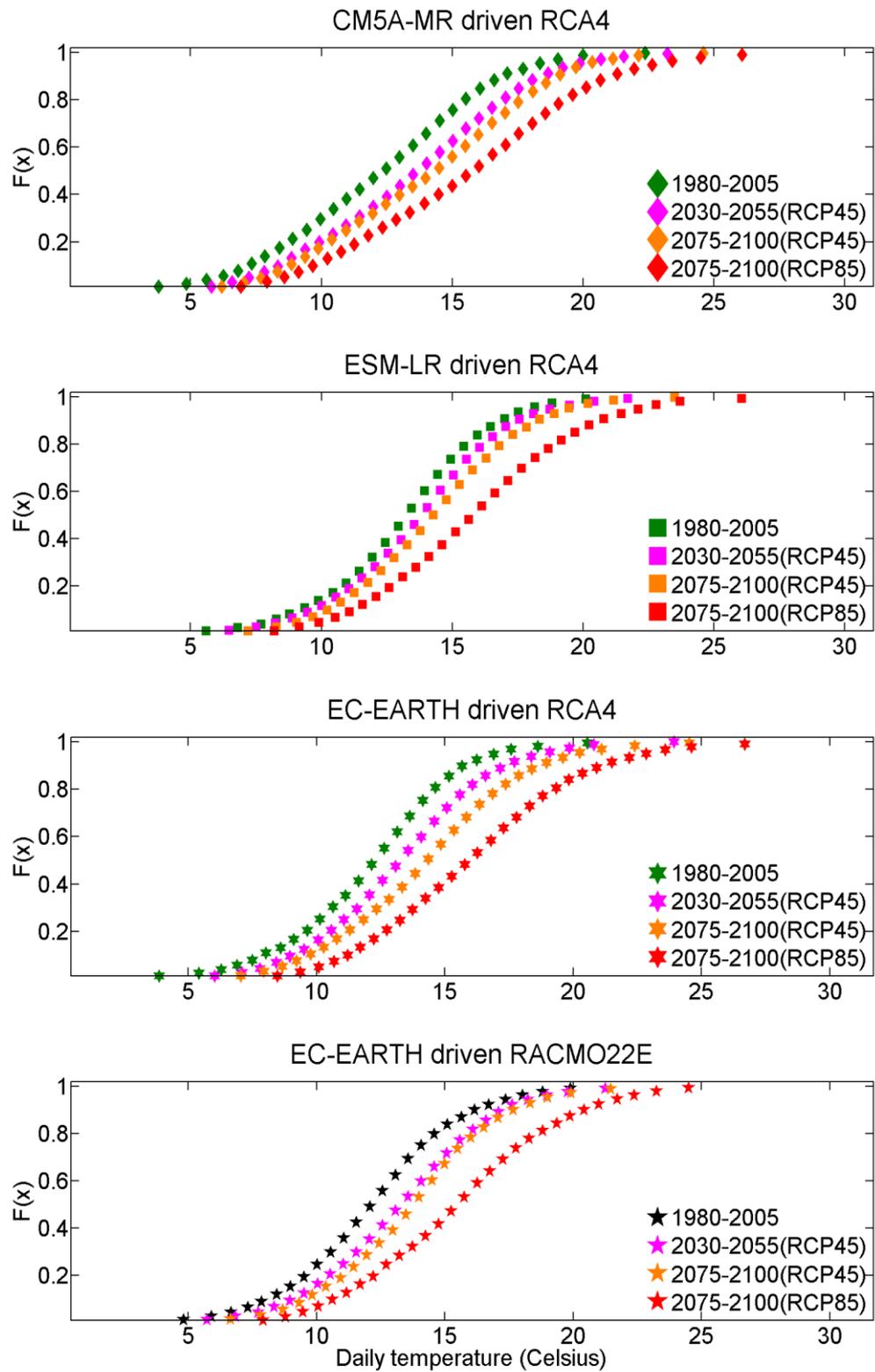


Figure 7.8 Simulated daily temperature distributions in summer based on four different GCM-driven RCM data sets for the station at Boscombe Down.

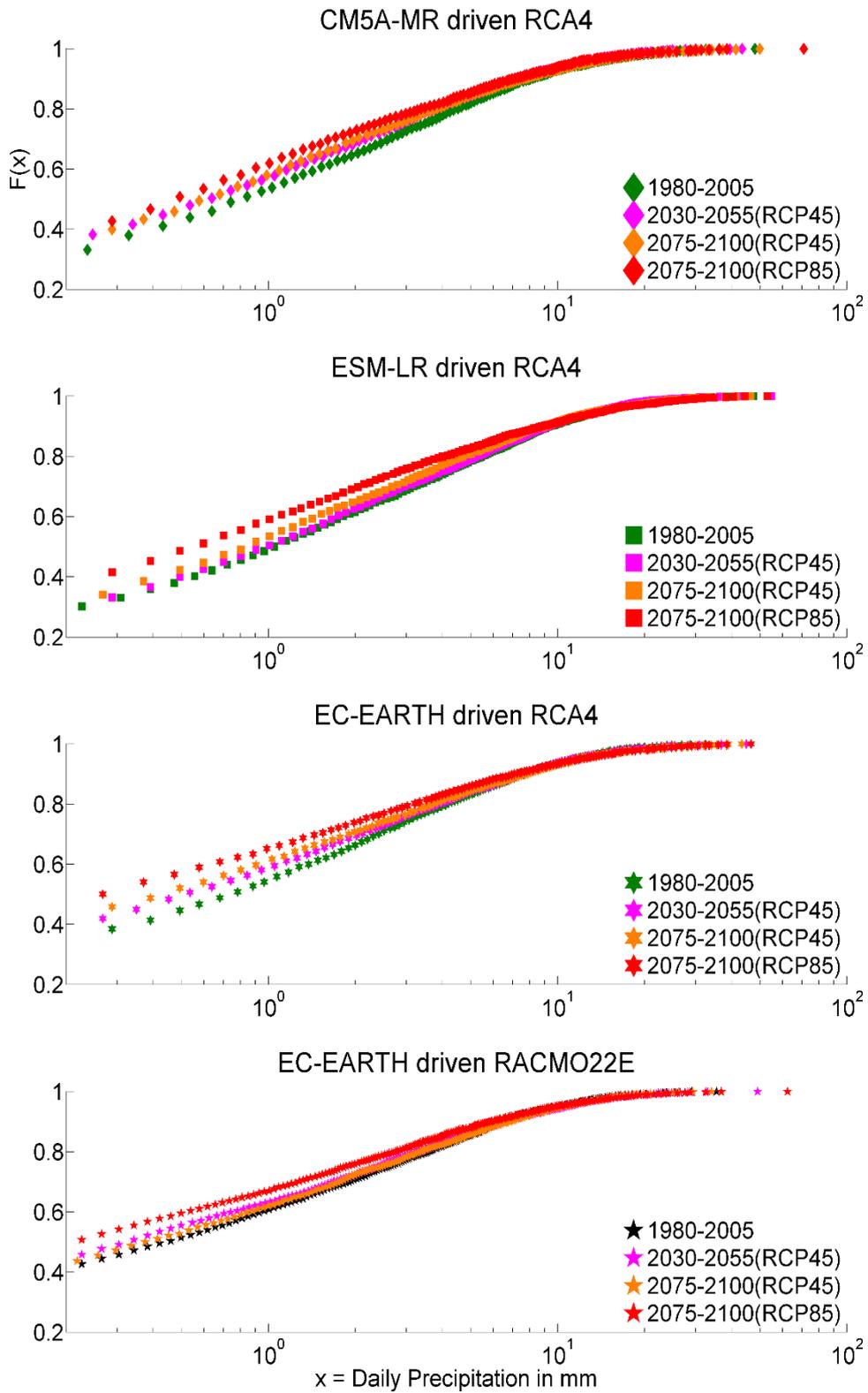


Figure 7.9 Simulated daily precipitation distributions in summer based on four different GCM-driven RCM data sets for the station at Boscombe Down.

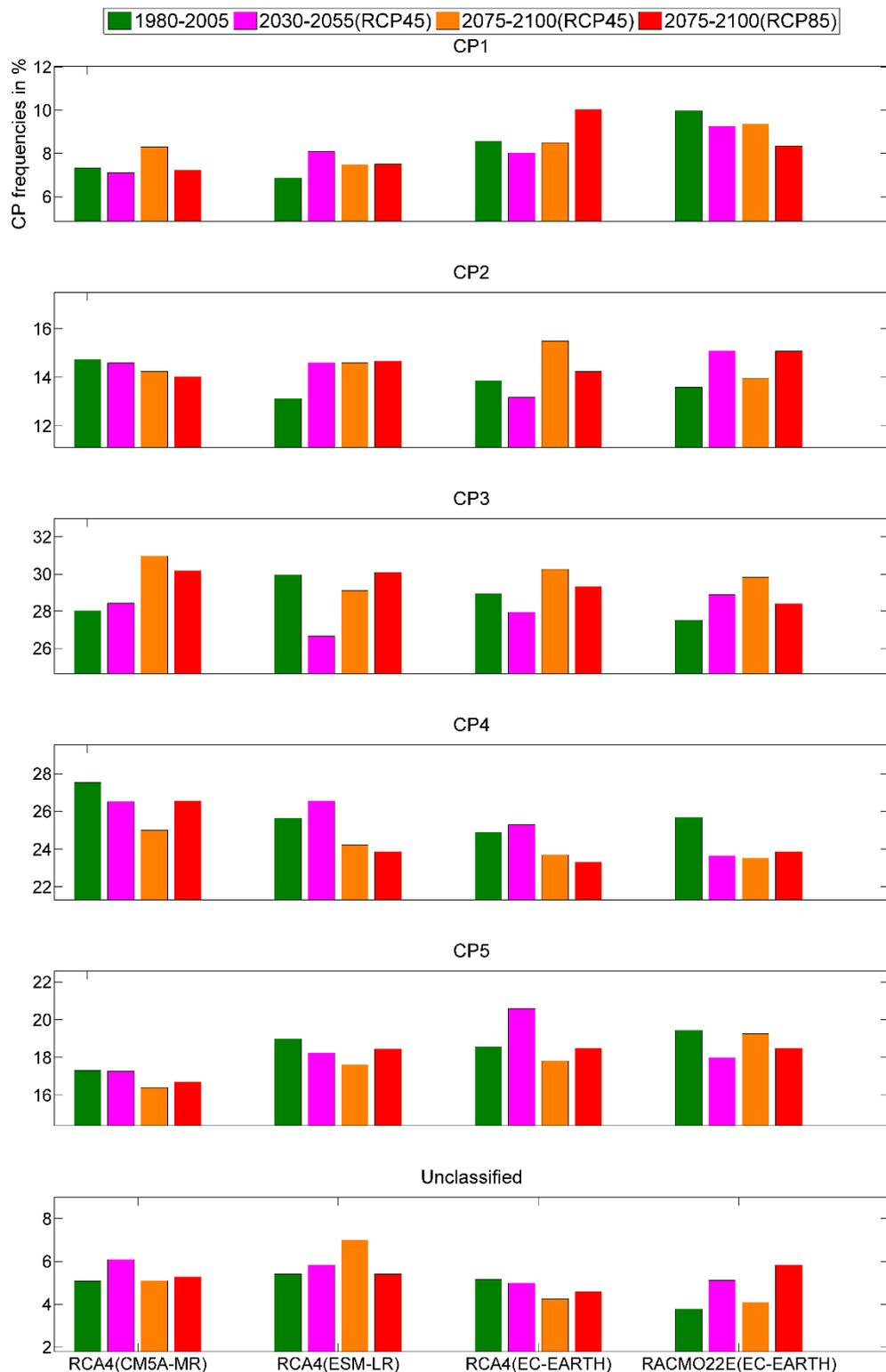


Figure 7.10 Simulated CP frequencies in summer based on four different GCM-driven RCM data sets.

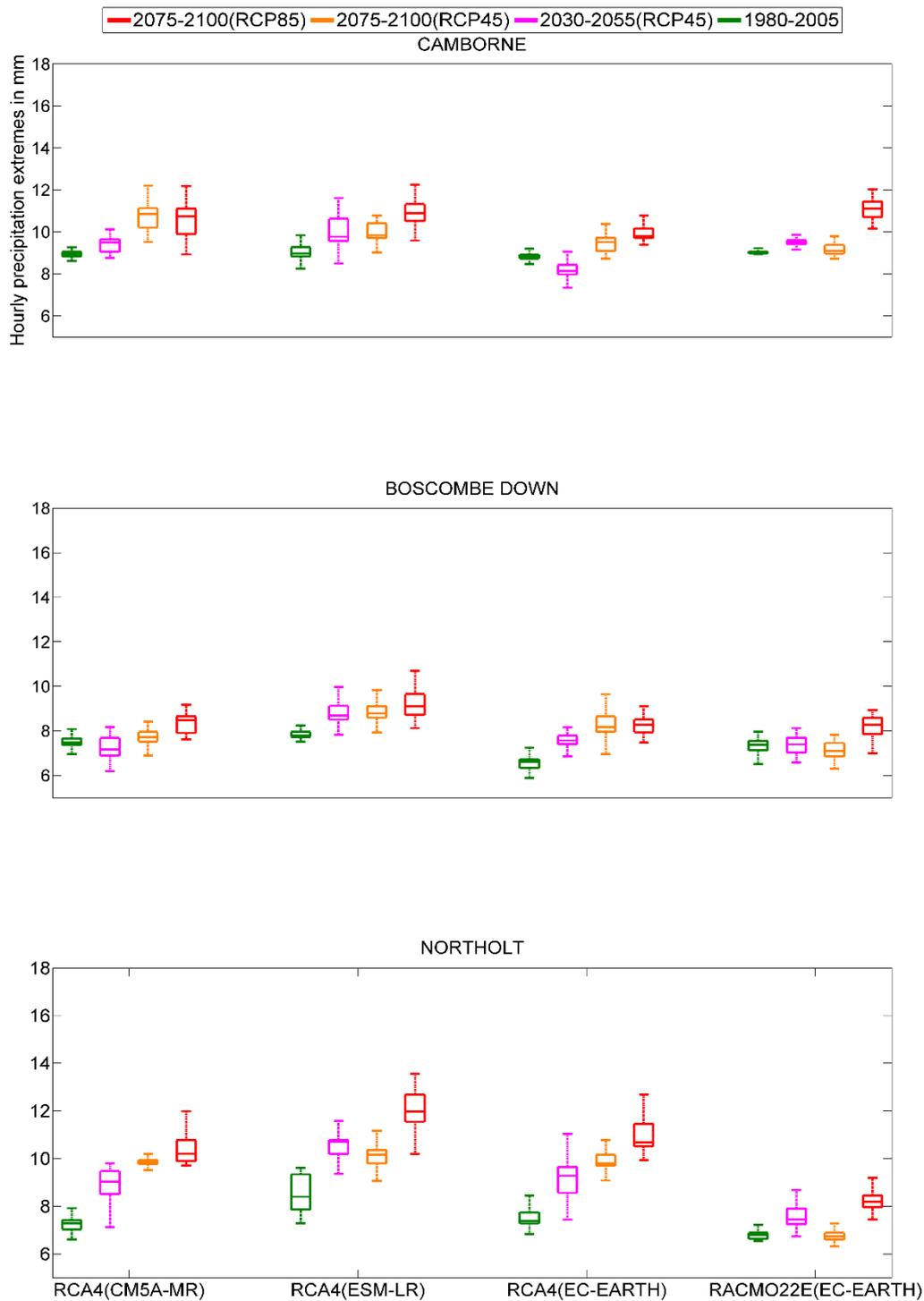


Figure 7.11 Estimated 99.5th percentiles of the daily maximum hourly precipitation based on GCM-driven RCM data. The percentiles are projected over 100 different bootstrapping samples for every individual GCM-driven RCM run. The edges of the boxes indicate the 25th and 75th percentiles derived from the 100 bootstrapping samples. The central mark within the boxes represents the median. Approach M1 is used to find the analogue days.

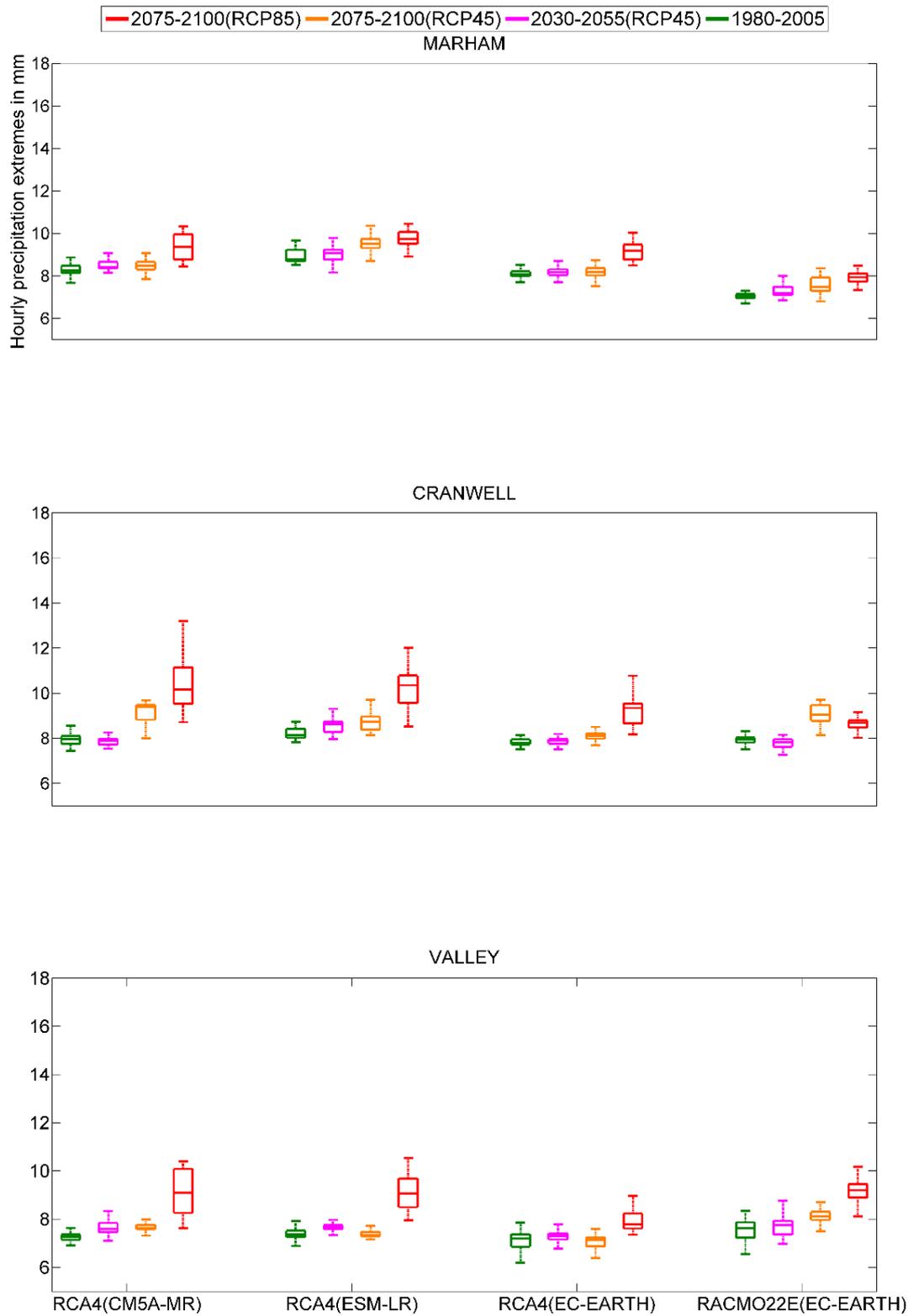


Figure 7.12 see Figure 7.11.

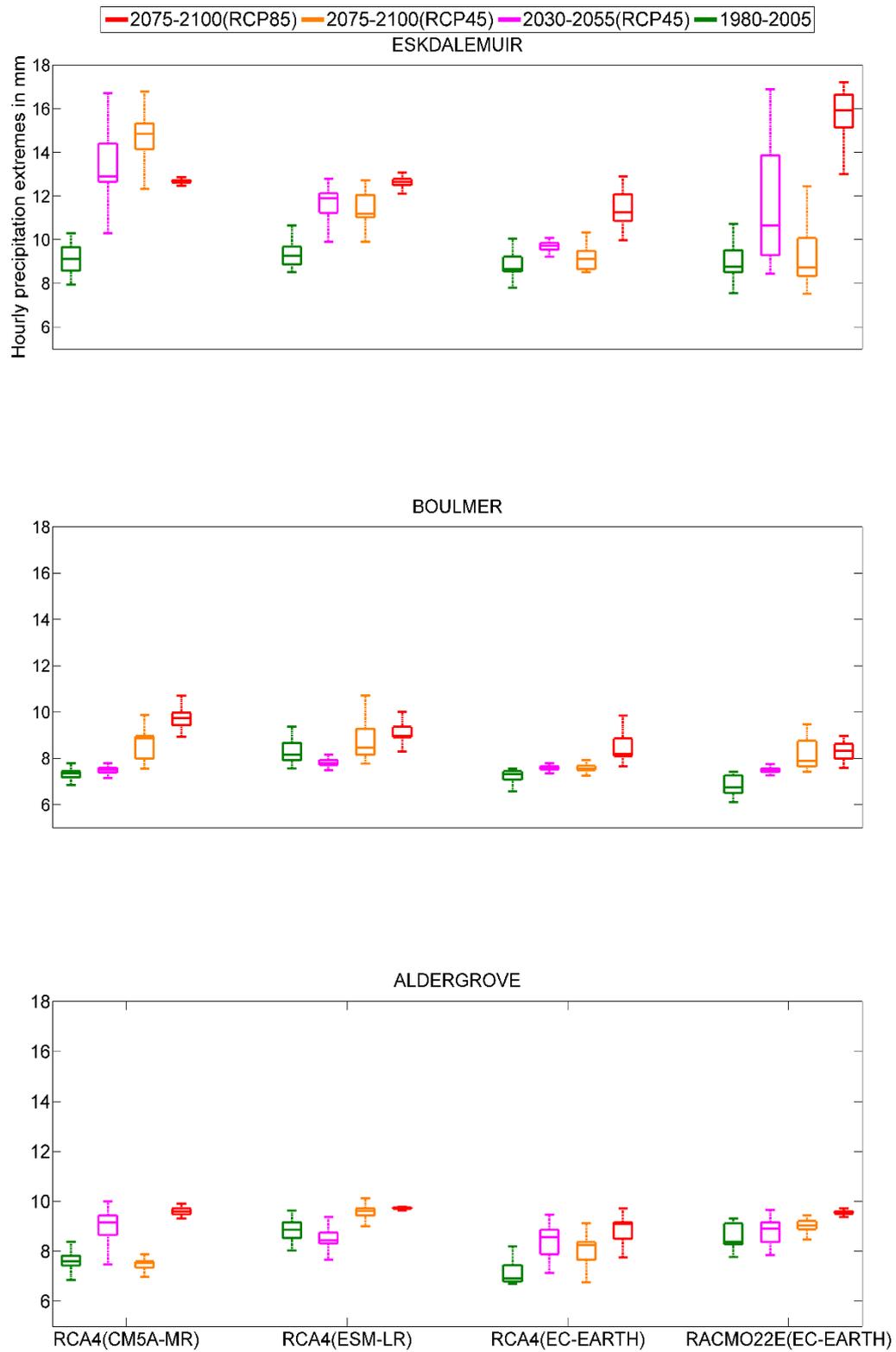


Figure 7.13 see Figure 7.11.

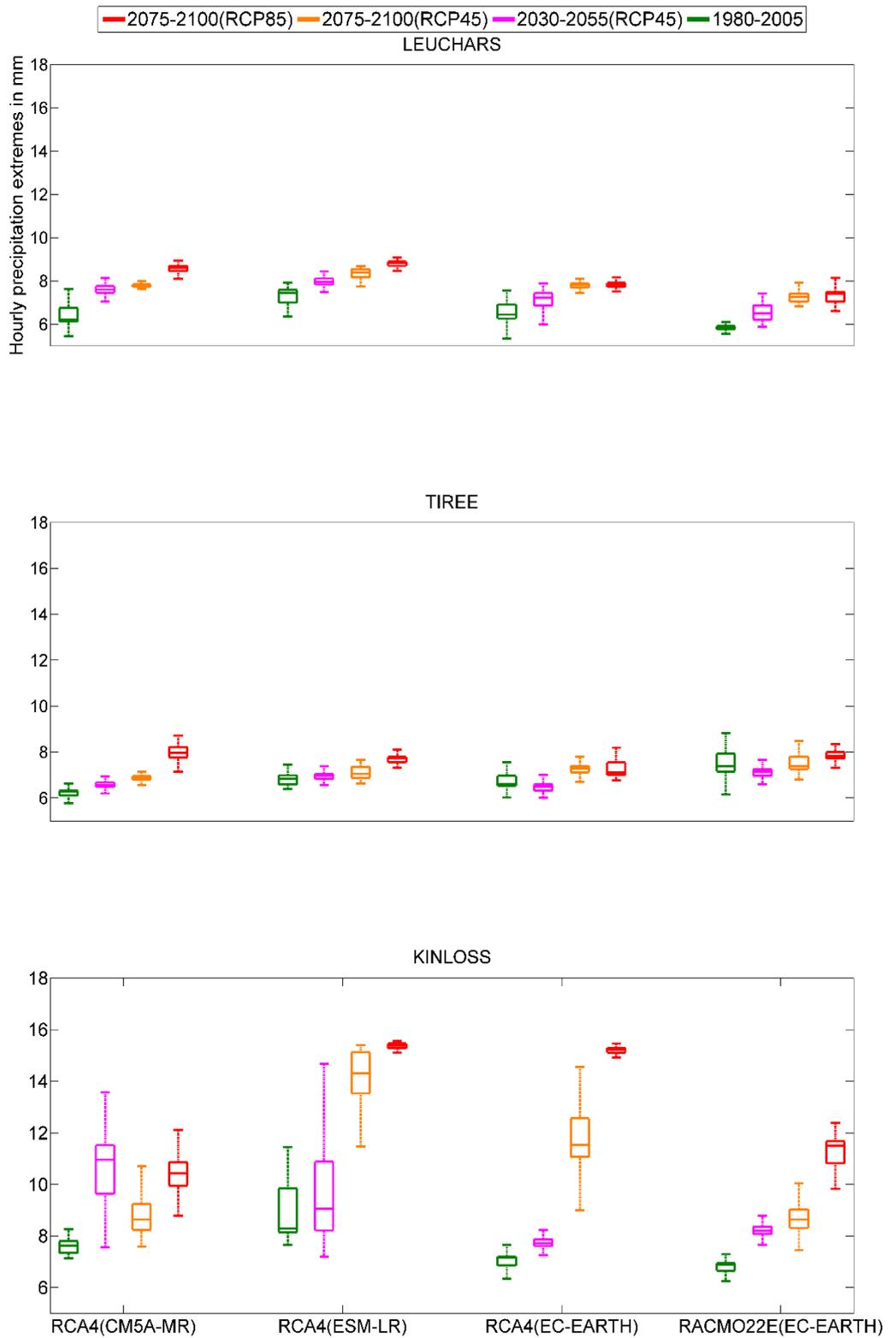


Figure 7.14 see Figure 7.11.

7.4 Summary

Chapter 7 focuses on the estimation of hourly precipitation derived from RCM predictors. The variables are the daily mean temperature, daily precipitation, and the associated CPs on each day. The statistical downscaling methodology described in Section 4.3 is applied. This chapter uses the same method and variables as Chapter 6, but the latter uses the observed predictors to estimate extreme hourly precipitation.

The daily temperature, precipitation and CPs derived from ERA-interim driven RCM runs on each day within the validation period are compared against the observed climate variables over the respective calibration periods, in order to find the analogue days. Four different approaches (M1, M2, M3 and M4) are applied to find the analogue day. Overall it is found that the RCM estimated hourly precipitation extremes show good agreement with the observed hourly precipitation extremes. The best performances in terms of extreme hourly precipitation estimates amongst all the twelve stations and the three validation periods are generated with the approach M1 that subsamples calibration days into seasons and CPs to find the analogue day.

In addition to the three different validation periods, the ten warmest summers between 1980 and 2009 are selected for each station to compare the estimated and observed extreme hourly precipitation. This is to assess the predictor-predictand relationship for a climate that is different to the one it is calibrated on. The approach M1 is again the best approach for all the twelve stations, hence it is used for the downscaling of hourly precipitation based on the GCM-driven RCM data set. It is found that the performance of the downscaling method over the ten warmest summers is very similar to the performances over the three validation periods and thus it can be concluded that the downscaling method is capable of performing under a warmer climate. However, future increases in temperature under emission scenario RCP8.5 are projected to be considerably higher than the historical temperature difference between the ten warmest summers in 1980-2009 and the calibration period. Therefore, the future estimates of hourly precipitation extremes still need to be treated with caution since the statistically downscaling process may not be fully sufficient to represent the entire complexity of changes under a considerably warmer climate.

In a second test to assess if the predictor-predictand relationship remains valid over time, the residuals time series between observed and estimated hourly precipitation extremes are assessed over the validation time period 1980-2009. It is shown that statistically significant trends in the residual time series can occur for certain stations indicating signs of nonstationarities in the statistical downscaling method in some cases. Considering the results

based on approach M1, statistically significant trends in the residual time series are found at two stations based on RCA4 RCM ERA-interim driven data and at one station based on RACMO22E RCM ERA-interim driven data.

In Section 7.2, estimates of hourly precipitation extremes derived from the four GCM-driven RCM data sets using the approach M1 are compared with the actual observed values over the reference period 1980-2005. It is found that estimates agree generally well with the observed precipitation extremes for all four different data sets. In terms of the RCM predictors, it is shown that the RCA4 runs overestimates daily precipitation and daily temperature tend to be underestimated for most of the stations by the RCA4 and the RACMO22E RCMs. However, the statistical downscaling process is based on normalized values of the predictor variables and therefore any existing biases in the RCM predictors are unlikely to have an adverse effect on the estimates of hourly precipitation. Regarding the large-scale atmospheric variability, it is found that the observed CP frequencies are realistically reproduced by the RCA4 and the RACMO22E RCMs.

In Section 7.3, hourly precipitation extremes are projected for two future periods 2030-2055 and 2075-2100. In terms of the future period 2075-2100, two different emission scenarios are considered. The emission scenario RCP8.5 represents a high emission scenario and therefore results in higher temperature increases than emission scenario RCP4.5. For the future period 2030-2055, only RCP4.5 is applied, because the respective changes in the climate are projected to be similar over this time period between the two emission scenarios (Kirtman et al., 2013). In terms of the RCM predictor variables, only daily temperature exhibits considerable future changes. The other two RCM predictors, daily precipitation and the CP frequencies, only show minor changes under a warming climate. It is shown, that hourly precipitation extremes increase for most of the stations under RCP4.5 and for all stations under RCP8.5. The increases in estimated hourly precipitation extremes are most pronounced for the future time period 2075-2100 under RCP8.5, for which also the temperature increases are the highest. Under those conditions, increases in hourly extremes can be as high as 111.9% at the station Kinloss based on the EC-EARTH driven RCA4. But there is a considerably wide range of changes depending on the stations, the future time periods, and the emission scenarios. Additional uncertainties in the projections of extreme hourly precipitation are induced by the four different GCM-driven RCMs. It also needs to be mentioned that for a given station, future time period, emission scenario and GCM-driven RCM, the projections may still vary considerably over the 100 different bootstrapping samples.

Chapter 8

Summary and Discussion

8.1 Summary

This thesis aimed to project site-specific hourly precipitation extremes over the UK by applying a CP-based statistical downscaling method.

For this purpose, hourly precipitation records from 530 different stations over the UK comprising the time period 1949-2012 were extracted and subject to multiple quality control measures. In theory, the quality controlled data set could also be used for other studies assessing hourly precipitation extremes. For example, it could provide the basis to identify different regions of extreme hourly precipitation characteristics. In a next step, a detailed regional trend analysis could be conducted to assess changes in hourly precipitation extremes over the last decades similarly to the study of Jones et al. (2014). Another potential use of the observed hourly data set could be the validation of high-resolution RCM results. These high-resolution models (1km-5km) are able to reproduce hourly precipitation extremes more realistically than the standard RCMs at 12km-50km resolution (Chan et al., 2014b; Kendon et al., 2014). A detailed validation process based on comprehensive and quality controlled hourly precipitation records could help to further improve the performances of these models.

A statistical downscaling method based on the concept of analogue days was developed. By means of analogue days, local precipitation characteristics can be reproduced on every time scale for which precipitation records are available. Different approaches of finding the analogue days were tested in order to optimise the performance of the statistical downscaling method. Two predictor variables (daily precipitation and temperature) conditioned on five (plus one unclassified) CPs were used in the process of finding the analogue days. In a next step, observed hourly precipitation extracted on the analogue days was perturbed based on precipitation duration – temperature relationships. Projections of changes in UK extreme hourly precipitation are given for twelve different stations using four different GCM-driven RCMs (at 12.5km resolution), two different future time periods and two different emission scenarios. Increases in extreme hourly precipitation (up to 112%) are strongest for the future time period 2075-2100 under emission scenario RCP8.5 when the rise in temperature is most pronounced. It needs to be noted that the changes in hourly precipitation extremes can vary

highly between the twelve stations, between the different GCM-driven RCM runs and over the 100 different bootstrapping samples within the same GCM-driven RCM run. This is at least partly caused by limitations in the RCM simulated predictors and in the statistical downscaling process (see Section 8.2). A potential development in the future would be to use RCM predictors at very high spatial resolution (1km-5km) within the statistical downscaling process. As mentioned above, high-resolution RCMs are able to represent hourly precipitation extremes more realistically compared to standard RCMs. However, these high-resolution models still misrepresent local convective showers (Chan et al., 2014a) and have a tendency to overestimate extreme hourly precipitation (Kendon et al., 2014). Therefore, linking the high-resolution RCM outputs with the statistical downscaling process in this thesis is likely to further improve the quality of extreme hourly precipitation projections.

In terms of the relationship between CPs and extreme hourly precipitation, the highest extreme precipitation probabilities were found for the two CPs that are associated with low pressure systems. The two CPs characterized by high pressure conditions over the UK exhibit relatively low extreme hourly precipitation probabilities but can occasionally cause very extreme precipitation events. It was suggested that those very extreme events may be a result of localized convective showers under high pressure conditions (Blenkinsop et al., 2015). Considering the relationship between extreme hourly precipitation and temperature, some studies (Haerter and Berg, 2009; Lenderink and van Meijgaard, 2008) found considerably higher increases in extreme hourly precipitation than the 7% increase per degree of warming suggested by the Clausius-Clapeyron relation. It was argued that those higher increases are due to a strong intensification of convective precipitation events with higher temperature (Lenderink and van Meijgaard, 2008; Molnar et al., 2014). In this thesis, for nine of the twelve stations the averaged increase rates in extreme hourly precipitation are higher than 7% °C⁻¹ but only four of them considerably exceed the CC relation.

Chapter 3 describes the various data sets. Observed hourly and daily precipitation at station level was extracted from the MIDAS Land Surface Stations database (NCAS British Atmospheric Data Centre, 2012). A number of criteria were set to eliminate extreme outliers, as well as suspicious and multiple available records. A total number of twelve stations were selected. They contain the most homogeneous and complete precipitation time series. They represent twelve of the 14 different UK extreme precipitation regions classified in Jones et al. (2014). The annual cycle of hourly precipitation extremes was assessed for the twelve selected stations. It was found that extreme hourly precipitation tends to be more intense between May and October for the majority of the stations. Two seasons were defined. Summer comprises the period between May and October, whereas winter spans from November to April. The climate

model data was extracted from the CORDEX data archive (CORDEX, 2014). It consists of two different 12.5km RCMs (RCA4 and RACMO22E) driven by ERA-interim data and by three different GCMs (CM5A-MR, ESM-LR and EC-EARTH).

The methodologies to derive the fuzzy rules based CP-classification and the statistical downscaling are explained in Chapter 4. Days of similar SLP patterns were categorised into a small number of characteristic circulation patterns (CPs). The definition of the fuzzy rules based CP-classification was based on four different objective functions, which optimise the quality of the classification to distinguish between dry and wet CPs. Information measures were applied to find the best applicable CP-classification for the downscaling process. The optimum classification consists of 5 (plus one unclassified) CPs.

The statistical downscaling method developed in this study can be regarded as a hybrid of an analogue and a regression-based method. It is similar to the analogue day method as applied in Willems and Vrac (2011). However, in contrast to Willems and Vrac (2011), the historical hourly precipitation records extracted on the analogue day were perturbed based on precipitation duration – temperature relationships in this study. The fact that daily precipitation and temperature conditioned on CPs were used as equally important, combined predictors is also unique to the applied method here. All the available days were first subsampled into the five (plus one unclassified) different CPs and two seasons or twelve calendar months. For each day over the validation period, the most similar day (analogue day) over the calibration period was found by comparing the daily precipitation and temperature values conditioned on five (plus one unclassified) CPs and two seasons (or twelve months). Four different approaches of finding the analogue day were proposed. The temperature variable is needed to represent changes of extreme hourly precipitation under a warmer climate, as extreme events are expected to intensify with higher temperature (Berg et al., 2013; Blenkinsop et al., 2015; Lenderink and van Meijgaard, 2008). Daily precipitation provides additional information as to whether and how much precipitation on that day actually takes place. This is important because high temperature in summer can lead to clear skies and dry conditions, but also have the potential of causing convective extreme precipitation (Beck and Bárdossy, 2013). In a next step, the daily maximum hourly precipitation on the analogue day was perturbed based on the relationships between precipitation duration and temperature. Those relationships were derived for different CPs and seasons. The basic assumption is that the duration of daily precipitation events decreases with higher temperature. It was shown that this assumption is mostly valid in summer under cyclonic conditions and under westerly airflow for those stations not in close proximity to the Sea. For example, 32 of the 36 relationships between precipitation duration and temperature conditioned on CP4 are

statistically significant and fulfil the RMSE (Root Mean Square Error) criterion. The statistical downscaling process to estimate and project hourly precipitation extremes was applied for the twelve stations selected in Chapter 3.

In Chapter 5, the results of the fuzzy rules based CP-classification over the UK were presented. In total, 180 different CP-classifications were defined based on varying input precipitation time series, calibration time periods and number of CPs. Each CP-classification was then assessed in terms of its skill in distinguishing between different precipitation patterns and in separating the SLP patterns. It was found that on average the CP-classifications based on daily precipitation series outperform those based on the series of the number of wet hours per day. Considering the different calibration periods, the CP-classifications based on a 20-year calibration time period performed slightly better on average compared to those derived from periods of ten years length. It was also shown that increasing the number of CPs has a positive effect on the performances of the CP-classifications up to ten CPs for three (wet day probability, precipitation intensity and separability of the SLP patterns) of the four information measures. For higher number of CPs (15 and 20), the performance of the CP-classifications improves only in terms of identifying distinguishable relationships between CP and extreme precipitation. Finding the optimal number of CPs is a trade-off between the performance of the CP-classification and the robustness of the statistical downscaling method. It was also shown that performances of the CP-classification fluctuate and thus certain CP-classifications with a small number CPs may outperform those CP-classification based on higher number of CPs. As a result, the final CP-classification used in the statistical downscaling method only consists of five (plus one unclassified) CPs, which are defined based on UK daily precipitation series over a 20 year calibration period (1960-1979).

Out of the five CPs, CP1 and CP4 are characterized by low pressure systems, CP2 and CP3 feature high pressure systems, and CP5 is associated with westerly airflow over the UK. The seasonal frequencies of the different CPs analysed based on observed SLP data were found to be similar as in previous studies using the Lamb Weather Types (Jones et al., 1993; Lamb, 1972). Cyclonic CPs (CP1 and CP4) exhibit the lowest frequency occurrences in winter, anticyclonic CPs (CP2 and CP3) are most likely to occur in summer and westerly CPs (CP5) occur most often in winter. The main difference lies in the fact that the fuzzy rules based CP-classification allows CPs with multiple SLP characteristics. For example, CP4 features cyclonic conditions in the North and south-westerly airflow in the southern part of the UK. It was also shown that the five (plus one unclassified) CPs derived from ERA-interim driven RCM data sets (CORDEX, 2014) agree well with the set of CPs based on observed NCAR gridded SLP data, both in terms of their SLP patterns and their seasonal frequencies.

Considering the relationships between CPs and extreme hourly precipitation, CP1 and CP4 show the strongest link with summer extreme hourly precipitation. CP1 and CP4 are both characterized by low pressure systems. Overall, the highest extreme precipitation probabilities in summer can be found for stations in the southern part of the UK conditioned on CP1. This circulation pattern may be linked with mesoscale convective complexes (MCCs). CP2 and CP3, both associated with high pressure conditions over the UK, exhibit relatively low extreme hourly precipitation probabilities. Nevertheless, both CPs can lead occasionally to very extreme precipitation events. It was suggested that those events may be linked with localized convective showers on days of high pressure conditions (Blenkinsop et al., 2015).

In Chapter 6 and Chapter 7, the estimated hourly precipitation extremes by the statistical downscaling method were validated for twelve different stations by calculating the 99.5th percentile error ($\overline{PE}_{99.5}$) between the estimated and observed distribution function averaged over 100 bootstrapping samples. This was done using predictors derived from observed (see Chapter 6) and ERA-interim driven RCM data sets (see Chapter 7). It was found that using ERA-interim driven RCM predictors, instead of observed predictors, has a negative effect on the statistical downscaling performance. This is probably due to the fact that the ERA-interim RCMs are not able to exactly reproduce the observed day to day weather sequences particularly in terms of precipitation. Nevertheless, the use of ERA-interim driven RCM predictors still results in overall high agreement between estimated and observed extreme hourly precipitation over the three different validation periods. It was therefore concluded, that the statistical downscaling method based on RCM predictors is able to provide reliable and robust estimates of hourly precipitation extremes. A reliable method is expected to attribute the “correct” exceedance probability to high precipitation values. Reliability can only be assessed based on observed data. A robust method simulates similar estimates for slightly perturbed data or for different time periods (Garavaglia et al., 2011). Four different approaches of finding the analogue day were compared. Overall, the best performances based on the observed and RCM simulated predictors were obtained using approach M1. For this approach, the $\overline{PE}_{99.5}$ between estimates derived from ERA-interim driven RCMs and observed hourly precipitation extremes amounts to 16.4% on average over the twelve selected stations and three independent validation periods. M1 subsamples all the precipitation records into two different seasons, instead of using the twelve calendar months as a subsample criterion. In order to find the analogue days, M1 uses daily precipitation and temperature as equally important predictors.

The stationarity assumption, that the projections of the statistical downscaling method remain valid in a warmer climate, was also assessed. For this purpose, the estimated 99.5th percentiles were validated over the ten warmest summers. In general, stationarity between large-scale atmospheric circulation and precipitation cannot be explicitly assumed. Instead, additional predictors, such as temperature (used in this thesis) or humidity, are required. Those predictors are able to represent changes in the water-holding capacity of the atmosphere, in order to capture the impact of a warmer climate on precipitation extremes. It was found that the performance of the downscaling method over the ten warmest summer is very similar to the performances over the three validation periods. It needs to be noted, however, that the temperature differences between the ten warmest summers and the colder calibration periods only represent a fraction of the projected temperature changes for the time period 2075-2100 under emission scenario RCP8.5. Another test to assess the stationarity assumption is to identify trends in the residual time series between observed and estimated annual maxima of hourly precipitation. For most stations, no signs of instationarities were found. Both stationarity tests thus suggest, using observed and RCM predictors, that the quality of the extreme hourly precipitation estimates is not downgraded under warmer conditions. However, the stationarity tests can only take into account the observed temperature differences, which are unlikely to fully represent the magnitude of future warming.

In a next step, hourly precipitation extremes are estimated based on GCM-driven RCM predictors. In this thesis, the data from two RCMs, namely RCA4 and RACMO22E, was used. The RCA4 RCM is driven by three different GCMs: CM5A-MR, ESM-LR and EC-EARTH. The RACMO22E RCM is only driven by the EC-EARTH GCM. The estimated 99.5th percentiles of hourly precipitation extremes agree well with the observed 99.5th percentiles over the reference period 1980-2005. Overall, the $\overline{PE}_{99,5}$ between estimates based on GCM-driven RCMs and observed extreme hourly precipitation is 10.25% on average over the twelve selected stations and the four different GCM-driven RCMs. This result, however, cannot be directly compared with the results derived from the ERA-interim RCMs. In contrast to the validation of the estimates based on ERA-interim driven RCM data, the validation of the estimates based on the GCM-driven RCMs was only conducted over one reference period, which overlapped with the calibration period of the statistical downscaling method.

Two different future time periods (2030-2055 and 2075-2100) were considered, in order to project hourly precipitation extremes. Regarding the period from 2075-2100, hourly precipitation extremes were projected for two different emission scenarios, RCP4.5 and RCP8.5. Estimates projected for the time period 2030-2055 were only based on emission

scenario RCP4.5. Hourly precipitation extremes increase for most of the stations under RCP4.5 (2030-2055 and 2075-2100) and for all stations under RCP8.5 (2075-2100) relative to the reference period 1980-2005. Unsurprisingly, the increases (up to 112%) are most pronounced for the future time period 2075-2100 under emission scenario RCP8.5 for which temperature increases are also the highest. Future changes in hourly precipitation extremes highly vary between the twelve stations. For example, at Tiree the average future increase over the four different GCM-driven RCMs for the future period 2075-2100 under RCP8.5 is only 12.1% compared to an average increase at Kinloss of 75.1%. Projections also vary considerably between different RCM runs and even within the same RCM run over the 100 different bootstrapping samples. At Kinloss, using the EC-EARTH driven RCA4 leads to an increase of 111.9% for the future time period 2075-2100 under RCP8.5, compared to increases of 36.9% based on the CM5A-MR driven RCA4. This is at least partly due to limitations in the RCM simulated predictors and in the statistical downscaling process as discussed in the following section. But it also needs to be mentioned that every extreme value analysis is by definition subject to uncertainties. As a consequence, the presented projections of hourly precipitation extremes should rather be interpreted as indicative of the magnitude and direction of future changes.

8.2 Limitations and potential of future developments

One main limitation of the methodology is the length and quality of the hourly precipitation records available to calibrate the statistical downscaling method. Only those observed hourly precipitation records can be used in the statistical downscaling process for the period for which ERA-interim driven RCM data is available as well. The ERA-interim driven RCM data is required for the normalization of the RCM predictor variables as described in Equation 4.15 and 4.16. ERA-interim driven RCA4 (RACMO22E) data is available from 1980 (1979) to 2010 (2012). The record length of the observed hourly precipitation highly varies between the different stations, but only few stations exist with high completeness of hourly precipitation records prior to 1980 up to the present date. Therefore, the selection of 1980 as the start year of the calibration period seems to be favourable in terms of RCM simulated and observed data sets. The end year of the calibration period was set to 2009, in order to have three different calibration periods of equal length (10 years) for the cross-validation of the statistical downscaling.

The restricted calibration period 1980-2009 leads to implications at several levels. First of all, it limits the pool of possible extreme hourly precipitation events which can be used for the

downscaling process. As a result, certain types of extreme events may not be represented over the calibration period as discussed in Section 6.3.

It is also important in terms of the stationarity assumption underlying the statistical downscaling method. It is likely that a longer calibration period leads to a better representation of those extreme events, which are expected to occur more often in a future warmer climate. But even if the calibration period was infinitely long, extreme precipitation events could occur in the future which have never been observed in the past. For this reason, the observed hourly precipitation on each analogue day was perturbed based on the relationships between precipitation duration and temperature in this thesis.

However, it could be argued that the application of the precipitation duration – temperature relationship is not physically meaningful. As discussed in Section 6.2, the linear regression slopes representing the relationships were defined over the entire sample size of each subsample, divided by each CP and season. This is because calibrating the precipitation duration-temperature relationships only on extreme events is not feasible due to very limited record lengths. But average weather conditions may not be relevant to extreme events and the response of the duration of a precipitation event to an increase in temperature may be very different under extreme weather conditions, such as convective thunderstorms. For example, (Lenderink and van Meijgaard, 2008; Molnar et al., 2014) suggested an intensification of convective precipitation events with higher temperature beyond the Clausius-Clapeyron equation due to complex physical mechanisms. Similarly, the change in precipitation duration of extreme events with higher temperature could be considerably different to the average climate response.

The limited calibration period also restricts the number of CPs, since the use of too many CPs could otherwise result in limited sampling variability of analogue days and limited records for establishing the relationships between the precipitation duration and temperature for each CP and season. On the other hand, the number of CPs needs to be large enough to represent the different characteristic types of atmospheric circulation over the UK and their specific relationship with precipitation. The final number of five (plus one unclassified) CPs is, therefore, the result of a trade-off between those two aspects. Nevertheless, it needs to be noted that using a higher number of CPs would lead to further improvements in the CP-precipitation relationships and in the separability of the atmospheric pressure patterns as discussed in Section 5.2.

One possibility to increase the sample sizes for each subsample of CPs and seasons would be the pooling of precipitation events over multiple stations. The MIDAS Land Surface Stations

database (NCAS British Atmospheric Data Centre, 2012) comprises hourly precipitation records at 530 different stations over the UK. For each of the twelve selected stations, additional records from nearby stations within a certain radius could be pooled together. Although those additional records are not as complete as the records used in this thesis, they still contain valuable information to incorporate into the downscaling process. As a result, an increase in the number of CPs and robust duration-temperature relationships could become feasible at the same time. It could also make the calibration of the duration-temperature relationships using only precipitation extremes possible. As discussed above, the perturbation of hourly precipitation extremes based on relationships that were defined over the entire sample size, which means over average weather conditions, is questionable. Instead, calibrating the duration-temperature relationships specifically for precipitation extremes is likely to be more representative of potential changes in the future.

The second limitation is related to the quality of the predictors. In general, any projection based on a statistical downscaling method can only be as good as the predictors that are simulated by the climate models. In Section 7.1, it was found that both ERA-interim driven RCMs underestimate the observed positive trend in temperature between 1980 and 2009. This may imply that future increases in the temperature variable derived from the two RCMs are also underestimated. As discussed in Section 7.2, RCA4 overestimates daily precipitation and both, RCA4 and RACMO22E, show a tendency to underestimate the observed daily temperature. It also needs to be mentioned that both RCMs rely on parameterization schemes to simplify the physical complexity of convective precipitation events and assumptions are required for the parameterization to remain valid under a warmer climate. This leads to uncertainties particularly in terms of precipitation extremes which affects the results of the statistical downscaling presented in this thesis. It is therefore advisable to increase the number of RCMs used in future studies to provide a more comprehensive picture of the uncertainties caused by the RCMs.

Another possible future development of the statistical downscaling method would be to use RCMs at very high spatial resolution (1km-5km). These high-resolution models are able to explicitly resolve physical processes of convection and thus represent convective precipitation extremes more realistically (Kendon et al., 2012). Nevertheless, Kendon et al. (2014) showed that extreme hourly precipitation tend to be overestimated in the 1.5km RCM over the UK. Applying the statistical downscaling process presented in this thesis could reduce the biases in the 1.5km RCMs projections and provide site-specific hourly precipitation extremes. However, due to computational costs, high-resolution RCM runs are limited and previous studies have been restricted to a single season (Hohenegger et al., 2008), selected events (Attema et al.,

2014) or small spatial domains (Chan et al., 2014a; Chan et al., 2014b; Chan, 2013; Kendon et al., 2012).

The third limitation lies in the selection of the predictors applied in this thesis. Daily precipitation and temperature conditioned on CPs may not be enough to describe the complex mechanisms sufficiently that lead to extreme precipitation events. Incorporating additional RCM predictor variables into the statistical downscaling method could therefore enhance the quality of the projected hourly precipitation extremes. Examples of potential variables are measures of humidity or the CAPE (Convective Available Potential Energy) index. Those variables provide information of the moisture flux in the atmosphere and could improve particularly the representation of convective activity under a warmer climate. In this context it is interesting to note that Berg et al. (2013) suggested a decrease in convective precipitation intensities for high temperatures due to a decline in relative humidity. Holley et al. (2014) recently developed a climatology of the CAPE index over the UK. Relating the results of this study with the predictors used in this thesis could help to understand the complex physical mechanisms behind the occurrence of extreme precipitation better and therefore improve the projections of the statistical downscaling method.

Finally, it needs to be mentioned that there is a current mismatch between what is defined as an extreme event for flood estimation purposes in engineering design and the definition of an extreme event in climate studies. For the estimation of floods, engineers are generally interested in extreme precipitation events such as the 1% annual exceedance probability (AEP) event, which will on average occur once every 100 years. On the other hand, climate studies tend to focus on the projection of “moderate” extremes such as the 99th percentile or the annual maximum precipitation events. In this thesis, the focus was on the 99.5th percentile of extreme hourly precipitation, which will on average occur at least once every year. It is therefore questionable whether the results of climate studies, including this one, can be extrapolated to the more extreme events that are often required by engineers for urban planning and the design of hydraulic flood prevention measures (Westra et al., 2014).

8.3 Concluding remarks

Notwithstanding the above mentioned limitations, this thesis is the first research attempt that classifies characteristic atmospheric circulation patterns for the UK using a fuzzy rules based CP-classification and downscales RCM predictors to site-specific UK extreme hourly precipitation events for two future time periods using two emission scenarios. For this purpose, a statistical downscaling method was developed, which combines daily precipitation

and temperature as equally important predictors to find similar weather conditions in the past (analogue day). In a next step, the hourly precipitation records on the analogue days were perturbed based on precipitation duration-temperature relationships.

Results vary considerably between different stations. The highest increases in extreme hourly precipitation were found at Northolt (38.3%), Eskdalemuir (45.8%) and Kinloss (75.1%) for the future time period 2075-2100 under the high emission scenario (RCP8.5) averaged over the four different GCM-driven RCMs. Eskdalemuir is the station with the highest elevation and Northolt is the station with the highest mean and extreme temperatures. It was argued that both stations can be linked with increased convective activity and concluded that increases in extreme hourly precipitation tend to be most pronounced for those stations that are highly affected by convective precipitation events.

For impact studies to assess the potential higher risk of flash floods in the future, the newly developed statistical downscaling method in this thesis is relatively simple to understand and can be readily applied. The method is able to provide future projections of hourly precipitation extremes at any site in the UK, for which observed hourly precipitation records are available. It could be also easily adapted to project multi-hour precipitation extremes depending on the flood forecasting context. For example, the Extreme Rainfall Alert (ERA), which is a warning system for surface water flooding, was issued for 1hr, 3hr and 6hr precipitation values (Hurford et al., 2012). In order to project precipitation extremes outside the UK, a re-definition of the CP-classification would be required. Two different stationarity tests were applied, indicating that the downscaling method remain valid in a future warmer climate. However, the stationarity assumption could only be partly assessed since future projections of temperature increases for the time period 2075-2100 under the emission scenario RCP8.5 highly exceed observed temperature differences in the past.

In terms of impact studies, it is important to note that extreme precipitation events occur within a broader meteorological context which should be considered to assess the severity of the resultant flood sufficiently. For example, UK extreme precipitation events over coastal areas are likely to be more severe on days when elevated tides and/or storm surges also occur (Svensson and Jones, 2004).

However, the aim of this thesis was to develop a statistical downscaling method to “only” project hourly precipitation extremes for the UK. It was found that changes in hourly extremes can vary strongly between different RCM runs and even within the same RCM run over the 100 different bootstrapping samples. This is at least partly due to limitations in the RCM simulated predictors and in the statistical downscaling process. As a consequence, there is a need to

include more RCMs to allow ensemble-based probabilistic projections of extreme hourly precipitation in the future. In terms of the statistical downscaling method, incorporating additional predictors could help to represent extreme hourly precipitation events more realistically and thus increase the reliability and robustness of the projections.

Bibliography

- Aarts, E. and Korst, J., 1988. Simulated annealing and Boltzmann machines. New York, NY; John Wiley and Sons Inc.
- Adler, R.F. et al., 2008. Relationships between global precipitation and surface temperature on interannual and longer timescales (1979-2006). *Journal of Geophysical Research*, 113: D22104.
- Alexander, L. and Jones, P., 2000. Updated precipitation series for the UK and discussion of recent extremes. *Atmospheric Science Letters*, 1(2): 142-150.
- Attema, J.J., Loriaux, J.M. and Lenderink, G., 2014. Extreme precipitation response to climate perturbations in an atmospheric mesoscale model. *Environmental Research Letters*, 9(1): 014003.
- Ban, N., Schmidli, J. and Schär, C., 2014. Evaluation of the convection-resolving regional climate modeling approach in decade-long simulations. *Journal of Geophysical Research: Atmospheres*, 119(13): 7889-7907.
- Ban, N., Schmidli, J. and Schär, C., 2015. Heavy precipitation in a changing climate: Does short-term summer precipitation increase faster? *Geophysical Research Letters*, 42(4): 1165-1172.
- Bárdossy, A., 2010. Atmospheric circulation pattern classification for South-West Germany using hydrological variables. *Physics and Chemistry of the Earth, Parts A/B/C*, 35(9): 498-506.
- Bárdossy, A., Duckstein, L. and Bogardi, I., 1995. Fuzzy rule-based classification of atmospheric circulation patterns. *International Journal of Climatology*, 15(10): 1087-1097.
- Bárdossy, A. and Pegram, G., 2011. Downscaling precipitation using regional climate models and circulation patterns toward hydrology. *Water Resour. Res.*, 47(4): W04505.
- Bárdossy, A., Stehlík, J. and Caspary, H.-J., 2002. Automated objective classification of daily circulation patterns for precipitation and temperature downscaling based on optimized fuzzy rules. *Climate Research*, 23(1): 11-22.
- Barry, R.G. and Chorley, R.J., 2009. *Atmosphere, weather and climate*. Routledge.
- Beard, P., 2010. Hourly precipitation observations over the UK: diurnal cycles and changes in intensity, Msc. thesis, University of East Anglia.
- Beck, C. and Philipp, A., 2010. Evaluation and comparison of circulation type classifications for the European domain. *Physics and Chemistry of the Earth, Parts A/B/C*, 35(9-12): 374-387.
- Beck, F., 2013. Generation of spatially correlated synthetic rainfall time series in high temporal resolution. Ph.D. thesis, University of Stuttgart.
- Beck, F. and Bárdossy, A., 2013. Indirect downscaling of hourly precipitation based on atmospheric circulation and temperature. *Hydrology & Earth System Sciences*, 17(12).
- Berg, P., Moseley, C. and Haerter, J.O., 2013. Strong increase in convective precipitation in response to higher temperatures. *Nature Geoscience*, 6(3): 181-185.
- Beuchat, X., Schaepli, B., Soutter, M. and Mermoud, A., 2011. Toward a robust method for subdaily rainfall downscaling from daily data. *Water Resources Research*, 47(9): W09524.
- Beuchat, X., Schaepli, B., Soutter, M. and Mermoud, A., 2012. A robust framework for probabilistic precipitations downscaling from an ensemble of climate predictions applied to Switzerland. *Journal of Geophysical Research*, 117(D3): D03115.
- Blenkinsop, S., Chan, S., Kendon, E., Roberts, N. and Fowler, H., 2015. Temperature influences on intense UK hourly precipitation and dependency on large-scale circulation. *Environmental Research Letters*, 10(5): 054021.

- Brigode, P. et al., 2013. Linking ENSO and heavy rainfall events over Coastal British Columbia through a weather pattern classification. *Hydrology and Earth System Sciences*, 17(4): 1455-1473.
- Buishand, T.A. and Brandsma, T., 2001. Multisite simulation of daily precipitation and temperature in the Rhine basin by nearest-neighbor resampling. *Water Resources Research*, 37(11): 2761-2776.
- Buonomo, E., Jones, R., Huntingford, C. and Hannaford, J., 2007. On the robustness of changes in extreme precipitation over Europe from two high resolution climate change simulations. *Quarterly Journal of the Royal Meteorological Society*, 133(622): 65-81.
- Burton, A., Fowler, H., Blenkinsop, S. and Kilsby, C., 2010. Downscaling transient climate change using a Neyman–Scott Rectangular Pulses stochastic rainfall model. *Journal of Hydrology*, 381(1): 18-32.
- Caliński, T. and Harabasz, J., 1974. A dendrite method for cluster analysis. *Communications in Statistics-theory and Methods*, 3(1): 1-27.
- Champion, A.J., Allan, R.P. and Lavers, D.A., 2015. Atmospheric rivers do not explain UK summer extreme rainfall. *Journal of Geophysical Research: Atmospheres*, 120(14): 6731-6741.
- Chan, S., Kendon, E., Fowler, H., Blenkinsop, S. and Roberts, N., 2014a. Projected increases in summer and winter UK sub-daily precipitation extremes from high-resolution regional climate models. *Environmental Research Letters*, 9(8): 084019.
- Chan, S.C. et al., 2014b. The value of high-resolution Met Office regional climate models in the simulation of multihourly precipitation extremes. *Journal of Climate*, 27(16): 6155-6174.
- Chan, S.C., Kendon, E.J., Roberts, N.M., Fowler, H.J. and Blenkinsop, S., 2016. Downturn in scaling of UK extreme rainfall with temperature for future hottest days. *Nature Geosci*, 9(1): 24-28.
- Chan, S.C.a.K., E. J. and Fowler, H. J. and Blenkinsop, S. and Ferro, C. A. T. and Stephenson, D. B., 2013. Does increasing the spatial resolution of a regional climate model improve the simulated daily precipitation? *Climate Dynamics*, 41(5-6): 1475-1495.
- Christiansen, B., 2007. Atmospheric circulation regimes: Can cluster analysis provide the number? *Journal of Climate*, 20(10): 2229-2250.
- Conway, D. and Jones, P.D., 1998. The use of weather types and air flow indices for GCM downscaling. *Journal of Hydrology*, 212–213(0): 348-361.
- Conway, D., Wilby, R. and Jones, P., 1996. Precipitation and air flow indices over the British Isle. *Climate Research*, 7(2): 169-183
- CORDEX, access date Oct 2014. CORDEX climate data archive. <http://cordex.dmi.dk/joomla/>
- Corte-Real, J., Qian, B. and Xu, H., 1998. Regional climate change in Portugal: precipitation variability associated with large-scale atmospheric circulation. *International Journal of Climatology*, 18(6): 619-635.
- Corte-Real, J., Quian, B. and Xu, H., 1999. Circulation patterns, daily precipitation in Portugal and implications for climate change simulated by the second Hadley Centre GCM. *Climate Dynamics*, 15(12): 921-935.
- Dale, M. et al., 2015. New climate change rainfall estimates for sustainable drainage. *Proceedings of the Institution of Civil Engineers. Engineering Sustainability*. doi: 10.1680/jensu.15.00030.
- DEFRA (Department for Environment, Food and Rural Affairs), 2006. Flood and Coastal Defence Appraisal Guidance FCDPAG3 Economic Appraisal – Supplementary Note to Operating Authorities – Climate Change Impacts. Defra, London, UK.
- de Lima, M.I.P., Santo, F.E., Ramos, A.M. and Trigo, R.M., 2014. Trends and correlations in annual extreme precipitation indices for mainland Portugal, 1941–2007. *Theoretical and Applied Climatology*, 119(1-2): 55-75.

- Deidda, R. et al., 2013. Regional climate models' performance in representing precipitation and temperature over selected Mediterranean areas. *Hydrology & Earth System Sciences*: 5041-5059.
- Dierer, S. and Schubiger, F., 2008. Evaluation of the Kain-Fritsch/Bechtold convection scheme. *COSMO Newsletter*, 7: 37-43.
- Documentation—Cy31r1, IFS. ECMWF, 2007. Part IV: Physical Processes, European Centre for Medium-Range Weather Forecasts. Shinfield Park, Reading, England.
- Environment Agency, 2011. Adapting to Climate Change: Advice for Flood and Coastal Erosion Risk Management Authorities. Environment Agency, London, UK.
- Esteban, P., Jones, P.D., Martín-Vide, J. and Mases, M., 2005. Atmospheric circulation patterns related to heavy snowfall days in Andorra, Pyrenees. *International Journal of Climatology*, 25(3): 319-329.
- Fernández-González, S. et al., 2012. Connection between NAO, weather types and precipitation in León, Spain (1948–2008). *International Journal of Climatology*, 32(14): 2181-2196.
- Fowler, H. and Ekström, M., 2009. Multi-model ensemble estimates of climate change impacts on UK seasonal precipitation extremes. *International Journal of Climatology*, 29(3): 385-416.
- Fowler, H. and Kilsby, C., 2003. Implications of changes in seasonal and annual extreme rainfall. *Geophysical Research Letters*, 30(13): 1720.
- Gangopadhyay, S., Clark, M. and Rajagopalan, B., 2005. Statistical downscaling using K-nearest neighbors. *Water Resources Research*, 41(2): W02024.
- Garavaglia, F. et al., 2010. Introducing a rainfall compound distribution model based on weather pattern sub-sampling. *HESS*, 14: 951-964.
- Garavaglia, F. et al., 2011. Reliability and robustness of rainfall compound distribution model based on weather pattern sub-sampling. *Hydrology and Earth System Sciences Discussions*, 15: 519-532.
- Gong, X. and Richman, M.B., 1995. On the application of cluster analysis to growing season precipitation data in North America east of the Rockies. *Journal of Climate*, 8(4): 897-931.
- Goodess, C., 2000. The construction of daily rainfall scenarios for Mediterranean sites using a circulation-type approach to downscaling, Ph.D. thesis, University of East Anglia.
- Goodess, C.M., 2012. How is the frequency, location and severity of extreme events likely to change up to 2060? *Environmental Science & Policy*, 27: 4-14.
- Goodess, C.M. and Palutikof, J.P., 1998. Development of daily rainfall scenarios for southeast Spain using a circulation-type approach to downscaling. *International Journal of Climatology*, 18(10): 1051-1083.
- Gregersen, I.B. et al., 2013. Assessing future climatic changes of rainfall extremes at small spatio-temporal scales. *Climatic Change*, 118(3-4): 783-797.
- Groth, S.L. and MacCracken, M.C., 1991. The use of general circulation models to predict regional climatic change. *Journal of Climate*; (United States), 4(3): 286-303.
- Gutiérrez, J., San-Martín, D., Brands, S., Manzanos, R. and Herrera, S., 2013. Reassessing Statistical Downscaling Techniques for Their Robust Application under Climate Change Conditions. *Journal of Climate*, 26(1): 171-188.
- Haberlandt, U., Belli, A. and Bárdossy, A., 2015. Statistical downscaling of precipitation using a stochastic rainfall model conditioned on circulation patterns—an evaluation of assumptions. *International Journal of Climatology*, 35(3): 417-432.
- Haerter, J.O. and Berg, P., 2009. Unexpected rise in extreme precipitation caused by a shift in rain type? *Nature Geoscience*, 2(6): 372-373.
- Hand, W.H., Fox, N.I. and Collier, C.G., 2004. A study of twentieth-century extreme rainfall events in the United Kingdom with implications for forecasting. *Meteorological Applications*, 11(1): 15-31.

- Hanel, M. and Buishand, T.A., 2010. On the value of hourly precipitation extremes in regional climate model simulations. *Journal of Hydrology*, 393(3): 265-273.
- Hardwick Jones, R., Westra, S. and Sharma, A., 2010. Observed relationships between extreme sub-daily precipitation, surface temperature, and relative humidity. *Geophysical Research Letters*, 37(22): L22805.
- Hartmann, D. et al., 2013. Observations: atmosphere and surface. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, Cambridge University Press, pp. 159-254.
- Haylock, M.R., Cawley, G.C., Harpham, C., Wilby, R.L. and Goodess, C.M., 2006. Downscaling heavy precipitation over the United Kingdom: a comparison of dynamical and statistical methods and their future scenarios. *International Journal of Climatology*, 26(10): 1397-1415.
- He, Y. et al., 2013. Flood inundation dynamics and socioeconomic vulnerability under environmental change, *Vulnerability of Water Resources to Climate*. Climate Vulnerability. Elsevier Sciences, pp. 241-255.
- Hess, P. and Brezowsky, H., 1969. *Katalog der Grosswetterlagen Europas*, 15. Deutscher Wetterdienst.
- Hohenegger, C., Brockhaus, P. and Schar, C., 2008. Towards climate simulations at cloud-resolving scales. *Meteorologische Zeitschrift*, 17(4): 383-394.
- Holley, D., Dorling, S., Steele, C. and Earl, N., 2014. A climatology of convective available potential energy in Great Britain. *International Journal of Climatology*, 34(14): 3811-3824.
- Hundecha, Y. and Bárdossy, A., 2005. Trends in daily precipitation and temperature extremes across western Germany in the second half of the 20th century. *International Journal of Climatology*, 25(9): 1189-1202.
- Hundecha, Y. and Bárdossy, A., 2008. Statistical downscaling of extremes of daily precipitation and temperature and construction of their future scenarios. *International Journal of Climatology*, 28(5): 589-610.
- Hurford, A., Parker, D., Priest, S. and Lumbroso, D., 2012. Validating the return period of rainfall thresholds used for Extreme Rainfall Alerts by linking rainfall intensities with observed surface water flood events. *Journal of Flood Risk Management*, 5(2): 134-142.
- Hurrell, J. and Trenberth, K., 2013. *The Climate Data Guide: NCAR Sea Level Pressure*. In: N.C.f.A. Research (Editor). <https://climatedataguide.ucar.edu/climate-data/ncar-sea-level-pressure>
- Huth, R., 1996. An intercomparison of computer-assisted circulation classification methods. *International Journal of Climatology*, 16(8): 893-922.
- Huth, R. et al., 2008. Classifications of Atmospheric Circulation Patterns. *Annals of the New York Academy of Sciences*, 1146(1): 105-152.
- IPCC, 2013. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 1535.
- Jacob, D. et al., 2014. EURO-CORDEX: new high-resolution climate change projections for European impact research. *Reg Environ Change*, 14(2): 563-578.
- Jenkinson, A. and Collison, F., 1977. An initial climatology of gales over the North Sea. *Synoptic Climatology Branch Memorandum*, 62: 18.
- Jones, C.G., Willén, U., Ullerstig, A. and Hansson, U., 2004. The Rossby Centre regional atmospheric climate model part I: model climatology and performance for the present climate over Europe. *AMBIO: A Journal of the Human Environment*, 33(4): 199-210.

- Jones, M.R., Blenkinsop, S., Fowler, H.J. and Kilsby, C.G., 2014. Objective classification of extreme rainfall regions for the UK and updated estimates of trends in regional extreme rainfall. *International Journal of Climatology*, 34(3): 751-765.
- Jones, M.R., Fowler, H.J., Kilsby, C.G. and Blenkinsop, S., 2013. An assessment of changes in seasonal and annual extreme rainfall in the UK between 1961 and 2009. *International Journal of Climatology*, 33(5): 1178-1194.
- Jones, P., Hulme, M. and Briffa, K., 1993. A comparison of Lamb circulation types with an objective classification scheme. *International Journal of Climatology*, 13(6): 655-663.
- Jones, P., Kilsby, C., Harpham, C., Glenis, V. and Burton, A., 2009. UK Climate Projections science report: Projections of future daily climate for the UK from the Weather Generator. University of Newcastle.
- Karl, T.R., Wang, W.-C., Schlesinger, M.E., Knight, R.W. and Portman, D., 1990. A method of relating general circulation model simulated climate to the observed local climate. Part I: Seasonal statistics. *Journal of Climate*, 3(10): 1053-1079.
- Kendon, E. and Clark, R., 2008. Reliability of future changes in heavy rainfall over the UK, BHS 10th National Hydrology Symposium.
- Kendon, E.J. et al., 2014. Heavier summer downpours with climate change revealed by weather forecast resolution model. *Nature Climate Change*, 4: 570-576.
- Kendon, E.J., Roberts, N.M., Senior, C.A. and Roberts, M.J., 2012. Realism of rainfall in a very high resolution regional climate model. *Journal of Climate*, 25: 5791-5806.
- Kendon, E.J., Rowell, D.P. and Jones, R.G., 2010. Mechanisms and reliability of future projected changes in daily precipitation. *Climate Dynamics*, 35(2): 489-509.
- Kharin, V.V., Zwiers, F.W., Zhang, X. and Hegerl, G.C., 2007. Changes in temperature and precipitation extremes in the IPCC ensemble of global coupled model simulations. *Journal of Climate*, 20(8): 1419-1444.
- Kilsby, C. et al., 2007. A daily weather generator for use in climate change studies. *Environmental Modelling & Software*, 22(12): 1705-1719.
- Kirtman, B. et al., 2013. Near-term climate change: projections and predictability. *Climate change*: 953-1028.
- Kistler, R. et al., 2001. The NCEP-NCAR 50-year reanalysis: Monthly means CD-ROM and documentation. *Bulletin of the American Meteorological Society*, 82(2): 247-267.
- Koutsoyiannis, D. and Onof, C., 2001. Rainfall disaggregation using adjusting procedures on a Poisson cluster model. *Journal of Hydrology*, 246(1): 109-122.
- Lamb, H.H., 1972. British Isles weather types and a register of the daily sequence of circulation patterns 1861-1971. *Geophysical Memoir 116*, HMSO, London (UK), 85 pp.
- Lamb, H.H., 1977. *Climate: present, past and future. Volume 2. Climatic history and the future.* Methuen & Company.
- Larsen, A.N., Gregersen, I.B., Christensen, O., Linde, J.J. and Mikkelsen, P.S., 2009. Potential future increase in extreme one-hour precipitation events over Europe due to climate change. *Water science and technology: a journal of the International Association on Water Pollution Research*, 60(9): 2205.
- Leahy, P.G. and Kiely, G., 2011. Short duration rainfall extremes in Ireland: influence of climatic variability. *Water resources management*, 25(3): 987-1003.
- Lee, T. and Jeong, C., 2014. Nonparametric statistical temporal downscaling of daily precipitation to hourly precipitation and implications for climate change scenarios. *Journal of Hydrology*, 510: 182-196.
- Lenderink, G., Buishand, A., Van Deursen, W. and others, 2007. Estimates of future discharges of the river Rhine using two scenario methodologies: direct versus delta approach. *Hydrology and Earth System Sciences Discussions*, 11(3): 1145-1159.
- Lenderink, G., Mok, H., Lee, T. and Van Oldenborgh, G., 2011. Scaling and trends of hourly precipitation extremes in two different climate zones--Hong Kong and the Netherlands. *Hydrology & Earth System Sciences*, 15(9): 3033-3041.

- Lenderink, G. and van Meijgaard, E., 2008. Increase in hourly precipitation extremes beyond expectations from temperature changes. *Nature Geoscience*, 1(8): 511-514.
- Little, M.A., Rodda, H.J. and McSharry, P.E., 2008. Bayesian objective classification of extreme UK daily rainfall for flood risk applications. *Hydrology and Earth System Sciences Discussions*, 5(6): 3033-3060.
- Madsen, H., Lawrence, D., Lang, M., Martinkova, M. and Kjeldsen, T., 2014. Review of trend analysis and climate change projections of extreme precipitation and floods in Europe. *Journal of Hydrology*, 519: 3634-3650.
- Maheras, P., 1989. Delimitation of the summer-dry period in Greece according to the frequency of weather-types. *Theoretical and Applied Climatology*, 39(3): 171-176.
- Maraun, D., 2013. Can Quantile Mapping be Used for Downscaling?, EGU General Assembly Conference Abstracts, pp. 13068.
- Maraun, D., Osborn, T. and Gillett, N., 2008. United Kingdom daily precipitation intensity: improved early data, error estimates and an update from 2000 to 2006. *International Journal of Climatology*, 28(6): 833-842.
- Maraun, D., Rust, H. and Osborn, T., 2009. The annual cycle of heavy precipitation across the United Kingdom: a model based on extreme value statistics. *International Journal of Climatology*, 29(12): 1731-1744.
- Maraun, D., Rust, H.W. and Osborn, T.J., 2010a. Synoptic airflow and UK daily precipitation extremes. *Extremes*, 13(2): 133-153.
- Maraun, D. et al., 2010b. Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user. *Reviews of Geophysics*, 48(3): RG3003.
- Melber, M. and Oswald, J., Personal communication. Wetterwarte Stoetten, Deutscher Wetterdienst (DWD). 2012
- Met Office, Weather extremes. Met Office. Web page.
<http://www.metoffice.gov.uk/climate/uk/extremes/>
- Mezghani, A. and Hingray, B., 2009. A combined downscaling-disaggregation weather generator for stochastic generation of multisite hourly weather variables over complex terrain: Development and multi-scale validation for the Upper Rhone River basin. *Journal of Hydrology*, 377(3): 245-260.
- Min, S.K., Zhang, X., Zwiers, F.W. and Hegerl, G.C., 2011. Human contribution to more-intense precipitation extremes. *Nature*, 470(7334): 378-381.
- Mishra, V., Wallace, J.M. and Lettenmaier, D.P., 2012. Relationship between hourly extreme precipitation and local air temperature in the United States. *Geophysical Research Letters*, 39(16).
- Molnar, P. and Burlando, P., 2005. Preservation of rainfall properties in stochastic disaggregation by a simple random cascade model. *Atmospheric Research*, 77(1): 137-151.
- Molnar, P. et al., 2014. Little evidence for super Clausius–Clapeyron scaling of intense rainstorm properties with air temperature. *Hydrol. Earth Syst. Sci. Discuss.*, 11(7): 8923-8948.
- NCAS British Atmospheric Data Centre, 2012. Met Office Integrated Data Archive System (MIDAS) Land and Surface Stations Data (1853-current). In: U.M. Office (Editor).
http://badc.nerc.ac.uk/view/badc.nerc.ac.uk_ATOM_dataent_ukmo-midas
- Nguyen, V.-T.-V., Nguyen, T.-D. and Cung, A., 2007. A statistical approach to downscaling of sub-daily extreme rainfall processes for climate-related impact studies in urban areas. *Water science and technology: water supply*, 7(2): 183-192.
- Ntegeka, V. and Willems, P., 2008. Trends and multidecadal oscillations in rainfall extremes, based on a more than 100-year time series of 10 min rainfall intensities at Uccle, Belgium. *Water Resources Research*, 44(7): W07402.

- Olsson, J., Uvo, C. and Jinno, K., 2001. Statistical atmospheric downscaling of short-term extreme rainfall by neural networks. *Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere*, 26(9): 695-700.
- Olsson, J., Willén, U. and Kawamura, A., 2012. Downscaling extreme short-term regional climate model precipitation for urban hydrological applications. *Hydrology research*, 43(4): 341-351.
- Onof, C. and Arnbjerg-Nielsen, K., 2009. Quantification of anticipated future changes in high resolution design rainfall for urban areas. *Atmospheric Research*, 92(3): 350-363.
- Orlowsky, B. and Seneviratne, S.I., 2012. Global changes in extreme events: regional and seasonal dimension. *Climatic Change*, 110(3): 669-696.
- Overland, J. and Hiester, T., 1980. Development of a synoptic climatology for the northeast Gulf of Alaska. *Journal of Applied Meteorology*, 19(1): 1-14.
- Pall, P., Allen, M. and Stone, D., 2007. Testing the Clausius–Clapeyron constraint on changes in extreme precipitation under CO2 warming. *Climate Dynamics*, 28(4): 351-363.
- Palutikof, J., Winkler, J., Goodess, C. and Andresen, J., 1997. The simulation of daily temperature time series from GCM output. Part I: Comparison of model data with observations. *Journal of Climate*, 10(10): 2497-2513.
- Paulson, K.S., 2010. Trends in the incidence of rain rates associated with outages on fixed links operating above 10 GHz in the southern United Kingdom. *Radio Science*, 45(1): RS1011.
- Philipp, A. et al., 2010. Cost733cat-a database of weather and circulation type classifications. *Physics and Chemistry of the Earth, Parts A/B/C*, 35(9-12): 360-373.
- Philipp, A. et al., 2007. Long-Term Variability of Daily North Atlantic–European Pressure Patterns since 1850 Classified by Simulated Annealing Clustering. *Journal of Climate*, 20(16): 4065-4095.
- Richman, M.B., 1981. Obliquely rotated principal components: An improved meteorological map typing technique? *Journal of Applied Meteorology*, 20(10): 1145-1159.
- Rodda, H.J., Little, M.A., Wood, R.G., MacDougall, N. and McSharry, P.E., 2009. A digital archive of extreme rainfalls in the British Isles from 1866 to 1968 based on British Rainfall. *weather*, 64(3): 71-75.
- Rousseeuw, P.J., 1987. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of computational and applied mathematics*, 20: 53-65.
- Rummukainen, M. et al., 2001. A regional climate model for northern Europe: model description and results from the downscaling of two GCM control simulations. *Climate Dynamics*, 17(5-6): 339-359.
- Samuelsson, P. et al., 2011. The Rossby Centre Regional Climate model RCA3: model description and performance. *Tellus A*, 63(1): 4-23.
- Schmith, T., 2008. Stationarity of regression relationships: Application to empirical downscaling. *Journal of Climate*, 21(17): 4529-4537.
- Seneviratne, S. et al., 2012. Changes in climate extremes and their impacts on the natural physical environment: An overview of the IPCC SREX report, Managing the risks of extreme events and disasters to advance climate change adaptation. Cambridge University Press, Cambridge, UK, pp. 109-230.
- Shabalova, M., van Deursen, W. and Buishand, T., 2003. Assessing future discharge of the river Rhine using regional climate model integrations and a hydrological model. *Climate Research*, 23(3): 233-246.
- SMHI, access date Apr 2015. Rossby Centre regional atmospheric model, RCA4. Web page. <http://www.smhi.se/en/research/research-departments/climate-research-rossby-centre2-552/rossby-centre-regional-atmospheric-model-rca4-1.16562>
- Sponsor, N., 2012. Improvement and extension of the coupling between RACMO2 and LOTOS-EUROS Technical report.

- Stehlik, J. and Bárdossy, A., 2003. Statistical comparison of European circulation patterns and development of a continental scale classification. *Theoretical and Applied Climatology*, 76(1-2): 31-46.
- Sunyer, M. et al., 2015. Inter-comparison of statistical downscaling methods for projection of extreme precipitation in Europe. *Hydrology and Earth System Sciences*, 19(4): 1827-1847.
- Svensson, C., Jakob, D. and Reed, D.W., 2002. Diurnal characteristics of heavy precipitation according to weather type at an upland site in Scotland. *International Journal of Climatology*, 22(5): 569-585.
- Svensson, C. and Jones, D.A., 2004. Dependence between sea surge, river flow and precipitation in south and west Britain. *Hydrology and Earth System Sciences Discussions*, 8(5): 973-992.
- Sweeney, J.C. and O'Hare, G.P., 1992. Geographical variations in precipitation yields and circulation types in Britain and Ireland. *Transactions of the Institute of British Geographers*: 448-463.
- Tebaldi, C., Hayhoe, K., Arblaster, J.M. and Meehl, G.A., 2006. Going to the extremes. *Climatic Change*, 79(3): 185-211.
- Trenberth, K.E., 1999. Conceptual framework for changes of extremes of the hydrological cycle with climate change, *Weather and Climate Extremes*. Springer, pp. 327-339.
- Trenberth, K.E. and Paolino Jr, D.A., 1980. The Northern Hemisphere sea-level pressure data set: Trends, errors and discontinuities. *Monthly Weather Review*, 108(7): 855-872.
- Van den Besselaar, E., Haylock, M., Van der Schrier, G. and Klein Tank, A., 2011. A European daily high-resolution observational gridded data set of sea level pressure. *Journal of Geophysical Research: Atmospheres* (1984–2012), 116(D11).
- Van Meijgaard, E. et al., 2008. The KNMI regional atmospheric climate model RACMO version 2.1., Technical Report, TR-302, Koninklijk Nederlands Meteorologisch Instituut.
- Vrac, M., Marbaix, P., Paillard, D. and Naveau, P., 2007. Non-linear statistical downscaling of present and LGM precipitation and temperatures over Europe. *Climate of the Past*, 3(4): 669-682.
- Wang, X.L. and Feng, Y., 2010. *RHtests V3 User Manual*, Toronto, Canada.
- Westra, S. et al., 2014. Future changes to the intensity and frequency of short-duration extreme rainfall. *Reviews of Geophysics*, 52(3): 522-555.
- Wigley, T., Lough, J. and Jones, P., 1984. Spatial patterns of precipitation in England and Wales and a revised, homogeneous England and Wales precipitation series. *Journal of Climatology*, 4(1): 1-25.
- Wilby, R., Tomlinson, O. and Dawson, C., 2003. Multi-site simulation of precipitation by conditional resampling. *Climate Research*, 23(3): 183-194.
- Wilby, R. and Wigley, T., 2000. Precipitation predictors for downscaling: observed and general circulation model relationships. *International Journal of Climatology*, 20(6): 641-661.
- Wilby, R.L., Dawson, C.W. and Barrow, E.M., 2002. SDSM—a decision support tool for the assessment of regional climate change impacts. *Environmental Modelling & Software*, 17(2): 145-157.
- Wilby, R.L. and Quinn, N.W., 2013. Reconstructing multi-decadal variations in fluvial flood risk using atmospheric circulation patterns. *Journal of Hydrology*, 487: 109-121.
- Wilby, R.L. and Wigley, T., 1997. Downscaling general circulation model output: a review of methods and limitations. *Progress in Physical Geography*, 21(4): 530-548.
- Wilby, R.L. et al., 1998. Statistical downscaling of general circulation model output: A comparison of methods. *Water Resources Research*, 34(11): 2995-3008.
- Willems, P., Arnbjerg-Nielsen, K., Olsson, J. and Nguyen, V., 2012. Climate change impact assessment on urban rainfall extremes and urban drainage: methods and shortcomings. *Atmospheric Research*, 103: 106-118.

- Willems, P. and Vrac, M., 2011. Statistical precipitation downscaling for small-scale hydrological impact investigations of climate change. *Journal of Hydrology*, 402(3): 193-205.
- Wilson, L.L., Lettenmaier, D.P. and Skillingstad, E., 1992. A hierarchical stochastic model of large-scale atmospheric circulation patterns and multiple station daily precipitation. *Journal of Geophysical Research*, 97(D3): 2791-2809.
- Yang, W., Bárdossy, A. and Caspary, H.-J., 2010. Downscaling daily precipitation time series using a combined circulation-and regression-based approach. *Theoretical and Applied Climatology*, 102(3-4): 439-454.
- Zhang, X., 2012. ETCCDI/CRD Climate Change Indices. Web page.
<http://etccdi.pacificclimate.org/index.shtml>
- Zorita, E. and Von Storch, H., 1999. The analog method as a simple statistical downscaling technique: comparison with more complicated methods. *Journal of climate*, 12(8): 2474-2489.