

# Defining dry/wet spells for point observations, observed area averages, and regional climate model gridboxes in Europe

Lars Barring<sup>1,2,\*</sup>, Tom Holt<sup>3</sup>, Maj-Lena Linderson<sup>1</sup>, Maciej Radziejewski<sup>4,5</sup>,  
Marco Moriondo<sup>6</sup>, Jean P. Palutikof<sup>7</sup>

<sup>1</sup>Department of Physical Geography and Ecosystems Analysis, GeoBiosphere Science Centre, Lund University, Sölvegatan 12, 223 62 Lund, Sweden

<sup>2</sup>Rosby Centre, Swedish Meteorological and Hydrological Institute, 601 76 Norrköping, Sweden

<sup>3</sup>Climatic Research Unit, University of East Anglia, Norwich NR4 7TJ, UK

<sup>4</sup>Faculty of Mathematics and Computer Science, Adam Mickiewicz University, ul. Matejki 48/49, 60-769 Poznań, Poland

<sup>5</sup>Research Centre for Agricultural and Forest Environment, Polish Academy of Sciences, Poznań, Poland

<sup>6</sup>Department of Agronomy and Land Management, University of Florence, Piazza S. Marco, 4-50121 Florence, Italy

<sup>7</sup>Hadley Centre, Met Office, Fitzroy Road, Exeter EX1 3PB, UK

**ABSTRACT:** A new method for optimising threshold values of dry/wet spells is evaluated. A set of indices is used to find the best threshold giving good correspondence between the frequency of dry/wet spells in Hadley Centre regional model (HadRM3) output, reference observations with pre-determined thresholds, and area-averaged observations. The analyses focus on selected model gridboxes in 3 different European climate regimes (Sweden, UK, Italy), where station data are available from several locations. In addition, a pan-European analysis using the European Climate Assessment (ECA) dataset is carried out. Generally, there is good agreement between point observations and the corresponding area average using the common thresholds of 0.1 or 1.0 mm with observational data. Applying the optimal thresholds on the model output is important, as it typically results in substantially better agreement between the simulated and observed series of dry/wet days. The fitted optimal pan-European dry/wet threshold is (1) 0.47 or 0.15 mm, depending on model version, for the observed point data threshold of 0.1 mm, and (2) 1.2 or 0.56 mm, depending on model version, for the threshold of 1.0 mm.

**KEY WORDS:** Dry spell · Wet spell · Regional climate model · Precipitation · Areal precipitation · Europe

*Resale or republication not permitted without written consent of the publisher*

## 1. INTRODUCTION

Prolonged dry spells result in considerable water shortages that impact agricultural yield and forest production. In contrast, prolonged wet spells are linked to waterlogged land surfaces, slope instability and, if associated with heavy precipitation, also cause flooding and soil erosion. Sequences of wet or dry days are one important factor in determining soil moisture conditions. This is particularly true for

regions that experience protracted dry periods. Dry spells are becoming an important topic, as climate models often indicate increasing frequency of summer droughts under global warming in regions where drought was previously a rare event (Aamlid et al. 2000, Seneviratne et al. 2002, Solberg 2004, Maracchi et al. 2005, Good et al. 2006). Conversely, devastating floods in Europe over recent years have increased the interest in wet spells and heavy precipitation events (Frei et al. 2000, Palmer & Räisänen 2002). Various

\*Email: lars.barring@smhi.se

indices are commonly used to determine the general characteristics of dry and wet spells (Frich et al. 2002, Klein Tank et al. 2002, Klein Tank & Können 2003, Good et al. 2006, this issue), because more detailed methods for the estimation of water availability and soil moisture amount depend on the context and an assortment of non-climatic factors for which data may not be readily available.

A threshold for determining dry and wet days is required when analysing spells of rainfall, but the frequency distribution of the length of dry and wet spells is highly skewed and very sensitive to the selected threshold. Recent studies have mainly used 2 thresholds for delineating dry/wet days in observational records. The standard observational threshold of 0.1 mm (Osborn 1997, Semenov & Bengtsson 2002) provides a good representation of precipitation conditions for some observational records (Dobi-Wantuch et al. 2000). However, a higher threshold of 1.0 mm has recently been introduced as it is more resistant to measurement errors related to light rainfall (Frei et al. 2003, Frich et al. 2002, Klein Tank & Können 2003, Peterson et al. 2001). The 2 thresholds can result in widely different spell lengths. This problem is exacerbated if indices of extremes, such as the longest wet/dry spell in a year, are derived from climate model output (Good et al. 2006). It cannot be assumed that use of the standard thresholds of 0.1 and 1.0 mm gives results that can be directly related to those obtained using the same thresholds with point rainfall data. In particular, climate models may produce arbitrarily small daily precipitation totals, often resulting in too many days with light precipitation (Frei et al. 2003). Thus, there is a need for a more rigorous determination of the threshold defining dry/wet days. Here we explore an approach (Barring et al. 2003) using several statistical parameters to find an optimal threshold.

This study focuses on the question: what is a suitable threshold for distinguishing dry and wet days in climate model output? The threshold for model precipitation should optimise the following criteria: (1) good fit to the observed proportion of dry/wet days, and (2) good fit to the frequency distribution of dry/wet spells of different lengths. The same problem arises when one wishes to relate observation-based areally-averaged rainfall, either from reanalysis products or directly derived from observations, to point observation series.

We address this problem by finding optimal thresholds for observed area-average series and gridbox series given the standard reference thresholds of 0.1 and 1.0 mm for the individual point series. We assess the validity of the threshold-fitting procedure by examining the relationship between the observed area average and the individual point series.

This explorative study considers the relationship between point observations, area-averaged observations and gridbox values from regional climate model (RCM) experiments. Thus there are 3 relationships to analyse

- Between the individual point observation series and their area-averaged series
- Between the point observations and RCM gridbox values
- Between the area-averaged observations and RCM gridbox values.

The analyses use the Hadley Centre RCM, HadRM3, used extensively within the 'Modelling the Impact of Climate Extremes' (MICE) project (Palutikof 2006, this issue), as well as in the associated PRUDENCE (Prediction of Regional scenarios and Uncertainties for Defining European Climate change risks and Effects) project (Christensen et al. 2005). The analyses concentrate on 3 particular regions where high-quality observational data are available. Even though local effects and site-specific conditions influence the point observations, we can use them as 'reference data'. The reason for this is that, for many applications, established methods such as impact models, statistical relationships, empirical rules-of-thumb etc. are constrained to using point observations because usually only 1 station record is available in a gridbox.

In most regions of Europe the station density is insufficient to permit a detailed rainfall analysis for every gridbox. Therefore, for most purposes it is desirable to have just 1 threshold for simulated dry/wet spells that gives a good relationship with observed dry/wet spells. Therefore, we try to find a common threshold that is applicable across the whole model domain, i.e. Europe, relating dry/wet spells from individual stations to their simulated gridbox counterparts.

## 2. DATA AND METHODS

### 2.1. RCM data

The simulated time-series of daily precipitation totals are taken from 2 versions of the Hadley Centre RCM that have a spatial resolution of  $0.44^\circ$  (~50 km), HadRM3H (Jones et al. 1995, Durman et al. 2001, see also Frei et al. 2003), and the newer version, HadRM3P (Hadley Centre 2002, Hudson & Jones 2002, Jones et al. 2004). At the boundary, the RCM are forced by the HadCM3/HadAM3 model hierarchy, (see for example Good et al. 2006). We focus here on model runs for the 'common' or 'control' period 1961–1990. For the 2 model versions, 3 ensemble members denoted *a*, *b* and *c*, are available, giving a total of 6 model experiments. In all experiments, 360 d years are used, giving time-series of 10 800 model days over the 30 yr period.

Table 1. Details of observational datasets.  $|\Delta z| > 500$  m: no. of stations excluded because of excessive altitude difference relative to RCM gridbox altitude;  $\bar{r}$ : average correlation (gives average inter-station correlation)

Region	Gridbox number	Gridbox centre		Stations within gridbox	Stations outside gridbox	Average station altitude	Gridbox altitude	$ \Delta z  > 500$ m	$\bar{r}$
		Latitude	Longitude						
Scania	4144	56.1140° N	13.0440° E	23	11	64	54	0	0.74
Hertfordshire	2425	51.7915° N	00.1574° W	23	0	79	88	0	0.88
Tuscany	3830	43.4107° N	10.5303° E	6	10	216	190	0	0.55
	3831	43.8465° N	10.5180° E	8	10	291	520	3	0.69
	3924	42.9277° N	11.1274° E	5	10	226	245	0	0.58
	3925	43.3948° N	11.0874° E	8	13	245	368	0	0.58
	3926	43.8696° N	11.0685° E	7	10	272	500	1	0.61
	4020	43.4226° N	11.6878° E	5	7	444	586	1	0.55

## 2.2. Precipitation station data

In analyses of dry and wet spells it is important to have high-quality observations, with the number of missing values kept to a minimum. Here we focus on 3 areas with high-quality high-density precipitation data (Table 1). The quality of the station networks was assessed using a range of descriptive statistics and graphs. The few missing values are filled in by linear interpolation from adjacent stations. The daily area-averaged total rainfall is estimated by the unweighted arithmetic mean of all stations used. The averaging method is not critical because we are not interested in the average precipitation amounts as such.

The first area draws on the dense network of precipitation gauges in Scania, southern Sweden (Linderson 2003) (Fig. 1a). We selected 34 stations with observational series for the period 1974–1990, giving a total number of observations of  $N = 6209$ . The stations are located in a relatively flat coastal region where the relief is  $< 200$  m, and dominated by farmland in the western part and forests in the eastern part. The inter-station correlations are consistently high ( $\bar{r} = 0.74$ , Table 1) and there is no discernible decay in the relationship with distance. The second area is located in Hertfordshire (just north of London) where 23 precipitation stations, all within the gridbox, are available (Fig. 1b). This dataset covers the period 1961–1977 ( $N = 6208$ ). The location is gently rolling farmland where the relief is  $< 200$  m. As for Scania, the inter-station correlations are high ( $\bar{r} = 0.88$ , Table 1) and do not decay with distance. The third area is located in Tuscany, central Italy. Here the station network is more sparse but distributed across several gridboxes (Fig. 1c) allowing replication over several adjacent gridboxes. This dataset covers the whole period 1961–1990 ( $N = 10957$ ). The region is located on the western slopes of the Apennine mountain range, and has a mixture of rolling hills and steep local relief, with a few peaks reaching above 2000 m. The average station height is about 280 m and the highest location is 1340 m.

A few station records were excluded (Table 1) to avoid excessive influence from orographic conditions that are not typical of the model gridbox. The criterion for removal is altitudinal difference of at least 500 m ( $|\Delta z| > 500$  m) from the gridbox altitude. Unlike the Scania and Hertfordshire stations, the inter-station correlations in Tuscany decay to about 0.5 at a distance of 70 km. Furthermore, in all gridboxes except Gridbox 3831, there were 1 or 2 stations with a substantially lower inter-station correlation (about 0.3 to 0.4). These reduced the overall average correlation ( $\bar{r} \approx 0.6$ , Table 1). The deviating correlations are due to the complex topography, and the substantially larger altitudinal variation, that influences the spatial distribution of precipitation. This is particularly the case during spring and summer when localised features such as showers and thunderstorms dominate the rainfall regime.

The Scania and Hertfordshire stations cover only 17 yr compared to 30 yr for the Italian stations. This has no important impact on the analyses because the former 2 locations represent humid climatic conditions where precipitation can be expected throughout the year. The maximum dry spell length is therefore shorter compared to what is observed in Tuscany.

In addition to detailed analyses in 3 regions where a dense station network is available, it is also useful to examine common features over the whole model domain. To do this, we selected a pan-European subset of data from the European Climate Assessment (ECA) dataset (Klein Tank et al. 2002). First, we chose only station time-series with no missing values during the period 1961–1990 ( $N = 10957$ ), which resulted in a subset of 109 stations. No additional quality control was carried out. Second, to exclude stations with orographic conditions atypical of the model gridbox, we again use the criterion  $|\Delta z| > 500$  m. In this way, a number of stations were rejected in the Alps and in the interior of Spain. The final subset (Fig. 1d) comprises 73 stations over the model domain.

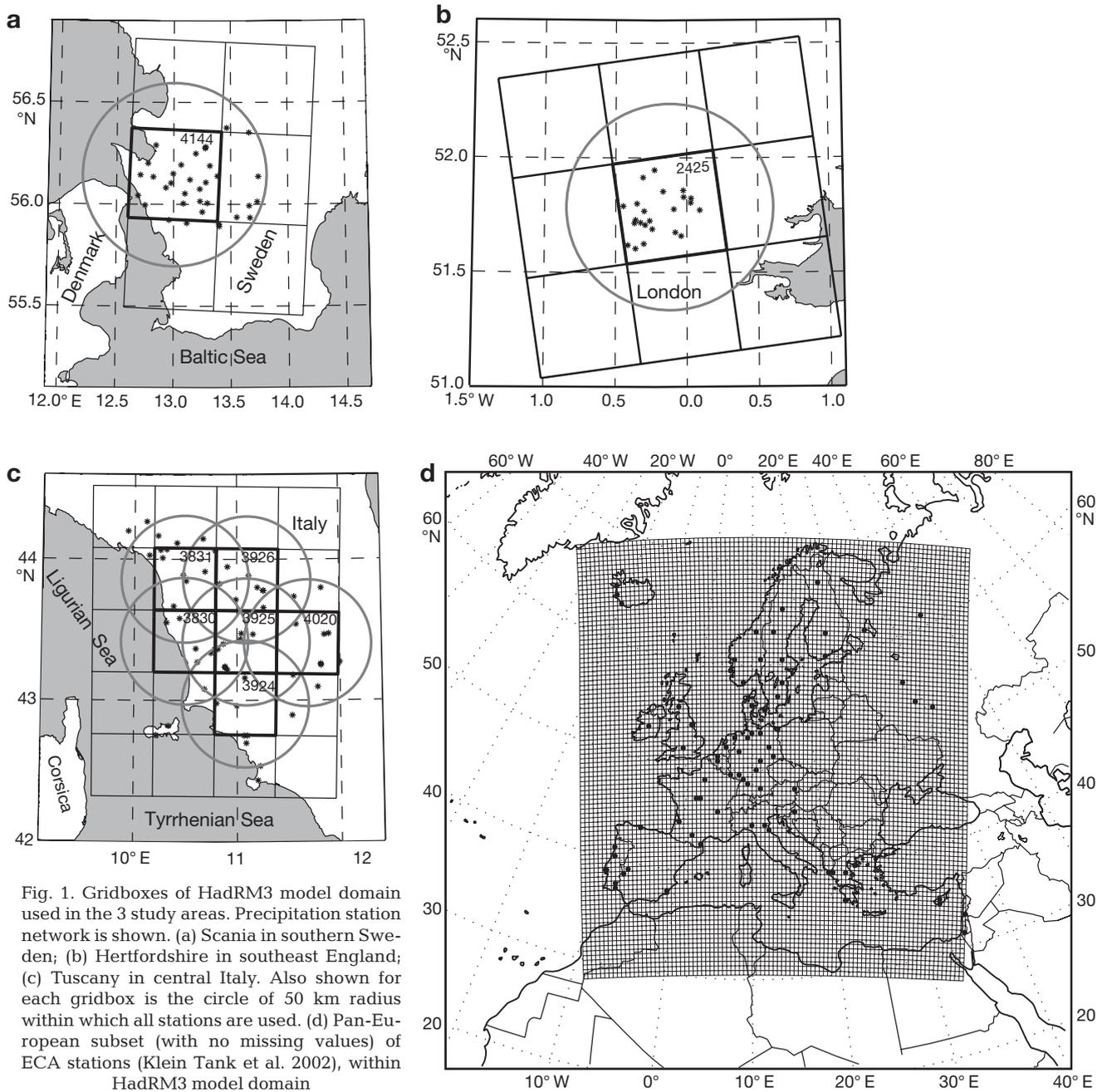


Fig. 1. Gridboxes of HadRM3 model domain used in the 3 study areas. Precipitation station network is shown. (a) Scania in southern Sweden; (b) Hertfordshire in southeast England; (c) Tuscany in central Italy. Also shown for each gridbox is the circle of 50 km radius within which all stations are used. (d) Pan-European subset (with no missing values) of ECA stations (Klein Tank et al. 2002), within HadRM3 model domain

### 2.3. Statistical measures for dry/wet spell comparisons

Various aspects of the spell characteristics are compared using the following statistical measures

- Maximum dry spell length across the entire data period ( $DS_{max}$ )
- The 99, 95, 90 and 75% quantiles (percentiles) of dry spell length across the entire data period, denoted as  $DS_{p99}$ ,  $DS_{p95}$ ,  $DS_{p90}$ ,  $DS_{p75}$
- Corresponding indices for wet spells (denoted as  $WS_{max}$ ,  $WS_{p99}$ ,  $WS_{p95}$ ,  $WS_{p90}$ ,  $WS_{p75}$ ).

For all indices based on point observations, we use the 2 thresholds suggested by previous studies: 0.1 and 1.0 mm. For the area average, either based on observations or model gridbox data, we look for an optimal corresponding threshold in the range from 0.0 to 4.0 mm in steps of 0.02 mm. Thus, for each point observation time-series, the value of the statistical measure is constant once the threshold is fixed at 0.1 or 1.0 mm. The corresponding value for the areally-averaged series is a function of the varying threshold. The optimal threshold for each index is the one that minimises

the absolute difference between the values of the measure for the observed point and the area-average time-series (Fig. 2a). The same procedure is used for comparing point observations to both observed and modelled area-averages. This typically results in different optimal thresholds because of different spell characteristics.

We also define the following 3 indices based on standard deviations to quantify the discrepancy between the distributions:

$$\begin{aligned} s_P &= \text{SD}[F(P_o) - F(P_m)] \\ s_{DS} &= \text{SD}[F(DS_o) - F(DS_m)] \\ s_{WS} &= \text{SD}[F(WS_o) - F(WS_m)] \end{aligned}$$

where

- $\text{SD}[\dots]$  is the standard deviation of the differences in cumulative frequency distribution  $F$  of the relevant variables

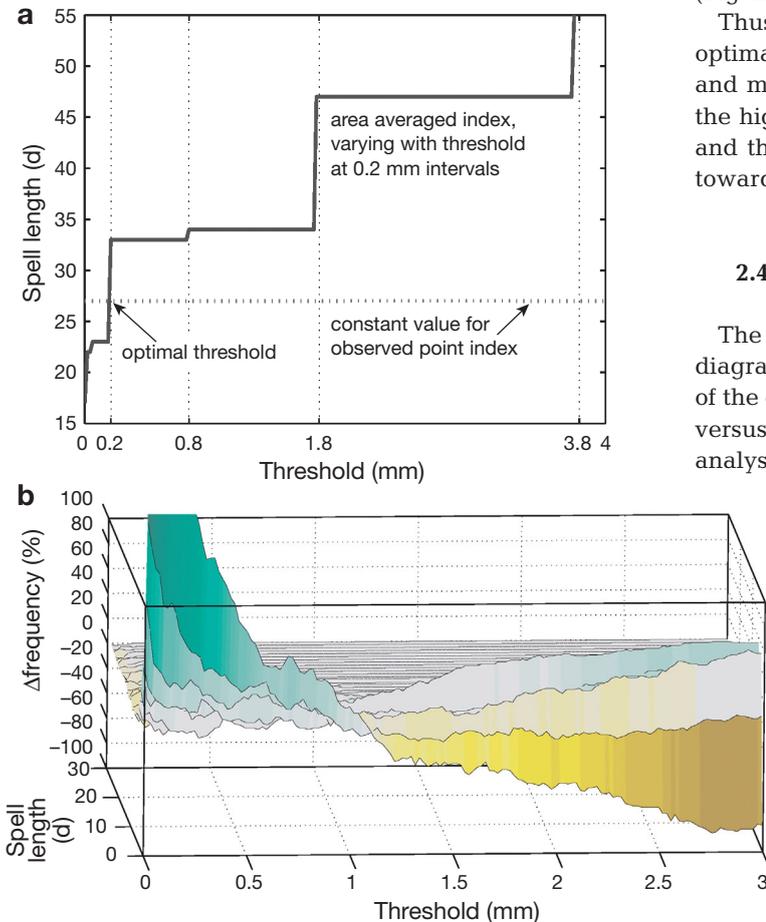


Fig. 2. Illustration of the general principle of index optimisation. (a) For each  $DS$  and  $WS$  measure the reference series has a specific value (dotted horizontal line) and the corresponding value varies depending on the threshold for the series that is being optimised. In this example, optimal threshold is near 0.2 mm. (b) Differences in cumulative frequency distribution of spell lengths used for the 3  $s$  indices (green = positive; yellow = negative differences from 0). Indices minimise the standard deviation of differences along the spell length dimension. In this example, the optimal threshold for the  $s$  index is  $\sim 1.2$  mm

- $P$  = daily precipitation amount (discretised into 0.1 mm steps according to the observational standard practice of recording no amounts  $< 0.1$  mm, and further processed in such a way that all values below the selected threshold are adjusted to 0)

- $DS$  = dry spell length

- $WS$  = wet spell length.

The subscript is either 'o' for point observations or 'm' for area-average data (observed or model gridbox).

Spell length is measured in units of whole days because precipitation is normally observed on a daily basis. Again, the frequency distribution for the point-observation series is fixed for each of the 2 thresholds, but the corresponding frequency distribution of area-average values depends on the threshold. The optimal thresholds for each of these 3 measures minimises the SD of the difference in frequency distributions (Fig. 2b).

Thus, we use 13 statistical measures for finding the optimal threshold. The measures based on percentiles and maximum values are metrics of what happens at the high (long spell) end of the frequency distribution and the measures based on SD are heavily weighted towards the low end of the frequency distribution.

## 2.4. Explanation of the graphical presentation

The graphical presentation of results uses 2 types of diagram: an array of box plots (e.g. Fig. 3c), and plots of the cumulative relative frequency of dry or wet days versus spell length (e.g. Fig. 3d). The procedure for analysing the frequency distribution and finding the optimal threshold is now described with reference to these 2 example plots.

In total, Fig. 3c shows 26 box plots, divided

into 2 parts: the left half refers to the 0.1 mm threshold, and the right half refers to the 1.0 mm threshold. Each half is further divided into 3 subsections, 1 for each group of measures ( $DS$ ,  $WS$  and  $s$ ). Overall median thresholds for each subgroup are shown as horizontal lines, and the overall fitted threshold is indicated by a dotted line. Each box plot represents the spread of optimised thresholds across the different stations. The 'box' covers the region between the lower and upper quartile, with the median indicated by the line within. Following the standard method of drawing boxplots, the lines ('whiskers') at each end of the box extend either to the lowest/highest value, or 1.5 times the inter-quartile range. Values outside the whiskers are identified by \*, thus allowing easy identification of stations being extreme/outliers relative to the majority of stations.

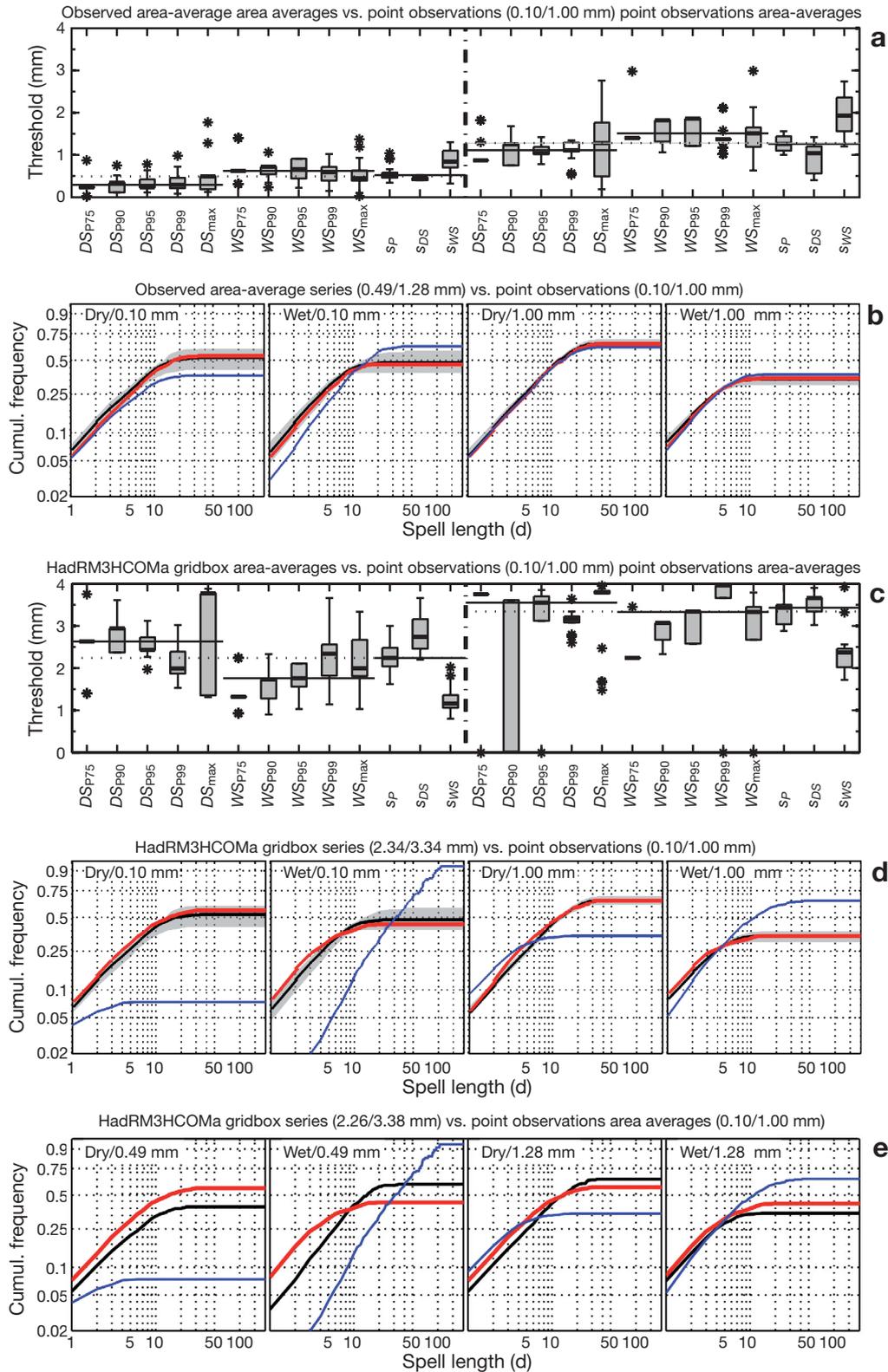


Fig. 3. Data for Scania and the HadRM3HCOMa experiment used as an example to show the steps in analysis. See Section 2.4 for further details. (a) Boxplots showing distribution of fitted thresholds for relating the observed area average to individual point series. Optimised thresholds (cf. Table 2) indicated by dotted lines. (b) Plots of cumulative relative frequency of dry (wet) days for the observed area average before (blue) and after (red) adjustment using the overall median of thresholds in boxplot (a). Median of individual point series (black line) and the envelope of individual observed series also shown (shaded region). Similarly, (c) and (d) show curves using simulated gridbox area-average series. Solid horizontal lines in (a,c) are overall median thresholds for each subgroup. (e) Simulated area average fitted to observed area average that in turn was adjusted to the point series (colours as in b,d). Simulated data are here fitted to 1 observational record; consequently, there is no envelope of individual station curves

The plots of cumulative relative frequencies (e.g. Fig. 3d) have 4 subplots: again, the left half refers to the 0.1 mm reference threshold, and right half refers to the 1.0 mm threshold. The  $x$ -axis gives spell length in number of days on a logarithmic scale, and the  $y$ -axis shows cumulative frequency of spells with axis scaling based on logarithms of the exponential distribution. This scaling is chosen to allow the main features of the plotted curves to be resolved. For each spell length  $N$  the curves show the frequency of dry (wet) days belonging to dry (wet) spells of length at most  $N$  days. Therefore, these are not simply plots of the cumulative frequency of spell lengths. The following information is plotted in the graphs

- Shaded area and black curve: the shaded area shows the envelope of the individual stations curves and the black curve within is the median. These are calculated as follows: for each spell length (1 to 360 d), the minimum, median and maximum are computed over all stations
- Blue curve: this corresponds to the RCM gridbox dry (wet) days based on an unoptimised threshold. For graphs based on the 0.1 mm reference threshold, any non-0 RCM precipitation value was treated as 'wet'. For graphs based on the 1.0 mm threshold, 1.0 mm was also the threshold for the RCM value
- Red curve: this is based on the RCM gridbox dry (wet) days after adjustment using the optimal threshold.

Ideally, therefore, in figures such as 3d, the red curve based on the optimal threshold should closely track the black curve of the median of the station values. Conversely, the blue curve, which is not based on an optimal threshold, is often far away from the median station values.

The cumulative frequency plots are designed to show essential features of the threshold optimisation procedure, rather than provide direct graphical information on the adjusted spell measures. The following information can be extracted. The curves typically level off to a constant  $y$ -level, which is the total proportion of dry (wet) days. By definition, for each pair of dry/wet plots, these 2 levels always add up to unity for curves of the same kind (black, blue, or red) The upper end of the shaded envelope in the dry spell graph goes with the lower end of the shaded envelope in the wet spell graph, i.e. the driest station is also the least wet station and vice versa. The  $x$ -axis point at which the levelling-off occurs is at the maximum spell length. No more dry (wet) spell days are cumulated beyond this point. Sometimes this point is clearly defined as a sharp bend in the curve, and sometimes it is more gradual. In the former case there are several spells of maximum length, in the latter case there is 1 spell of maximum length and several spells that are nearly as

long. If a curve does not level off to a constant  $y$  value it means that there are spells longer than 200 d (which is the  $x$ -axis limit). Because the total cumulative frequency of a pair of graphs (dry and wet) always adds up to unity, we can derive information about other cumulative frequencies by combining the total cumulative frequency of dry (wet) days with the relative frequency of some selected spell length in the wet (dry) graph.

Taking the 'Dry/0.1mm' graph of Fig. 3d as an example: the median of the station records (black line) suggests that typically 25% (i.e. lower quartile) of all days belong to dry spells of length at most 5 d. The corresponding percentile of the unadjusted gridbox curve (blue) is never reached because the model is too wet. According to the unadjusted model data, no dry days occur in spells longer than 5 d. The adjusted gridbox data (red) are very close to the observed median value. This indicates that the adjustment brings the gridbox data into close agreement with typical point observation data with respect to dry/wet spell frequencies. If we use 1.0 mm as the reference threshold, we find that about 5% of all days occur in wet spells lasting  $>4$  d; the total proportion of dry days is about 70% plus 25% for wet spells  $\leq 4$  d, leaving about 5% for wet spells longer  $>4$  d (in this particular example, all curves almost coincide at this point).

## 2.5. Procedure for analysis and adjustment of frequency distributions

Fig. 3 provides an example of the steps in the analysis for the Scania stations. We compare the area-averaged observations with the  $n$  individual series ( $n = 34$  in this example) that make up the area average. Fig. 3a shows a typical box plot of the spread of fitted thresholds for the different measures. There is variation in the fitted thresholds, even within an area of about  $50 \times 50 \text{ km}^2$ . In this example, the overall optimal threshold for the observed area average is 0.48 mm when the point observation series are thresholded by 0.1 mm, and 1.28 mm when the corresponding point observation threshold is 1.0 mm. The resulting adjustment to the frequency distribution of observed area-averaged dry/wet spells is shown in Fig. 3b; the fit is very good in this particular example as the black median line almost overlaps the fitted red line. The scope for improving the frequency distribution is limited because the area-average is already close to the corresponding frequency distribution of point observations. The same procedure is repeated, but using the modelled gridbox area average from a HadRM3H model run. In this case the box plot (Fig. 3c) shows a much greater spread and the median thresholds are higher. In Fig. 3d, it is clear

that using this particular gridbox time-series without any adjustment results in a substantial bias. For the reference threshold of 0.1 mm, about 95 % of all simulated wet days without any threshold (blue curve in graph 'Wet/0.1 mm') occur in spells of at least 5 d and about 50 % of all wet days occur during spells lasting 30 d or longer. Moreover, the proportion of wet days is over 90 % before adjustment. After adjustment (red line) using a fitted threshold of 2.24 mm, the situation is much improved and the red line is now close to the median observation curve (black).

For the reference threshold of 0.1 mm we compare point observations with unadjusted model data without a threshold (rather than applying the same threshold to the model data), because this corresponds to the 'no action' alternative of just using the data as it is. Although the situation before adjustment improves substantially when the reference threshold is 1.0 mm, there is still a substantial bias in the simulated time-series. Here we apply the reference threshold to both point observations and unadjusted model data because using this threshold involves a deliberate choice. By applying the fitted threshold of 3.34 mm to the model time-series, almost all bias is removed.

In the cumulative frequency diagrams (e.g. Fig. 3d), when the blue curve begins to the right (left) of the black reference curve in a wet (dry) spell diagram, the simulated gridbox area average is too wet. That is, there are too many days with precipitation above the reference threshold. This results in the fitted threshold being higher than the reference value. If the opposite is true, i.e. the blue curve begins to the left (right) of the black reference curve in a wet (dry) spell diagram, the simulated time-series is too dry, and the fitted threshold will be lower than the reference value. Because the  $y$ -axes of the frequency diagrams are truncated and scaled to focus on long spells, this can sometimes occur outside the limits of the diagram.

An alternative approach for obtaining a fitted threshold is to assume that the gridbox area average should, at least in theory, be better related to the observed area average than to point data. This approach is shown in Fig. 3e, which provides the same information as Fig. 3d with the exception that the observations are now represented by 1 line only (the area average) and thus there is no shaded area. Both approaches are used and the results are presented below.

## 2.6. Pan-European analysis

The method for deriving the site-specific fitted thresholds requires several rain gauges within a gridbox. To find a pan-European optimised threshold, a slightly different approach is required. In this case only 1 sta-

tion is available in each of a few gridboxes (Fig. 1d). The threshold fitting is thus carried out for each pair of station/gridbox time-series and for the 13 statistical measures. Also, here we use the 2 reference thresholds of 0.1 mm and 1.0 mm for the observed time-series. The fitted threshold is the overall median of the individually optimised thresholds for all 73 station time-series and 13 statistical measures.

## 2.7. Bias assessment

From the plot of simulated dry/wet spell frequencies (e.g. Fig. 3d) it is clear that there may be a substantial bias in the model data when it comes to the proportion of dry/wet days, and thus in the extreme spell measures. It is further evident that the suggested method for fitting an optimised threshold may provide a considerable bias reduction. It is useful to analyse the bias in a way that allows assessment of the bias remaining after the threshold fitting. Here, we take an empirical approach and calculate the proportional reduction in the maximum absolute distance between the point observation reference curve and the modelled gridbox average curve. The threshold is optimised to reduce the median bias across the stations, which may result in the bias increasing for a few stations. Thus, we count the number of stations for which the bias reduction (i.e. improvement) is better than 20, 40, 60 and 80 %, and also the number of stations having an increased bias (worsening) of 20, 40, 60 and 80 %.

The bias is calculated separately for each station, and for wet and dry spells. It is defined as the maximum absolute distance between a station curve (not displayed) and the blue curve (unadjusted bias) or red curve (adjusted bias). For each station, the effect of the adjustment is then calculated as the change of bias expressed as a percentage of the unadjusted bias.

## 3. RESULTS

### 3.1. Observed area average versus point observations

The fitted threshold for adjusting observed area-averages to the stations series (cf. Fig. 3a,b) is shown in Table 2. As the area average is always at least as wet as the individual stations series, the thresholds are higher than the corresponding reference point observation thresholds. The closer the agreement between the different point series, the closer the optimised and reference threshold will be. The thresholds listed in Table 2 clearly show that Hertfordshire has the lowest optimised thresholds and Tuscany has the highest thresholds, with Scania somewhere in between. This

Table 2. Fitted thresholds for relating observed area-averaged precipitation to point observations having a threshold of 0.1 or 1.0 mm

Site	0.1	1.0
Scania	0.49	1.28
Hertfordshire	0.27	1.06
Tuscany 3830	0.98	2.26
Tuscany 3831	1.14	2.28
Tuscany 3924	1.17	1.99
Tuscany 3925	1.1	1.95
Tuscany 3926	1.33	2.3
Tuscany 4020	0.83	2.04

ordering follows both the station spacing—and the inter-station correlation—in the 3 regions, and also the general precipitation climate. In Hertfordshire, precipitation is to a large extent organised by large-scale weather systems. Scania belongs to the same general climatic zone but has a somewhat more continental climate. Tuscany belongs to the Mediterranean climatic zone with a pronounced summer drought and a higher frequency of local showers.

This pattern is further supported by the bias statistics in Table 3. For both Scania and Tuscany, the adjustment results in a clear reduction in bias, but for Hertfordshire the results are inconsistent; with the 0.1 mm threshold, the bias is reduced for the large majority (87 %) of stations, whereas with the 1.0 mm threshold, the bias increases for the majority (54 %) of stations. For the reference threshold of 0.1 mm, the optimised threshold provides an improvement (i.e. 65 % reduction in bias). For the 1.0 mm threshold the bias increases by 22 %, even though the unadjusted maxi-

Table 3. Change in bias produced using the optimised threshold for adjusting observed area-average series to point observation series for the 3 selected regions with reference thresholds of 0.1 and 1.0 mm. Dry and wet spell parts are pooled together, as are 6 Tuscany gridboxes. Proportion (%) of all stations (n depends on gridbox) that experiences a 20, 40, 60 and 80 % decrease/increase is shown. F: cumulative frequency; Max  $\Delta F$ : the maximum value across dry/wet of the median bias (over individual stations) before and after adjustment; percentage change: in maximum bias

		Scania		Hertfordshire		Tuscany	
		0.1	1.0	0.1	1.0	0.1	1.0
Reduction	$\geq 20\%$	94	60	87	17	94	88
	$\geq 40\%$	91	47	85	2	89	80
	$\geq 60\%$	85	22	72	0	79	66
	$\geq 80\%$	22	0	17	0	41	18
Increase	$\geq 20\%$	3	7	13	54	5	9
	$\geq 40\%$	3	7	9	11	3	9
	$\geq 60\%$	3	4	4	4	1	9
	$\geq 80\%$	3	3	2	2	0	7
Max $\Delta F$ unadjusted		0.15	0.04	0.07	0.02	0.18	0.10
Max $\Delta F$ adjusted		0.04	0.04	0.03	0.02	0.05	0.04
Percentage change		-75	-13	-65	28	-68	-45

mum bias is already small (0.02). Heterogeneities in local physiography, such as orography and land/sea contrasts etc., contribute to the differences in thresholds between the different stations within each region.

### 3.2. Simulated gridbox averages versus observed area average or point observations

Table 4 shows the fitted thresholds for all gridboxes and all model experiments, as well as the ensemble averages of the individual members *a*, *b* and *c* (see Section 2.1). Columns labelled 'Point' show the result of fitting simulated gridbox averages directly to the point observations (cf. Fig. 3c,d). Columns labelled 'Area' show the result when fitting the simulated area average series to the observed area average series that, in turn, were adjusted to the point series using the optimised threshold of Table 2. In almost all cases the first approach is more successful (as in Fig. 3). In principle, the second approach has the advantage of adjusting a simulated area average to an observed area average. However, analyses show that, in practice, this approach does not work well. The reason is that the procedure has to find a best fit between the simulated series and only 1 observed series, which has been fitted to the point observation series without any constraints induced by the structure of the simulated series. In the first approach, where the fit is made between the simulated series and several point observation series, more freedom is allowed when finding the median threshold. This is the main reason why fitting a threshold by relating the simulated area average directly to point observations generally performs better than taking the extra step of first fitting a threshold for the observed area average to point observations.

From Table 4 it is evident that the 2 model generations differ with respect to how they simulate the wet/dry day structure of precipitation. The thresholds are generally lower for the newer HadRM3P model version, as is illustrated by the lower thresholds shown in Fig. 4a (compared with Fig. 3c). This indicates that HadRM3P matches the observed frequency of light precipitation events more closely than the older HadRM3H model and, therefore, results in an overall closer agreement between the unadjusted spell frequency curves (blue) and observed spell frequencies (black curves) in Fig. 4b (compared to in Fig. 3d). This is particularly noticeable with respect to the 1.0 mm reference threshold.

Table 4. Fitted thresholds for relating model time series to observations in study regions and for Europe as a whole, at 0.1 and 1.0 mm reference thresholds. Avg: ensemble average; Point: simulated time series directly fitted to point observation series; Area: simulated series fitted to adjusted observed area-average precipitation using the threshold given in Table 2 (see Section 3.2 text for further details). nd: no data

Gridbox	Ensemble member	HadRM3H				HadRM3P				
		Point		Area		Point			Area	
		0.1	1.0	0.1	1.0	0.1	1.0	0.1	1.0	
Scania	<i>a</i>	2.24	3.34	2.26	2.38	1.20	2.07	1.18	1.23	
	<i>b</i>	2.50	2.61	2.20	2.50	1.31	1.96	1.29	1.49	
	<i>c</i>	2.18	3.38	1.73	1.98	1.28	1.99	1.20	1.34	
	avg	2.31	3.11	2.06	2.29	1.26	2.01	1.22	1.35	
Hertfordshire	<i>a</i>	0.79	1.40	1.05	1.21	0.01	0.40	0.02	0.14	
	<i>b</i>	0.77	1.33	0.90	1.09	0.01	0.32	0.03	0.10	
	<i>c</i>	0.58	1.25	0.99	1.05	0.02	0.39	0.03	0.10	
	avg	0.71	1.33	0.98	1.12	0.01	0.37	0.03	0.11	
Tuscany 3830	<i>a</i>	0.08	0.00	1.46	0.45	1.01	1.70	0.33	0.40	
	<i>b</i>	0.00	0.00	0.46	0.51	0.99	1.46	0.36	0.40	
	<i>c</i>	0.14	0.00	1.37	1.09	0.94	1.38	0.34	0.39	
	avg	0.07	0.00	1.10	0.68	0.98	1.51	0.34	0.40	
Tuscany 3831	<i>a</i>	2.39	1.36	1.52	1.68	0.46	1.23	0.33	0.35	
	<i>b</i>	2.41	0.00	2.42	2.06	0.58	0.97	0.38	0.41	
	<i>c</i>	2.33	0.12	2.58	2.29	0.52	0.99	0.29	0.37	
	avg	2.38	0.49	2.17	2.01	0.52	1.06	0.33	0.38	
Tuscany 3924	<i>a</i>	2.23	2.78	1.52	1.60	3.03	3.66	2.13	2.38	
	<i>b</i>	2.27	2.77	1.65	1.82	2.82	3.67	2.16	2.42	
	<i>c</i>	2.05	2.62	1.46	1.73	2.34	3.56	1.56	1.98	
	avg	2.18	2.72	1.54	1.72	2.73	3.63	1.95	2.26	
Tuscany 3925	<i>a</i>	1.89	2.58	1.40	1.40	0.00	0.00	0.00	0.00	
	<i>b</i>	1.93	2.66	1.44	1.44	0.00	0.00	0.00	0.00	
	<i>c</i>	1.86	2.44	1.29	1.32	0.00	0.00	0.00	0.00	
	avg	1.89	2.56	1.38	1.39	0.00	0.00	0.00	0.00	
Tuscany 3926	<i>a</i>	1.96	2.68	1.50	1.61	1.23	2.16	0.87	1.09	
	<i>b</i>	2.07	2.85	1.48	1.60	1.49	2.08	0.94	1.11	
	<i>c</i>	2.04	2.60	1.41	1.77	1.18	2.03	0.81	0.96	
	avg	2.02	2.71	1.46	1.66	1.30	2.09	0.87	1.05	
Tuscany 4020	<i>a</i>	0.49	1.40	0.33	0.39	1.16	2.58	1.05	1.12	
	<i>b</i>	0.53	1.44	0.40	0.42	1.54	2.52	1.11	1.15	
	<i>c</i>	0.49	1.21	0.25	0.29	1.10	2.25	0.86	0.86	
	avg	0.50	1.35	0.33	0.37	1.27	2.45	1.01	1.04	
<b>Overall median</b>		1.95	1.94	1.45	1.52	1.06	1.83	0.60	0.64	
<b>Overall average</b>		1.51	1.78	1.38	1.40	1.01	1.64	0.72	0.82	
Europe	<i>a</i>	0.47	1.14	nd	nd	0.17	0.56	nd	nd	
	<i>b</i>	0.49	1.25	nd	nd	0.16	0.58	nd	nd	
	<i>c</i>	0.45	1.20	nd	nd	0.16	0.54	nd	nd	
	avg	0.47	1.20	nd	nd	0.16	0.56	nd	nd	

We now consider how well the 2 models, and the adjustment procedure, perform in terms of the bias. Table 5 provides a comprehensive view of the overall performance of each model when the 3 ensemble members, and the dry and wet spells, are taken together in each region. The table shows that the adjustment procedure generally removes more bias with the HadRM3H model. For example, the adjustment procedure produces at least a 60% reduction of the bias in the HadRM3H data for the Scania site when the reference threshold is 1.0 mm. This occurs irrespective of which ensemble member is used with any of the 34 stations for both dry and wet spells. The corresponding bias reduction for the newer HadRM3P model is only 49%. This is explained by the fact

that HadRM3P has a lower initial bias for unadjusted threshold (0.21, compared to 0.32 for HadRM3H), leaving less room for improvement. If the lower reference threshold is used, the performance of the adjustment is very similar, even though the fitted threshold values are higher for the older HadRM3H model. This is further illustrated in Fig. 5. The adjusted HadRM3H series (red curves in Fig. 5b) are very close to the median of the station series, and for HadRM3P (Fig. 5d) the unadjusted simulated and observed spell length frequency distribution are virtually overlapping. In the latter case, using the optimised threshold results in a poorer fit.

In Tuscany, the variations between the different ensemble members are larger than for the other 2 regions

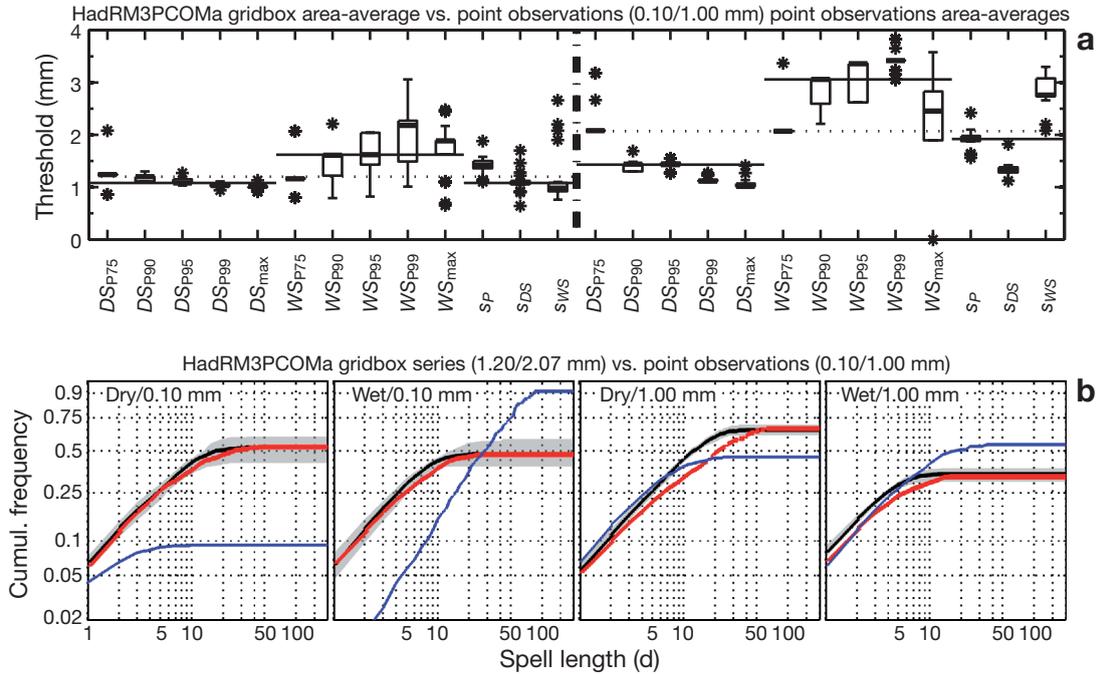


Fig. 4. Same as in Fig. 3c & d but for the newer model version HadRM3P, here illustrated with the HadRM3PCOMa experiment

(Table 4). Also, the variation between the different gridboxes in the region is substantial. For a few combinations of gridboxes and model experiments, the method is not able to provide any stable improvement. For Tuscan Gridboxes 3830 and 3831 and the HadRM3H model, the results are inconsistent in that a lower threshold is produced for the 1.0 mm reference threshold compared to the 0.1 mm reference threshold. A closer scrutiny of the boxplots (not shown) reveals that the individual thresholds for the *DS* and *WS* indices are inconsistent and vary widely between the stations. Low thresholds for *DS* indices in combination with high thresholds for *WS* indices indicate that both simulated long dry spells and long wet spells are too

long compared to the observed. This suggests a more complicated difference between the simulated and observed dry/wet spell structure than just too many light precipitation events. In addition, the fact that the optimised 1.0 mm thresholds are lower than those for 0.1 mm indicates that spell length structure is very sensitive to light rainfall events. For all 3 HadRM3P runs and Tuscan Gridbox 3925, the situation is different in that the adjustment method consistently produces an optimised threshold of 0.0 mm. The boxplots (not shown) reveal that the model is consistently too dry and the method thus accepts all simulated precipitation days by setting the optimised threshold to 0.0 mm. One reason for the large variation between ensemble members and adja-

Table 5. Change in bias produced using the optimised threshold to adjust gridbox data to individual station series for the 2 model versions over the 3 selected regions with reference thresholds of 0.1 and 1.0 mm; as Table 3, but with 3 ensemble runs pooled together

	HadRM3H								HadRM3P								
	Scania		Hertfordshire		Tuscany		Europe		Scania		Hertfordshire		Tuscany		Europe		
	0.1	1.0	0.1	1.0	0.1	1.0	0.1	1.0	0.1	1.0	0.1	1.0	0.1	1.0	0.1	1.0	
Reduction	≥20%	100	100	100	67	81	66	70	29	100	98	0	0	79	59	22	0
	≥40%	100	100	100	46	81	55	47	8	100	82	0	0	76	51	3	0
	≥60%	97	100	96	28	79	48	29	2	99	49	0	0	63	43	0	0
	≥80%	87	64	36	0	66	31	13	0	92	19	0	0	39	23	0	0
Increase	≥20%	0	0	0	0	1	3	17	17	0	0	0	0	3	6	26	0
	≥40%	0	0	0	0	1	2	13	11	0	0	0	0	2	4	11	0
	≥60%	0	0	0	0	1	1	12	7	0	0	0	0	2	3	5	0
	≥80%	0	0	0	0	1	0	10	3	0	0	0	0	2	2	2	0
Max ΔF unadjusted		0.45	0.32	0.26	0.08	0.64	0.44	0.22	0.11	0.45	0.21	0.10	0.13	0.47	0.39	0.15	0.13
Max ΔF adjusted		0.06	0.09	0.07	0.06	0.64	0.44	0.14	0.10	0.06	0.10	0.10	0.13	0.47	0.39	0.16	0.13
Percentage change		-88	-72	-74	-21	0	0	-35	-9	-86	-48	0	0	0	0	-6	0

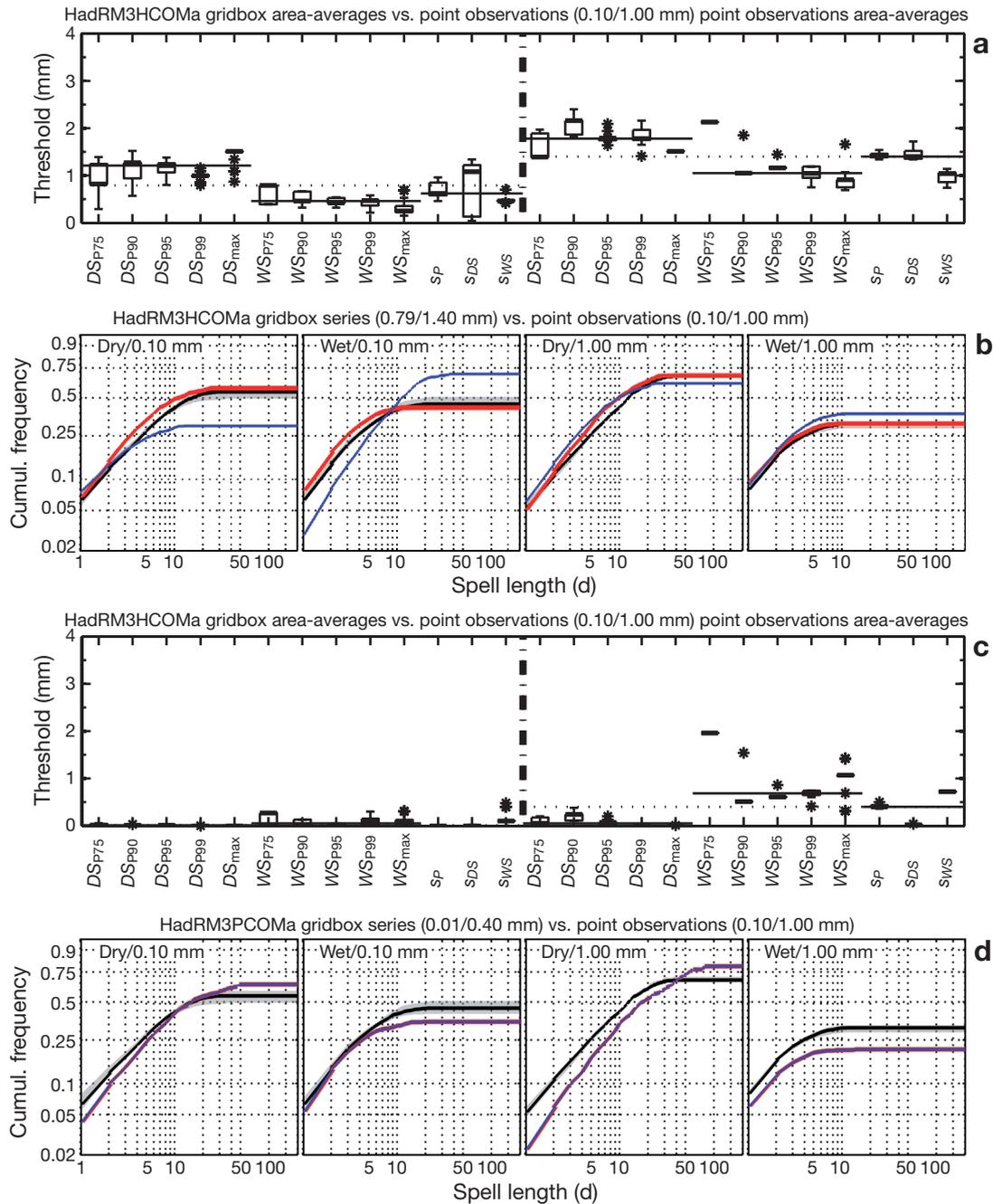


Fig. 5. Comparison of the results from the 2 model versions for the Hertfordshire gridbox. As in Fig. 3c,d but (a,b) the HadRM3HCOMa experiment and (c,d) the HadRM3PCOMa experiment. Note that in panel (d) the perfect fit making the unadjusted curve (blue) overlap with the observed median curve (black), deteriorates when applying the threshold to obtain an 'adjusted' curve (red)

cent gridboxes is the less dense station network. In addition, the pronounced relief makes the statistics sensitive to random variations between ensemble members and their relation to the high variability within and between the relatively few station series (i.e. low inter-station correlation). For the 2 gridboxes where the method fails to find a consistent threshold, the land/sea contrast provides an additional complication.

Interestingly, in comparison with HadRM3H, the adjustment method typically results in lower fitted thresholds for the HadRM3P ensemble members (Table 4), with a larger occurrence of fitted thresholds lower than the corresponding reference value. This means that HadRM3P is too dry more often than HadRM3H. That is, HadRM3P produces too few days with precipitation above the selected reference threshold.

### 3.3. Pan-European thresholds

The bottom part of Table 4 summarises the pan-European thresholds, and Fig. 6 shows typical box plots and frequency diagrams. Because the adjust-

ment operates on each station/gridbox pair individually, the cumulative frequency plots show only the envelope and median of all the 73 curve triples. The large span between the minimum and maximum of the envelopes in the frequency diagrams is a result

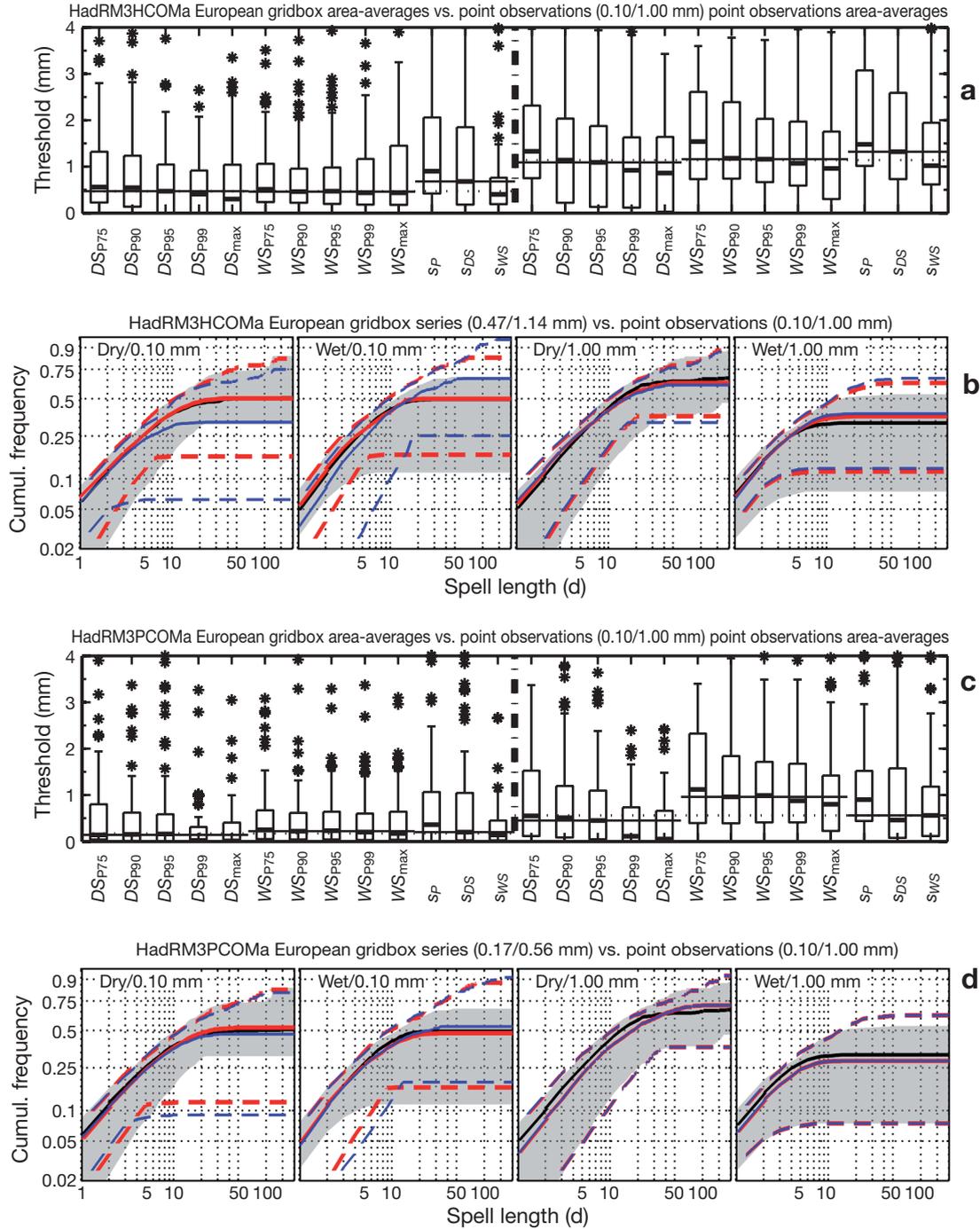


Fig. 6. Results for pan-European thresholds for (a,b) HadRM3HCOMa and (c,d) HadRM3PCOMa. Boxplots (a,c) show spread of individual station thresholds; cumulative frequency diagrams (b,d) show median (solid line), lower and upper envelope (both dashed) for observed point precipitation (black), unadjusted (blue) and adjusted (red) model gridbox average precipitation. In (d), the blue curves are hardly visible because they coincide with red curves, meaning that the adjusted threshold is very low and does not result in noticeable change

of very different precipitation regimes over Europe, from the normally wet northwestern regions (Iceland, the British Isles and Norway), to the seasonally very dry Mediterranean region. The optimised thresholds do not differ much (at most 0.11 mm) between different ensemble members. For HadRM3H (Fig. 6a,b) the optimised threshold is higher than the corresponding reference threshold, indicating that this model is on average too wet over Europe. Applying the optimised threshold to HadRM3H gridbox data, using reference threshold 0.1 mm (1.0 mm), results in improvements (Fig. 6b) at 70% (29%) of the 109 stations, but also gives a poorer fit at 17% of the stations for both thresholds (Table 5). The lower reference threshold results in larger bias, which typically can be reduced by the adjustment method. This picture changes when looking at the results for the newer HadRM3P model (Fig. 6c,d), for which the optimised thresholds are consistently lower (Table 4) than with HadRM3H, resulting in no improvement (Table 5). For the 0.1 mm reference threshold, the optimised thresholds are only slightly larger, 0.16 to 0.17 mm. For the 1.0 mm reference threshold, the optimised thresholds are in the range of 0.54 to 0.58 mm, suggesting that the HadRM3P model is on average drier than observations over Europe. The overall impression from Fig. 6 is that the optimised threshold provides an improvement with HadRM3H data but not with HadRM3P data. However, the bias after adjustment (Table 5) of HadRM3H data is of the same magnitude as before (and after) the adjustment of HadRM3P data.

#### 4. CONCLUSION

Analyses of dry/wet days in time series of observed precipitation and climate model output normally use subjectively-selected rainfall thresholds (often 0.1 mm or 1.0 mm) to define dry/wet days. The empirical method described here provides an objective alternative for finding this threshold. The method can be used to relate an observed area average to the individual point observation series, or to relate simulated gridbox series to point observations. We use data from 2 versions of the Hadley Centre RCM HadRM3, which have been extensively used in the MICE project. The method is applied comprehensively to observational data from 3 case study regions where dense rain gauge networks are available and, in less detail, to more limited station data across Europe.

The method provides a simple approach for reaching a good agreement between either observed area average precipitation or simulated gridbox values and point observations. The technique is to search for a

threshold defining dry and wet days that optimises the agreement between selected statistical measures related to the cumulative frequency distributions of dry and wet spell lengths. The statistical measures used were chosen specifically to analyse various aspects of the spell length frequency distribution.

The method works well for finding optimised thresholds when the climate model overestimates the number of wet days by frequently simulating light rainfall (i.e. is 'too wet'). However, if the model is 'too dry', the procedure cannot improve the situation because it cannot 'invent' wet days. Thus, an optimal threshold lower than the reference threshold does not necessarily lead to a better agreement in the structure of dry/wet spells between simulated and observed point series. As illustrated with 2 of the Tuscany gridboxes, the simulated dry/wet spell structure is not always well simulated when local factors such as topography and land/sea contrast are important for the precipitation process. This suggests that the method could be useful for (1) more detailed considerations of how precipitation processes and the hydrological cycle are parameterised in atmospheric models, and (2) indicating when and why models are 'too wet' or 'too dry' relative to observations.

For HadRM3H, the ensemble average of optimised thresholds is 0.47 mm (1.20 mm) for the reference threshold of 0.1 mm (1.0 mm) in European point observation series. For HadRM3P, the corresponding ensemble average of the optimised thresholds is 0.16 mm (0.56 mm). For HadRM3P, these values do not improve the fit to any appreciable extent. Over Europe, although being too dry, the more recent HadRM3P still produces precipitation data that are in better agreement with observations with respect to the dry/wet spell structure.

Selecting one single optimised threshold for the whole of Europe is a compromise: for some stations the fit will improve, and for other stations the fit will deteriorate. Nevertheless, the technique provides an objective method for optimising this compromise. Current practice is to find a threshold for the whole study domain in a subjective way, or even to simply apply one of the standard thresholds without giving much thought to its validity. The results presented here suggest that there are substantial regional differences in optimal threshold. The systematic variation in regional differences, as well as possible seasonal variations, can be explored to further improve the fit between model and observations.

The overall conclusion is that this empirical method offers several advantages over subjective selection of a threshold. These promising results indicate that extending the analyses using other observational datasets and model data could be very rewarding.

*Acknowledgements.* This work was carried out within the European project MICE, contract no. EVK2OCT2001–0018. M.R. was supported by the Foundation for Polish Science. Model data were supplied by the Climate Impacts LINK project (funded by the UK Department of the Environment, Food and Rural Affairs) on behalf of the Hadley Centre and UK Met Office. Station data for Hertfordshire were downloaded from the British Atmospheric Data Centre (BADC) web site: <http://badc.nerc.ac.uk/index.html>. We are grateful to 2 anonymous reviewers for their constructive comments.

## LITERATURE CITED

- Aamlid D, Torseth K, Venn K, Stuanes AO, Solberg S, Hylen G, Christophersen N, Framstad (2000) Changes of forest health in Norwegian boreal forests during 15 years. *For Ecol Manage* 127:103–118
- Barring L, Linderson ML, Nilsson C, Stjernquist I, Jönsson AM (2003) Sensitivity of dry/wet spell indices to biases in HadRM3 gridcell precipitation vs. observed areal averages and point values. Abstract #13795, EGS-AGU-EUG Joint Assembly, Nice
- Christensen JH, Carter TR, Rummukainen M (2006) Evaluating the performance and utility of regional climate models: the PRUDENCE project. *Clim Change* (in press)
- Durman CF, Gregory JM, Hassell DC, Jones RG, Murphy JM (2001) A comparison of extreme European daily precipitation simulated by a global and regional climate model for present and future climates. *QJR Meteorol Soc* 127: 1005–1015
- Dobi-Wantuch I, Mika J, Szeidl L (2000) Modelling wet and dry spells with mixture distributions. *Meteorol Atmos Phys* 73:245–256
- Frei C, Davies HC, Gurtz J, Schär C (2000) Climate dynamics and extreme precipitation and flood events in Central Europe. *Integr Assess* 1:281–299
- Frei C, Christensen JH, Deque M, Jacob D, Jones RG, Vidale PL (2003) Daily precipitation statistics in regional climate models: evaluation and intercomparison for the European Alps. *J Geophys Res* 108:4124, doi:10.1029/2002JD002287
- Frich P, Alexander LV, Della-Marta P, Gleason B, Haylock M, Klein Tank AMG, Peterson T (2002) Observed coherent changes in climatic extremes during the second half of the twentieth century. *Clim Res* 19:193–212
- Good P, Barring L, Giannakopoulos C, Holt T, Palutikof J (2006) Non-linear regional relationships between climate extremes and annual mean temperatures in model projections for 1961–2099 over Europe. *Clim Res* 31:19–34
- Hadley Centre (2002) The Hadley Centre regional climate modelling system: PRECIS update 2002. Met Office Hadley Centre, Exeter
- Hudson DA, Jones RG (2002) Regional climate model simulations of present-day and future climates of Southern Africa, Hadley Centre Tech Note 39. Met Office Hadley Centre, Exeter
- Jones RG, Murphy JM, Noguer M (1995) Simulation of climate change over Europe using a nested regional climate model. 1. Assessment of control climate, including sensitivity to location of lateral boundaries. *QJR Meteorol Soc* 121:1413–1449
- Jones RG, Noguer M, Hassell DC, Hudson D, Wilson SS, Jenkins GJ, Mitchell JFB (2004) Generating high resolution climate change scenarios using PRECIS. Met Office Hadley Centre, Exeter
- Klein Tank AMG, Können GP (2003) Trends in indices of daily temperature and precipitation extremes in Europe, 1946–99. *J Clim* 16:3665–3680
- Klein Tank AMG, Wijngaard JB, Können GP, Böhm R and 35 others (2002) Daily dataset of 20th-century surface air temperature and precipitation series for the European Climate Assessment. *Int J Climatol* 22:1441–1453
- Linderson ML (2003) Spatial distribution of meso-scale precipitation in Scania, southern Sweden. *Geogr Ann* 85A:183–196
- Maracchi G, Sirotenko O, Bindi M (2005) Impacts of present and future climate variability on agriculture and forestry in the temperate regions: Europe. *Clim Change* 70:117–135
- Osborn TJ (1997) Areal and point precipitation intensity changes: implications for the application of climate models. *Geophys Res Lett* 24:2829–2832
- Palmer TN, Räisänen J (2002) Quantifying the risk of extreme seasonal precipitation events in a changing climate. *Nature* 415:512–514
- Palutikof J (2006) Preface from the MICE project co-ordinator. *Clim Res* 31:1
- Peterson TC, Folland CK, Gruza G, Hogg W, Mokssit A, Plummer N (2001) Report on the activities of the Working Group on Climate Change Detection and Related Reporters 1998–2001, Reports WCDMP-47, WMO-TD 1071. World Meteorological Organisation, Geneva
- Semenov VA, Bengtsson L (2002) Secular trends in daily precipitation characteristics: greenhouse gas simulation with a coupled AOGCM. *Clim Dyn* 19:123–140
- Seneviratne SI, Pal JS, Eltahir EAB, Schär C (2002) Summer dryness in a warmer climate: a process study with a regional climate model. *Clim Dyn* 20:69–85
- Solberg S (2004) Summer drought: a driver for crown condition and mortality of Norway spruce in Norway. *For Pathol* 34:93–104

*Submitted: April 6, 2005; Accepted: February 25, 2006*

*Proofs received from author(s): June 9, 2006*