Climatic Change (2007) 85:381–406
DOI 10.1007/s10584-007-9263-2

On complex extremes: flood hazards and combined high spring-time precipitation and temperature in Norway

Rasmus E. Benestad · Jan Erik Haugen

Received: 26 October 2004 / Accepted: 14 March 2007 / Published online: 28 June 2007

Abstract Extreme weather and climatic events can have detrimental effects on society. The coincidence of several factors, themselves not necessarily extreme, can have similar adverse implications, such as a combination of high spring-time temperatures and heavy rainfall. A combination of high temperature and heavy precipitation during spring can produce flooding when run-off due to snow-melt adds to river discharge from the rainfall. Such combined events are often referred to as 'complex extremes' (IPCC, Climate change 2001: impact, adaptation and vulnerability. Summary for policymakers. WMO, Geneva, Switzerland, p. 7, 2001) . A likely effect of a climate change is a shift in the frequency of both extremes in traditional sense as well as in complex extreme events. Results from a global climate model were downscaled through a higher-resolution nested regional climate model in order to obtain more realistic descriptions of regional climatic features in Norway. Empirically-based joint frequency distributions (two dimensional histograms) were used to study shifts in the frequency of complex extremes. A slight shift in the joint frequency distributions for spring-time temperature-and-rainfall was detected in downscaled results with HIRHAM from a transient integration with the ECHAM4/OPYC3 climate model following the IS92a emission scenario. The analysis involved values that spanned between ordinary and extreme values of the bivariate distributions complicating the estimation of a representative confidence interval as the data fall in the zone between different types of behaviour. The results from HIRHAM were spatially interpolated and compared with station observations, and substantial biases were revealed, however, the apparent model discrepancy is largely due to great small-scale variability due to a complex physiography. The general temporal trends predicted by the model are nevertheless realistic.

R. E. Benestad (✉) · J. E. Haugen
The Norwegian Meteorological Institute, P.O. Box 43, 0313 Oslo, Norway
e-mail: rasmus.benestad@met.no

1 Introduction

1.1 Background

An Intergovernmental Panel on Climate Change (IPCC) workshop on 'Changes in Extreme Weather and Climate Events' was held in Beijing, China, 11–13 June 2002 (IPCC 2002) where the importance of climate and weather extremes (henceforth referred to as 'extremes') was highlighted. The central question was how these extremes will change in the future as a result of a global warming. Easterling et al. (2000) have suggested that extreme events can become more frequent as a result of an anthropogenic climate change, and in many areas there have been reports of enhanced changes in the upper tail of the distribution of observed precipitation (IPCC 2001).

Traditionally, extremes in climatological contexts have been studied in terms of magnitude (variability), return periods of single events (probability distributions), or exceedence where one parameter is greater than a given threshold value. It is generally acknowledged that it is difficult to determine trends of very rare events and that conventional statistical tests are often inappropriate for testing trends in these (Frei and Schär 2001).

Despite the difficulties of detecting trends in extreme events, there is some empirical evidence for changes in extreme values over the 20th century. Frei and Schär (2001) analysed trends in heavy daily precipitation in order to investigate whether long-term changes in these events have contributed to the increase of the seasonal means. They examined records of event counts, which can be shown to follow a binomial distribution, and used logistic regression to provide a means for trend analysis. One important finding was that a large number of Swiss stations had a statistical significant trend at the 5% level in intense precipitation events for winter and autumn in the eastern regions characterised as the 'pre-Alpine flatland.' Schmidli and Frei (2005) documented statistical for significant positive trends at the 5% level for a number of statistics related to the winter-time precipitation strength and occurrence in Switzerland. López-Díaz (2003) proposed a non-parametric test for trends in rare events based on logistic regression, and applied this to minimum annual temperature from the Villanubla observatory (1939–2000) and precipitation from San Fernando (1839–2000). He examined the absolute annual minimum temperature (one per year) over 1939–2000 and found that the highest values for annual minimum temperature occurred towards the end of this interval, but no corresponding trend was detected for lowest minimum temperature. His analysis also pointed to more dry autumns becoming more frequent in the recent years while it has become more rainy in winter. Benestad (2003a, 2004) analysed the occurrence of record-breaking events, and found that record-high temperatures are broken more often that would be expected if the temperature records were stationary (independent and identically distributed, commonly referred to as 'iid' in the statistical literature). Hence the upper tails of the temperature distributions are being stretched towards higher values. Frich et al. (2002) found coherent spatial patterns over northern hemisphere land areas and Australia with increased frequency of warm summer nights and decrease in number of frost days. Thus, a significant proportion of the global land area has been increasingly affected by changes in climatic extremes during the second half of the 20th century.

Global climate models (GCMs) are commonly used to make projections about future climatic changes. Modelling work by Meehl and Tebaldi (2004) suggested consequences like more frequent and intense heat waves and longer duration directly from GCM results. Hegerl et al. (2004) concluded from two different GCM simulations that the detection of changes in seasonal mean temperature cannot be substituted for the detection of changes in extremes. They found more pronounced changes in simulated extreme precipitation than for the changes in the corresponding seasonal means. Palmer and Räisänen (2002) used a multi-model ensemble of GCMs to infer a reduction in the return period of extreme precipitation. Although the GCMs give a realistic description of large-scale climatic variability, they are not designed to describe local and regional scales necessary for studying extreme weather and climatic events. It is nevertheless possible to derive information about the small scales through downscaling of the GCM results. There are two downscaling methodologies: dynamical and empirical-statistical downscaling. Murphy (1999) and Hanssen-Bauer et al. (2003) have suggested that the two different types in general are associated with similar levels of skill for describing mean values, although the different approaches have different strengths and weaknesses.

As with studies on historical observations, there are relative few studies in which downscaling has been used to study extreme weather or climatic events for the future. Schubert (1998) used a perfect prognosis downscaling technique applied to SLP to study local extreme temperature changes, and found that the dramatic changes in extreme temperatures seen in the GCM were not reproduced by the downscaling when using SLP as predictor. Hence, the changes in extreme temperatures were purported not to be driven by circulation changes but rather by changes in radiative properties of the atmosphere. Regression type models often assume Gaussian variables and tend to neglect the noise term, and it is therefore often the case that these types of statistical models are not suitable for reproducing the tails of a distribution (von Storch 1999). Senior et al. (2002) also examined extremes in terms of precipitation in a climate model and concluded that a warmer climate may lead to a more intense hydrological cycle and reduced return periods for intense rainfall. Imbert (2003) applied an analog model to results (time slices for "1980–1999" and "2030–2049") produced by a Regional Climate Model (RCM), whose boundary descriptions were prescribed by a GCM, to study extreme daily mean temperature and 24-h accumulated precipitation respectively for 91 stations in Norway. Her analysis indicated a general shift in the location of daily winter temperature distribution, but with no systematic change in the variance. Although analog models do not suffer from the shortcoming of regression models, this type of models cannot extrapolate values beyond the range of the historical sample for cases where the upper tails of the distribution is stretched and the variable are non-iid. Hayhoe et al. (2004) applied statistical downscaling to the NCAR PCM (Washington et al. 2000) and HadCM3 (Gordon et al. 2000) global climate models following the IPCC SRES B1 and A1fi emission scenarios to infer changes in extreme temperature over California. They obtained increased extreme temperatures, both in frequency and in magnitude.

The literature on climatic extremes tend to focus on univariate events, such as temperature or precipitation. There are few studies on more complex type events, such as a combination of high temperature and heavy precipitation (IPCC 2002). For some impact studies, such complex type events may be more interesting than

ordinary univariate analysis. A motivation behind this study is that a combination of high spring-time temperatures and heavy rainfall in Scandinavia can lead to severe flooding due to the additional run-off from heavy rain on top of the contribution from Schmidli and Frei (2005). Flooding has implications for the local society (i.e. disruptions, damage to infrastructure), the economy (i.e. insurance) as well as the ecology (e.g. increase risk of pollution). Such complex events may be analysed in terms of joint distributions, from which probabilities and return-periods can be derived.

Yue (1999, 2001) and Yue and Rasmussen (2002) carried out bivariate frequency analysis on flooding events and applied joint probability distributions to describe the likelihood of combined events such as flood peaks and durations. They argued that there are two types of joint return periods: (1) $T(x, y)$ associated with the case that either x, y or both are exceeded ($X > x$ or $Y > y$ or $[X > x$ and $Y > y]$), and (2) $T'(x, y)$ which represent events where both x and y are exceeded ($X > x$ and $Y > y$). An analysis on risk of spring-time flooding, i.e.events consisting of a combination of high temperature and heavy rainfall, has a similar 2-dimensional character as that of Yue and Rasmussen.

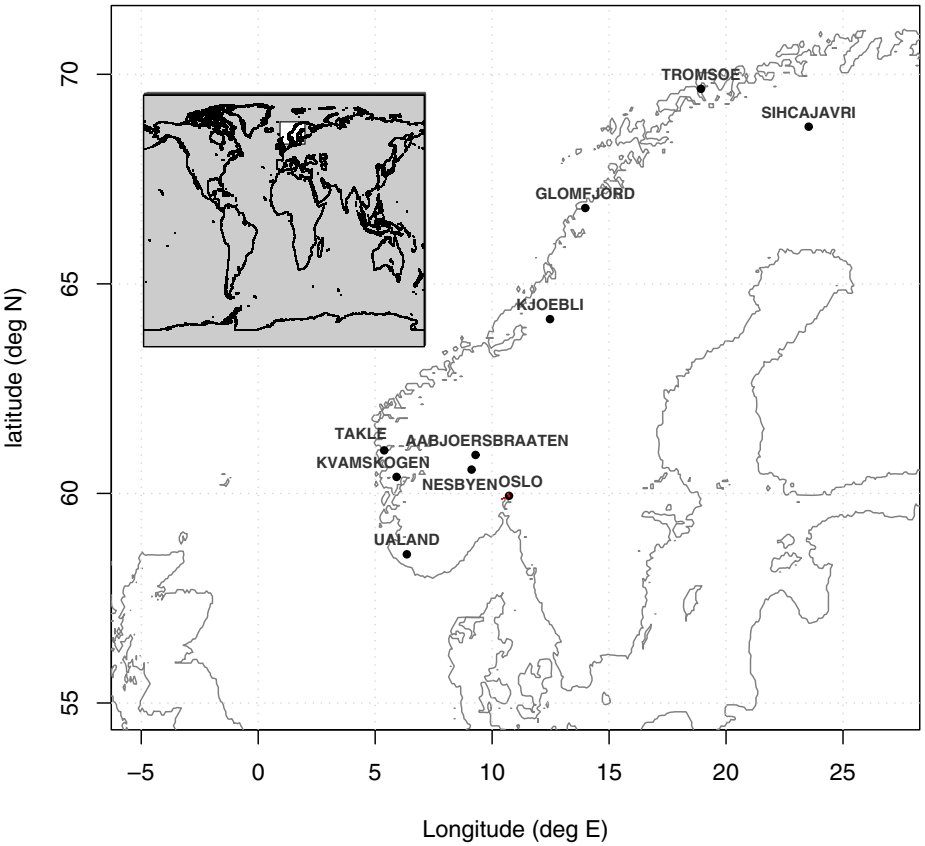

Fig. 1 Map showing the location of the climate stations

1.2 Structure of the paper

The data used in this study are described in Section 2 and the methodology in Section 3 as well as the Appendix. These are followed by a presentation of the results in Section 4, and a discussion and a summary of conclusions. In Section 5, we provide explanations for why bivariate Gumbel distributions may not be appropriate for complex events of the type that we consider here. This paper also provides a documentation of HIRHAM in terms of how well it represents high spring-time temperatures and heavy rainfall. Hence, the question of how well the results from the model experiment apply to the real world will also be discussed.

2 Data

Daily observations of 2-m temperature and precipitation for the selected locations in Norway were taken from the climate station data base ("Klimadatavarehuset") from the Norwegian Meteorological Institute. The criterion for selecting these stations was a combination of high data quality, the representation of a variety of local climate types, and that the stations have been used in other projects and thus provide a useful means for relating to other studies. The location of these stations are shown in on the map in Fig. 1 and further details about the stations are listed in Table 1. A combination of high spring-time temperature and heavy rainfall over a few days can cause severe spring floods, and therefore both daily values and 5-day means are analysed. It is important to keep in mind that extreme rainfall amounts predicted by the RCM is not equivalent to that of the station observations, as the former represent an areal mean of one grid box whereas the latter is a point measurement of a quantity with complicated small-scale spatial variations.

Downscaled model data were obtained from the HIRHAM model (Christensen et al. 1996), subsequently interpolated to the station coordinates for comparison with observations. The HIRHAM simulations were carried out with 50 km horizontal resolution and 19 vertical levels. The RCM integration domain covered Europe, North Atlantic and Greenland (The RCM has a rotated spherical grid,

Table 1 Detail about the station records used in this study

Location	Station number	Longitude	Latitude	Altitude	Days	Interval
Oslo–Blindern	18700	10.721°E	59.943°N	94	5,244	1950–2006
Ualand–Bjuland	43500	6.354°E	58.547°N	196	2,637	1968–1997
Takle	52860	5.385°E	61.027°N	38	4,692	1956–2006
Kvamskogen	50300	5.913°E	60.393°N	408	4,597	1957–2006
Kjøbli i snåsa	70850	12.473°E	64.159°N	195	4,876	1954–2006
Åbjørsbråten	23160	9.290°E	60.918°N	639	4,849	1954–2006
Nesbyen–Skoglind	24880	9.122°E	60.569°N	167	2,468	1977–2005
Glomfjord	80700	13.981°E	66.810°N	39	4,537	1954–2006
Tromsø	90450	18.928°E	69.654°N	100	5,242	1950–2006
Sihcajavri	93900	23.534°E	68.750°N	382	4,842	1954–2006

The number of days in the last column only includes the spring season (March–May) and gives an indication of the number of data points involved

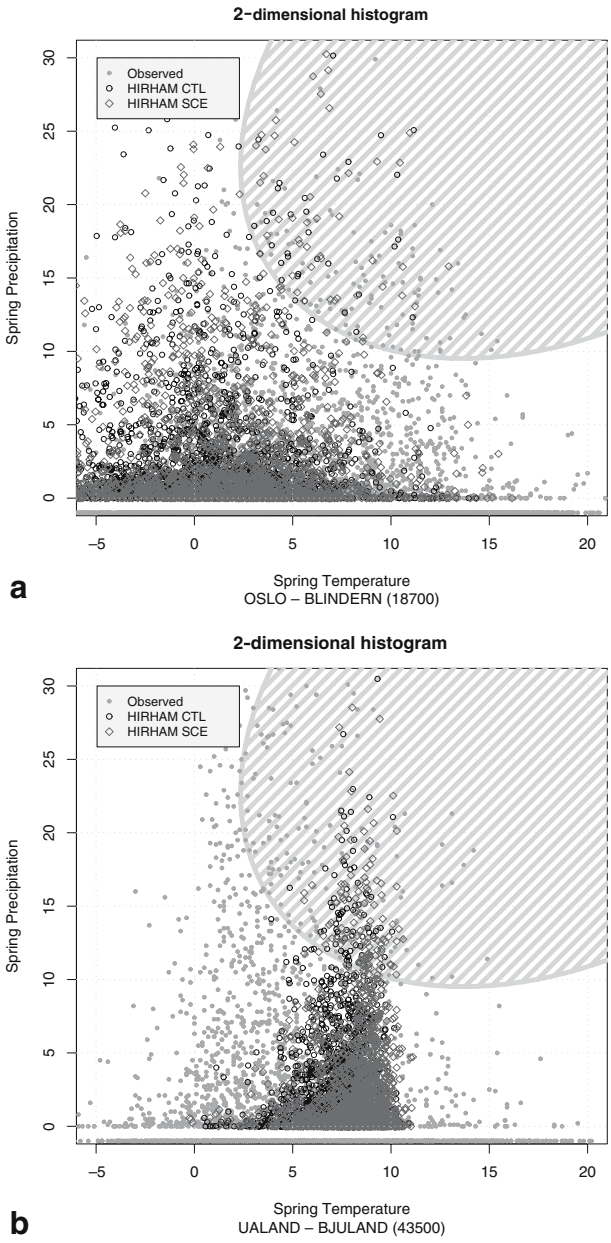

Fig. 2 **a** Scatter plots for daily spring-time (March–May) temperature and 24h rainfall recorded at **a** Oslo–Blindern and **b** Ualand–Bjuland for the period 1950–2004. Corresponding values interpolated from CTL and SCE derived using bilinear interpolation. The *grey* hatched elliptical region in the *upper right corner* is used as a criterion for counting the number of 'extreme' events (**b**)

and the domain can be seen in Fig. 5). Daily values of 2 m-temperature and total precipitation from 20-year time-slices (present-day control run and scenario period) were extracted for the analysis. The control integration will henceforth be referred to as 'CTL' (time slice "1980–1999") while the scenario run will be denoted 'SCE' (time slice "2030–2049"). The RCM was forced with 12-hourly output from the coupled ECHAM4/OPYC3 GCM (Roeckner et al. 1992; Oberhuber 1993)

at the lateral boundaries. Sea surface temperature and ice-cover were specified from the GCM. The ECHAM4/OPYC3 model (transient climate response of 1.4; effective climate sensitivity of 2.6; Houghton et al 2001, p. 539) has a T42 spatial resolution (which gives 128 × 64 grid points), a constant flux adjustment, and the model results were from the GSDIO integration that follows the IS92a emission scenario that also includes the indirect effects of industrial aerosols as well as tropospheric ozone. The parameterization of physical processes in HIRHAM uses the same code as in the driving GCM except for some modification to account for higher resolution. The HIRHAM model uses the same dynamical model as the numerical weather model (HIRLAM) used by the Norwegian Meteorological Institute for operational forecasts, although these differ in terms of parameterisations (HIRHAM uses the same parameterisation as ECHAM4) and that the routine numerical model output is subjected to a Kalman filter in order to compensate for systematic biases.

3 Method

Several simple approaches for detecting shifts in the joint probabilities are explored: (1) by comparing the contours of the joint frequency distributions, (2) by counting the number of events that fall into the category high precipitation and high temperature, and (3) by analysing cuts through the 2-D joint distributions.

Empirically-based two-dimensional (2-D) joint distributions were constructed following the method proposed by Yue and Rasmussen (2002). The counts of events falling into the various bins (the bin size was taken as 1/30 of the ranges in both of the two dimensions) were taken to represent the frequency of occurrence of the category of events (according to Eq. 1 in the Appendix), where n_{ij} is taken as the number of points in a temperature-rainfall scatter plots within a set of temperature-rainfall ranges. Approach (1) consisted of comparing contours of n_{ij} for SCE and CTL, whereas approach (2) involved counting the number of times a given criterion is fulfilled. Here, complex extremes can also be defined as the number of data points in a scatter-plot that fall within the predefined region denoting high temperature and high precipitation. The predefined regions in this case were ellipsoids, as a rectangular shape could result in an under-count of points representing high temperature (precipitation) and moderate precipitation (temperature) that are just outside the region whereas points near the lower corner that barely fall within the rectangle would be counted. It is easy then to use this complex extreme count from different time intervals to infer whether there are changes in the frequencies of complex events.

Since we are interested in the question of recurrence of high rainfall combined with high temperatures, crude 'semi-univariate' frequency distributions – approach (3) – for the complex events may be constructed by interpolating the values in the 2-D bins along the straight cuts trough the bivariate distribution. The choice of the cut through the 2-dimensional distribution was subjective to the extent of the its precise location as the objective here was to obtain a rough representation of the frequency of days with high temperature and heavy precipitation. Here we are only looking for tendencies in terms of projected future changes, and the ad hoc choice can be justified from the differences between the contours representing CTL and SCE in a 2-D plot (1): the tendencies are representative for changes in the contours.

The cuts through the 2-D joint distributions can be expressed as $F(x|y_0)$ where y_0 is not constant but varies with x. The key question this paper tries to answer is whether there is a shift in the probabilities of the combination of warm spring days and heavy precipitation in a climate scenario from a single state-of-the-art GCM, rather than trying to estimate the true return values for the real world. For this reason, we have not attempted to adjust the interpolated temperatures according to correct altitude, as this kind of correction will not affect the estimated change. Likewise, we have not adjusted the interpolated precipitation. The histograms taken along the cuts are also complemented with the analysis of the counts of events in the predefined region (2) exploring a slightly different part of the data space.

Here, statistical extreme value modelling (as given by Eqs. 2–11 in the Appendix) will merely be used as a basis for estimating confidence intervals. Monte Carlo type simulation and binomial modelling were used for estimating the confidence interval for the counts falling in the predefined region. More details about the methods are provided in the Appendix.

4 Results

Figure 2 shows daily temperature-rainfall scatter plots for Oslo and Ualand respectively. In the Oslo case, the cloud of observed values appears to have similar location and spread as that of the modelled values. The grey-hatched elliptical region in the upper right corner of the figures marks the predefined region which was used to identify 'extreme' complex events. Data points within this region are counted in Table 2 (henceforth referred to as count rates), and a comparison between the number of points from the CTL and the SCE integrations provides a basis for studying changes in the frequency of complex extremes. Corresponding counts from the station records are also shown in order to assess the similarity between the model and empirical values. There are some differences in magnitude between CTL and the station data, especially for Takle and Kjøbli (daily means), but it is important to keep in mind that the model values represent combinations of grid box areal means whereas the observations represent point measurements.

For the 5-day means, the comparison between the empirical and CTL counts reveals more substantial differences. Keeping these discrepancies in mind, one can look at the differences between CTL and SCE in order to make a tentative projection. These suggest that a global warming may entail an increase in the frequency of combined high temperature and precipitation on a 24-h basis. For events of longer durations, such as 5 days, the projected increase is less systematic since the projected changes for Ualand, Takle and Kvamskogen are negative.

The estimation of confidence intervals for the counts is discussed in the Appendix. Different results obtained using a variety of approaches suggest substantial uncertainty associated with these estimates, however, it was concluded that Monte Carlo simulations following a Gaussian/Gamma bivariate distribution provided the most realistic description of a null-distribution representative for CTL. The Gamma function usually provides a good approximation of empirical frequency distributions (histograms) of observed daily rainfall amounts (Benestad et al. 2005). According to this model, a count rate change (SCE-CTL) of 0.015 (0.019 → 0.026; columns 4 and 5 in Table 2) for Oslo is outside the range of the 95% confidence range of CTL

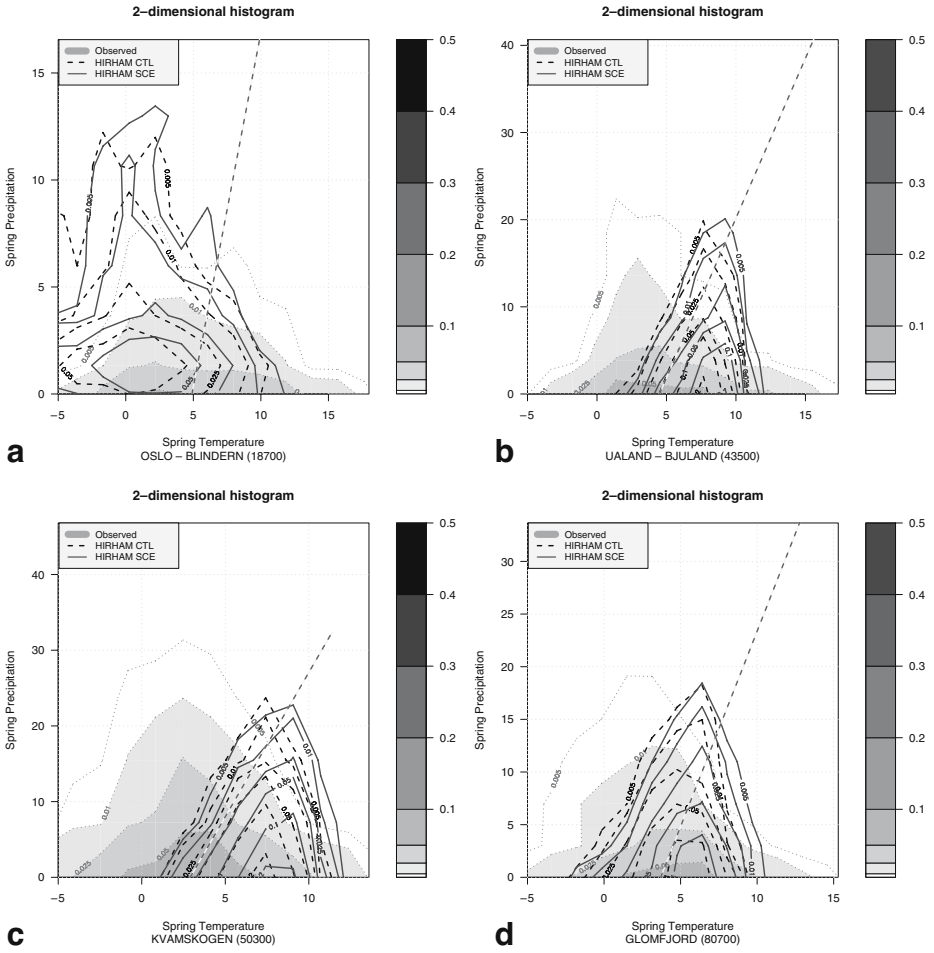

Fig. 3 *Contour lines* of the joint frequency distribution estimated from the cloud of points in Fig. 3 (using the expression in Eq. 1 from the Appendix) as suggested by Yue and Rasmussen (2002). The panels show the bivariate distributions for **a** Oslo–Blindern, **b** Ualand–Bjuland, **c** Kvamskogen, and **d** Glomfjord. The *straight dashed lines* mark the cross-sections used for making 1-D histograms, and was subjectively chosen so that they provided a representation of wet and warm conditions by intercepting the points $\max[0, x_{0.50}]$, 0 and $2x_{0.75}, 3y_{0.95}$. The bin size was determined by dividing the range of temperature and precipitation into 30 intervals respectively, and therefore varied from location to location

(0.009–0.021; Table 6 in the Appendix, under 'M-C all' row and the 'Gauss./Gamma' column)) for the daily results. The confidence intervals varied with the station, however, the changes for Ualand (0.040 → 0.063), Nesbyen (0.015→ 0.032) and Tromsø (0.003 → 0.022) as seen in columns 4 and 5 of Table 2 also exceed the difference between the range of estimated confidence intervals (Table 6: 0.029–0.048, 0.016–0.031, and 0.003–0.012 respectively). In addition to using the confidence interval to assess whether the projected change in the number of events was significant, we also used the same analysis to assess the difference between the observations

Table 2 The number of events divided by length of series in the observed record (third column), CTL results (fourth) and the SCE results (fifth)

Station number	Location	Obs. freq.	CTL freq.	SCE freq.
Daily means:				
18700	Oslo	0.021	0.019	0.026
43500	Ualand	0.058	0.040	0.063
52860	Takle	0.116	0.034	0.046
50300	Kvamskogen	0.054	0.046	0.062
70850	Kjøbli	0.007	0.022	0.027
23160	Åbjørsbr.	0.008	0.008	0.019
24880	Nesbyen	0.010	0.015	0.032
80700	Glomfjord	0.033	0.031	0.038
90450	Tromsø	0.002	0.003	0.022
93900	Sihcajavri	0.001	0.005	0.008
5-day means:				
18700	Oslo	0.043	0.111	0.136
43500	Ualand	0.214	0.186	0.175
52860	Takle	0.286	0.094	0.094
50300	Kvamskogen	0.223	0.239	0.217
70850	Kjøbli	0.037	0.128	0.192
23160	Åbjørsbr.	0.020	0.078	0.142
24880	Nesbyen	0.020	0.117	0.208
80700	Glomfjord	0.175	0.194	0.233
90450	Tromsø	0.020	0.033	0.119
93900	Sihcajavri	0.002	0.031	0.064

The tendency described by these results point to an increased frequency of the complex extreme events for the future. The confidence limits for these are discussed in the Appendix.

(third column in Table 2) and the CTL. The estimated confidence intervals based on Monte Carlo fitted to all data ('M-C all') span all the CTL counts rates for daily means, except for Sihcajävri (CTL freq. is 0.005 for daily means in Table 2, but c.i. for CTL is 0.000–0.003 according to Table 4).

The plots in Fig. 3 show examples of joint frequency distributions for both real observations as well as RCM results interpolated to the location for Oslo, Ualand, Kvamskogen and Glomfjord. The shaded contours show the analysis for the observations whereas the contour lines indicate the model values. The 2-D distribution for Oslo shows that the RCM tends to produce a rainfall distribution with a somewhat thicker upper tail (note the contour intervals are not evenly spaced). A comparison between the distributions for CTL and SCE, suggests a slight shift toward higher temperature and higher rainfall. A similar shift can be seen in the other three locations (panels b–d). These results are consistent with the increase of counts given by Table 2.

The scatter of temperature-precipitation points and edfs derived from the RCM are consistent with the observations from Glomfjord (Fig. 4d). Similar edfs for Ualand–Bjuland (b) and Kvamskogen (c) suggests that the good agreement between the observed and modelled curves is coincidental, as the RCM has a tendency to overestimate the frequency of combined high spring-time temperature and rainfall in these locations. The fortuitous similarity for Glomfjord can also inferred from the

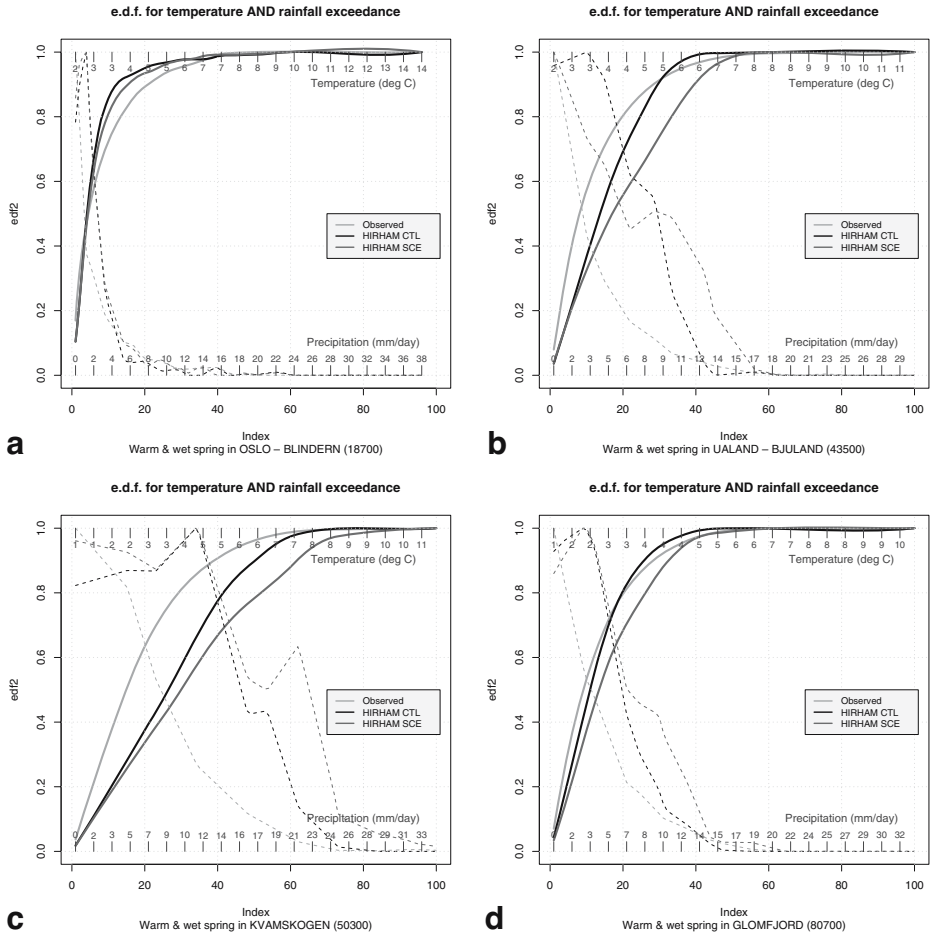

Fig. 4 Histograms (*thin lines*) and edf (*thick lines*) for temperature-rainfall events estimated along the *thick grey-and-black dashed lines* in Fig. 4. The panels show the bivariate distributions for **a** Oslo–Blindern, **b** Ualand–Bjuland, **c** Kvamskogen, and **d** Glomfjord

contours in Fig. 3. In all the cases, a comparison between the upper parts of the distributions (upper tail of distribution) indicates that the downscaled ECHAM4/OPYC3 scenario projects a shift towards more days with combined high rainfall and high temperature.

Figure 5 shows projected changes in the 95-percentile return period for temperature and precipitation according to HIRHAM. In both cases, the return period is reduced, indicating that both extreme temperatures and precipitation are more frequent in the scenario than in CTL. Over the Norwegian mainland, the 95-percentile return factor increases more with respect to precipitation than for temperature, suggesting that the most important factor is the precipitation. However, the bulk of the shift in temperature appears to be more substantial than for precipitation in

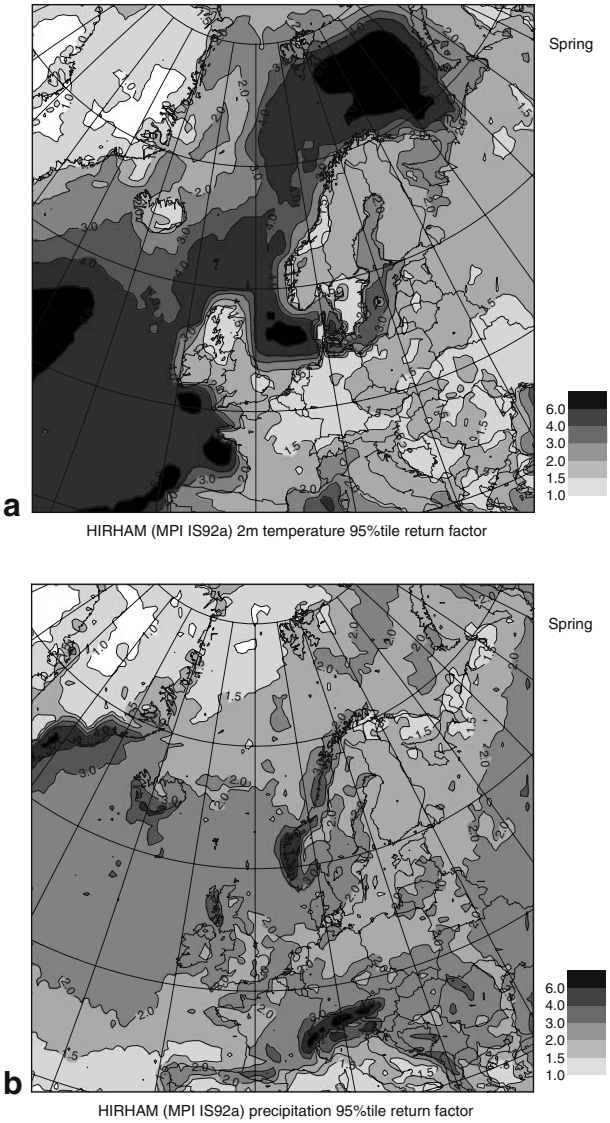

Fig. 5 Maps of changes in the ratio of 95% return period R for 2-m temperature (**a**) and precipitation (**b**). Here $R = 1$ means same return period as in CTL whereas $R = 2$ a doubling in frequency (i.e. the return period is halved)

Fig. 3, i.e. the change in extremes cannot be inferred directly from the change in the mean (Hegerl et al. 2004).

Another point concerning the question whether these results can be applied to the real world involves decadal variability. The short time slices of 20 years are strongly affected by internal fluctuations on inter-annual and decadal time scales that are unrelated to the global warming. A time series plot can be used to assess the contribution of internal variability to the detected shifts in the frequency distributions (illustrated in Fig. 6). Such internal (chaotic) variations influence the magnitude of sampling fluctuations for short data series. The exact temporal evolution of the variables may depend on the GCM's initial conditions, however, our knowledge of

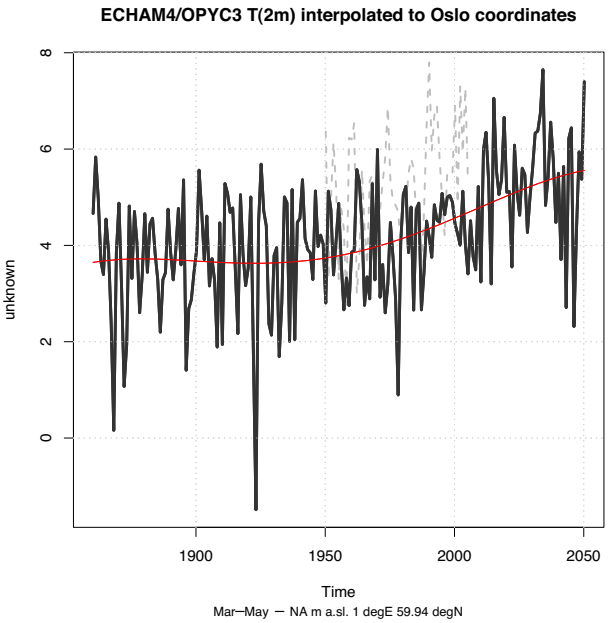

Fig. 6 Time series of interpolated 2m-temperature from the ECHAM4/OPYC3 model (*thick dark*i) and T(2m) temperature measured at Oslo–Blindern (*dashed light grey*)

the state of the atmosphere for the past is incomplete. Hence, the trajectory of a single scenario is associated with a degree of uncertainty (Benestad 2001). Robust result can only be obtained from a multi-scenarios approach, which would imply large samples and hence reduce such sampling fluctuations.

Hayhoe et al. (2004) inferred from a global warming study that the snow-pack in California may be reduced, which again has cascading impacts on run-off. Dankers

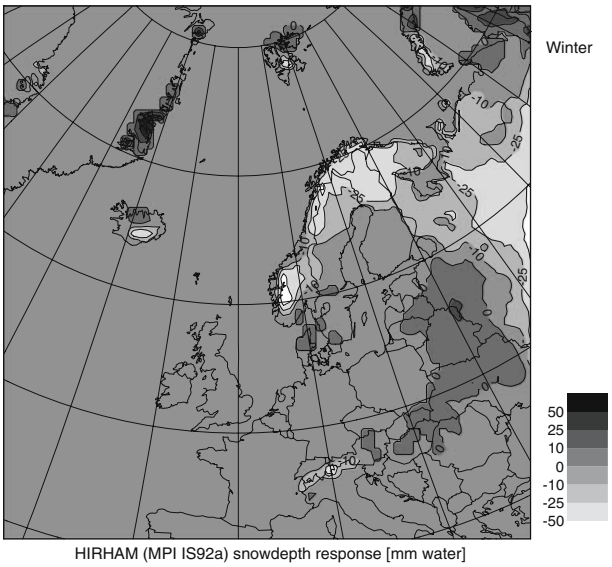

Fig. 7 Maps of changes in snow-depth (snow water equivalent in mm) projected by HIRHAM

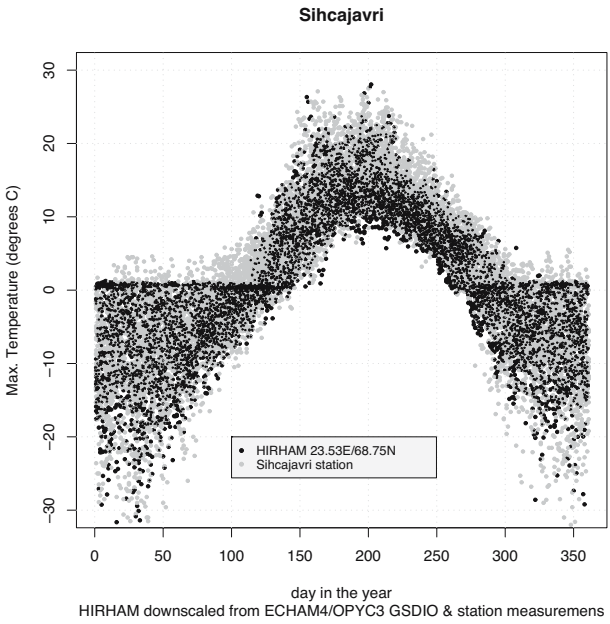

Fig. 8 Daily maximum temperatures through out the year for observations at Sihcajävri and simulations interpolated for the same location (CTL). The comparison suggests that the RCM results are subject to systematic errors in spring-time as the modelled values tend to be restricted to negative values while the observations also indicate positive values

and Christensen (2005) also suggested that a future warming may alter the timing and amount of the annual peak river discharge. The snow-melt, according to their results, is projected to start earlier and increase the risk of winter flooding while reducing spring and summer stream-flows. Similar kind of changes may take place in Norway, and it is important to keep this in mind when adopting these findings to the real world. The RCM in this study projects a reduction in snow-depth in the mountain regions of Norway (Fig. 7), despite an increase in the snow fall during winter. A substantial reduction of the snowpack can reduce the risk of spring-time flooding, however, the snow does not disappear altogether according to the HIRHAM results, suggesting that a combined event of high spring-time temperature and heavy precipitation still represents a risk of flooding.

It is important to keep in mind that the RCM description of the snow-extent may be influenced by systematic biases (Wood et al. 2004) associated with details in the description of the local scales. Figure 8 shows results that have implications for the snow-cover, comparing the maximum temperature for Sihcajävri with corresponding values interpolated from the RCM. It is evident that there is a systematic cold bias in the RCM spring-time maximum temperature: the simulated maximum temperatures have a tendency to be artificially confined to below zero in spring, and the RCM under-estimates daily spring-time temperature extremes. Such biases will affect the snow-melting process, and this problem appears to be typical for the Arctic regions for this RCM. The difference between the maximum temperature from the station data and the RCM may reflect the fact that whereas the former represents a point measurement the latter describes interpolated values from area averages.

An estimation of confidence intervals often requires a reliable estimate of probabilities, and in this case our objective was to obtain a statistical model for a bivariate distribution that was representative for both the extremes as well as ordinary values.

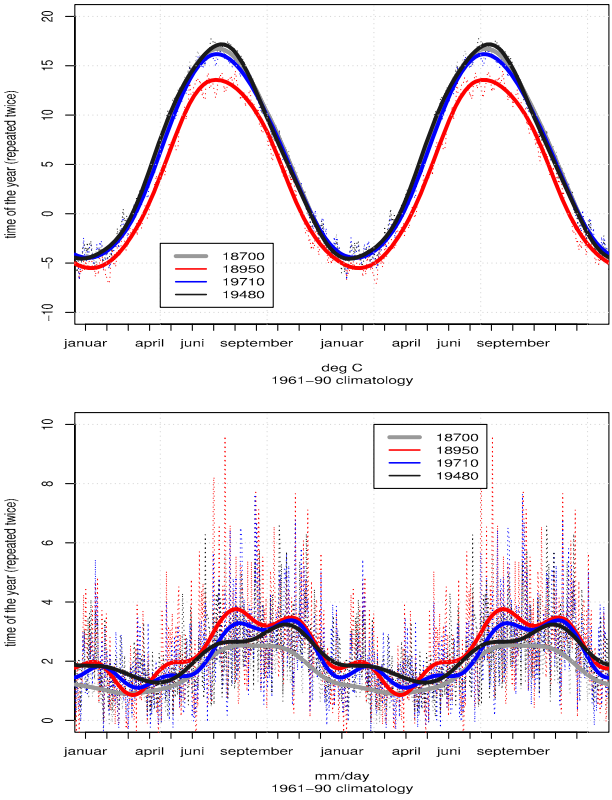

Fig. 9 Comparison of the 1961–1990 climatologies of locations near Oslo: Blindern–Oslo (18700; 94 m.a.s.l.), Tryvann–Oslo (18950; 514 m.a.s.l.), Asker (19710; 163 m.a.s.l.), and Dønski (19480; 59 m.a.s.l.). All of which lie within a radius of 12 km. The large difference between nearby stations is not particular for the Oslo-region, but is typical for most of Norway

The Gumbel distribution did not provide as close fit to the edf of the entire batch as the Gaussian/Gamma distribution functions. The estimation of Gumbel-parameters are typically obtained using the annual maximum (or the r greatest values each year) to provide a better representation of the upper tails of the distribution functions, and therefore the analysis was repeated for a subset of the data consisting of the five highest values each year ($5 \times 20 = 100$ data points). The greatest daily temperature and precipitation were not strongly correlated, which implied selecting different days for the two variables, however, it is not required that these events have to correspond (Eq. 6 in the Appendix). In our case, we were not interested in just the upper tail of the distributions of each variable as complex extremes may also involve more ordinary events that in combination becomes extreme: The probability of two independent extreme events taking place simultaneously is much lower than two more ordinary events that combined give rise to severe conditions. Moreover, complex extremes may involve values that are not extreme in isolation, and thus lie in the transition zone between ordinary and extreme values. Thus, the use of Gaussian/Gamma distribution functions may be more appropriate to use for studying risks associated with combined high snow melt and heavy precipitation than bivariate Gumbel distributions.

Two different approaches were employed to estimate 95% confidence intervals: (1) based on Binomial distribution utilising estimates of probability of CTL points

falling inside the elliptic region, and (2) based on Monte Carlo simulations using different statistical models for the null-distribution. The Monte Carlo method based on a combination of Gaussian and Gamma distributions gave the most appropriate confidence intervals, but the large spread in the results illustrated the difficulties associated with statistically modelling the transition zone between ordinary and extreme values in bivariate distributions.

5 Discussion

When using climate models in making projections for the future, it is critical to examine how well the models describe the real world. It is therefore necessary to assess shortcomings and uncertainties associated with the downscaling of such extremes and document systematic biases before model results can be related to the real world. This paper serves as a documentation of how well the HIRHAM model represents extreme daily spring-time temperature and precipitation. The over-representation of extreme events in climate models may result in erroneous conclusions about the future climate, even when based on the difference between SCE and CTL. It is believed that the RCM data is not sufficiently representative of the real world to provide accurate estimates of return values for such complex events in the real world, as Skaugen et al. (2002) have reported systematical biases in the HIRHAM regional climate model when integrated with ERA-15 boundary conditions. In areas with complex physiography, RCMs are not capable of resolving local-scale climate differences that are typical for Norway (Wood et al. 2004). Figure 10 shows differences in the annual cycle of temperature and precipitation for four stations within a radius of 12 km near Oslo, comprising an altitude range of 59–514 m a.s.l., and it is clear that such variation cannot be represented by a model with a 50×50 km^2 spatial resolution and a smoothed topography. The RCM produces a warm bias and too little temperature variability. Temperature biases can be explained by different altitude used in the RCM (Table 3) as the RCM uses a smoothed topography where elevations reflect an area average rather than the station altitude and the RCM data used here have not been adjusted for this height difference. Part of the warm biases at Kvamskogen and Glomfjord can therefore be explained by differences in elevation. Furthermore, a 50×50 km^2 spatial resolution is not sufficiently high for representing the local precipitation, explaining some of the discrepancies between the CTL and the observations.

Skaugen et al. (2002) have already noted discrepancies between results from HIRHAM and corresponding empirical data, and suggested a simple adjustment scheme for making the RCM results more in line with actual observations. It is not certain whether such modifications will give a sufficient description of the tails of the distributions. Alternatively, more sophisticated method involving a linear empirical-statistical downscaling model (Benestad and Hanssen-Bauer 2003) to refine temperature scenarios from HIRHAM may be used to fit general extreme value models, but again it is important to ensure that such statistical models give a good description of the tails of the distribution (many models, such as regression, are not suitable for modelling extremes).

Although the HIRLAM model tends to produce biases when interpolated down to local climate, the fact that the local variability is strongly influenced by large scales (which is the general requirement for downscaling) enables forecasters to use the

Table 3 Altitudes corresponding to the station locations in the RCM

Location	z_{RCM} (m)	$z_{\mathrm{RCM}} - z_{\mathrm{station}}$ (m)	ΔT (°C)
Oslo	237	143	+0.93
Ualand	271	75	+0.49
Takle	285	247	+1.61
Kvamskogen	492	84	+0.55
Kjøbli	431	236	+1.53
Åbjørsbr.	890	251	+1.63
Nesbyen	851	684	+4.45
Glomfjord	364	325	+2.11
Tromsø	227	127	+0.83
Sihcajavri	414	32	+0.21

The second column lists the model altitudes, the third column the difference between the model altitude and actual station elevation, and the fourth column gives an estimate of the corresponding temperature bias (RCM – station) based on the lapse rate −0.655°C/100 m.

model for making skillful forecasts. Dankers and Christensen (2005) used a similar model (HIRHAM4) for studying hydrological response in northern Fennoscandia to a global warming scenario. It is often assumed that the model-observation differences are systematic errors that cancel when subtracting the results from a control run from a scenario run, and that these model-observation differences may be neglected. However, SCE-CTL difference may result in an overestimate in probabilities if the model yields an edf which approaches unity much slower than seen in the observations.

Furthermore, temperatures and precipitation may be non-stationary under a changing climate (Benestad 2003a, 2004), making it difficult to interpret fitted extreme value distribution functions. Hence, the view that return value analysis along the lines outlined by Yue and Rasmussen (2002) will not produce reliable results in this case.

6 Conclusions

In conclusion, the results with ECHAM4/OPYC3 (GSDIO) and the HIRHAM models generally point towards more frequent high-rainfall-high-temperature events, hence a greater risk for spring-time flooding given that the snow-extent does not change significantly. These results therefore seem to follow the trends reported by Hayhoe et al. (2004) and Senior et al. (2002), i.e. more extreme high temperature and precipitation. It is important to keep in mind that one single regional climate scenario may not capture the correct time evolution of climatic parameters due to long-term chaotic internal variability (Benestad 2003b, 2001) and model shortcomings. More robust conclusions must be based on ensembles of different climate projections and different models. In this study, we were particularly interested in whether the RCM gives a good representation of the temperature and precipitation in spring. The model-observation comparisons presented here, in terms of complex extremes, suggest substantial mismatch between the modelled and empirical values. RCMs should therefore not be applied uncritically to studies of climate impacts. The comparison

between the results from the RCM and the observations suggest that the HIRHAM is not able to capture the small-scale spatial dependencies of temperature and precipitation associated with unresolved sharp and complex topographical gradients.

The objective of this paper, however, is to present an intuitive and simple analysis, as it is questionable whether a more theoretical and rigorous treatment in terms of mathematical modelling of return values is going to lead to more robust conclusions or new knowledge. A theoretical and rigorous mathematical treatment cannot reduce the uncertainties in this case.

Acknowledgements This work was done under the Norwegian Regional Climate Development under Global Warming (RegClim) programme and was supported by the Norwegian Research Council (Contract NRC-No. 120656/720), and the Norwegian Meteorological Institute. The analysis was carried out using the R (Ellner 2001; Gentleman and Ihaka 2000) data processing and analysis language, which is freely available over the Internet (URL http://www.R-project.org/). We are also grateful for valuable comments from four anonymous reviewers.

Appendix

Details on statistical modelling of joint distributions

Yue and Rasmussen (2002) proposed using a so-called Gringorten type formula for constructing empirical joint frequency distributions:

$$f(q_i, v_j) = \frac{n_{ij}}{N + 0.12}, \tag{1}$$

where n_{ij} is the number of occurrences of the combination of q_i and v_j and N is the total number of data. We use the Gringorten formulae for the same reason as Yue and Rasmussen (2002): it provides an unbiased estimate for EV1 quantiles. However, we also compared the Gringorten with the more basic $f'(q_i, v_j) = n_{ij}/N$, but the choice of formula used for constructing empirical joint frequency distributions did not influence our conclusions (not shown). The Gringorten joint empirical distribution function (edf) was then expressed as:

$$F(q, v) = \frac{\sum_{j=1}^{q} \sum_{j=1}^{v} n_{ij} - 0.44}{N + 0.12}. \tag{2}$$

According to Yue and Rasmussen (2002), the cumulative distribution function (cdf) for a univariate Gumbel distribution can be written as

$$F_Z(z) = \exp\left[-\exp\left(-\frac{z-u}{\alpha}\right)\right], \qquad (Z = X, Y), \tag{3}$$

where u is the 'location' and α the 'scale' parameters. The moments estimators (Wilks 1995) for these are:

$$\alpha = \sqrt{6}S/\pi, \tag{4}$$

$$u = M - 0.577\alpha, \tag{5}$$

and S and M are the sample standard deviation and mean of the data used to fit the statistical model (i.e. the annual maximum value) respectively. In addition to these, maximum likelihood estimators were also derived using a package for the R environment called 'ismev' (which contains S functions for computations in Coles (2001) ported to R by Alec Stephenson; available from http://cran.r-project.org). The estimated values for these Gumbel parameters are given in Table 4. The joint cdf can then be written as

$$F(x, y) = \exp\left(-\left[(-\ln F_X(x))^m + (-\ln F_Y(y))^m\right]^{1/m}\right), (m \geq 1). \tag{6}$$

This expression was used to obtain the bivariate distributions for temperature and precipitation. The parameter m is a measure of the correlation ρ between the two variables: $m = 1/\sqrt{(1-\rho)}$ (Yue and Rasmussen 2002). In this case, the precipitation is far from Gaussian, and hence the estimate using a standard Pearson correlation on untransformed data, as done here, will be biased. The joint return period associated with the event that both x and y are exceeded can then be expressed in the form

$$T'(x, y) = \frac{1}{1 + F(x, y) - F_X(x) - F_Y(y)}. \tag{7}$$

The joint bivariate probability was estimated as $1/T'(x, y)$ for both the $F(x, y)$ fitted to the entire population and the upper tail only. An estimation probabilities for an event where temperature exceeds 10°C and precipitation is greater than 22 mm/day (i.e. a rectangle shaped predefined region as opposed to an ellipsoid) using the different statistical models and different estimators suggested generally similar values for models fitted within the upper tail of the distribution, but more substantial differences for when the entire data batch was used (Table 5). The expression for bivariate pdf for $F(x, y)$ in Eq. 6 is:

$$f(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y}, \tag{8}$$

and involves complex expressions which are different for the bivariate Gumbel, Gamma and Gaussian distributions. However, this function can easily be estimated numerically for any distribution type by using an approximation of $F_Z(z)$ based on an exponential series,

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^r}{r!} + \cdots, \tag{9}$$

as well as other types of functions. We used a multiple regression analysis to fit a number (here 8) of polynomials to $F(x, y)$:

$$\hat{F}(x', y') \approx a_0(y') + a_1(y')x' + a_2(y')x'^2 + \cdots + a_8(y')x'^8 \tag{10}$$

Here, we used a linear transform for y and x so that with $0 < x' < 1$ and $0 < y' < 1$. The first derivative for Eq. 10 can now easily solved using the same approach as Benestad (2003b):

$$\frac{\partial \hat{F}(x', y')}{\partial y'} = \int \hat{f}(x, y)dx = a_1(y') + 2a_2(y')x' + \cdots + 8a_8(y')x'^7 \tag{11}$$

The operation described in Eq. 11 was applied again in the x' dimension and the coordinates were rescaled to x, y in order to arrive at the final pdf. An example of the results are shown in Fig. 10, and the numerical differentiation was applied to the bivariate Gumbel as well as Gaussian/Gamma distributions. The pdfs derived from Eq. 11 were used to obtain a theoretical estimate for an event corresponding to a point within the ellipsoid (here denoted with the symbol $\in$), and was calculated a s

$$Pr(\in) = \sum_{\in} \hat{f}(x, y) / \sum_{\text{all}} \hat{f}(x, y), \tag{12}$$

summing the probability density over the ellipsoid and dividing by the total sum. The estimated probabilities are listed in Table 5. If all the data was used to fit the statistical distribution, the Gaussian/Gamma models yielded systematically and substantially higher probability estimates than the Gumbel models. When only the upper tail was used for fitting the models, the correspondence between the models improved (the count rate for the five greatest values per year in this analysis did not take into account correlation between temperature and precipitation, as the five warmest spring days were selected irrespective of the amount of precipitation and vice versa).

The expected counts x within the ellipsoid was modelled using a binomial distribution given the probability p:

$$\Pr(X = x) = \frac{N!}{x!(N-x)!} p^x (1-p)^{N-x}. \tag{13}$$

where $\Pr(\in)$ from Eq. 12 was used as an estimate for p, and $\hat{f}(x, y)$ was solved for assuming Gumbel as well as combined Gaussian/Gamma distributions (Table 6). Both moments and maximum likelihood estimators were used for the Gumbel distribution which was fitted to both the entire batch of data and the five greatest values respectively. The confidence intervals were only derived from the results of CTL.

In a second independent approach to derive confidence intervals, estimates of the location u and scale α parameters (Table 4) were used with a Gumbel-distributed random numbers in a series of Monte Carlo simulations for estimating confidence intervals (Table 6). Two sets of Monte Carlo simulations were performed assuming bivariate Gumbel distributions: one based on moments estimators and one using maximum likelihood estimators. The synthetic temperature and precipitation series were assumed to be uncorrelated in the Monte Carlo simulations, and the autocorrelation was assumed to be zero. Scatter plots between daily temperature and precipitation indicate a weak correlation between these elements (eg Fig. 2). A third set of Monte Carlo simulations were conducted using random Gaussian-distributed

data for temperature and stochastic Gamma-distributed values for daily precipitation instead of Gumbel-distributed data (Table 6). All the Monte Carlo simulations were performed for statistical models fitted using the entire data population. One difference between the Monte Carlo approach and the confidence interval derived from binomial distribution is that the former assumes no correlation between the temperature and precipitation whereas the latter to some degree takes correlation into account through the parameter m in Eq. 6, albeit using a biased value.

The different approaches to estimating the confidence intervals produced different ranges which were not internally consistent. The Gumbel-based results agreed well when only the five greatest annual values were used for fitting the statistical models, but produced inconsistent results when the entire data batch was used. In fact, the consistency between the different types of statistical models were improved if only the five greatest annual values were used, although the binomial approach systematically shifted the interval to higher values. However, a distribution derived from the five greatest annual values was not generally representative for CTL counts in Table 2. It is only the Monte Carlo simulations based on Gaussian/Gamma distributions that produce confidence intervals which are consistent with the CTL counts in Table 2.

Tables for the Appendix

Table 4 Location, shape and scale estimates for the Gumbel, Gaussian and Gamma distributions using moments as well as maximum likelihood estimators (MLE)

Location	Method	u	α	Mean/Scale	Stdv/Shape
Temperature:					
Oslo	Moments	3.305	−1.637	0.27	4.24
	MLE	4.355	−1.868		
Ualand	Moments	1.094	6.614	7.24	1.4
	MLE	1.782	6.49		
Takle	Moments	1.108	6.406	7.04	1.42
	MLE	1.765	6.289		
Kvamskogen	Moments	1.252	6.124	6.85	1.61
	MLE	1.997	5.998		
Kjøbli	Moments	3.106	0.810	2.6	3.98
	MLE	4.135	0.614		
Åbjørs.	Moments	3.08	−2.661	−0.88	3.95
	MLE	4.089	−2.867		
Nesbyen	Moments	3.084	−1.885	−0.11	3.96
	MLE	4.033	−2.081		
Glomfjord	Moments	1.555	3.630	4.53	1.99
	MLE	2.424	3.460		
Tromsø	Moments	2.709	−0.207	1.36	3.47
	MLE	3.599	−0.415		
Sihcajavri	Moments	4.357	−5.527	−3.01	5.59
	MLE	5.66	−5.823		
Precipitation:					
Oslo	Moments	3.959	1.439	6.924	0.538
	MLE	2.591	1.862		

Table 4 (continued).

Location	Method	u	α	Mean/Scale	Stdv/Shape
Ualand	Moments	2.943	1.173	4.963	0.579
	MLE	2.054	1.432		
Takle	Moments	3.163	1.458	5.013	0.655
	MLE	2.333	1.689		
Kvamskogen	Moments	3.252	1.412	5.29	0.622
	MLE	2.343	1.672		
Kjøbli	Moments	3.801	1.665	6.159	0.626
	MLE	2.621	2.046		
Åbjørs.	Moments	4.648	2.776	6.511	0.838
	MLE	3.687	3.007		
Nesbyen	Moments	5.073	2.792	7.401	0.773
	MLE	3.907	3.090		
Glomfjord	Moments	3.227	1.443	5.183	0.638
	MLE	2.344	1.704		
Tromsø	Moments	2.82	0.847	5.289	0.468
	MLE	1.773	1.215		
Sihcajavri	Moments	2.018	1.044	3.032	0.728
	MLE	1.375	1.260		

The total sample ($N = 1,800$) was used for these parameter estimations. Derived for CTL.

Table 5 Estimated probabilities for events in the ellipsoid ($\Pr(\in)$) derived from $\hat{f}(x, y)$ and probability of temperature exceeding 10°C simultaneously as precipitation is greater than 20 mm/day (derived as $1/T'$ and Eq. 7)

Location	$\Pr(\in)/\Pr(X > 10, Y > 20)$ (Gauss/Gamma)	$\Pr(\in)/\Pr(X > 10, Y > 20)$ moments	$\Pr(\in)/\Pr(X > 10, Y > 20)$ MLE
All data			
Oslo	0.217 / 0.0024	0.054 / 0.0029	0.015 / 0.0023
Ualand	0.335 / 0.0036	0.033 / 0.0019	0.004 / 0.0018
Takle	0.355 / 0.0026	0.035 / 0.0022	0.018 / 0.0029
Kvamskogen	0.311 / 0.0068	0.043 / 0.0029	0.008 / 0.0034
Kjøbli	0.156 / 0.0117	0.034 / 0.0048	0.014 / 0.0039
Åbjørs.	0.123 / 0.0017	0.052 / 0.0029	0.044 / 0.0054
Nesbyen	0.197 / 0.0024	0.092 / 0.0041	0.068 / 0.0069
Glomfjord	$0.247 / 8\times10^{-4}$	0.022 / 0.0011	0.006 / 0.0017
Tromsø	0.085 / 0.0016	$0.007 / 8\times10^{-4}$	$0.000 / 3\times10^{-4}$
Sihcajavri	$0.036 / 0.0012$	$0.002 / 3\times10^{-4}$	$-0.001 / 1\times10^{-4}$
5 per year			
Oslo	0.468 / 0.103	0.686 / 0.0888	0.681 / 0.0968
Ualand	0.948 / 0.049	0.781 / 0.0487	0.775 / 0.0591
Takle	0.942 / 0.0206	0.546 / 0.0309	0.541 / 0.0373
Kvamskogen	0.669 / 0.0896	0.800 / 0.0745	0.805 / 0.0895
Kjøbli	0.433 / 0.3175	0.704 / 0.1994	0.739 / 0.215

Table 5 (continued).

Location	$\Pr(\in)/\Pr(X > 10, Y > 20)$ (Gauss/Gamma)	$\Pr(\in)/\Pr(X > 10, Y > 20)$ moments	$\Pr(\in)/\Pr(X > 10, Y > 20)$ MLE
Åbjørs.	0.052 / 0.0339	0.607 / 0.0507	0.605 / 0.0479
Nesbyen	0.086 / 0.0509	0.730 / 0.0641	0.724 / 0.0622
Glomfjord	$0.668 / 1\times10^{-4}$	0.681 / 0.0031	0.677 / 0.0069
Tromsø	0.535 / 0.0199	0.357 / 0.0196	0.350 / 0.0225
Sihcajavri	0.536 / 0.0674	0.137 / 0.0273	0.118 / 0.0232

Here $\Pr(\in)$ and $\Pr(X > 10°\text{C}, Y > 20$ mm/day) represent different aspects of the analysis, and $\Pr(X > 10°\text{C}, Y > 20$ mm/day) is provided to give an indication of the robustness of the results. The results are shown for fitting of all data ($N = 1,800$) as well as the five greatest values per year ($N = 100$). Note that the count rates for the five greatest annual values do not correspond with those derived using the entire batch and that take into account the actual correspondence between temperature and precipitation. The probabilities here were derived from the CTL results.

Table 6 95% Confidence intervals derived through Monte Carlo simulations (M-C) and binomial distribution (Bin.) based on Gaussian/Gamma distribution (column 2), Gumbel with moments (3), and Gumbel with maximum likelihood estimation (4)

Location		Gauss./Gamma	Moments	MLE
Daily				
Oslo	M-C all	0.009 – 0.021	0.008 – 0.019	0.002 – 0.010
	Bin. all	0.197 – 0.237	0.044 – 0.066	0.010 – 0.021
Ualand	M-C all	0.029 – 0.048	0.022 – 0.037	0.004 – 0.012
	Bin. all	0.312 – 0.358	0.025 – 0.042	0.001 – 0.007
Takle	M-C all	0.026 – 0.044	0.018 – 0.033	0.004 – 0.012
	Bin. all	0.332 – 0.378	0.026 – 0.045	0.012 – 0.025
Kvamskogen	M-C all	0.038 – 0.058	0.030 – 0.050	0.009 – 0.021
	Bin. all	0.288 – 0.334	0.033 – 0.053	0.004 – 0.012
Kjøbli	M-C all	0.020 – 0.036	0.015 – 0.030	0.006 – 0.015
	Bin. all	0.139 – 0.174	0.026 – 0.044	0.009 – 0.020
Åbjør.	M-C all	0.010 – 0.022	0.011 – 0.023	0.011 – 0.023
	Bin. all	0.107 – 0.138	0.041 – 0.063	0.035 – 0.055
Nesbyen	M-C all	0.016 – 0.031	0.017 – 0.032	0.017 – 0.031
	Bin. all	0.178 – 0.216	0.078 – 0.106	0.056 – 0.081
Glomfjord	M-C all	0.019 – 0.035	0.013 – 0.026	0.004 – 0.013
	Bin. all	0.225 – 0.269	0.015 – 0.030	0.003 – 0.010
Tromsø	M-C all	0.003 – 0.012	0.001 – 0.007	0.000 – 0.003
	Bin. all	0.071 – 0.100	0.003 – 0.012	0.000 – 0.001
Sihcajavri	M-C all	0.000 – 0.003	0.000 – 0.002	0.000 – 0.001
	Bin. all	0.027 – 0.045	0.000 – 0.004	NA

The confidence limits derived from the binomial distribution were based on probabilities $\Pr(\in)$ from Table 5. Derived for CTL.

Figure for the Appendix

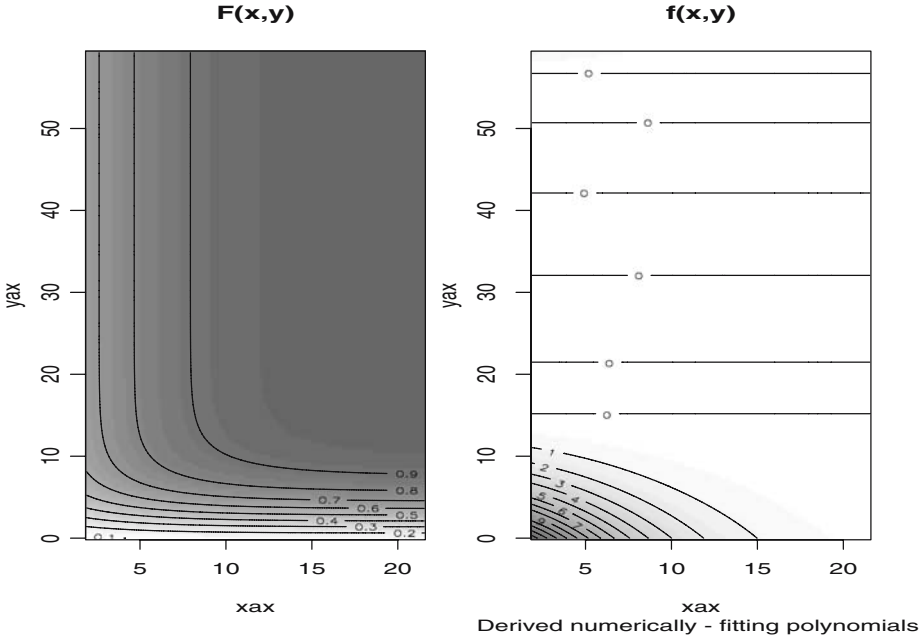

Fig. 10 *F* (x, y) for Oslo-Blindern temperature and precipitation after Yue and Yue and Rasmussen (2002) (right) and the numerically derived bivariate pdf (f(x, y)) (left). Here, *F* (x, y) was derived from the maximal likehood estimators (Table 4)

References

Benestad RE (2001) The cause of warming over Norway in the ECHAM4/OPYC3 *GHG* integration. Int J Climatol 21:371–387

Benestad RE (2003a) How often can we expect a record-event?. Clim Res 25:3–13

Benestad RE (2003b) What can present climate models tell us about climate change?. Clim Change 59:311–332

Benestad RE, Hanssen-Bauer I (2003) Empirical-based refinement of dynamically downscaled temperature scenarios in southern Norway. KLIMA 07/03. The Norwegian Meteorological Institute, P.O. Box 43 Blindern, 0313 Oslo, Norway (www.met.no)

Benestad RE (2004) Record-values, non-stationarity tests and extreme value distributions. Glob Planet Change 44:11–26, doi:10.1016/j.gloplacha.2004.06.002

Benestad RE, Achberger C, Fernandez E (2005) Empirical-statistical downscaling of distribution functions for daily precipitation, Climate 12/2005. The Norwegian Meteorological Institute, P.O. Box 43 Blindern, 0313 Oslo, Norway (www.met.no)

Christensen JH, Christensen OB, Lopez P, van Meijgaard E, Botzet M (1996) The HIRHAM4 regional atmospheric climate model. DMI Sci. Rep. No. 96-4. Danish Meteorological Institute, Lyngbyvei 100, DK-2100 Copenhagen

Coles SG (2001) An introduction to statistical modeling of extreme values. Springer, London, UK

Dankers R, Christensen OB (2005) Climate change impact on snow coverage, evaporation and river discharge in the sub-arctic tana basin, Northern Fennoscandia. Clim Change 69:367–392

Easterling DR, Meehl GA, Parmesan C, Changnon SA, Karl TR, Mearns LO (2000) Climate extremes: observations, modeling, and impacts. Science 289: 2068–2074

Ellner SP (2001) Review of R, Version 1.1.1. Bull Ecol Soc Am 82:127–128
Frei C, Schär C (2001) Detection of trends in rare events: theory and application to heavy precipitation in the Alpine region. J Climate 14:1568–1584
Frich P, Alexander LV, Della-Marta P, Gleason B, Haylock M, Tank AMG, Klein, Peterson T (2002) Observed coherent changes in climatic extremes during the second half of the twentieth century. Clim Res 19:193–212
Gentleman R, Ihaka R (2000) Lexical scope and statistical computing. J Comput Graph Stat 9: 491–508
Gordon C, Cooper C, Senior CA, Banks H, Gregory JM, Johns TC, Mitchell JFB, Wood RA (2000) The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments. Clim Dyn 16:147–168
Hanssen-Bauer I, Førland EJ, Haugen JE, Tveito OE (2003) Temperature and precipitation scenarios for Norway: comparison of results from dynamical and empirical downscaling. Clim Res 25:15–27
Hayhoe K, Cayan D, Field CB, Frumhoff PC, Maurer EP, Miller NL, Moser SC, Schneider SH, Cahill KN, Cleland EE, Dale L, Drapek R, Hanemann RM, Kalkstein LS, Lenihan J, Lunch CK, Neilson RP, Sheridan SC, Verville JH (2004) Emission pathways, climate change, and impacts on California. Proceedings of the National Academy of Sciences, vol 101, pp 12422–12427. National Academy of Science, Washinton, DC
Hegerl GC, Zwiers FW, Stott PA, Kharin VV (2004) Detectability of anthropogenic changes in annual temperature and precipitation extremes. J Climate 17:3683–3699
Houghton JT et al (2001) Climate change 2001: the scientific basis. International Panel on Climate Change. WMO, Geneva, Switzerland (www.ipcc.ch)
Imbert A (2003) The analog method applied to downscaling of climate scenarios. KLIMA 08/03. met.no, P.O. Box 43 Blindern, 0313 Oslo, Norway (www.met.no)
IPCC (2001) Climate change 2001: impact, adaptation and vulnerability. Summary for policymakers. WMO, Geneva, Switzerland
IPCC (2002) Changes in extreme weather and climate events. Beijing, China http://www.ipcc.ch/pub/support/, (11–13 June)
López-Díaz J (2003) A nonparametric test for trends in the occurrence of rare events. J Climate 16:2602–2614
Meehl GA, Tebaldi C (2004) More intense, more frequent, and longer lasting heat waves in the 21st century. Science 305:994–997
Murphy J (1999) An evaluation of statistical and dynamical techniques for downscaling local climate. J Climate 12:2256–2284
Oberhuber JM (1993) Simulation of the Atlantic circulation with a coupled sea ice-mixed layer isopycnal general circulation model. Part 1: model description. J Phys Oceanogr 22:808–829.
Palmer TN, Räisänen J (2002) Quantifying the risk of extreme seasonal precipitation events in a changing climate. Nature 415:512–514
Roeckner E, Arpe K, Bengtsson L, Dümenil L, Esch M, Kirk E, Lunkeit F, Ponater M, Rockel B, Sausen B, Schlese U, Schubert S, Windelband M (1992) Simulation of present-day climate with the ECHAM model: impact of model physics and resolution. Tech. rept. 93. Max Planck-Institute für Meteorologie, Hamburg
Schubert S (1998) Downscaling local extreme temperature change in south-eastern Australia from the CSIRO MARK2 GCM. Int J Climatol 18:1419–1438
Senior CA, Jones RG, Lowe JA, Durman CF, Hudson D (2002) Prediction of extreme precipitation and sea-level rise under climate change. Philos Trans R Soc Lond A 360:1301–1311
Skaugen TE, Hanssen-Bauer I, Førland EJ (2002) Adjustment of dynamically downscaled temperature and precipitation data in Norway. KLIMA 20/02. The Norwegian Meteorological Institute, PO Box 43 Blindern, 0313 Oslo, Norway (www.met.no)
Schmidli J, Frei C (2005) Trends of heavy precipitation and wet and dry spells in Switzerland during the 20th century. Int J Climatol 25:753–771
von Storch H (1999) On the use of "inflation" in statistical downscaling. J Climate 12:3505–3506
Washington WM, Weatherly JW, Meehl GA, Semtner AJ, Bettge TW, Graig AP, Strand WG, Arblaster J, Wayland VB, James R, Zhang Y (2000) Parallel climate model (PCM) control and 1% per year CO_2 simulations with a 2/3 degree ocean model and a 27 km dynamical sea ice model. Clim Dyn 16:755–774
Wilks DS (1995) Statistical methods in the atmospheric sciences. Academic, Orlando, FL
Wood AW, Leung LR, Sridhar V, Lettenmaier DP (2004) Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. Clim Change 62:189–216

Yue S (1999) The Gumbel mixed model for flood frequency ananlysis. J Hydrol 226:88–100
Yue S (2001) Comment on Bivariate extreme value distributions: an application of the Gibbs sampler to the ananlysis of floods. In: Adamson PT, Metcalfe AV, Parmentier B (eds) Water Resources Research, vol 37, pp 1107–1110
Yue S, Rasmussen P (2002) Bivariate frequency analysis: discussion of some useful concepts in hydrological application. Hydrol Process 16:2881–2898