

Downscaling precipitation extremes

Correction of analog models through PDF predictions

Rasmus E. Benestad

Received: 6 January 2009 / Accepted: 1 June 2009 / Published online: 1 July 2009
© The Author(s) 2009. This article is published with open access at Springerlink.com

Abstract A new method for predicting the upper tail of the precipitation distribution, based on empirical–statistical downscaling, is explored. The proposed downscaling method involves a re-calibration of the results from an analog model to ensure that the results have a realistic statistical distribution. A comparison between new results and those from a traditional analog model suggests that the new method predicts higher probabilities for heavy precipitation events in the future, except for the most extreme percentiles for which sampling fluctuations give rise to high uncertainties. The proposed method is applied to the 24-h precipitation from Oslo, Norway, and validated through a comparison between modelled and observed percentiles. It is shown that the method yields a good approximate description of both the statistical distribution of the wet-day precipitation amount and the chronology of precipitation events. An additional analysis is carried out comparing the use of extended empirical orthogonal functions (EOFs) as input, instead of ordinary EOFs. The results were, in general, similar; however, extended EOFs give greater persistence for 1-day lags. Predictions of the probability distribution function for the Oslo precipitation indicate that future precipitation amounts associated with the upper percentiles increase faster than for the lower percentiles. Substantial random statistical fluctuations in the few observations that make up the extreme upper tail implies that modelling of these is extremely difficult, however. An extrapo-

lation scheme is proposed for describing the trends associated with the most extreme percentiles, assuming an upper physical bound where the trend is defined as zero, a gradual variation in the trend magnitude and a function with a simple shape.

1 Introduction

Downscaling It is often necessary to have a description of the local climate and how it may change in the future in order to work out the best adaption strategies, given a changing climate. State-of-the-art global climate models (GCMs) still have a spatial resolution that is too coarse to provide a good description of local processes that are important in terms of natural hazards, agriculture, built environment or local ecological systems. Thus, in order to get useful information on a local scale, it is necessary to *downscale* the GCM results (Benestad et al. 2008). One particular challenge is then to derive reliable statistics for future local precipitation and estimate the impacts of a global climate change.

There are two main approaches for downscaling GCM results, (1) dynamical (also referred to as ‘numerical’ downscaling) and (2) empirical–statistical downscaling (also referred to as just ‘statistical’ or ‘empirical’ downscaling), and there is a large volume of downscaling-related publications in the scientific literature (Christensen et al. 1997, 2007; Timbal et al. 2008; Frei et al. 2003; Chen et al. 2005; Linderson et al. 2004; Fowler et al. 2007; von Storch et al. 2000; Hanssen-Bauer et al. 2005; Salathé 2005; Kettle and Thompson 2004; Penlap et al. 2004; Benestad et al. 2008).

Dynamical downscaling involves running an area-limited high-resolution regional climate model (RCM)

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

with (large-scale) variables from a GCM as boundary conditions. The RCMs tend to be expensive to run and may not provide realistic local conditions on a very small spatial scale (Benestad and Haugen 2007).

In empirical–statistical downscaling (ESD), a wide range of models and approaches have been used, and this approach may be divided into several sub-categories: (a) linear models (e.g. regression or canonical correlation analysis) (Huth 2004; Busuioc et al. 1999, 2006; Bergant and Kajfež-Bogataj 2005; Bergant et al. 2002), (b) non-linear models (Dehn 1999; van den Dool 1995; Schoof and Pryor 2001) or (c) weather generators (Semenov and Barrow 1997). The choice of ESD model type should depend on which climatic variable is downscaled, as different variables have different characteristics that make them more or less suitable in terms of a given model.

RCMs and ESD complement each other, as these represent two independent approaches for deriving information on a local scale, given a large-scale situation. Thus, the different approaches tend to have different strengths and weaknesses (Benestad 2007a). Comparisons between the two approaches also suggest that ESD models can have as high skill as RCMs for the simulation of precipitation (Haylock et al. 2006; Hanssen-Bauer et al. 2005).

RCMs have been used to downscale the precipitation, but these models tend to exhibit systematic biases and only provide rainfall averages over a grid box size area (typically greater than $\sim 10 \times 10 \text{ km}^2$). In mountainous areas and regions with complex topography, the RCMs do not provide a representative description of climate variables with sharp spatial gradients, and it is necessary to re-scale the output in order to obtain a realistic statistical distribution (Engen-Skaugen 2004; Skaugen et al. 2002). A re-scaling, giving the right mean level or standard deviation, assumes that the RCM reproduces the true shape for the statistical distribution, as well as the correct wet-day frequency in order for other percentiles to also be correct, as all the percentiles are scaled by the same factor. However, RCMs may not give a correct representation of the whole statistical distribution (Benestad and Haugen 2007), and such post-processing adjustments may, hence, yield a reasonable mean level and spread, but this does not guarantee a realistic description of the upper tails of the distribution and, hence, extremes.

There are also a number of possible caveats associated with running RCMs: (1) ill-posed solution associated with boundary conditions (lateral boundaries, as well as lack of two-way coupling with the ocean and land surfaces), (2) inconsistencies associated with different ways of representing sub-grid scale

processes (parameterisation schemes), (3) up-scaling effects where improved simulations and better resolution of small-scale processes (e.g. clouds and cyclones) have implications for the large-scale environment or (4) general systematic model errors.

ESD models also have a number of limitations, and these models are based on a number of assumptions: (a) that the statistical relationship between the predictor and predictand is constant, (b) that the predictor carries the climate change ‘signal’, (c) that there is a strong relationship between the predictor and predictand and (d) that the GCMs provide a good description of the predictor. Furthermore, linear models tend to provide predictions with reduced variance (von Storch 1999).

Distributions for daily precipitation Extreme amounts of daily precipitation are tricky to model, due to the fact that the downpour may have a very local character, in addition to the precipitation amount being zero on dry days and having a non-Gaussian statistical distribution on wet days. It is common to assume that the 24-h precipitation follows a gamma distribution¹ which is described by the shape (α) and scale (β) parameters (Wilks 1995).

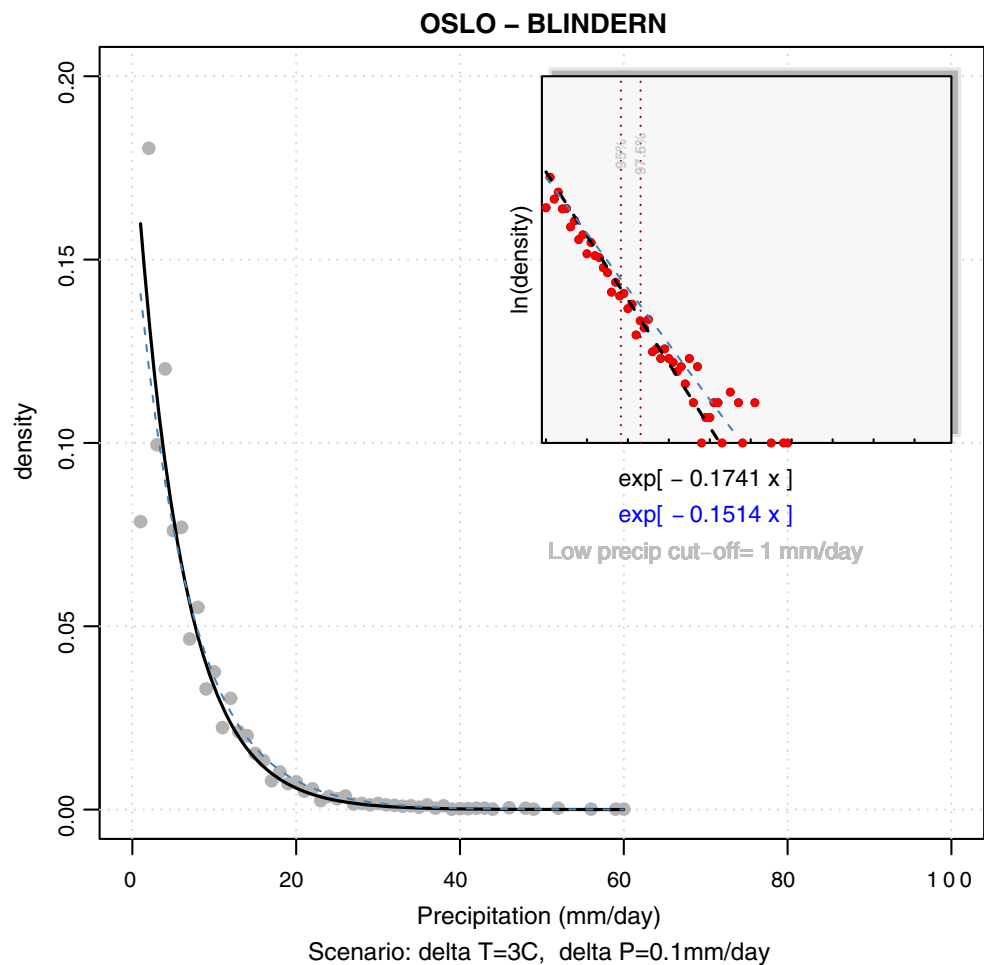
Extremes and the upper tails of the distribution Extreme modelling such as generalised extreme value (GEV) modelling and general Pareto distribution (GPD) are commonly used for modelling extremes because they provide a sophisticated basis for analysing upper tails of the distribution. The gamma distribution, on the other hand, provides a good fit to the entire data sample, but not necessarily to the upper tail. These models² for the statistical distributions, however, involve two or more parameters (scale and shape) to be fitted, requiring large data samples in order to get good estimates.

Furthermore, models of the statistical distribution tend to assume a constant probability density function (PDF), but these are inadequate for when the PDF changes over time (Benestad 2004b). It is, nevertheless possible to use a simple single-parameter model that provides an approximate description of the statistical distribution, which is both easier to fit to the data and for which the parameter exhibits a dependency on the mean local climate conditions. In other words, it is possible to predict changes in the PDF, given a change in the mean conditions.

¹ $f(x) = \left(\frac{x}{\beta}\right)^{\alpha-1} \frac{\exp[-x/\beta]}{\beta\Gamma(\alpha)}$, $x, \alpha, \beta > 0$.

² Here, the term ‘model’ is used loosely: referring to either GCMs, RCMs, ESD or a theoretical statistical distribution such as gamma, exponential, GEV or GPD.

Fig. 1 The PDF for Oslo–Blindern wet-day precipitation-amount and a log-count histogram (*insert*). The *two red dotted vertical lines* in the *insert* mark the 95 and the 97.5 percentiles, respectively. The *dashed blue lines* and *blue text* describe the PDF for a future climate 3°C warmer than present and with 2 mm/day higher mean precipitation. The PDF assumes an exponential distribution with m derived from Eq. 2



Approximate statistical distributions The PDF for the wet-day, 24-h precipitation amount (P_w) can be approximated with an exponential distribution (Benestad 2007b)³ according to:

$$f(x) = me^{-mx}, \quad (1)$$

where $m > 0$ and x represents the wet-day precipitation P_w . The best-fit between a PDF based on Eq. 1 and the empirical data is shown in Fig. 1, indicating an approximate agreement between the data points and the line for all but the most extreme values. The dashed red vertical lines in the insert mark the 95 and the 97.5 percentiles ($q_{0.95}$ and $q_{0.975}$), and it is evident that these are associated with levels where the exponential distribution provides a reasonable description of the frequency. Moreover, Eq. 1 provides a good approximation of the frequency as long as the values are not far out in the upper tail of the distribution.

³The parameter m in this context corresponds to $-m$ in Benestad (2007b).

Benestad (2007b) argued that the geographical variations in the rainfall statistics suggest that the character of the distribution (in this case m) can be predicted, given the mean local temperature, \bar{T} ; all-days (wet and dry) precipitation amounts, \bar{P}_a ; and geographical parameters.

Analog models Linear ESD models may provide a good description of climate variables that have a Gaussian distribution (Benestad et al. 2008) but perform poorly for 24-h precipitation. Analog models, involving re-sampling of historical values depending on the state of large-scale conditions, have been used to provide realistic scenarios for the 24-h precipitation (Dehn 1999; Fernandez and Saenz 2003; van den Dool 1995; Timbal et al. 2003, 2008; Timbal and Jones 2008; Wilby et al. 2004).

Zorita and von Storch (1999) argued that more sophisticated non-linear ESD models do not perform better than the simple analog model. However, such models imply fundamental shortcomings due to the fact that they cannot predict values outside the range of

observed values (Imbert and Benestad 2005). The analog model may distort the upper tail if all the values greater than the historical maximum value are attributed the bin associated with the maximum value of the statistical distribution (PDF) of the past observations. Alternatively, the values exceeding the range of the past may be distributed over several bins or not counted at all. In any case, the analog model will not give a reliable description of the upper tail of the distribution.

Imbert and Benestad (2005) suggested a method for dealing with new values outside the range of the historical sample by shifting the whole PDF according to a trend predicted by linear ESD models. Such a shift in the location of the statistical distribution does not resolve the problem with the distorted upper tail, however.

Another caveat with the analog model may be that they may not preserve the time structure. Here, the time structure refers to the characteristic way precipitation changes over time, reflected in aspects such as persistence, duration of wet and dry spells, and the probability of a rainy day following a dry day. The time structure will henceforth be referred to as the precipitation ‘chronology’.

On the one hand, the chronology of the predicted response is given by the evolution in the large-scale situation (e.g. sea level pressure, henceforth referred to as ‘SLP’), but on the other, there is often not a straightforward, one-to-one relationship between the weather type and the local precipitation, and similar large-scale circulation may be associated with different rainfall amounts. It is possible to use weather types and statistics for each category to carry out Monte-Carlo type simulations, also referred to as ‘weather generators’ (Soltani and Hoogenboom 2003; Semenov and Brooks 1999; Semenov and Barrow 1997), but uncertainties associated with the transition between the types may also affect the timing of precipitation events.

The philosophy of the proposed method Here, a new strategy for downscaling 24-h precipitation is presented, which aims at both providing a representative description of the PDF and yielding a plausible chronology of both wet and dry days with precipitation amounts. Moreover, the general idea of the proposed method is to determine the PDF for wet-day 24-h precipitation, and then generate time series with the same wet-day PDF and a realistic chronology. This is possible by combining different techniques with different advantages. Most of the techniques employed here build on previous work, as the prediction of precipitation PDFs was done by Benestad (2007b), while the analog model

used for producing a realistic chronology is based on Imbert and Benestad (2005). However, here, the time structure is improved by introducing information about 1-day evolution of the large-scale situation.

The outline of this paper is as follows: A methods and data section, followed by the results, a discussion and the conclusions. Here, the focus will be on the 24-h precipitation amounts in Oslo, Norway.

2 Methods and data

2.1 Methods

Implementation All the analyses and data processing described herein are carried out within the R-environment (version 2.8.0) (Ihaka and Gentleman 1996; Gentleman and Ihaka 2000; Ellner 2001), and is based on the contributed packages called `clim.pact`⁴ (version 2.2-26) (Benestad et al. 2008; Benestad 2004a) and `anm`⁵ (version 1.0-5). The R-environment and R-packages are open source and freely available from <http://cran.r-project.org>.

Definitions Here, the term ‘trend’ is used to mean the mean rate of change over the interval for which there is data. An ordinary linear regression (OLR) $\hat{y} = \alpha t + \beta$ is used to estimate the rate of change α , with y being the variable analysed and t the time (here the year). The notation \hat{y} is used to refer to the estimate of y (which, for instance, can be the 95-percentile). More complex trends, such as polynomial trends (Benestad 2003b), are also shown for illustration, but all the trend analyses discussed below will refer to linear trend models and the trend gradient α .

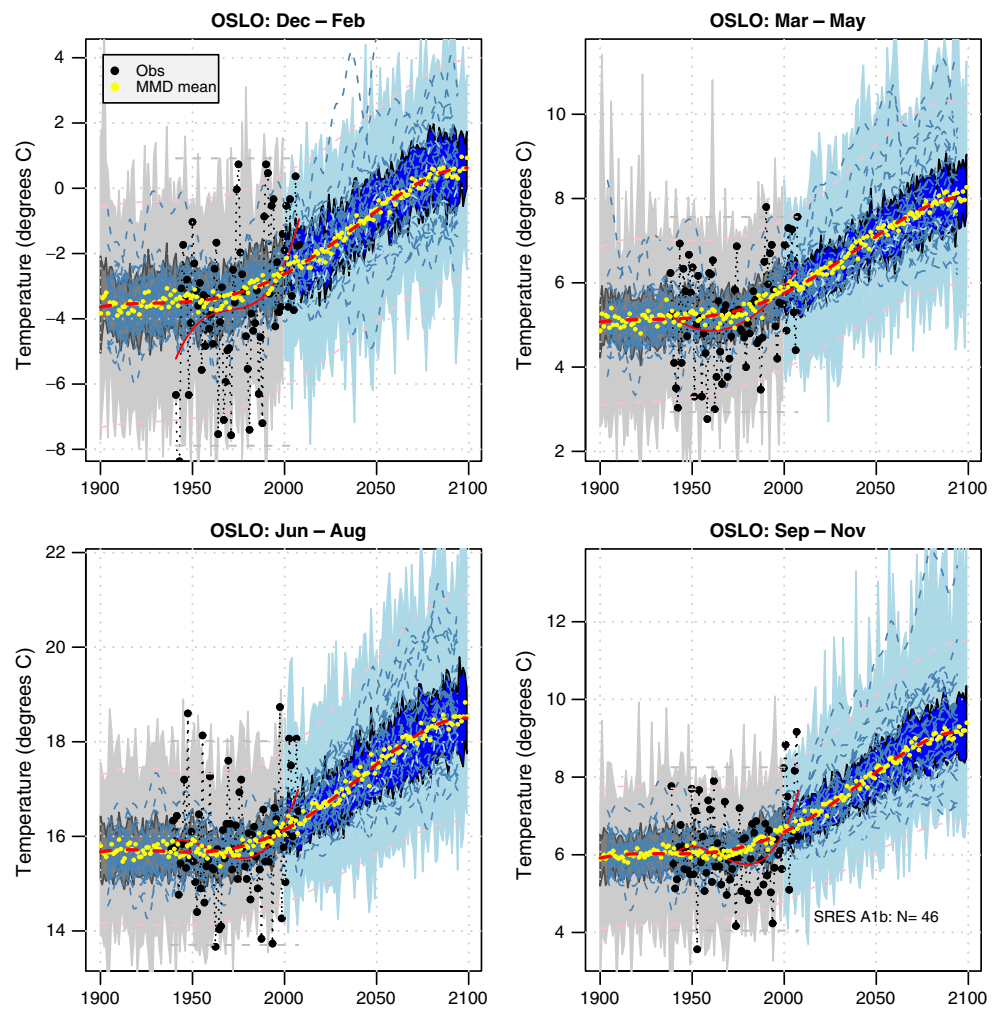
The notation \bar{x} will be used when referring to a temporal mean, be it over the entire interval, 5-year running means, or individual months. For instance, \bar{T} is used here to represent either the mean temperature with respect to the whole data record or monthly mean temperature. Wet-days were taken to be days with rainfall greater than 1 mm, and will be referred to as ‘ P_w ’, as opposed to ‘all-days’ precipitation ‘ P_a ’.

The terms ‘quantile’ and ‘percentile’ are used here, both referring to the value of ranked data that is greater than a p proportion of the data sample, and will, in general, be represented by the symbol q_p . Thus, the 95-percentile ($q_{0.95}$) is the ranked value that is greater than 95% of all of the values in the data sample.

⁴<http://cran.r-project.org/web/packages/clim.pact/index.html>

⁵<http://cran.r-project.org/src/contrib/Descriptions/anm.html>

Fig. 2 Plume plot for Oslo–Blindern (station nr. 18700), showing the time evolution of the observed values (*black*), the twentieth century simulations (*grey*) and the future scenarios (*blue*). The *light shading* shows the minimum–maximum range for the ensemble, and the *darker shading* marks the inter-quantile range (25%–75%). The *yellow symbols* mark the ensemble mean values and the *thick red-dashed line* is the polynomial trend fit to these. The *solid red line* shows a polynomial fit to the observations. The *thin dashed pink lines* show best-fit polynomial to the 5 and 95 percentiles, and the *dashed blue lines* show 10-year low-pass filtered (Gaussian filter) of the individual runs



2.1.1 Analysis stages

The ESD approach The downscaling approach involved a four-stage process:

- Determining monthly mean temperature, \bar{T} , and precipitation totals for *all*⁶ days, \bar{P}_a .
- Using \bar{T} and \bar{P}_a to predict the PDF for the 24-h wet-day precipitation, P_w .
- Generating time series with realistic chronology.
- Re-calibrate the time series from *stage c* to ensure they have a ‘correct’ PDF.

Stages a–c build on older published work, while stage d introduces new concepts to downscaling. The first three stages are described here to provide a complete picture of the analysis, but the analyses in stages a–c were also carried out from scratch in order to use the most up-to-date data and to tailor the analysis to

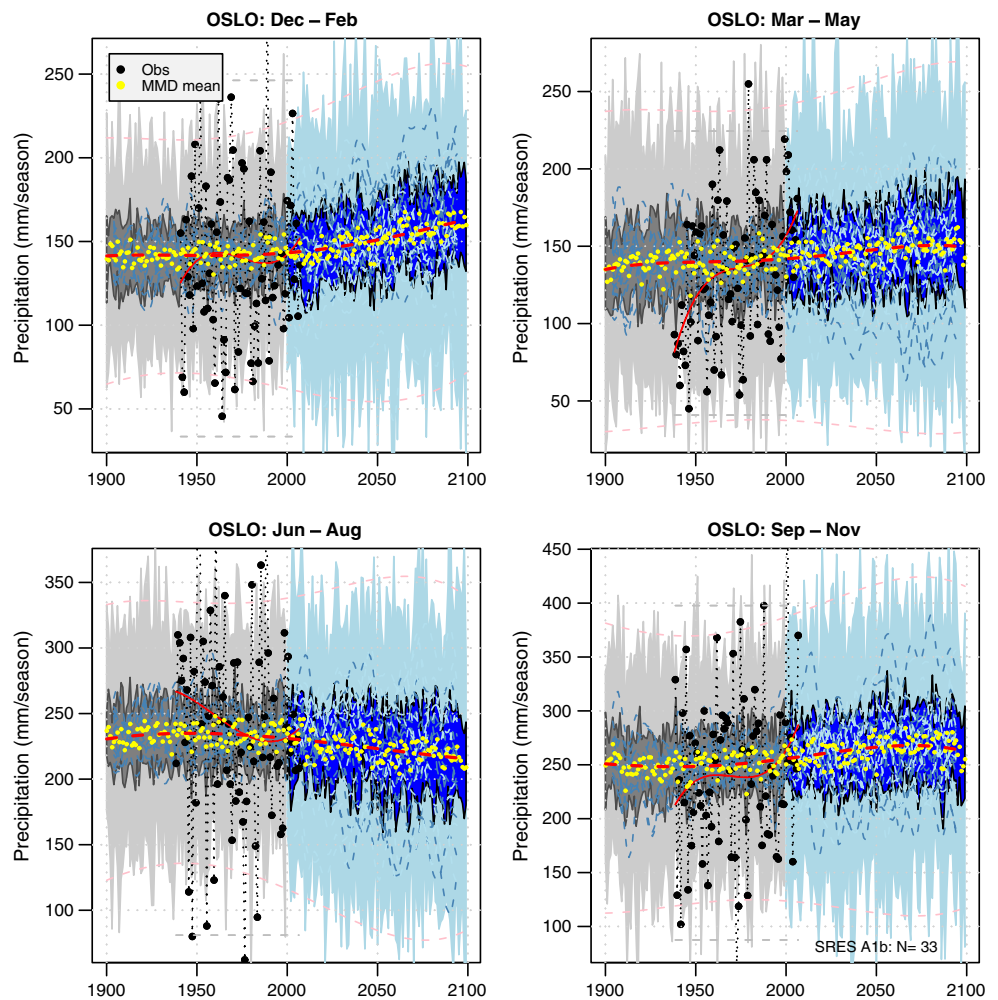
Oslo–Blindern. There were two kinds of ESD analyses in stages a–c: (1) performed on monthly mean predictors to derive monthly mean temperature and monthly precipitation totals (\bar{T} and \bar{P}_a) and (2) performed on daily predictors to derive 24-h precipitation for wet and dry days (P_a). For the former, the ESD involved linear multiple regression, whereas the latter involves a non-linear analog model (Benestad et al. 2008). Both were based on common empirical orthogonal functions (EOFs) for relating observed to simulated predictors (Benestad et al. 2008; Benestad 2001).

Stage 1 The ESD for monthly mean temperature and precipitation totals (shown in Figs. 2 and 3) followed the strategy described in Benestad (2005a), but involved a larger multi-model ensemble, as more runs had become available from the CMIP3 data set (Meehl et al. 2007a). This stage provided \bar{T} and \bar{P}_a for *stage 2*.

Stage 2 The PDF for the wet-day precipitation $f(x) = me^{-mx}$ was determined by predicting the parameter m

⁶Wet and dry.

Fig. 3 Same as Fig. 2, but for seasonal precipitation. Daily precipitation amounts were obtained by dividing the seasonal precipitation totals by 90



according to the multiple regression analysis proposed by Benestad (2007b), using quadratic expressions for both temperature and precipitation:

$$\begin{aligned} \hat{m} = & -(0.247 \pm 0.012) + (2.7 \pm 1.5) \times 10^{-3} [^{\circ}\text{C}]^{-1} \bar{T} \\ & + (0.024 \pm 0.004) [\text{mm}/\text{day}]^{-1} \bar{P}_a \\ & + (3.1 \pm 0.1) \times 10^{-4} [^{\circ}\text{C}]^{-2} \bar{T}^2 \\ & + (1.7 \pm 0.9) \times 10^{-3} [\text{km}]^{-0.5} \sqrt{d} \\ & - (2.3 \pm 1.1) \times 10^{-5} [\text{km}]^{-1} x. \end{aligned} \quad (2)$$

In Eq. 2, the values of \bar{T} and \bar{P}_a were taken to be the 5-year moving average of the station measurements or the ensemble mean of the downscaled monthly \bar{T} and \bar{P}_a from stage 1. The remaining terms were geographical parameters (constant terms for a given location), where d was the distance from the coast (units: km), and x represented eastings (units: km from the 10°E meridian).

Any percentile of an exponential distribution (Eq. 1) can be estimated according to:

$$\hat{q}_p = -\ln|1 - p|/m, \quad (3)$$

m being the best-fit linear slope of the exponential distribution (Benestad 2007b). Figure 1 shows the exponential PDF together with the observations, and the insert shows a histogram for the logarithm of the counts. A good fit between the PDF and the data is characterised by similar linear character in the logarithmic histogram (insert). The PDF $f(x)$ predicted in stage 2 was used in stage 4.

Stage 3 The third stage involved generating time series that give a chronology with dry days, wet days and duration of wet spells that is consistent with the large-scale situation. In this case, the large-scale situation could be represented by the SLP since the objective of stage 3 was to get a good chronology, rather than the exact precipitation amounts. To get a representative

prediction of the precipitation amounts, it is important to include information about the atmospheric water content too. This was already accounted for through the downscaling of future mean climatic conditions and the prediction of the wet-day PDF in *stages 3–2*, and a combination of these results with the chronology would take place in *stage 4* to ensure both a reasonable chronology and a realistic wet-day PDF.

The generation of precipitation chronologies that were consistent with the large-scale situation involved the use of analog models. However, this type of model may suffer from a time inconsistency when there is no one-to-one relationship between the SLP pattern and the local precipitation. It may nevertheless be possible to reduce this problem by including a description of the day-to-day evolution of the large-scale situation. Here, an analog model that incorporated information about the evolution of the large-scale circulation pattern was compared with a more traditional set-up.

The more traditional analog model set-up used ordinary EOFs (Lorenz 1956) as input, and will be referred to as the ‘EOF-model’. The analog model is described in Imbert and Benestad (2005) and implemented using the R-package *anm*.

The new analog model set-up that incorporated information about the 1-day evolution of the large-scale weather situation used extended EOFs (von Storch and Zwiers 1999) as predictors.⁷ Extended EOFs (EEOFs), which put more weight on more persistent features, were estimated from the same daily SLP data, but with a 1-day lag (each EEOF consisted of the SLP of two consecutive days). The analog model based on these will henceforth be referred to as the ‘EEOF-model’.

The generation of daily local P_a with the EEOF- and EOF-models was based on eight leading (extended) common EOFs of the observed gridded daily SLP (at noon) to identify the weather pattern to which the local precipitation could be attributed. The model training was done for the interval 1957–2001, and the predictor domain for the analog model was 25°W–20°E/47°N–67°N. The analog model used SLP (extended) common EOFs weighted by their eigenvalue (Imbert and Benestad 2005).⁸ The results from the EEOF- and EOF-models were used in *stage 4*.

Stage 4 The last stage of the method involved a combination of the PDF from *stage 2* and the chronology from *stage 3*. Since statistical distributions for the analog

model results are expected to be biased in terms of their upper tails, a transformation was applied to the data to give the predicted statistical distribution $f(x)$ based on Eqs. 1 and 2. This transform will henceforth be referred to as *re-calibration*, involving a local quantile–quantile mapping for the wet-day precipitation (P_w) only.

The basis for the re-calibration is as follows: If the biased statistical distribution returned by the analog model is $g(x)$ (here, x represents wet-day precipitation P_w), then the probabilities for days with precipitation amounts less than x are given by the cumulative distributions, $Pr(X < x) = G(x)$ for the analog model results and $Pr(X' < x') = F(x')$ for the predicted PDF. The cumulative probability functions are defined as follows: $G(x) = \int_{-\infty}^x g(X)dX$ and $F(x') = \int_{-\infty}^{x'} f(X')dX'$ (Wilks 1995). Then, x can be transformed to x' , assuming that the quantile of one distribution corresponds to the quantile of the other: $F(x') = G(x) \rightarrow x' = F^{-1}[G(x)]$.

Figure 4 illustrates the local quantile–quantile transform. Here, the results for the quantile–quantile plots between the results from the analog model in *stage 3* and the exponential PDFs predicted in *stage 2* are shown as symbols, with the x-axis representing the quantiles associated with $f(x)$ and the y-axis the quantiles for the analog model results $g(x)$. The transform then involved reading off the x-value of the points corresponding to the given y coordinate.

2.1.2 Trend analysis for percentiles

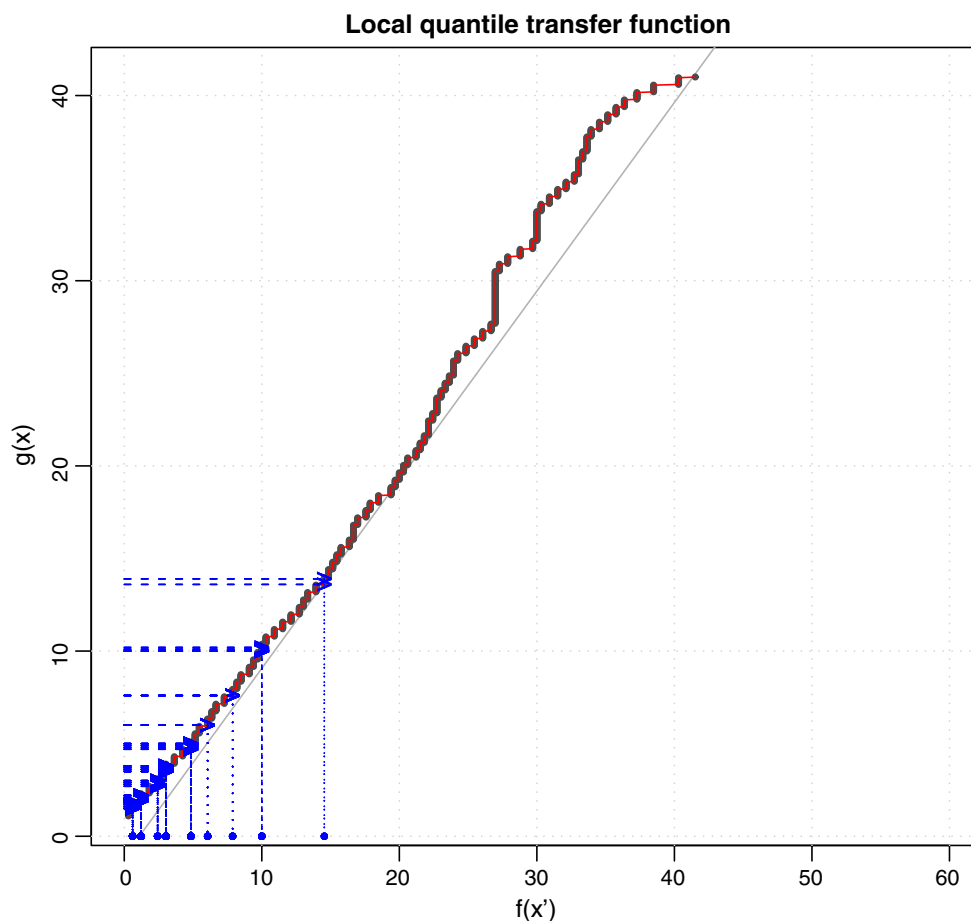
Extrapolation of high-percentile trends Benestad (2007b) argued that the estimates for the percentiles from Eq. 3 are only valid for ‘moderate extremes’, i.e. invalid for the far part of the upper tail (e.g. q_p , where $p > 0.975$). This can also be seen in Fig. 1, where the fit between the data and the PDF is good for most of the data, except for the very highest and most infrequent values. The high-tail clutter seen in Fig. 1 furthermore suggests that trends for the upper extreme percentiles (q_p where $p > 0.99$) may be meaningless, due to small samples with a strong presence of random statistical fluctuations.

Given an upper bound of physically possible precipitation P_* beyond which P_w cannot exceed, the trend α beyond this level can be assumed to be zero because the associated probabilities will always be zero above this limit, thus constraining the trend α to zero for very high P_w . On the other hand, one may also argue that the trends are undefined for values beyond P_* , and the implications may be that the trend α does not converge to zero near the upper limit.

⁷In both of these cases, common EOFs were used as a basis for the analysis, i.e. ‘ordinary’ common EOFs and extended common EOFs.

⁸But no ‘adjustment’.

Fig. 4 An illustration of the quantile–quantile mapping of stage 2). The y -axis marks the original values, and the transformed data are derived by taking the x -coordinate of the curve with a corresponding y -value (here, a few example values are marked in blue)



Assuming that the trend α does converge to zero, a plausible scenario (educated guess) can be made for α in the most extreme amounts based on a number of objective assumptions: (a) there is no trend in precipitation amounts exceeding the present-time potential maximum precipitation (PMP), which, for Oslo-Blindern, is estimated by to be 214 mm/day (Alfnes and Førland 2006);⁹ (b) the trend varies smoothly with the precipitation level q_p ; and (c) the function describing the trend has the simplest possible shape. The assumption of α being a smoothly varying function of q_p is reasonable as long as the climate change in question involves a gradually changing PDF. The last assumption is inspired by the principle of Occam's razor.¹⁰

It is likely that the PMP will change in the future as a consequence of a global warming, since higher temperature will favour an increase in the water holding capacity of the air. Hence, the sensitivity to the upper

limit was explored by varying P_* between the maximum observed precipitation (59.8 mm/day) and $2 \times$ PMP, thus spanning a range of values that exceeds any realistic confidence interval. This kind of extrapolation is similar to Benestad (2005b), but here, a cubic spline (Press et al. 1989) was used to interpolate between the zero point at the upper bound, and the trend estimates $\hat{\alpha}$ were derived for the percentiles $q_{0.70}$ – $q_{0.97}$.

Additional Monte-Carlo simulations A set of additional Monte-Carlo simulations were conducted to explore possible causes for the clutter in the extreme part of the upper tail of the distribution. In this case, the Monte-Carlo simulations involved using a random number generator to produce stochastic numbers following a gamma distribution with similar characteristics to the actual observations, and hence involved fitting the gamma distribution to the data. There are different ways of obtaining best-fit parameters for the gamma distribution, involving 'moment estimators' or 'maximum likelihood estimators'. Wilks (1995) recommends using the so-called 'maximum likelihood estimators', but in this case, it was found that the two estimators

⁹Referred to in the report as the 'British M5 method'. The Hershfield method gives 193 mm, but here, the largest value is used.

¹⁰http://en.wikipedia.org/wiki/Occam's_razor

provided a similar description of the histogram for P_w , and hence, the simpler moments estimators were used:

$$\hat{\alpha} = \frac{(\overline{x_R})^2}{s^2}, \quad (4)$$

$$\hat{\beta} = \frac{s^2}{\overline{x_R}}, \quad (5)$$

where the variable $\overline{x_R}$ represents the mean value for wet days only, and s is the standard deviation for the wet-day precipitation P_w . The moment estimators in Eqs. 4–5 were henceforth used for fitting the gamma distributions representative for P_w at Oslo–Blindern.

2.2 Data

Historical data The station data were, in this case, taken from the Norwegian Meteorological Institute’s climate data archive (‘KlimaDataVareHuset’¹¹). The large-scale predictors used to calibrate the ESD models, however, were taken from the European Centre for Medium-range Weather Forecasts (ECMWF) ERA40 re-analysis data (Bengtsson et al. 2004; Simmons et al. 2004; Simmons and Gibson 2000). Both monthly mean 2-m temperature and monthly totals of precipitation were used as predictors, as discussed in Benestad (2005a). The 24-h SLP data were also taken from the ERA40.

Model data The predictors used for making local climate scenarios involved the CMIP3 data set (Meehl et al. 2007a, b) from the Program for Climate Model Diagnosis and Intercomparison (PCMDI) archives.¹² The ESD analysis for monthly mean temperature and precipitation therefore involved a large set of different GCMs (22 GCMs/50 runs for temperature and 21 GCMs/43 runs for precipitation; further details provided in Benestad 2008a), but excluded some of the GCMs performing poorly. The weeding of poorly performing GCMs resulted in an ensemble of 46 members for temperature and 33 for precipitation. Further details of this ESD analysis are given in Benestad (2008a).

Here, the GCM simulations followed the SRES A1b emission scenario (Solomon et al. 2007) in addition to the simulations for the twentieth century (‘20C3M’). The predictors included the monthly mean 2-m temperature and monthly mean precipitation, as in Benestad (2005a). The GCM precipitation was scaled to match the physical units used in the ERA40.

The ESD for the 24-h (daily) precipitation only involved the ECHAM5 GCM (Keenlyside and Latif 2002; Giorgetta et al. 2002; Meehl et al. 2007b) sea-level pressure (SLP). The choice of ECHAM5 was, to some extent, arbitrary; however, it has been shown to describe realistic features in the SLP field such as cyclones (Benestad 2005b). Daily precipitation amounts were also retrieved from ECHAM5 and were interpolated to 10.7207°E/59.9427°N (the coordinates for Oslo–Blindern) using a bi-linear interpolation scheme from the R-package *akima* (version 0.4-4).

3 Results

3.1 Stage 1: downscaled monthly data

Figures 2 and 3 show a comparison between seasonal ESD results from *stage 1* (based on the monthly values) and corresponding observations. The observed temperature (black symbols; Fig. 2) is within the envelope of ESD-results, suggesting that the ESD analysis for monthly mean temperature is consistent with the true values in terms of the mean level, variability and the time evolution.

A comparison between the ESD-results for the seasonal precipitation totals, P_a , and observations suggest that the downscaled results mainly fall within the corresponding ESD-envelope. However, the variability is not as well captured for the precipitation as for the temperature. Nevertheless, Figs. 2 and 3 show that the ESD is able to give a realistic reproduction of these local climate variables.

It is important to note that the ESD results for the past are independent of the actual observations, as these were derived with GCM simulations for the past rather than using the calibration data (ERA40). Thus, the comparison between the observations and the ESD results over the twentieth century constitutes an independent test of skill.

3.2 Stage 2: downscaled PDFs

The dashed blue line in Fig. 1 illustrates how the PDF $f(x)$ may change in the future. In this case, the PDF was predicted from *stage 2* by taking the ensemble mean temperature and precipitation difference between 1961–1990 and 2081–2100 ($\Delta T = 3\text{ K}$ and $\Delta P_a = 0.1\text{ mm/day}$) of the downscaled CMIP3 data. It is important to assess the method’s ability to predict the changes in the PDF, which can be done by looking at changes in the past.

¹¹<http://eKlima.met.no>

¹²<http://esg.llnl.gov:8443/index.jsp>

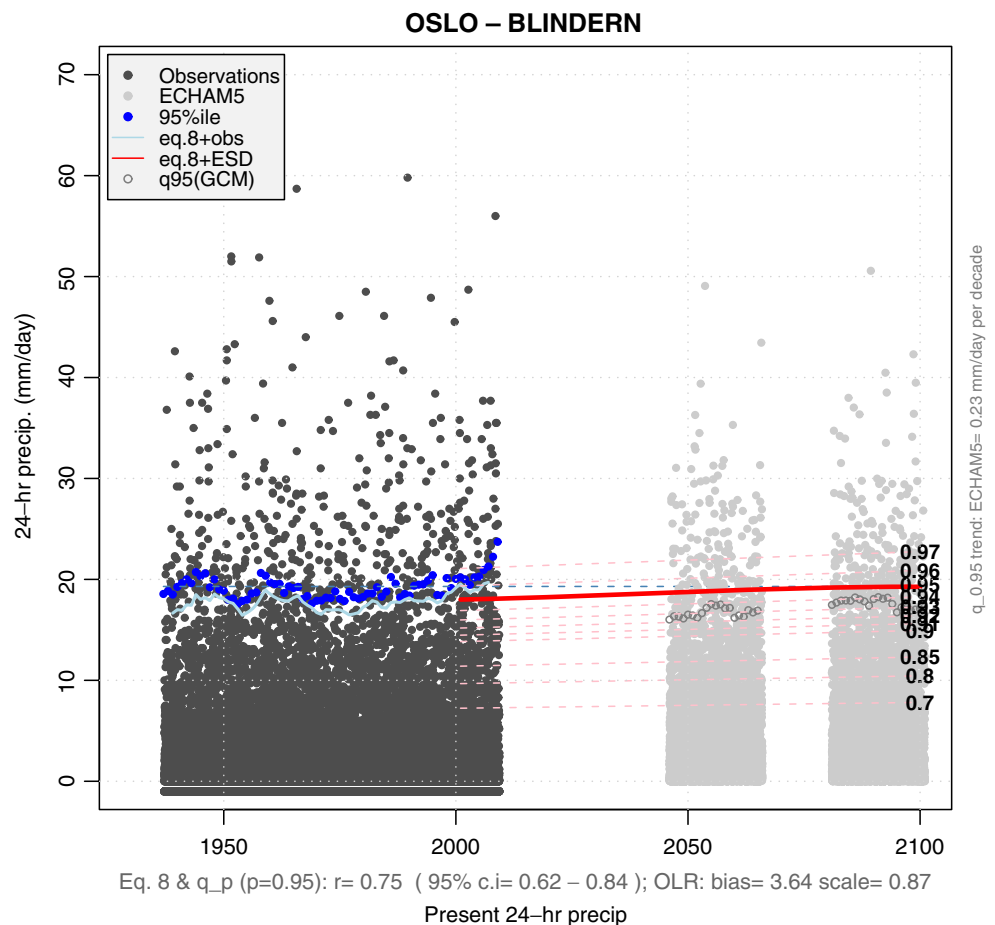
Percentiles of the past Figure 5 shows the 24-h precipitation measured at Oslo–Blindern (dark grey), and corresponding interpolated daily precipitation from ECHAM5 (light grey). The evolution of the wet-day $q_{0.95}$, estimated directly from the data, is shown as dark blue symbols. Estimates for $\hat{q}_{0.95}$, derived using Eqs. 2 and 3 and taking \bar{T} and \bar{P}_a directly from the station measurements as input, are shown as a light blue line. The historical values of $q_{0.95}$ and $\hat{q}_{0.95}$ were estimated by taking the 95-percentile of wet days only (P_w) over a 5-year sliding window.

A good correspondence between the dark blue symbols and the light blue line in Fig. 5 suggests that Eqs. 2 and 3 provided skillful predictions of wet-day $q_{0.95}$. Thus, the comparison between $q_{0.95}$ and $\hat{q}_{0.95}$ constitutes an independent test of the simple single-parameter distribution model described in Eq. 1. The correlation between $q_{0.95}$ and $\hat{q}_{0.95}$ was 0.79, with a p value of 8×10^{-16} , but the level of statistical significance was, in reality, lower due to the presence of auto-correlation. Furthermore, the derived values $\hat{q}_{0.95}$ had a low bias of ~ 2.3 mm/day.

The dashed light blue line shows the $q_{0.95}$ for all wet days P_w from the station measurements. The 95-percentile for the interpolated data from the ECHAM5 GCM is also shown (dark grey open circles), and the level of GCM-based $q_{0.95}$ was lower than both the estimates based on the observations ($q_{0.95}$) and $\hat{q}_{0.95}$ derived with Eqs. 2 and 3.

Historical trends in percentiles The value for $q_{0.95}$ estimated directly from the past precipitation measurements (dark blue) exhibits some variations over time, with a recent upturn since the 1980s, but the best-fit linear trend for the entire period is also positive. The positive trend for Oslo is representative for the rest of the country: Out of a total of 62 sites in Norway with more than 50 years of daily data, 20 cases had an estimated linear trend in $q_{0.95}$ that was negative (not tested for statistical significance at the station level) and 42 positive. The significance of the nationwide results can be tested by taking a null-hypothesis of 50% chance for either sign and using

Fig. 5 Daily precipitation at Oslo–Blindern (dark grey) plotted with $q_{0.95}$ (blue) and percentiles estimated according Eq. 3 (light blue) based on 5-year sliding windows. Also shown is the $q_{0.95}$ estimated for the entire set of observations P_w (dashed blue horizontal line) and trends of \hat{q}_p based on ESD-results (red solid and pink dashed lines). The grey symbols mark interpolated P_w from ECHAM5, and the percentiles for ECHAM5 $q_{0.95}$ are highlighted with different symbols/shading



binomial distribution for the null-hypothesis.¹³ Thus, the probability of getting 20 cases or less with one sign is 0.04% (i.e. statistically significant at the 1% level). The data may not be homogeneous, however, and some series contained jumps. If such errors were to affect the analysis either way on equal terms, then the binomial distribution should be unaffected, as p should still be 0.5. There is one caveat, though, that all sites have undergone changes in the instrumentation that, over time, has improved the capture of extreme precipitation amounts.

Future trends in percentiles If part of the trend in $q_{0.95}$ is due to an ongoing global warming caused by an enhanced greenhouse effect (Hegerl et al. 2007), then the future trends in $q_{0.95}$ should bear some relation with those of the past, albeit with different magnitudes.

To make a projection for the future, the downscaled monthly temperature and precipitation from *stage 1* were used as input in Eq. 2 from *stage 2* in order to estimate m and, hence, used in Eq. 3 to predict the wet-day percentiles for the 24-h wet-day precipitation amount. Moreover, the empirical estimates of \bar{T} and \bar{P}_a were replaced by ensemble mean values of the downscaled CMIP3 GCM results (shown in Figs. 2 and 3) to make projections for the local future climate.

Linear trend fits of the projected change in the percentiles $\hat{q}_{0.70}$ – $\hat{q}_{0.97}$ are shown as pink linear curves in Fig. 5, and the red curve marks the projected $q_{0.95}$. The projected values for $q_{0.95}$ (red) indicate levels below that estimated directly from the observations $q_{0.95}$ (dark blue symbols) in the beginning of the twenty first century, but slightly higher than the all-period level (blue dashed) towards the end of the century. Note that the observations are completely independent of the predictions based on the ensemble mean of the downscaled CMIP3 results. Thus, the comparable levels seen in the predictions and the observations suggest that the predictions of *stage 2* are also reasonable when downscaled results are used as input for the analysis.

The scenarios for the future suggest a further increase due to warmer and wetter conditions (Fig. 5), but the higher percentiles are projected to increase faster than the lower percentiles. The exercise was repeated by replacing \bar{T} or \bar{P}_a in Eq. 2 with the present mean values, respectively (not shown). The results based on variable \bar{P}_a and constant \bar{T} indicated that most (~90%) of the increase in $\hat{q}_{0.95}$ could be associated with higher temperatures, while the projected precip-

itation increase by itself only accounts for ~10% of the increase.

3.3 Stage 3: deriving precipitation chronology

Analog modelling While the changes in 5-year-running \bar{T} and \bar{P}_a can provide a basis for predicting the PDF for wet-day precipitation P_w , the monthly mean ESD does not provide any description of how the 24-h precipitation amounts may vary from day to day. Furthermore, the PDF only describes the probabilities for wet-day precipitation P_w , and many applications require realistic time series with wet and dry days. The analog model is, in principle, capable of reproducing wet and dry sequences and amounts, albeit biased.

Figures 6 and 7 show the 24-h precipitation for Oslo–Blindern derived from the EOF- and EEOF analog models, respectively. The observations are shown in grey, and the raw results from the analog models of *stage 3* in red. It is difficult to distinguish the results from the EOF and EEOF models merely from these time series plots, which suggests that the two model-strategies in general produce similar results.

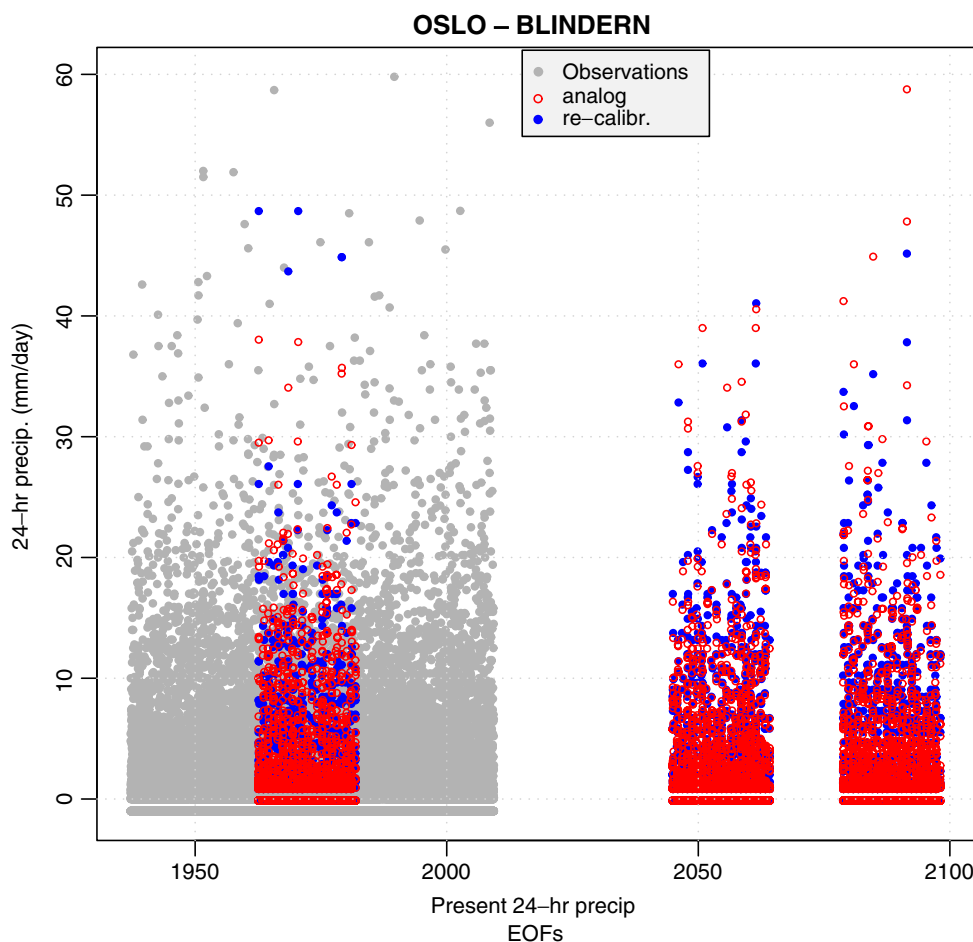
Time structure Figure 8 compares the auto-correlation function of the observations (black columns) with EOF- and EEOF-model results (filled circles). The analysis was done for the common interval: 1961–1980. Since the data did not have a Gaussian distribution, the confidence limits shown in the figure are not unbiased. The lagged correlation for the results derived with the EEOF model was, nevertheless, slightly higher than for the EOF model for a 1-day lag, but both showed weaker 1-day persistence than the observations. For lags of 2–10 days, the two models produced similar auto-correlations to the observations, and stronger persistence for lags greater than 10 days. The PDFs from *stage 2* were used in conjunction with these analog model results to derive a daily time series of the precipitation in *stage 4*.

3.4 Stage 4: re-calibration

The re-calibrated results from *stage 4* are shown in blue in Fig. 8, and it is difficult to tell from this figure whether the re-calibrated results had more days with heavy precipitation. However, a quantile–quantile plot (Fig. 9) can reveal systematic differences not easily seen in the time series plots. Figure 9 compares the results from the analog model and re-calibration with the observations for the common interval 1961–1980. The figure also shows a comparison between the raw analog and re-calibrated results for the future (insert).

¹³ $Pr(K = k) = \frac{N!}{k!(N-k)!} p^k (1-p)^{N-k}$, where $p = 0.5$, and $N = 62$

Fig. 6 Observed precipitation at Oslo–Blindern (grey) and downscaled using the EOF-model (red) of stage 3. The blue symbols show the re-calibrated results from stage 4



The quantile–quantile plot suggests that only the upper percentiles (x-axis: $P_W > 20$ mm/day) of re-calibrated results are shifted for the 1961–1980 period, and that both raw analog results and the re-calibration show a close match with the observed distribution at lower percentiles (x-axis: $P_W < 20$ mm/day).

The analog model overestimated the higher quantiles of the precipitation amounts for the past, but the re-calibration produced values in the extreme upper tail that were both higher and lower than the observed values, depending on the type of analog model. The re-calibration of EOF-model results produced values mainly greater than those observed for the most extreme part of the tail (red diamonds for predicted $P_W > 40$ mm/day), whereas, the re-calibration of the EEOF-model results adjusted the extreme upper tail to lower values (blue triangles for predicted $P_W > 30$ mm/day).

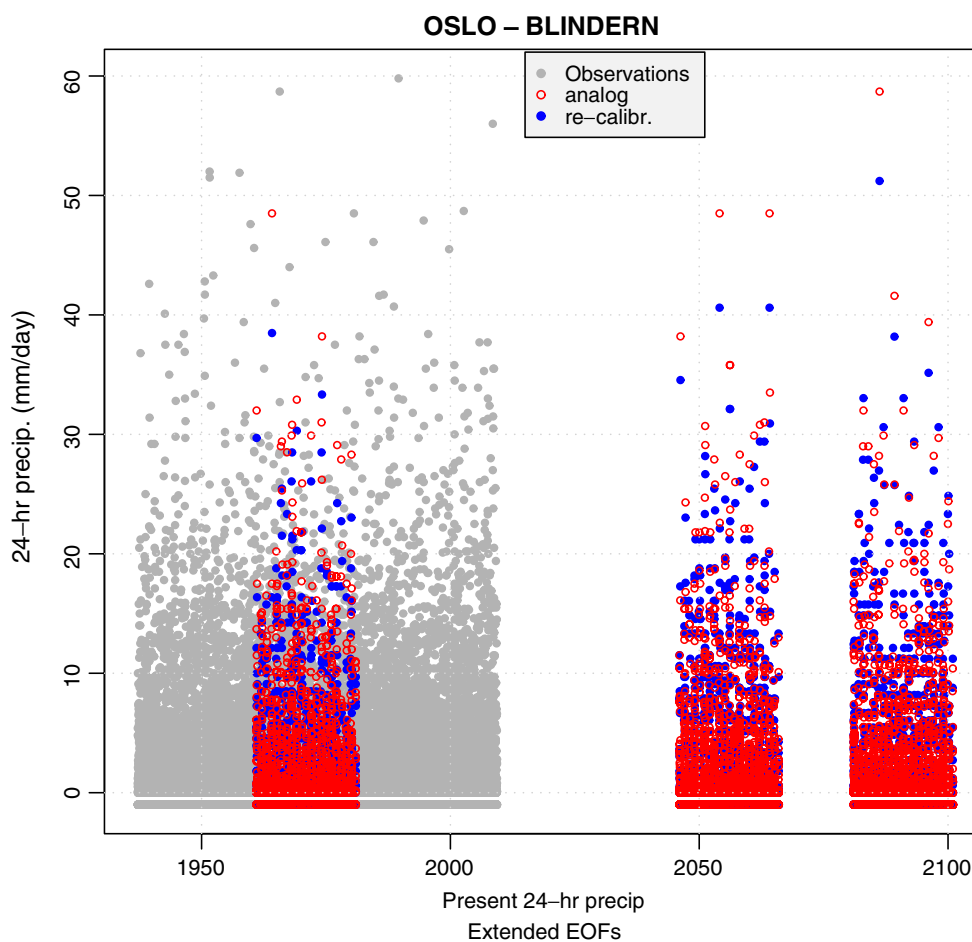
Although the raw analog model results and the re-calibrated results for the common 1961–1980 period were off the diagonal for $P_W > 20$ mm/day, it is important to keep in mind that the values of the highest percentiles were uncertain due to sampling fluctuations

and discontinuities in $g(x)$ (Fig. 4). The results presented in Fig. 9 were consistent with the argument that the upper tail is distorted.

The situation was less ambiguous for the future (2081–2100), where the raw analog model results for both the EOF and EEOF models suggested greater values in the upper tail of the distribution (insert), as the most extreme upper quantiles of the re-calibrated distribution $f(x')$ tended to have lower values than corresponding quantiles in the raw analog model results $g(x)$. For more moderate values of P_w , the re-calibration suggested a slight increase with respect to the analog model results.

Figure 10 compares the empirical distribution functions (EDFs) (Folland and Anderson 2002; Jenkinson 1977) for the analog model results (red and blue), the re-calibrated data (pink and steel blue), the observations (yellow), precipitation interpolated from ECHAM5 (dark green dashed), and the predicted cumulative distribution function (CDF; $F(x) = 1 - e^{-mx}$ grey shaded area). All the data represent the 2081–2100 interval, except for the observations. The precipitation

Fig. 7 Same as Fig. 6, but using the EEOF model



interpolated directly from ECHAM5 had statistical distributions that were closer to the present-day climate than the predicted exponential distribution, whereas the raw analog results from *stage 3* produced probabilities lying between the present-day distribution and $F(x)$ predicted for the future. The re-calibrated ESD results, on the other hand, were both closer to the predicted $F(x)$.

The raw analog model results for 2081–2100 (red and blue solid) suggested that the probability for exceeding the present wet-day 95-percentile ($q_{0.95}$ over 1961–1980 = 18.6 mm/day) was similar to the present ($\text{Pr} \sim 0.05$). The predicted CDF, on the other hand, suggests that the probability for $P_w \geq 18.6$ mm/day was 0.06 and that the projected wet-day 95-percentile for the period 2081–2100 was $q_{0.95} = 19.8$ mm/day. The re-calibrated results had not taken into account the low bias seen in the beginning of the twenty first century (Fig. 5), which was likely to affect the results for the 2081–2100 period. The analysis suggested that the statistical distributions were, in general, similar for the EOF and EEOF models.

Figure 11a provides a comparison between the statistical distributions of observations, interpolated daily precipitation from ECHAM5, the results from the analog models and the calibrated results. The linear line marks best-fit $f(x) = me^{-mx}$ and shows that the simple exponential PDF yielded a good approximation for percentiles lower than $q_{0.97}$, but that the precipitation had a fatter upper tail for $P_w > q_{0.97}$ than predicted by the exponential distribution.

3.4.1 Temporal consistency for the re-calibrated results

In order to assess whether the re-calibration affected the chronology, the auto-correlation analysis was repeated and the results were compared with the observations and the raw analog model results in Fig. 8 (open circles with faint shadings). The auto-correlations for the EEOF-model results were relatively insensitive to the re-calibration, while the re-calibration of the EOF-model results had a tendency of lowering the 1-day persistence.

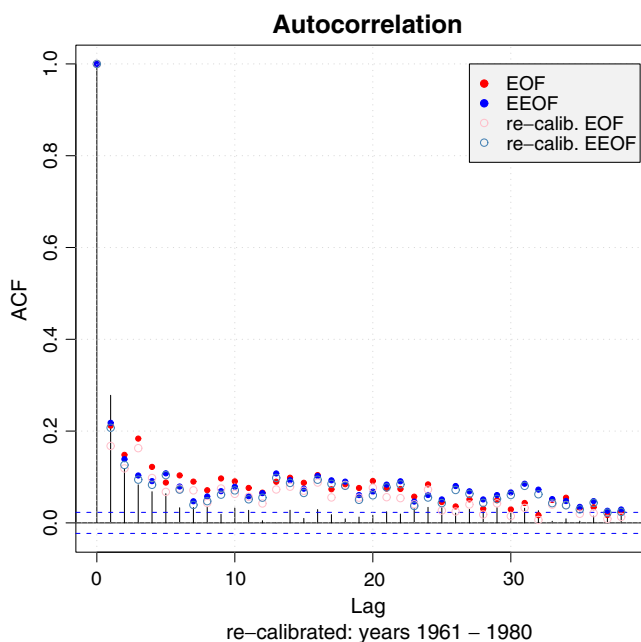


Fig. 8 Autocorrelation functions for the observations (*black lines*), the raw analog model results (*filled circles*) and for the re-calibrated results (*open circles*). The results derived from the EOF model are shown in *red shades* while the EEOF-model results are in *blue*

3.5 The extreme upper tail

Upper tail clutter Figures 1 and 11a both exhibit a clutter of points at the very high end of the statistical distribution, for which no single formulae can provide a good description. This clutter may be due to simple sampling fluctuations or be caused by the presence of different physical processes, such as random position/catch, different large-scale conditions (cyclone-related, frontal or convective precipitation) or different micro-physical processes (cold or warm cloud environment, warm or cold initiation or the effect of different entrainment processes (Rogers and Yau 1989; Blyth et al. 1997)).

In order to see if the high-tail clutter could simply be due to plain sampling fluctuations, a set of Monte-Carlo simulations was carried out involving the generation of synthetic time series with a prescribed gamma distribution using best-fit scale and shape parameters (according to Eqs. 4–5). The results from the Monte-Carlo simulations are shown in Fig. 11b, and these results exhibited a similar high-tail scatter to the real precipitation. Thus, one explanation of the high-end clutter is pure randomness, although this does not rule out the possibility of various physical processes having different effects on the precipitation statistics.

The stochastic explanation for the extreme upper tail clutter implies that trends cannot be determined

for the most extreme events for which there are only a few observations. The data sample of such high amounts becomes too small for proper trend analysis, and increasing statistical fluctuations make the trend estimates difficult to define. Furthermore, since Monte-Carlo simulations with a constant PDF also produced similar clutter, it is interesting to explore the effect a changing PDF will have for the theoretical percentiles. A scheme for extrapolating the trends α from lower and ‘well-behaved’ theoretical percentiles may be possible, if the data behave according to the three assumptions stated in Section 2.1.2: that there is an upper limit P_* where α converges to zero, that α is a function that varies smoothly with the percentiles q_p and that the shape of the function is simple.

The most extreme percentiles Figure 12 shows how the trend estimates $\hat{\alpha}$ vary with the percentile (black symbols) and how these relate to the trend in the mean precipitation. Also shown is the range of the values measured to date (grey shaded region) and an extrapolation of the trends, based on the assumptions explained in Section 2.1.2.

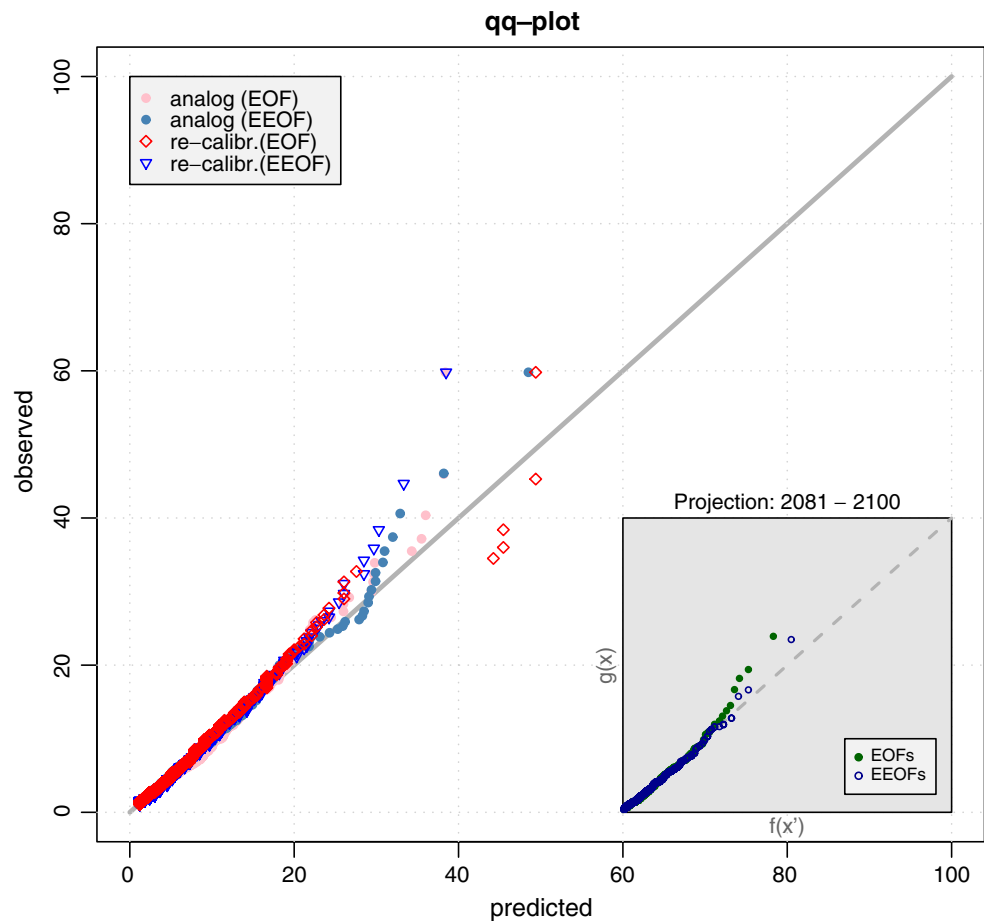
The extrapolated trends $\hat{\alpha}$ were sensitive to the level chosen to be the upper bound, but a crude confidence analysis was carried out by repeating the exercise with P_* set to the maximum observed value and $2 \times$ the PMP, respectively (hatched region in Fig. 12). The extrapolation gave trends in the high percentiles that, with the unlikely exception of upper bound set to near the present maximum value, exceeded the trend in \bar{P}_a (0.01 mm/day per decade) and the trend in ECHAM5 $q_{0.95}$ (0.23 mm/day per decade).

3.6 Discussion

The four-stage ESD Here, a four-stage non-linear method is suggested for downscaling the precipitation, providing a realistic description of the statistical distribution, as well as the chronology of the precipitation events. The advantage of this approach over neural nets (Schoof and Pryor 2001; Haylock et al. 2006) is that the latter is more of a ‘black box’, while the present approach provides transparency for the actions in every step. Such re-calibrated time series can then be used as inputs in, e.g. hydrological or other climate-impact models where both the statistical distribution and the chronology matter. It is also possible to apply such a re-calibration directly to GCM or RCM output, although both RCM and GCMs may over-estimate the number of wet days (Benestad et al. 2008, p. 37).

In general, the re-calibration gave greater heavy precipitation amounts for the future than the raw results

Fig. 9 Quantile–quantile plots of the raw analog model results (*filled circles*) and the corresponding re-calibrated results (*open symbols*). The re-calibration of the past data was based on the PDF shown in Fig. 1 (*black curve*), whereas the re-calibration for the future (*insert*) was based on the PDF for 2081–2100 (*blue curve* in Fig. 1). The *x-axis* represents the predicted \hat{q}_p from stage 4 based on Eqs. 1 and 2, whereas the *y-axis* represents the quantiles directly inferred from the empirical data q_p . The *large plot* shows a comparison between analog/re-calibrated results for the past (1961–1980) and corresponding observations, whereas the *insert* shows results for the future, comparing raw analog results with re-calibrated results



from the analog models and the precipitation interpolated directly from the GCM. However, the quantiles \hat{q}_p derived through ESD exhibit a low bias in the beginning of the twenty first century, and lead to an underestimation of the extreme upper percentiles (Fig. 9), and the re-calibration produced lower values than the analog models for most extreme percentiles.

The fact that the analog model underestimated the 24-h precipitation amounts suggests that the increase in the precipitation was not primarily due to an increase in the frequency of weather types associated with more precipitation. It is important to also account for changes in the atmospheric moisture and temperature, which was done indirectly in the proposed approach through the prediction of the PDF based on both \bar{T} and \bar{P}_a . Most of the increase in the upper percentiles could be ascribed to the increase in the local temperature.

Exponential distribution can be used as an approximate description of the statistical distribution for the wet-day precipitation P_w . It was shown here in an independent test that there is a good correlation between $\hat{q}_{0.95}$ estimated directly from the data and predicted by Eqs. 2 and 3, but Benestad (2007b) has also provided

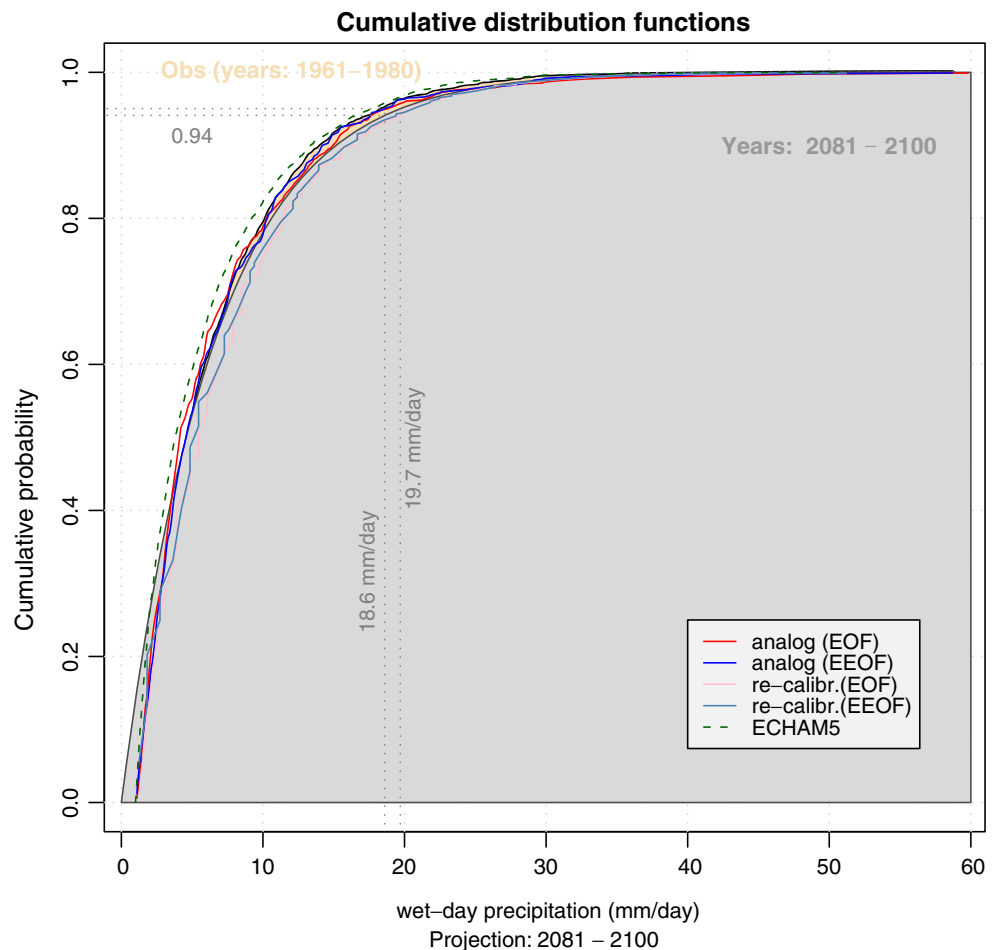
an independent validation of the model in terms of geographical distribution of $\hat{q}_{0.95}$. It was shown that the mean $\hat{q}_{0.95}$ level was in better agreement with corresponding empirical percentile than the percentiles estimated from ECHAM5 simulated precipitation, interpolated to the same location.

Another observation is that the gamma distribution and $f(x) = me^{-mx}$ diverged at low quantiles, but provided similar frequencies at the high tail (Fig. 11b). Thus, the probability estimates and return values should not be too sensitive to the choice of distribution if the analysis is limited to the range where the gamma and the exponential distributions converge.

EEOFs and temporal consistency The comparison between ordinary and EEOFs done here was not exhaustive, and it is possible that either or both can be improved further by choosing a different predictor domain, a different number of EOFs, using different predictors or using mixed predictor types (Benestad et al. 2007, 2002).

High percentiles The four-stage ESD approach merely provided approximate results and gave valid results

Fig. 10 CDF and empirical distribution functions (EDFs)(Jenkinson 1977). The PDF-adjusted values lie on top on the CDF $F(x) = 1 - e^{-mx}$ ($m > 0$)



only for the moderately high percentiles. An extrapolation scheme was proposed for making scenarios for the trends for the higher percentiles. It was shown that simple random behaviour combined with a gamma distribution was sufficient to provide a high-tail clutter. The statistical fluctuations in the extreme upper tail clutter imply that it is not meaningful to apply trend analysis to the extreme upper percentiles associated with this clutter.

A better way to analyse the trends in the most extreme values, however, is to examine the occurrence of record-breaking events and apply an IID test (Benestad 2008b, 2004b, 2003a). Another solution, adopted here, is to derive trend estimates for the most extreme percentiles q_p ($p > 0.975$) of the fitted PDFs by making a number of assumptions: (1) that the trend magnitude, α , varies smoothly with the q_p ; (2) that α is zero for precipitation amounts greater than a given upper boundary P_* and (3) that the function describing α in terms of percentile is of a simple form. This extrapolation scheme can also be applied directly to RCM and GCM results. However, the trends derived for these extreme

percentiles were highly uncertain and sensitive to the level for which the zero-trend was imposed. It is also possible that trends for such high percentiles cannot be defined. Thus, the extreme percentile trends should merely be regarded as ‘plausible’ scenarios, and should be associated with a high degree of uncertainty.

The trend α was slightly lower than the trend in $q_{0.95}$ estimated directly from ECHAM5 over two shorter time slices. However, the ESD results are also not directly comparable with ECHAM5, as the former represent the ensemble mean of a large number of state-of-the-art GCMs, of which ECHAM5 is only one member.

Sources of uncertainty The projection of future percentiles involved uncertainties from a number of sources: (a) the future may not follow the assumed emission scenario (here SRES A1b), (b) the GCM simulations may involve systematic errors and biases, (c) shortcomings associated with the ESD, and (d) approximations implied using the simple single-parameter model $f(x) = me^{-mx}$ (Eqs. 1, 3 and 2) and limitations in $\hat{q}_{0.95}$. In addition, the extrapolated trends involved

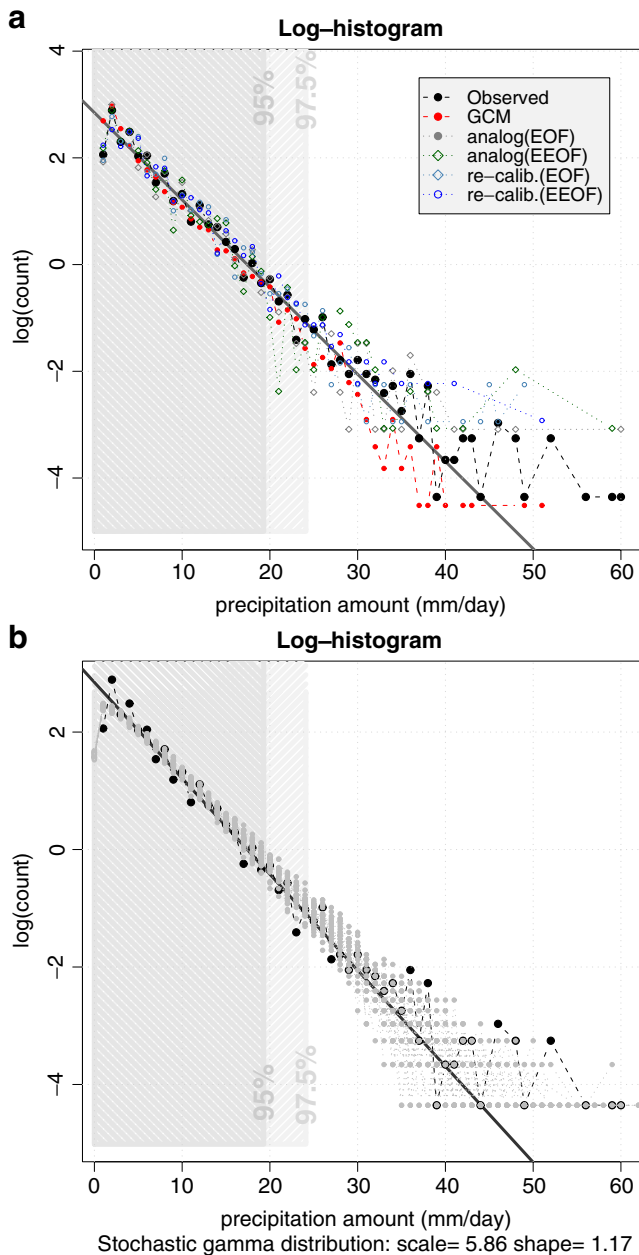


Fig. 11 Log count histograms for the observed precipitation (black) and the simulated precipitation. **a** Daily interpolated precipitation from ECHAM5 (red), as well as the results from the analog modelling and the re-calibration. The time interval is not the same for all the different data sets. **b** Similar diagnostics, but now, the simulations have been replaced by random data with a fixed best-fit gamma distribution

further uncertainties associated with the assumptions about the function describing the trend and the question of how the PMP may change under a global warming.

The first type of uncertainties associated with future emissions was difficult to evaluate, but the others were possible to assess. The evaluation of the ESD results

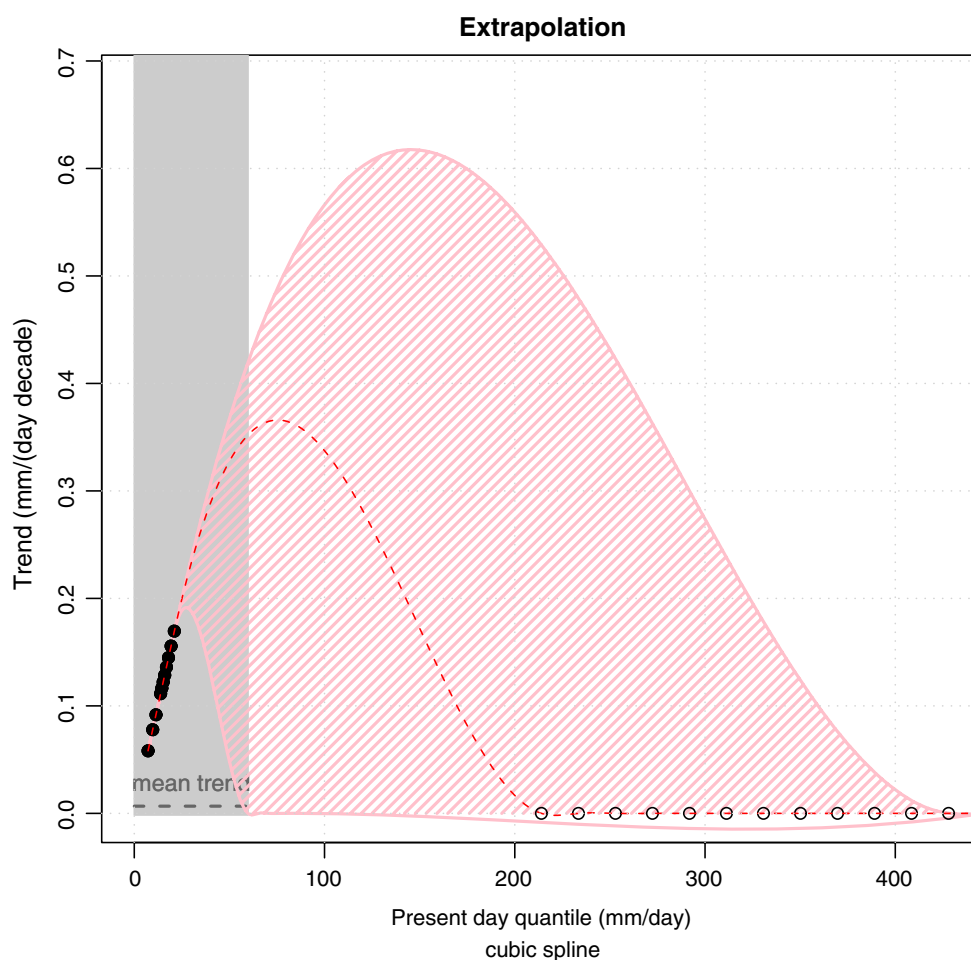
for the past, based on GCM simulation of the twentieth century (Figs. 2–3), suggested that shortcomings associated with b–c have not caused discrepancies for the past and that the combination GCMs and ESD gave realistic solutions. As long as these do not break down in a warmer climate, the ESD results for \bar{T} and \bar{P}_a should be valid for the future too. Since the regression models reproduced less than 100% of the variability, especially for precipitation (Benestad et al. 2007), it is expected that the ESD results, to some degree, will underestimate the future local mean climate variables.

The slope m estimated according to Eqs. 2 and 3 has been developed for 49 different locations in Europe and validated over 37 independent sites by Benestad (2007b). It is important to keep in mind that Eq. 2 is only valid for the European region. The time evolution and mean level for the predicted value for $q_{0.95}$ showed a good agreement with the empirical values for Oslo–Blindern (Fig. 5), albeit with a low bias of 2.3 mm/day. The coefficients in Eq. 2 were associated with uncertainties, expressed as standard errors, that may account for some of the bias: the uncertainty in the constant, temperature and precipitation terms may explain about 1% each, but additional uncertainty was also introduced through the geographical parameters. It is also possible that the statistical relationship expressed in Eq. 2 becomes invalid under a future climate.

The extrapolation was probably associated with the greatest uncertainties, as it was based on three assumptions. It is also possible that the trends α are undefined trends for $q_p > P_*$, which would mean that the upper values do not necessarily converge towards zero.

Physical interpretations An interesting observation is that the exponential distribution seen in Fig. 1 exhibits a character that resembles scaling laws seen elsewhere in nature (Malamud 2004). The initiation of rain is thought to often involve a stage of collision and coalescence (Rogers and Yau 1989), which may give rise to a stochastic avalanche-type process where an initial cloud drop population grows exponentially. Such a stochastic view is consistent with a pronounced growth in the upper percentiles with the mean precipitation level, as more drops will imply higher probabilities for such avalanche events, as well as favour larger-scale events. In addition, an increase in the percentile with higher temperature can be explained in terms of higher moisture-holding capacity in warmer air, and that convective processes are more likely to occur during warmer conditions. However, a warmer climate may also result in a change in the likelihood for warm (collision and coalescence) and cold initiation of rain (involving freezing processes).

Fig. 12 The linear trend slopes derived for the different percentiles $q_{0.70}$ – $q_{0.97}$ (closed circle), assumed upper bound (open circle) and a cubic spline interpolation between these points (red curve). The grey shaded area marks the range of the observed precipitation amounts and the pink hatched region marks a crude confidence region estimated by replacing the upper boundary by the maximum observed value and $2 \times$ PMP, respectively



Implications The re-calibrated ESD results indicate a stronger increase in the heavy precipitation over Oslo than a corresponding analysis based on traditional analog models, but the analog model suggested higher values for the most extreme percentiles. Since the extreme upper tail is associated with a small statistical sample and substantial sampling fluctuations, it also involves a higher degree of uncertainty.

The trends in the upper percentiles have implications for the probabilities and design values, suggesting a higher probability for severe 24-h precipitation events. Heavy 24-h precipitation can be a challenge for drainage systems and presents a problem in terms of damage to property, infrastructure, or agriculture. The results presented here represent projections that are independent of dynamical downscaling.

Further applications This technique may be used to downscale wind (Pryor et al. 2005a, b) or other climate parameters such as cloudiness. It is also possible to generate maps for extreme precipitation and probabilities associated with these (Prudhomme and Reed 1999; Benestad 2007b), but it may be more tricky to

ensure spatial consistency over larger distances. Over smaller regions, however, the analog model can provide a description for a set of sites in close proximity of each other. There is a limit to the size of the area represented by the predictands and the ability of the analog model for making spatially consistent scenarios, since there is a trade-off between predictor area and performance that remains to be explored.

Future work In order to further elucidate the limitations associated with downscaling 24-h precipitation, wet-day distributions derived from RCMs, driven with present-day boundary conditions, should be used to identify similar statistical relationships as the empirical data represented by Eq. 2. Then, if the RCM reproduces the observed relationships, the same exercise should be done for a future climate, and a comparison between the statistical regression analyses for the different time slices can give an indication of whether Eq. 2 breaks down under a different climate.

A next step could also be to see if similar extrapolation of percentiles can be used to make plausible scenarios for hourly precipitation. The 24-h precipitation

can be compared with similar analyses for 48 h (2 days), 96 h (4 days) and so on, and a similar extrapolation for time scales as shown in Fig. 12 can be applied to trend estimates associated with the different time scales, albeit using a linear fit rather than a cubic interpolation in order to make a prediction of the trends for 12-h, 6-h and 3-h precipitation. However, precipitation has a diurnal cycle that may invalidate such an extrapolation, so thorough tests are required to see if it is possible to extrapolate to shorter time scales. The extrapolated values can then be compared with hourly pluviometric data. Further tests can be carried out with Monte-Carlo techniques, using a gamma distribution fitted to data on an hourly scale, by estimating averages over different time scales to test the extrapolation.

An interesting question is whether the trends can be taken as a function of both percentile and time scales, e.g. a bi-variate function. Alternatively, different relationships between trend and percentile for different time scales should have an explanation in terms of physical processes and statistics.

3.7 Conclusions

ESD of monthly mean temperature and monthly precipitation totals have been used to derive the slope parameter m for the exponential distribution $f(x) = me^{-mx}$ for the wet-day 24-h precipitation amount. The way m is projected to change in the future has a direct effect on percentiles. Higher percentiles are projected to increase more rapidly than the lower ones.

A set of analog models has been used to provide chronology scenarios for daily precipitation events. However, analog models in isolation do not yield a reliable statistical distribution of the precipitation amounts, particularly near the upper tail.

A re-calibration, involving a quantile–quantile mapping, was performed using the PDF $f(x)$ predicted from the downscaling of the monthly data. The raw results from the analog models overestimated the higher quantiles of the precipitation amounts for the past, but re-calibration produced values in the extreme upper tail that were both higher and lower than the observed values, depending on the type of analog model. However, the re-calibration of the analog model results suggested a systematic increase in the probability for heavy precipitation events in the future, except for the most extreme upper percentiles. It is important to keep in mind, however, the high degree of statistical fluctuations associated with the most extreme values.

The predictors for the analog model in this study have involved both ‘ordinary’ and extended common EOFs. Both provided credible statistical distributions,

but the latter resulted in stronger 1-day persistence for the re-calibrated results than more traditional common EOFs.

The conclusions that can be drawn from experiments with Monte-Carlo simulations for gamma-distributed random variables is that it is meaningless to try to estimate trends in the most extreme upper percentiles, as statistical fluctuations are too pronounced. However, moderately high percentiles (q_p ; $p < 0.975$) tend to follow the exponential PDF, and by making a number of assumptions, it is possible to make scenarios for the highest percentiles, albeit with a high degree of uncertainty. The established statistical relationship and the ESD results suggest a stronger increase in the frequency of the most extreme precipitation events than for more moderate precipitation.

Acknowledgements This work has been supported by the Norwegian Research Council (Norclim #178246) and the Norwegian Meteorological Institute. The climatological data archive is maintained and quality controlled by ‘Seksjon for Klimadata’ in the Climate Department of the Norwegian Meteorological Institute. Their work is invaluable. 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/>). I acknowledge the international modeling groups for providing their data for analysis, the Program for Climate Model Diagnosis and Intercomparison (PCMDI) for collecting and archiving the model data, the JSC/CLIVAR Working Group on Coupled Modelling (WGCM) and their Coupled Model Intercomparison Project (CMIP) and the Climate Simulation Panel for organising the model data analysis activity and the IPCC WG1 TSU for technical support. The IPCC Data Archive at Lawrence Livermore National Laboratory is supported by the Office of Science, U.S. Department of Energy.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Alfnes E, Førland E (2006) Trends in extreme precipitation and return values in Norway 1900–2004. Report 02-2006, met.no
- Benestad R (2001) A comparison between two empirical downscaling strategies. *Int J Climatol* 21:1645–1668. doi:10.1002/joc.703
- Benestad R (2003a) How often can we expect a record-event? *Clim Res* 25:3–13
- Benestad R (2003b) What can present climate models tell us about climate change? *Clim Change* 59:311–332
- Benestad R (2004a) Empirical–statistical downscaling in climate modeling. *Eos* 85(42):417
- Benestad R (2004b) Record-values, non-stationarity tests and extreme value distributions. *Glob Planet Change* 44:11–26. doi:10.1016/j.gloplacha.2004.06.002

- Benestad R (2005a) Climate change scenarios for northern europe from multi-model ipcc ar4 climate simulations. *Geophys Res Lett* 32:L17,704. doi:[10.1029/2005GL023401](https://doi.org/10.1029/2005GL023401)
- Benestad R (2005b) Storm frequencies over fennoscandia—relevance for bark beetle outbreak. *reglim results. Climate* 20, met.no
- Benestad R (2007a) Changes in daily mean, maximum and minimum temperatures. *Climate* 9, met.no, www.met.no
- Benestad R (2007b) Novel methods for inferring future changes in extreme rainfall over northern europe. *Clim Res* 34:195–210. doi:[10.3354/cr00693](https://doi.org/10.3354/cr00693)
- Benestad R (2008a) Empirical-statistical downscaled arctic temperature & precipitation series. *Climate* 10, met.no, www.met.no
- Benestad R (2008b) A simple test for changes in statistical distributions. *Eos* 89(41):389–390
- Benestad R, Haugen J (2007) On complex extremes: flood hazards and combined high spring-time precipitation and temperature in norway. *Clim Change* 85:381–406. doi:[10.1007/s10584-007-9263-2](https://doi.org/10.1007/s10584-007-9263-2)
- Benestad R, Chen D, Hanssen-Bauer I (2008) Empirical-statistical downscaling. World Scientific, Singapore
- Benestad R, Hanssen-Bauer I, Førland E (2002) Empirically downscaled temperature scenarios for svalbard. *Atmos Sci Lett* 3(2–4):71–93. doi:[10.1006/asle.2002.0051](https://doi.org/10.1006/asle.2002.0051)
- Benestad R, Hanssen-Bauer I, Førland E (2007) An evaluation of statistical models for downscaling precipitation and their ability to capture long-term trends. *Int J Climatol* 27:649–665. doi:[10.1002/joc.1421](https://doi.org/10.1002/joc.1421)
- Bengtsson L, Hodges K, Hagemann S (2004) Sensitivity of the era40 reanalysis to the observing system: determination of the global atmospheric circulation from reduced observations. *Tellus* 56A:456–471
- Bergant K, Kajfež-Bogataj L (2005) N-pls regression as empirical downscaling tool in climate change studies. *Theor Appl Climatol* 81:11–23
- Bergant K, Kajfež-Bogataj L, Črepinšek Z (2002) The use of eof analysis for preparing the phenological and climatological data for statistical downscaling—case study: the beginning of flowering of the Dandelion (*Taraxacum officinale*) in Slovenia. In: Mrvar A, Ferligoj A (eds) *Developments in Statistics*, FDV, pp 163–174
- Blyth A, Benestad R, Krehbiel P, Latham J (1997) Observations of supercooled raindrops in new mexico summertime cumuli. *J Atmos Sci* 54(4):569–575. doi:[10.1175/1520-0469\(1997\)054](https://doi.org/10.1175/1520-0469(1997)054)
- Busuioc A, Giorgi F, Bi X, Ionita M (2006) Comparison of regional climate model and statistical downscaling simulations of different winter precipitation change scenarios over romania. *Theor Appl Climatol* 86:101–123
- Busuioc A, von Storch H, Schnur R (1999) Verification of gcm-generated regional seasonal precipitation for current climate and of statistical downscaling estimates under changing climate conditions. *J Clim* 12:258–272
- Chen D, Achberger C, Räisänen J, Hellström C (2005) Using statistical downscaling to quantify the gcm-related uncertainty in regional climate change scenarios: a case study of swedish precipitation. *Adv Atmos Sci* 23:54–60
- Christensen J, Machenhauer B, Jones R, Schär C, Ruti P, Castro M, Visconti G (1997) Validation of present-day regional climate simulations over europe: Lam simulations with observed boundary conditions. *Clim Dyn* 13:489–506. doi:[10.1007/s003820050178](https://doi.org/10.1007/s003820050178)
- Christensen J, Hewitson B, Busuioc A, Chen AXG, Held I, Jones R, Kolli R, Kwon WT, Laprise R, na Rueda VM, Mearns L, Menéndez C, Räisänen J, Rinke A, Sarr A, Whetton P (2007) Regional climate projections. In: *Climate change: the physical science basis*. Cambridge University Press, New York
- Dehn M (1999) Application of an analog downscaling technique to the assessment of future landslide activity—a case study in the italian alps. *Clim Res* 13:103–113
- Ellner S (2001) Review of R, version 1.1.1. *Bull Ecol Soc Am* 82:127–128
- Engen-Skaugen T (2004) Refinement of dynamically downscaled precipitation and temperature scenarios. *Climate* 15/2004, met.no, www.met.no
- Fernandez J, Saenz J (2003) Improved field reconstruction with the analog method: searching the cca space. *Clim Res* 24:199–213
- Folland C, Anderson C (2002) Estimating changing extremes using empirical ranking methods. *J Climate* 15:2954–2960
- Fowler H, Blenkinsop S, Tebaldi C (2007) Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. *Int J Climatol* 27:1547–1578
- Frei C, Christensen J, Déqué M, Jacob D, Jones R, Vidale P (2003) Daily precipitation statistics in regional climate models: Evaluation and intercomparison for the european alps. *J Geophys Res* 108(D3). doi:[10.1029/2002JD002287](https://doi.org/10.1029/2002JD002287)
- Gentleman R, Ihaka R (2000) Lexical scope and statistical computing. *J Comput Graph Stat* 9:491–508. <http://www.amstat.org/publications/jcgs/>
- Giorgetta M, Manzini E, Roeckner E (2002) Forcing of the quasi-biennial oscillation from a broad spectrum of atmospheric waves. *Geophys Res Lett* 29. doi:[10.1029/2002GL014756](https://doi.org/10.1029/2002GL014756)
- Hanssen-Bauer I, Achberger C, Benestad R, Chen D, Førland E (2005) Statistical downscaling of climate scenarios over scandinavia: A review. *Clim Res* 29:255–268
- Haylock M, Cawley G, Harpham C, Wilby R, Goodess C (2006) Downscaling heavy precipitation over the united kingdom: a comparison of dynamical and statistical methods and their future scenarios. *Int J Climatol* 26:1397–1416
- Hegerl G, Zwiers FW, Braconnot P, Gillett N, Luo Y, Orsini JM, Nicholls N, Penner J, Stott P (2007) Understanding and attributing climate change. In: *Climate change: the physical science basis*. Cambridge University Press, New York
- Huth R (2004) Sensitivity of local daily temperature change estimates to the selection of downscaling models and predictors. *J Climate* 17:640–652
- Ihaka R, Gentleman R (1996) R: a language for data analysis and graphics. *J Comput Graph Stat* 5(3):299–314
- Imbert A, Benestad R (2005) An improvement of analog model strategy for more reliable local climate change scenarios. *Theor Appl Climatol* 82:245–255. doi:[10.1007/s00704-005-0133-4](https://doi.org/10.1007/s00704-005-0133-4)
- Jenkinson A (1977) The analysis of meteorological and other geophysical extremes. Tech. rep. 58, U.K. Met.Office Synoptic Clim. Branch
- Keenlyside N, Latif M (2002) Enso hindcast performance of the mpi-om/echam5 coupled model. EGS XXVII General Assembly, Nice, abstract #1150
- Kettle H, Thompson R (2004) Statistical downscaling in European mountains. *Clim Res* 26:97–112
- Linderson ML, Achberger C, Chen D (2004) Statistical downscaling and scenario construction of precipitation in Scania, Southern Sweden. *Nord Hydrol* 35:261–278
- Lorenz E (1956) Empirical orthogonal functions and statistical weather prediction. Sci. rep. 1. Department of Meteorology, MIT
- Malamud B (2004) Tails of natural hazards. *Physics World* 17:31–35

- Meehl G, Covey C, Delworth T, Latif M, McAvaney B, Mitchell J, Stouffer R, Taylor K (2007a) The wcrp cmip3 multimodel dataset: a new era in climate change research. *Bull Am Meteor Soc* 88:1383–1394
- Meehl G, Stocker T, WDidlingstein, Gaye A, Gregory J, Kitoh A, Knutti R, Murphy J, Noda A, Raper S, Watterson I, Weaver A, Zhao ZC (2007b) Global climate projections. In: *Climate change: the physical science basis*. Cambridge University Press, New York
- Penlap E, Matulla M, von Storch H, Kamga F (2004) Downscaling of gcm scenarios to assess precipitation changes in the little rainy season (March–June) in cameroon. *Clim Res* 26:85–96
- Press W, Flannery B, Teukolsky S, Vetterling W (1989) *Numerical recipes in Pascal*. Cambridge University Press, Cambridge
- Prudhomme C, Reed D (1999) Mapping extreme rainfall in a mountainous region using geostatistical techniques: a case study inn scotland. *Int J Climatol* 19:1337–1356
- Pryor S, School J, Barthelmie R (2005a) Climate change impacts on wind speeds and wind energy density in northern europe: empirical downscaling of multiple aogcms. *Clim Res* 29:183–198
- Pryor S, School J, Barthelmie R (2005b) Empirical downscaling of wind speed probability distributions. *J Geophys Res* 110:D19,109. doi:10.1029/2005JD005899
- Rogers R, Yau M (1989) *A short course in cloud physics*, 3rd edn. Pergamon, Oxford
- Salathé E (2005) Downscaling simulations of future global climate with application to hydrologic modelling. *Int J Climatol* 25:419–436
- Schoof J, Pryor S (2001) Downscaling temperature and precipitation: a comparison of regression-based methods and artificial neural networks. *Int J Climatol* 21:773–790
- Semenov M, Barrow E (1997) Use of a stochastic weather generator in the development of climate change scenarios. *Clim Change* 35:397–414
- Semenov M, Brooks R (1999) Spatial interpolation of the lars-wg stochastic weather generator in great britain. *Clim Res* 11:137–148
- Simmons A, Gibson J (2000) The era-40 project plan. ERA-40 project report series 1, ECMWF. www.ecmwf.int
- Simmons A, Jones P, da Costa Bechtold V, Beljaars A, Källberg P, Saarinen S, Uppala S, Viterbro P, Wedi N (2004) Comparison of trend and variability in cru, era40 and ncep/ncar analyses of monthly-mean surfaceair temperatures. ERA40 Project report series 18, ECMWF
- Skaugen T, Hanssen-Bauer I, Førland E (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
- Solomon S, Quin D, Manning M, Chen Z, Marquis M, Averyt K, Tignotr M, Miller H (eds) (2007) *Climate Change: the physical science basis*. In: *Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, New York
- Soltani A, Hoogenboom G (2003) Minimum data requirements for parameter estimation of stochastic weather generators. *Clim Res* 25:109–119
- van den Dool H (1995) Constructed analogue prediction of the east central tropical pacific through spring 1996. *NOAA: Experimental Long-Lead Forecast Bulletin* 4:41–43
- von Storch H (1999) On the use of “inflation” in statistical downscaling. *J Climate* 12:3505–3506
- von Storch H, Zwiers F (1999) *Statistical analysis in climate research*. Cambridge University Press, Cambridge
- von Storch H, Hewitson B, Mearns L (2000) Review of empirical downscaling techniques. In: T I, BAK H (eds) *Regional Climate Development Under Global Warming*. General technical report 4. http://regclim.met.no/rapport_4/presentation02/presentation02.htm
- Timbal B, Jones D (2008) Future projections of winter rainfall in southeast australia using a statistical downscaling technique. *Clim Change* 86:165–187
- Timbal B, Dufour A, McAvaney B (2003) An estimate of future climate change for western france using a statistical downscaling technique. *Clim Dyn* 20:807–823. doi:10.1007/s00382-002-0298-9
- Timbal B, Hope P, Charles S (2008) Evaluating the consistency between statistically downscaled and global dynamical model climate change projections. *J Climate* 21(22):6052–6059
- Wilby R, Charles S, Zortia E, Timbal B, Whetton P, Mearns L (2004) Guidelines for use of climate scenarios developed from statistical downscaling methods. In: *Supporting material of the Intergovernmental Panel on Climate Change, task group on Data and Scenario Support for Impacts and Climate Analysis*
- Wilks D (1995) *Statistical methods in the atmospheric sciences*. Academic, Orlando
- Zorita E, von Storch H (1999) The analog method as a simple statistical downscaling technique: comparison with more complicated methods. *J Climate* 12:2474–2489