

Robust assessment of future changes in extreme precipitation over the Rhine basin using a GCM

S. F. Kew, F. M. Selten, G. Lenderink, and W. Hazeleger

Royal Netherlands Meteorological Institute, De Bilt, The Netherlands

Received: 3 November 2010 – Published in Hydrol. Earth Syst. Sci. Discuss.: 26 November 2010

Revised: 6 March 2011 – Accepted: 30 March 2011 – Published: 8 April 2011

Abstract. Estimates of future changes in extremes of multiday precipitation sums are critical for estimates of future discharge extremes of large river basins. Here we use a large ensemble of global climate model SRES A1b scenario simulations to estimate changes in extremes of 1–20 day precipitation sums over the Rhine basin, projected for the period 2071–2100 with reference to 1961–1990.

We find that in winter, an increase of order 10%, for the 99th percentile precipitation sum, is approximately fixed across the selected range of multiday sums, whereas in summer, the changes become increasingly negative as the summation time lengthens. Explanations for these results are presented that have implications for simple scaling methods for creating time series of a future climate. We show that the dependence of quantile changes on summation time is sensitive to the ensemble size and indicate that currently available discharge estimates from previous studies are based on insufficiently long time series.

1 Introduction

Estimates of future changes in multiday precipitation extremes are critical for estimates of future discharge extremes occurring once every 100–1000 yr, yet they are often based on the order of just 30 yr of global climate model simulations (Shabalova et al., 2003; Kay et al., 2006; Dankers et al., 2007) or 90 yr at best (Lenderink et al., 2007a). The precipitation input for discharge models is commonly generated by high resolution regional climate models (RCMs), due to the need to resolve small scale processes. Global climate models (GCMs), however, are required to supply the boundary conditions and effectively impose the large scale flow and its

variability on the RCM simulations. If future discharge estimates have been based on too few years of data, there is a risk that the natural variability of the climate has not been adequately sampled (Selten et al., 2004) and the impact of changes in large-scale circulation on extreme precipitation may have been mis-represented.

Global warming-induced changes in circulation regimes (e.g. Ulbrich and Christoph, 1999; Gillett et al., 2003; Hu and Wu, 2004; Yin, 2005; Pinto et al., 2007; Brandefelt and Körnich, 2008) and atmospheric moisture content (Trenberth, 1999) are expected to affect the intensity, frequency and relative persistence of extreme precipitation events and dry spells (Frei et al., 2000; Van Ulden and van Oldenborgh, 2006; Van den Hurk et al., 2007; Meehl et al., 2007). In summer, for example, sequences containing long dry spells followed by intense precipitation (Lenderink et al., 2009), might become more common. This could cause fractional changes in multiday precipitation extremes (relevant for catchment-scale discharge) to differ in amplitude, or even in sign, from fractional changes in single-day extremes. That would have implications for the delta-change technique (Lenderink et al., 2007a), a method that uses mean changes in climate parameters to transform historical precipitation sequences, by multiplication with a scale factor, to future time series for input to hydrological models.

Here we will study changes in extreme multiday precipitation over the Rhine catchment area in a very large GCM ensemble, originating from the ESSENCE project (Sterl et al., 2008). This ensemble consists of 17 integrations from 1950 to 2100 with an identical model (ECHAM5/MPI-OM) forced by A1b emissions. In this ensemble we are optimally able to distinguish the signal due to climate change from natural variability. Note that a dynamical downscaling of such a large ensemble using nested RCM simulations is currently computationally very expensive and beyond the scope of this study. The following questions are addressed: How do changes in n -day precipitation extremes depend on



Correspondence to: S. F. Kew
(sarah.kew@knmi.nl)

n? Are there significant differences between single-day and multiday precipitation extremes? How large does an ensemble need to be to distinguish climate change from natural variability? The paper is structured as follows: description of the ensemble (Sect. 2), methods (Sect. 3), comparison of GCM results with present-day climate observations (Sect. 4), climate change results (Sect. 5) and concluding remarks (Sect. 6).

2 Study area and data

2.1 The Rhine basin

The Rhine basin (Fig. 1) covers an area of 185 000 km² shared between 9 different countries. The main river, the longest in western Europe, is about 1300 km in length and passes through a range of landscapes, originating in the Swiss Alps, cutting through highlands to the North and branching out in several deltas in the Netherlands before joining the North Sea. The annual mean discharge (1901–2000) at Lobith (Fig. 1) is 2200 m³ s⁻¹ and current defences in the Netherlands are designed to withstand a 1 in 1250-yr flood event with a discharge of 16 000 m³ s⁻¹. It is expected that, as global temperatures rise, the mean discharge of the Rhine will increase in winter, due to increased precipitation and earlier snow melt, and decrease in summer due to reduced precipitation and increased evaporation (e.g. Hurkmans et al., 2010). Such changes will impact the seasonal likelihood of flooding and increase restrictions on river transport in low discharge periods.

2.2 ESSENCE dataset

The ESSENCE dataset (Sterl et al., 2008) is a 17-member ensemble simulation for the years 1950–2100, generated from the ECHAM5/MPI-OM coupled climate model which has a horizontal resolution of T63 and 31 vertical hybrid atmospheric levels, and is forced by the SRES A1b scenario (Nakićenović et al., 2000). The different ensemble members are formed by perturbing the initial state of the atmosphere, with ocean conditions unchanged.

Figure 1 displays the ESSENCE grid over the Rhine basin. There are eight (shaded) ESSENCE grid cells that notably overlap the basin (on the order of 20% or greater of their area is part of the basin) and these are taken to represent the Rhine basin in the ESSENCE dataset.

The 8-cell domain representing the Rhine basin is divided into three zonal regions, the North Rhine (2 cells), the Central Rhine (4 cells) and the Alpine Rhine (2 cells), which are treated separately. This choice is motivated by the possible differences in the precipitation distribution following the flow of the river from the south to the north of the domain, meridional gradients in temperature and topography, and reported North-South gradients in the modeled mean precipitation response to climate change (see Fig. 11.5 of chapter 11

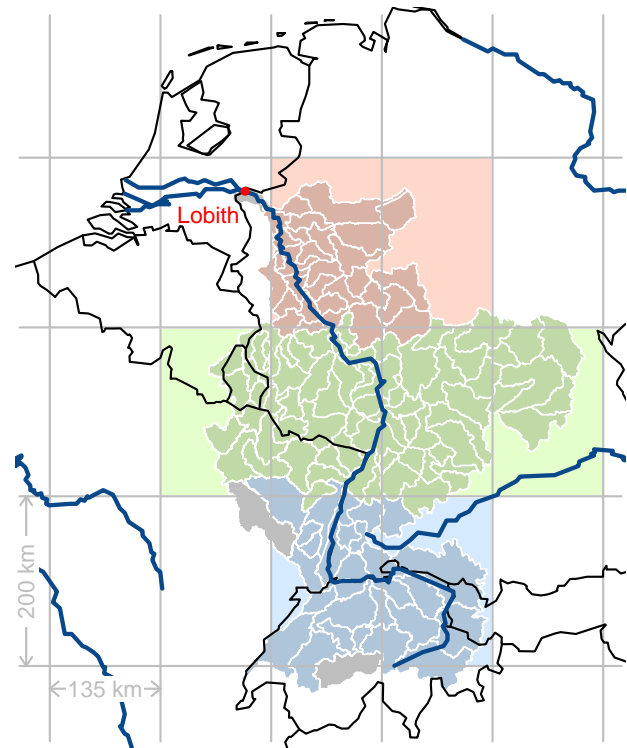


Fig. 1. The Rhine basin and ESSENCE grid. The basin is represented by 8 cells: 2 in the North Rhine (pink), 4 in the Central Rhine (green) and 2 in the Alpine Rhine (blue) region. CHR-OBS subbasins are outlined in white and shaded according to which region they are assigned for upscaling. Main waterways are traced in blue.

in the IPCC 4AR report; Christensen et al., 2007). Splitting the domain will also provide multiple output sets for comparison and thus an indication of the consistency of the results and their sensitivity to location.

2.3 CHR-OBS dataset

A historical set of precipitation observations issued by the International Commission for the Hydrology of the Rhine basin (CHR) will be used to gauge the model performance. The CHR dataset, recently named CHR-OBS, comprises area-averaged daily precipitation sums for the 134 Hydrologiska Bryåns Vattenbalansavdelning (HBV) model subbasins (contoured in white in Fig. 1) of the Rhine catchment for the period spanning January 1961–December 1995. Details on the development of this dataset are given in the (German) CHR technical report (Sprokkereef, 2001) and a brief (English) summary can be found in e.g. Terink et al. (2010).

We upscale the CHR data to the approximate size of the three zonal regions (North Rhine, Central Rhine, Alpine Rhine) by area-averaging the daily totals for the group of sub-basins whose centers lie within the boundaries of a particular region (Fig. 1).

3 Methodology

Time series of the area averaged ESSENCE daily precipitation for the three regions are produced for two 30-yr time slices: a control period, December 1961–November 1991, and a future period, December 2070–November 2100. A wet-day threshold of 0.1 mm is applied, i.e. values below 0.1 mm are set to zero and thereby treated as dry days. With 17 members, this gives a total of $30 \times 17 = 510$ simulated years for each 30-yr period.

We investigate seasonal differences by comparing results for summer (JJA) and winter (DJF). Time series of n -day precipitation sums or “accumulation intervals” ($n = 1, 2, 5, 10, 20$) centered on each day in a season are formed. Whilst consecutive multiday day sums overlap and thus are not independent, the increased sample size allows an improved estimation of the form of the distribution.

A range of quantiles for each n -day accumulation interval and season are assessed. While we focus on the extreme quantiles (q_{99}) of the distribution, we also present results for intermediate quantiles (q_{50} , q_{90} and q_{95}) so that one can gain insight into the robustness of the results. The relative percentage change of the future precipitation quantile, q^f , with respect to the control period quantile, q^c , is evaluated i.e.

$$\Delta q = 100 \left(\frac{q^f - q^c}{q^c} \right). \quad (1)$$

We determine quantiles for two different distributions:

- The *full* season of sums including dry events (n -day sum is zero), for which quantiles are easily inverted into return values.
- The seasonal distribution excluding dry events, i.e. a multiday equivalent of the *intensity* distribution. The term “intensity” is usually used to refer to the mean amount of rainfall on wet days.

For a 10-day sum, method *a* provides an answer to the question “what sum of rain can we expect over a randomly selected 10-day period in the future compared to now?”. Method *b* provides an answer to the question “what sum of rain can we expect in a 10-day period in the future compared to now, on the condition that we know at least some rain falls in that period?”. In a practical sense, this question might be of importance if an amount of rain exceeding the wet-day threshold is forecast or if current and future 10-day periods with precipitation-favorable weather regimes were selected for comparison.

Note that for *a*, the set of individual days included is fixed across the different multiday sums, permitting a fair inter-comparison of the changes in precipitation quantiles for different accumulation intervals. For *b*, a direct inter-comparison is impeded by the removal of a decreasing fraction of dry days (and thus allowing another factor to vary) as the

accumulation interval n increases. At large n , there are few dry sums and *a* and *b* yield practically the same quantiles. Results from method *b* are presented here nevertheless as they provide complementary insight into predicted changes to multiday precipitation. The relative change in single-day sum intensities in the future with respect to the control period may also be compared to values in the literature.

Bootstrapping is used to estimate confidence intervals of Δq for the 17-member ensemble and also for a range of simulated smaller ensembles. A new 30-yr time series for a single ensemble member, e.g. during the DJF control period, is generated as follows: For each year of the control period in turn, one member out of the 17 ESSENCE ensemble members is randomly selected (with replacement) and the entire DJF season for that member and year forms one year of the time series. In this way, 17^{30} different arrangements are possible for each season and timeslice. A 3-member ensemble, for example, is then simulated as a collection of 3 such randomly constructed sequences. We create 10 000 samples of each ensemble size, for each season and timeslice in this manner. Quantiles for the n -day sums are estimated from each sample and the 95% confidence interval is taken as the bootstrap. Note that there is no subseason mixing in this procedure in order to preserve the autocorrelation of daily precipitation series. On the other hand, seasonal precipitation series in neighboring years are assumed to be independent (there is no significant autocorrelation of seasonal quantiles at a lag of 1 yr or beyond).

4 Comparison with observational data

In this section we compare the ESSENCE data for the control period (1962–1991, but including December 1961 for the winter season) to upscaled observations from the CHR-OBS dataset. The wet-day threshold of 0.1 mm is also applied to the upscaled observations.

Figure 2 presents ESSENCE and CHR-OBS probability density functions (PDFs) of 1-, 10- and 20-day sums for the North Rhine region during JJA and DJF of the 30-yr control period. The dry-event frequency is included as a separate “zero” column to the left of the PDF within each panel. In JJA, a reasonable match in the 1-day intensity distributions (Fig. 2a) can be seen by the near alignment of their quantiles (q_{50} and q_{99} shown by solid vertical lines). The two q_{50} for the full distribution (dashed vertical lines) are not well aligned due to the model’s larger dry-day frequencies. As n increases, the dry event frequency must decrease, and thus the intensity distribution converges into the full distribution (Fig. 2b–c). The model’s excess of dry 1-day sums have been mixed into wet multiday sums and consequently the PDF is shifted left towards lower values with respect to the observations. In DJF we see the opposite tendency with n . The single-day intensity PDF corresponds closely to the observations but the model has a larger wet-day frequency than the

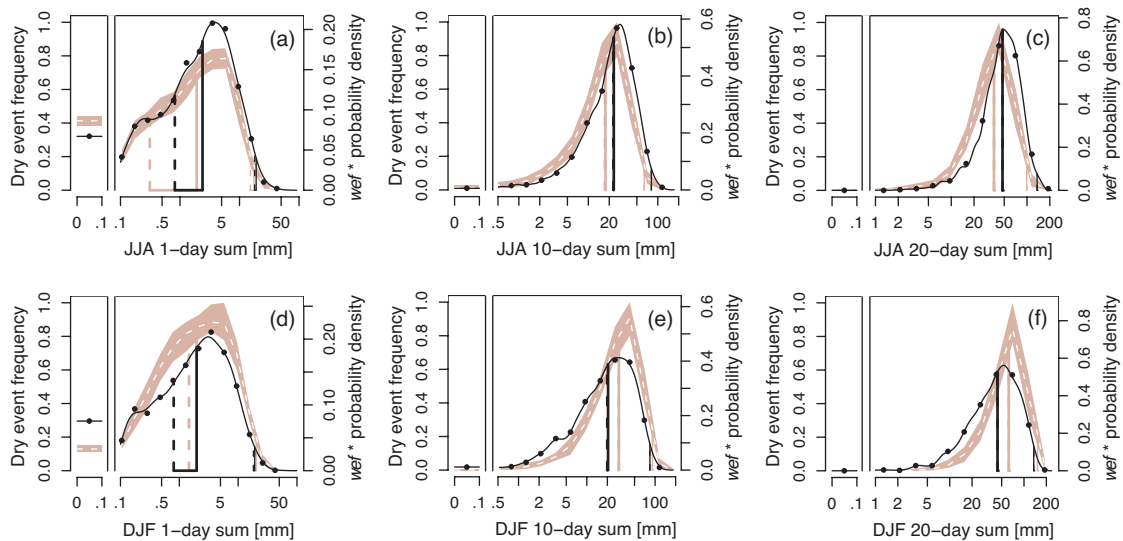


Fig. 2. Validation of ESSENCE against CHR-OBS: PDFs for the North Rhine region during the control period (1961/1962–1990/1991), in JJA (top row) and in DJF (lower row) for 1-, 10- and 20-day precipitation sums (left-right). The color shading envelops the 95% range of the probability density attained from individual ensemble members, dashed white shows the mean. Black dots show CHR-OBS binned observations and the black curve is an empirical fit (kernel density estimate using Gaussian smoothing) giving the CHR-OBS probability density. The frequency of dry events (separate column, left of each PDF) plus the integrated PDF of wet events (scaled by the wet event frequency, wef) together sum to unity. Vertical lines mark the locations of the 50% (thick) and 99% (thin) quantiles for the intensity PDF (solid) and the full distribution that includes the dry events (dashed). Note that the counting measure used for binning is the logarithm of the precipitation sum.

observations and this causes the multiday PDF to be shifted to higher values. In addition, the multiday PDF is narrower for ESSENCE.

Equivalent figures for the Central and Alpine Rhine regions can be found in the Supplement. For the Central Rhine region, the agreement is remarkably good in summer, (observed frequencies fall mostly within the shaded envelope of ensemble results) and is similar to the North Rhine region in the winter. The Alpine Rhine region exhibits the strongest bias in (low) intensities in summer, whilst a better centered but too narrow a PDF is seen in winter.

With regard to meridional tendencies, both datasets give larger intensities in the south compared to the north (summer and winter) but only the CHR-OBS show north-south trends in wet-day frequency. In ESSENCE, poorly resolved topography will surely take its toll and is likely the reason why the Central and Alpine Rhine distributions differ to a greater extent in the observational data than in the model data, and also why the Alpine Rhine is much drier in summer in the model than in the observations.

Overall, ESSENCE demonstrates reasonable behavior at the Rhine basin scale. The absolute quantile values however cannot be directly relied upon without correcting for model bias. We will report on relative changes between the control and future period, which stem directly from differences in the forcing or internal variability of the model ensemble. Relative changes remain unaltered under a bias correction of a

multiplicative error in both the control and future intensities. In the case of an additive (or a combined additive and multiplicative) error, the amplitude of the relative changes would be affected but the sign would remain the same.

5 Results

5.1 Dependence of relative quantile changes on accumulation interval

The relative quantile changes Δq for the North Rhine region's summer and winter are presented in Fig. 3 as a function of accumulation interval n for both the full distribution (left panels) and the intensity distribution (right panels).

Looking first at q_{99} of the full distribution, the most extreme quantile considered (Fig. 3a), we note contrasting behavior for the different seasons. In summer, a non-trivial dependence on increasing accumulation interval is observed: Δq_{99} is positive at 5.5% for the single-day sum, but turns negative for the 5-day sum, reaching -6.5% for 20-day sums. In the winter, Δq_{99} is positive across the board, between 6 and 10%, and remains relatively uniform across the range of multiday sums within the estimated confidence intervals for the 17-member ensemble.

For q_{95} and lower quantiles of the full distribution (Fig. 3c, e and g), the summer Δq turns negative for all accumulation

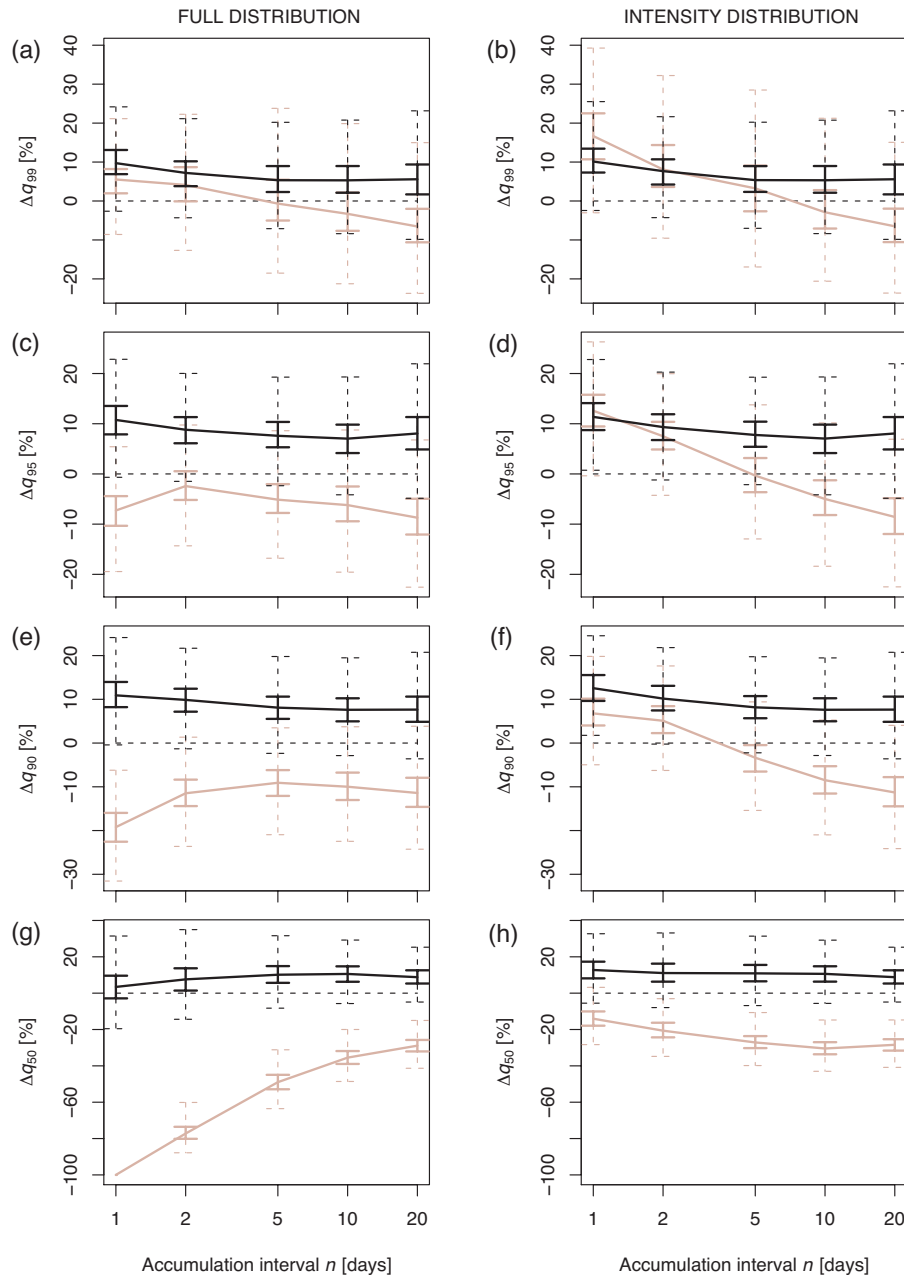


Fig. 3. The projected changes in quantiles (top to bottom: q_{99} , q_{95} , q_{90} , q_{50}) of 1–20 day precipitation sums expected by 2070–2100 with respect to 1961–1991, for the North Rhine region. Left: full distribution quantiles. Right: intensity distribution quantiles. Results for JJA (colored) and DJF (black) are shown together in the same panel. Error bars indicate 95% confidence intervals given by bootstrapping (10 000 samples) on 17 ensemble members (solid) or 1 member (dashed).

intervals n , and by q_{50} the dependency on n is even reversed, i.e. the fractional quantile change in the 1-day sum is far more negative than for the 20-day sum. The winter Δq remains relatively uniform and positive across the accumulation periods for all quantiles.

What is the cause of the difference in the dependence of Δq on n between the summer and winter in the North Rhine region? A uniform Δq , as we see in winter, can

be expected if the distributions of wet-day frequency and wet-period duration remain the same while the intensity of rain days changes. Indeed, the winter intensity distribution (Fig. 3b) shows the same magnitude for Δq_{99} as the full distribution (Fig. 3a), indicating that the relative change must be due almost entirely to an increase in event intensity, whilst the wet-day frequency remains largely unchanged. Note that the intensity distribution at $n = 1$ is independent of the

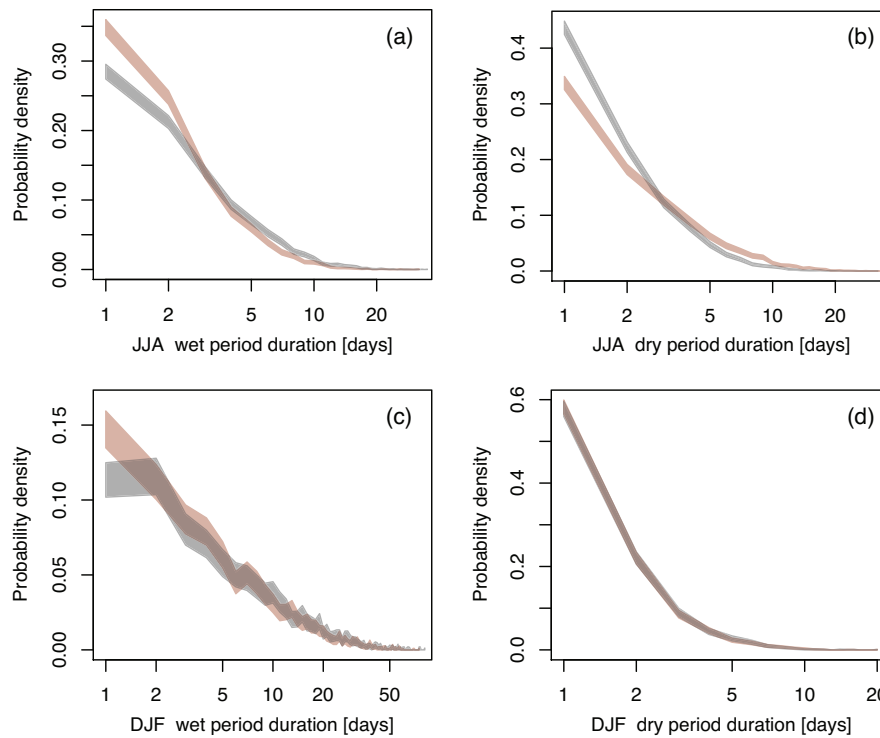


Fig. 4. Probability density functions for wet (left) and dry (right) period durations in JJA (top row) and DJF (lower row) for the North Rhine region. Shaded regions show 95% confidence intervals from bootstrapping for the 1961/1962–1990/1991 (control) time slice (gray) and the 2070/2071–2099/2100 (future) time slice (color).

wet-day frequency and therefore any difference in Δq between the full and intensity distribution at $n = 1$ is due to the change in wet-day frequency. We also find that the PDF of wet and dry spell durations in winter does not significantly change (Fig. 4c–d).

In summer, the North Rhine’s dependence of Δq on increasing accumulation interval is caused by a combination of increased extreme single-day intensities and reduced wet-day frequencies. Two aspects of the full distribution’s behavior would be present with a reduced wet-day frequency alone: (i) the 1-day sum’s lowest quantiles decrease, leaving high quantiles hardly affected – simply a consequence of raised probabilities at the “dry” end of the PDF (compare the magnitude of the difference between left hand and right hand panels of Fig. 3 for high and low quantiles at $n = 1$), and (ii) Δq converges towards the mean precipitation change as the summation interval lengthens. Together, (i) and (ii) lead to the positive n -dependence of Δq seen for low quantiles.

The added impact of increased intensities of extremes is to create a non-trivial n -dependence, whereby Δq is positive for 1-day extremes but negative for multiday extremes. The 1-day intensity distribution (Fig. 3b) shows there is a stronger positive Δq_{99} of 16.7% compared to 5.5% for the full distribution. The increase in intensity is large enough to hold

Δq_{99} for the full distribution positive for small n , off-setting the negative contribution from a reduced wet-day frequency.

The composition of 20-day summer extremes in both the control and future periods is such that around 80% of the sums satisfying the q_{99} threshold contain at least one day satisfying the respective q_{99} thresholds for single-day extremes (not shown). In other words, in many cases, it is the same event that makes a 1-day sum and 20-day sum extreme, and not persistence of moderate rainfall alone. An increase of dry/drier days mixed into the 20-day sum in between the extreme(s) must be the reason for the future decrease in multiday extremes. The PDF of summer wet and dry spell durations supports this showing that wet spells are projected to become shorter and dry spells longer (Fig. 4a–b).

Differences are seen between the three regions of the basin (Figs. S3, S4, Supplement). The summer dependence of Δq_{99} on n is strongest for the North Rhine. In the Central and Alpine regions Δq_{99} is negative for all n , being most negative furthest south. In the winter, the magnitudes of Δq_{99} are similar for the North and Central Rhine ($\sim 8\%$) but increase to around 15%, for the Alpine Rhine.

It is also of interest to inspect the transient simulated evolution in the seasonal cycle of monthly mean wet-day frequency and intensity (see Fig. 5 for the North Rhine and the

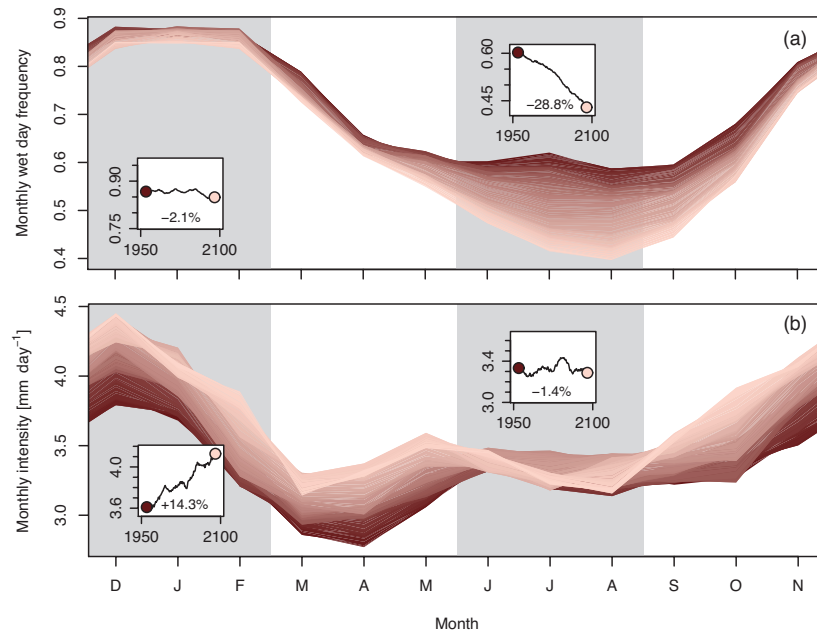


Fig. 5. Projected evolution of the seasonal cycle in wet-day frequency **(a)** and in intensity **(b)** from the beginning (dark shading) to the end (light shading) of the ESSENCE period, for the North Rhine region. Monthly values were averaged over a sliding window of 21 yr. Insets show the ensemble mean temporal evolution for DJF and JJA.

Supplement for the other regions). It is clear to see that in summer, the change in wet-day frequency is the dominating factor, whereas in winter it is a change in intensity that will modulate the quantile changes. For the Central and Alpine regions, a decrease in JJA mean intensity and wet-day frequency takes effect, consistent with the more negative Δq_{99} towards the south. These negative trends undergo acceleration during the second half of the ESSENCE simulation (see insets to figures in the Supplement). Further, it is projected that the seasonal cycles change form. In Fig. 5a, for example, it appears that for early years, the wet-day frequency follows a plateau from May to September, yet at the end of the simulation, the number of wet days continues to decrease until August. The cause of this non-linearity still needs to be investigated but we expect it can be attributed to feedbacks from an extended period of drying out of the soil.

5.2 Sensitivity to ensemble size

The non-trivial dependency of Δq on n seen for the North Rhine region in summer was detected using a 17 member ensemble. Current discharge estimates are based on much smaller datasets providing between 30 and 90 yr of integration for each timeslice (equivalent to 1–3 ensemble members here). In this section we simulate smaller ensembles using the bootstrap method to see if they are also capable of reproducing the non-trivial behavior and a climate signal that is significantly different from zero.

For the North Rhine region, Fig. 6 shows the 63% and 95% confidence intervals of Δq_{99} for 1-day and 20-day sums estimated from 10 000 samples for each ensemble size, in summer and winter. In summer, Fig. 6a, around 240 yr (8 members) are needed to detect the Δq_{99} signal as significantly different from zero. The different behavior of Δq_{99} for the 1- and 20-day sums is also separable at this point but the overlapping confidence bands further left suggest that this might not be the case for smaller ensembles. In Fig. 7a we display the direct relationship between the 1-day and 20-day Δq_{99} signal for each of the bootstrapped samples used in Fig. 6a for ensembles of sizes 1 and 3. The peak in the density of scattered points lies in the lower right hand quadrant, where Δq_{99} for the 1-day sum is positive and Δq_{99} for the 20-day sum is negative. However, a small fraction of points lie in the opposite quadrant, showing that, for small ensembles, even the opposite dependency of Δq_{99} on n can be attained.

In winter, Fig. 6b, for the 17-member ensemble, there is a small difference in Δq_{99} for the 1 and 20-day sums but this difference is not significant and is not distinguishable for smaller ensembles (Fig. 7b). The Δq_{99} for 1-day sums is significantly different from zero for an ensemble with 2 or more members, whereas, for 20-day sums, around 9 members are required for the same level of confidence.

Note the large range including both positive and negative values of Δq_{99} that might be obtained if just 30 yr of integration (1 member) or even 90 yr (3 members) are used. Confidence intervals for Δq , estimated using 10 000 1-member simulations, are also added (dashed error bars) to Fig. 3.

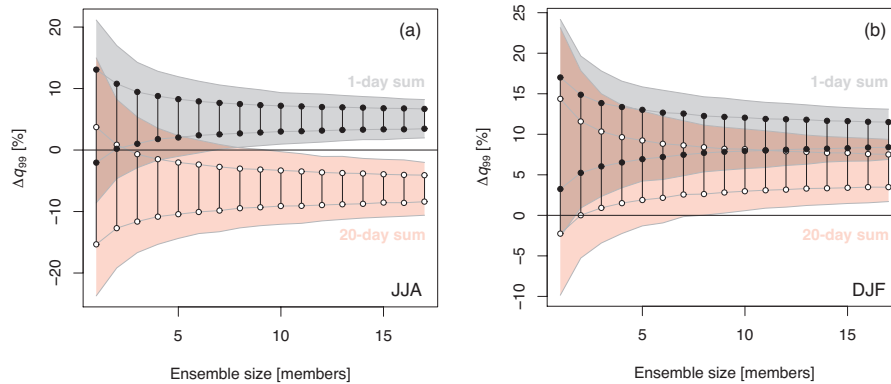


Fig. 6. Sensitivity of Δq_{99} to ensemble size for 1 day sums (gray shading, black dots) and 20 day sums (color shading, white dots) in JJA (a) and DJF (b), North Rhine. Confidence intervals of 63% and 95%, approximately corresponding to 1σ and 2σ standard deviations, are shown by the vertical extent of the battons and shaded areas respectively. The distribution includes dry events.

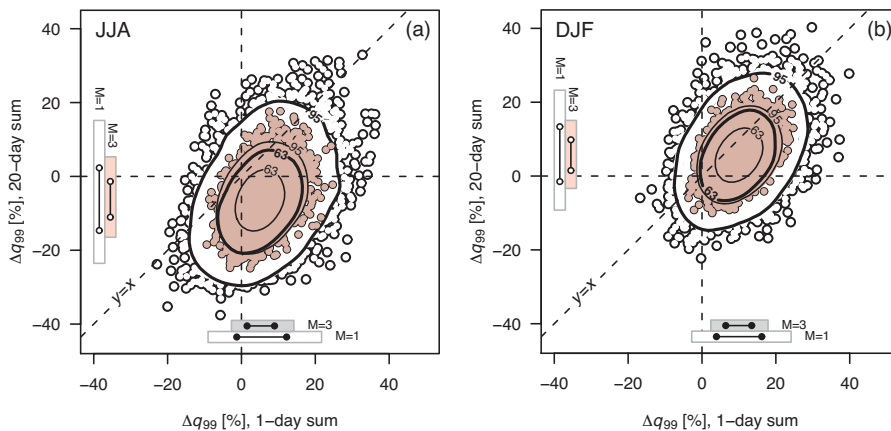


Fig. 7. Density plots for JJA (a) and DJF (b) showing the relationship between the 1-day sum- and 20-day sum-values of Δq_{99} found in the bootstrapped samples of Fig. 6 for ensembles with $M=1$ (white points) and $M=3$ (color points) members. Density contours enclosing 63% and 95% of the data cloud are drawn (thick for $M=1$, thin for $M=3$).

They illustrate the magnitude of uncertainty associated with 30 yr of input of large scale boundary conditions to hydrological models.

The size of ensemble necessary to distinguish an externally forced signal depends on the strength of the signal as well as the magnitude of internal variability. Towards the south of the basin, smaller ensembles are sufficient to distinguish the multiday response (Δq_{99}) from zero, as the signal strengthens while internal variability is of the same magnitude on the scale of the basin (Figs. S7, S8, Supplement).

6 Summary and discussion

For the first time, the relative importance of natural variability in precipitation extremes and the signal due to climate change was studied systematically in a very large, 17-member GCM ensemble of one global climate model. We

focused specifically on future changes in the upper quantiles of multiday precipitation and their dependence on the accumulation interval, on the scale of the Rhine basin.

The dependence of extremes on the accumulation interval is limited to the summer season and is strongest in the North of the basin, where one-day sum extremes increase by around 6% and 20-day sums decrease by a similar degree. This result has implications for the delta-change downscaling technique. In its simplest form, the delta-change method applies a single multiplication factor (consistent with mean changes in climate parameters) to transform a historical time series into a future scenario for input into hydrological models. Such an approach would result in a change of the same sign for both single and multiday precipitation quantiles, so would not be capable of reproducing the results here. A more complex transformation is required; certainly one which first takes the change in wet-day frequency into account (e.g. Van den Hurk et al., 2007), and at best in a highly controlled manner (akin

to some bias correction methods, e.g. Te Linde et al., 2010). The quantile scaling technique (Shabalova et al., 2003; Leander and Buishand, 2006) using an exponential in place of a linear transformation to adjust the intensity of the remaining wet-day amounts can be used to achieve a more appropriate future variance. Even with these adjustments, there is no guarantee that the constructed future time series will include an appropriate, uncertainty-spanning range of changes in the sequences of events, e.g. long dry periods followed by intense rain. These can only be captured and assessed by a realistic handling/modeling of the changes in large-scale circulation regimes and surface-atmosphere feedbacks.

On the other hand, in winter, relative changes of the quantiles are positive and are modulated mainly by increased intensities. The simple delta-change technique could be adequate for modeling basin-scale changes to the winter precipitation. Ensemble mean wet-day frequencies and the distribution of wet and dry period durations remain basically unchanged. Note that this is the case despite a mean circulation change in ESSENCE. The ensemble mean shows an increase in westerly flow during winter (and an increase easterly flow in summer) over the period of the ESSENCE integrations (not shown here). Also, the majority of GCMs, including ECHAM5/MPI-OM from which ESSENCE is derived, show an increase in westerly flow in winter (Van Ulden and van Oldenborgh, 2006) in the vicinity of the Rhine basin. Thus for the model and emission scenario used here, the circulation change that does occur does not impact the wet event frequency or duration much, although within individual transient realisations, circulation and precipitation extremes may be linked within the natural climate variability.

The availability of a large ensemble permitted the dependence of uncertainty due to sample size to be estimated for a range of ensemble sizes. For the North Rhine, it was seen that, for the model and scenario used, on the order of 8 ensemble members (240 yr of integration per time slice) or more were needed to distinguish the climate change signal in extremes of multiday precipitation sums from natural climate variations and their dependence on accumulation period. The length of the integrations required depend on the local signal strength relative to the local background variability and on the time of year.

Finally, we would like to emphasize two limitations of our study. Firstly, the coarse resolution of the GCM could be a limitation, in particular when considering the smaller scale extremes occurring in summer. Commonly, downscaling with RCMs (or statistical methods) are employed to provide high resolution information to discharge models. However, it is well known that, on the larger scale, the climate change signal in RCM downscaling is largely determined by the response in the GCM climate scenario integration (e.g. Déqué et al., 2007). Secondly, this study has been performed with only one GCM using only one emission scenario. While this enabled us to neglect uncertainty due to the GCM model formulation and future emissions, and therefore easily separate

the climate change signal from natural variability, this is also obviously a limitation. For instance, the mean response in precipitation is rather low in ESSENCE (North Rhine DJF: +8.5%, JJA: -23.7%) in comparison to other GCM A1b integrations (e.g. Lenderink et al., 2007b, Fig. 1), and therefore the role of natural variability is, in relative terms, large. Thus, results with a different GCM, other greenhouse gas emissions, or other time periods could be different. However, we are convinced that qualitatively our results are robust, and contain a clear warning that natural variability is an important part of the response in (multiday) precipitation extremes seen in global and regional climate model simulations. Therefore, current estimates of discharge changes, which are often based on relatively short periods of 30 yr, could be subject to inadequate sampling of large-scale variability and should be treated with caution.

Supplementary material related to this article is available online at:

<http://www.hydrol-earth-syst-sci.net/15/1157/2011/hess-15-1157-2011-supplement.pdf>.

Acknowledgements. The ESSENCE project, lead by Wilco Hazeleger (KNMI) and Henk Dijkstra (UU / IMAU), was carried out with support of DEISA, HLRS, SARA and NCF (through NCF projects NRG-2006.06, CAVE-06-023 and SG-06-267). We thank the DEISA Consortium (co-funded by the EU, FP6 projects 508830/031513) for support within the DEISA Extreme Computing Initiative (www.deisa.org). The authors thank Andreas Sterl (KNMI), Camiel Severijns (KNMI), and HLRS and SARA staff for technical support. We acknowledge the international Commission for the Hydrology of the Rhine basin (CHR) for providing the CHR precipitation data through KNMI. Thanks to Boris Orlovsky for helpful comments on the manuscript.

Edited by: S. Attinger

References

- Brandefelt, J. and Körnich, H.: Northern Hemisphere Stationary Waves in Future Climate Projections, *J. Climate*, 21, 6341–6353, 2008.
- Christensen, J. H., Hewitson, B., Busuioc, A., Chen, A., Gao, A., Held, I., Jones, R., Kolli, R. K., Kwon, W.-T., Laprise, R., Magaña Rueda, V., Mearns, L., Menéndez, C. G., Räisänen, J., Rinke, A., Sarr, A., and Whetton, P.: Regional Climate Projections, in: *Climate Change 2007: The Physical Science Basis, Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by: Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K. B., Tignor, M. and Miller, H. L., Cambridge University Press, Cambridge, UK and New York, NY, USA, 2007.
- Dankers, R., Christensen, O. B., Feyen, L., Kalas, M., and de Roo, A.: Evaluation of very high-resolution climate model data for simulating flood hazards in the Upper Danube Basin, *J. Hydrol.*, 347, 319–331, 2007.

- Déqué, M., Rowell, D. P., Lüthi, D., Giorgi, F., Christensen, J. H., Rockel, B., Jacob, D., Kjellström, E., de Castro, M., and van den Hurk, B.: An intercomparison of regional climate simulations for Europe: assessing uncertainties in model projections, *Climatic Change*, 81, 53–71, 2007.
- Frei, C., Davies, H. C., Gurtz, J., and Schär, C.: Climate dynamics and extreme precipitation and flood events in Central Europe, *Integrated Assessment*, 1, 281–299, 2000.
- Gillett, N., Graf, H., and Osborn, T.: Climatic change and the North Atlantic Oscillation, *Geophysical Monograph*, 134, 193–209, 2003.
- Hu, Z. and Wu, Z.: The intensification and shift of the annual North Atlantic Oscillation in a global warming scenario simulation, *Tellus A*, 56, 112–124, 2004.
- Hurkmans, R., Terink, W., Uijlenhoet, R., Torfs, P., Jacob, D., and Troch, P. A.: Changes in streamflow dynamics in the Rhine basin under three high-resolution regional climate scenarios, *J. Climate*, 23, 679–698, 2010.
- Kay, A. L., Jones, R. G., and Reynard, N. S.: RCM rainfall for UK flood frequency estimation, II. Climate change results, *J. Hydrol.*, 318, 163–172, 2006.
- Leander, R. and Buishand, T. A.: Resampling of regional climate model output for the simulation of extreme river flows, *J. Hydrol.*, 332, 487–496, 2006.
- Lenderink, G., Buishand, A., and van Deursen, W.: Estimates of future discharges of the river Rhine using two scenario methodologies: direct versus delta approach, *Hydrol. Earth Syst. Sci.*, 11, 1145–1159, doi:10.5194/hess-11-1145-2007, 2007a.
- Lenderink, G., van Ulden, A., van den Hurk, B., and Keller, F.: A study on combining global and regional climate model results for generating climate scenarios of temperature and precipitation for the Netherlands, *Clim. Dynam.*, 29, 157–176, doi:10.1007/s00382-007-0227-z, 2007b.
- Lenderink, G., van Meijgaard, E., and Selten, F. M.: Intense coastal rainfall in the Netherlands in response to high sea surface temperatures: analysis of the event of August 2006 from the perspective of a changing climate, *Clim. Dynam.*, 32, 19–33, 2009.
- Meehl, G. A., Covey, C., Delworth, T., Latif, M., McAvaney, B., Mitchell, J. F. B., Stouffer, R. J., and Taylor, K. E.: The WCRP CMIP3 multi-model dataset: A new era in climate change research, *B. Am. Meteorol. Soc.*, 88, 1383–1394, 2007.
- Nakićenović, N., Alcamo, J., Davis, G., de Vries, B., Fenhann, J., Gaffin, S., Gregory, K., Grubler, A., Jung, T. Y., Kram, T., La Rovere, E. L., Michaelis, L., Mori, S., Morita, T., Pepper, W., Pitcher, H. M., Price, L., Riahi, K., Roehrl, A., Rogner, H.-H., Sankovski, A., Schlesinger, M., Shukla, P., Smith, S. J., Swart, R., van Rooijen, S., Victor, N., and Dadi, Z.: Special Report on Emissions Scenarios: A Special Report of Working Group III of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, UK, 599 pp., 2000.
- Pinto, J. G., Ulbrich, U., Leckebusch, G. S., Spanghel, T., Reyers, M., and Zacharias, S.: Changes in storm track and cyclone activity in three SRES ensemble experiments with the ECHAM5/MPI-OM1 GCM, *Clim. Dynam.*, 29, 195–210, 2007.
- Selten, F. M., Branstator, G. W., Dijkstra, H. A., and Kluiphuis, M.: Tropical origins for recent and future Northern Hemisphere climate change, *Geophys. Res. Lett.*, 31, L21205, doi:10.1029/2004GL020739, 2004.
- Shabalova, M. V., van Deursen, W. P. A., and Buishand, T. A.: Assessing future discharge of the river Rhine using regional climate model integrations and a hydrological model, *Clim. Res.*, 23, 233–246, 2003.
- Sprokkereef, E.: A hydrological database for the Rhine basin/Eine hydrologische Datenbank für das Rheingebiet, Technical report, International Commission for the Hydrology of the Rhine Basin, Arnhem, The Netherlands, 2001 (in German).
- Sterl, A., Severijns, C., Dijkstra, H., Hazeleger, W., van Oldenborgh, G. J., van den Broeke, M., Burgers, G., van den Hurk, B., van Leeuwen, P. J., and van Velthoven, P.: When can we expect extremely high surface temperatures?, *Geophys. Res. Lett.*, 35, L14703, doi:10.1029/2008GL034071, 2008.
- Te Linde, A. H., Aerts, J. C. H., Bakker, A. M. R., and Kwadijk, J. C. J.: Simulating low-probability peak discharges for the Rhine basin using resampled climate modeling data, *Water Resour. Res.*, 46, W03512, doi:10.1029/2009WR007707, 2010.
- Terink, W., Hurkmans, R. T. W. L., Torfs, P. J. J. F., and Uijlenhoet, R.: Evaluation of a bias correction method applied to downscaled precipitation and temperature reanalysis data for the Rhine basin, *Hydrol. Earth Syst. Sci.*, 14, 687–703, doi:10.5194/hess-14-687-2010, 2010.
- Trenberth, K. E.: Conceptual framework for changes of extremes of the hydrological cycle with climate change, *Climatic Change*, 42, 327–339, 1999.
- Ulbrich, U. and Christoph, M.: A shift of the NAO and increasing storm track activity over Europe due to anthropogenic greenhouse gas forcing, *Clim. Dynam.*, 15, 551–559, 1999.
- Van den Hurk, B., Tank, A. K., Lenderink, G., Ulden, A., van Oldenborgh, G. J., Katsman, C., Brink, H., Keller, F., Bessembinder, J., Burgers, G., Komen, G., Hazeleger, W., and Drijfhout, S.: New climate change scenarios for the Netherlands, *Water Sci. Technol.*, 56, 27–33, 2007.
- Van Ulden, A. P. and van Oldenborgh, G. J.: Large-scale atmospheric circulation biases and changes in global climate model simulations and their importance for climate change in Central Europe, *Atmos. Chem. Phys.*, 6, 863–881, doi:10.5194/acp-6-863-2006, 2006.
- Yin, J. H.: A consistent poleward shift of the storm tracks in simulations of 21st century climate, *Geophys. Res. Lett.*, 32, L18701, doi:10.1029-2005GL023684, 2005.