

Interactive comment on “Robust assessment of future changes in extreme precipitation over the Rhine basin using a GCM” by S. F. Kew et al.

S. F. Kew et al.

sarah.kew@knmi.nl

Received and published: 6 March 2011

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

Final response to anonymous referees 1 and 2

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

C5201

We wish to thank both reviewers for their comments which have highlighted parts of our manuscript requiring further clarification. In our response below, all page and line numbers refer to the HESSD document.

Response to reviewer 1

- *Discuss limitations of approach, given that it is based on 1 GCM and 1 emission scenario only.*

The final paragraph of the manuscript has been extended to discuss the above limitations and others:

‘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

C5202

global and regional climate model simulations. Therefore, current estimates of discharge changes, which are often based on relatively short periods of 30 years, could be subject to inadequate sampling of large-scale variability and should be treated with caution.'

- P9044, L10: *Confusing use of term 'scaling' or 'simple scaling'.*

Both reviewers have mentioned that our use of the term 'scaling' is confusing. We agree to substitute it with 'relative percentage change in ... with respect to the control period' or similar.

We use the term 'scale' or 'scaling' now only in conjunction with spatial scales (e.g. at the scale of the Rhine basin) or when discussing the delta-change statistical downscaling technique.

All further points raised by reviewer 1 are addressed in the author comment AC C4740 in the interactive discussion. In the supplement to this final comment we include 2 figures to which we refer in point 4 of AC C4740. Fig. S1 is the equivalent of Fig. 2 with the east-most cell of the North Rhine region removed from the ESSENCE data, and Fig. S2 is the equivalent of Fig. A1 with the west-most cell of the Central Rhine region removed from the ESSENCE data.

Response to reviewer 2

- *A key question that this kind of study should address is: Are the RCMs in re-analysis mode able to reproduce extremes observed in the past? If not, how can be justified that these models are able to predict extremes in the future? It is also argued in the summary that this ensemble with 17 members is large enough to capture the climate signal, specially the extremes. This assumption is quite strong in my opinion. I suggest to quantify extreme statistics from the*
C5203

interpolated data and the RCM results and to check whether they are able to reproduce statistics such as maximum monthly precipitation, frequency of wet and dry spells, number of days per year with precipitation large a threshold, etc. Otherwise it is non-scientific to accept this hypothesis. Without these empirical evidence it is difficult to accept the conclusions of this study.

It is indeed a very important question how RCMs (driven by analyzed boundaries) simulate extremes in the past. In particular, this is important when driving discharge models with the output from the RCM simulations. Yet, this study is not about RCM simulations. It is about GCM simulations, and it aims to answer a different question. The effective scientific question asked is 'What is relative importance of climate change versus natural variability', and this question is answered using one large ensemble of only one GCM. We mention the role of RCMs in conjunction with GCMs twice: firstly in the introduction (P9044, L19-23) and secondly in the discussion (P9056 L4-6). In both instances we refer to RCMs in general and not to results emanating from our study.

We made the following modifications in the introduction and discussion to make it clearer that we are dealing with multiple integrations from one GCM and that the downscaling of such a GCM ensemble with an RCM is not available:

P9045, L13-16. '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.'

P9054, L23-26. '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.'

We have also added a reference to a rephrased version of P9056, L4–6, to emphasize that the RCM results mentioned refer to the literature and not our study:

'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).'

The 17-member ensemble is indeed assumed large enough to capture the *simulated* climate signal. This assumption is statistically verified using the simulated data (e.g. Fig. 6 shows that with 17 members the relative change is significantly different from zero, but with 1 member this is not the case). Whether or not the simulated signal is the true climate signal is a different question which does not yet have a good answer. Reproducing the observed climate statistics with a climate model is not a guarantee that the change in statistics is modeled realistically and, vice-versa, errors in reproducing the current statistics do not preclude that the model is not capable of simulating a realistic response (relative changes).

Nevertheless, in response to the reviewer's request for further verification statistics, we provide in a supplement (for the archive) the wet and dry spell probability density functions obtained from the upscaled observations for comparison with the GCM data from the control and future periods (Fig. S.3). Providing these statistics for an RCM is beyond the scope of this paper. Information regarding the percentage of days surpassing a certain threshold is contained at least implicitly in Fig. 2 for the 50% and 99% quantiles (as the area under each *wef*-scaled curve to the right of the respective quantile). The 99% quantile is the threshold we focus on in the paper and the most extreme quantile considered. The 99% quantile is approximately the same as the mean seasonal precipitation maximum (1 season is approximately 100 days).

C5205

- *Another issue that should be consider in more detail is the comparison of the coarse RCM output with the interpolated data based on point observations. In my opinion, a statistical downscaling is needed this in this case.*

If we exchange 'RCM' for 'GCM' in this comment, we agree that a downscaling is necessary in order to interpret GCM results at the resolution of station data. Here, however, we are working at the resolution of the GCM, seeking results relevant at the scale of the Rhine basin. Instead of downscaling the GCM, we upscale the observations (not directly from point observations but from area-averages for the HBV sub-basins) to approximately match the resolution of the GCM.

We stated that the coarse resolution of the GCMs is a limitation (P9056 L4-5), implying that we would have used higher resolution global model data if it had been available. If we *were* to dynamically downscale the GCM ensemble, we would expect the large-scale variability we observe in this study to be transmitted into large-scale variability in the RCM simulations and thus to have a significant impact on the variability of the interior dynamics.

- *The term scaling is misleading. I suggest to remove it from the manuscript. This term refers to a percentage increase w.r.t. reference period. Just call this statistic it what it is.*

See the response to reviewer 1 above.

- *Provide more information regarding the CHR-OBS data set. e.g. resolution, time scale, interpolation method, etc. It would be useful for the reader to know how this dataset was developed and by whom.*

Information on the temporal and spatial resolution of CHR-OBS is provided in our HESSD manuscript (P9047, L6-8):

C5206

'The CHR dataset ... provides area-averaged daily precipitation sums for the 134 ... subbasins of the Rhine catchment'.

We also stated in the HESSD manuscript that the dataset was issued (i.e. developed) by the CHR (P9047, L4-5) and cited the (German) technical report by Spokkereef describing the dataset (P9047, L9).

However, we concede that this part of the manuscript could be clearer and have now adapted it (see next point below). In recognition that the Spokkereef reference is in German and thus not ideal for most readers, we have added a further reference to a brief but English description of the development including the interpolation methodology.

- *P9047 L7. Sentence not clear. Is it necessary to know that this data set was used for forcing a HBV model in another study?*

We will amend P9047 L6-9 as follows:

'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 (Spokkereef, 2001) and a brief (English) summary can be found in e.g. Terink et al. (2010).'

By repositioning the Spokkereef reference and changing 'provides' to 'comprises', it should now be clear that our mention of the HBV basins is not a reference to the use of the dataset in another study but rather a description of the spatial structure of the data. It is a small detail which indeed is not critical to the message, but on the other hand indicates that the data used is of a quality suitable for input to hydrological models and could be a helpful pointer for those familiar with the HBV model.

C5207

- *Indicate the bootstrap procedure to estimate the confidence interval of Δq . How are the 30-y time slice found?*

The bootstrap procedure for creating the 30-year timeslice samples is described in the methodology section on P9049, L4-10. It is likely that the reviewer overlooked this description due to our formulation being overly concise. We propose to expand the description as follows:

'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-year 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 year or beyond).'

- *PDFs in fig2 it is indicated that 'the black curve is a fit' does not indicate which theoretical PDFs were used.*

The fit is empirical not theoretical. We will change the relevant part of the caption (P9060) to '... and the black curve is an empirical fit (kernel density estimate using gaussian smoothing) giving the CHR-OBS probability density'.

C5208

- *The vertical lines indicate quantiles in which dry events have been either included or not. Why? Which is the purpose for that?*

We give quite a full explanation on P9048 L7 – P9049 L3. Basically, we need to look at the distribution that includes dry days in order to make a fair comparison of the changes in the quantiles for different n -day sums. The distribution excluding dry days is also presented because it provides insight into the effect of intensity changes on the full-distribution quantiles and because wet-day quantile changes are also found in the literature (e.g. Frei et al., 2006).

- *This also indicate that there are large discrepancies in the PDFs for pre. sums large than 10 days.*

We have explained in the HESSD text why there are discrepancies in the PDFs for the longer summation intervals on P9049 L26 – P9050 L4, describing panels (b) and (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 observations and this causes the multiday PDF to be shifted to higher values.’

In other words, when it rains, the model’s probability distribution is reasonable, but biases in the rainy-day frequency introduce biases in the probability distribution for longer summation periods.

- *Panel a and d also exhibit large completely different allocations of the probability density. Please comment on that taken into account that your are using reanalysis data.*

C5209

We disagree that there are completely different allocations of the probability density in panels (a) and (d). The peaks for the observations and model distributions occur in about the same position (precipitation sum) for the upscaled observations and model. The positions of the wet-distribution quantiles for the two datasets also correspond well, i.e. the distributions show similar spread, which can also be seen by eye. We find the agreement quite remarkable considering the coarse resolution of a GCM.

It is unclear what the reviewer means by ‘re-analysis data’. We do not use re-analysis products in our study, nor do we mention ‘re-analysis’ in the manuscript. The observations we used were subbasin-averages and model runs were not used for interpolation. Our model data here are climate model runs and are not constrained by observations (except at initialisation). In their ‘general comments’ the reviewer says that our ensemble simulations run from 1950-2010. Typing 2010 instead of 2100 is a typing error easily made, but in case of a misunderstanding we point out that our model simulations run until 2100 and thus beyond the period for which observations are available.

- *Fig. 4 should indicate the period for which this analysis was done. Comparison in the validation period is necessary.*

The caption to Fig. 4 does specify the analysis period for the two differently shaded curves to be the control period and future period, both of which are defined in the the main text. We have now added the years explicitly in response to this comment, and also the analysis region (North Rhine), which we noticed was missing. In making this amendment we noticed that we had mistakenly used 1950/1951–1979/1980 instead of 1961/1962-1990/1991 as the reference period. The error was made in Fig. 4 only and has now been corrected, with no consequences for the conclusions drawn.

We provide an extra figure in the supplement to this comment to show the observations against the model data (Fig. S.3). It can be seen that the observations correspond well to the model control period for the JJA dry period duration. The correspondence for

C5210

the JJA wet period durations is not as good but the observed PDF is closer to the control period PDF than to the future period PDF. For DJF wet and dry durations, the model control and future period PDFs do not differ significantly. The observed PDF corresponds well to the model PDFs for wet period duration but is not a good match for the dry durations. It should be borne in mind however that the (17 times shorter) observed series is noisy with respect to the model, we have made no assessment of measurement uncertainty, and that our study focuses on presenting and explaining *relative* quantile changes between the two time slices of the model.

- *Technical corrections*

This manuscript require also the following corrections: 1. Fig2 is very confusing. The double axes make it difficult to read. I suggest to separate dry and wet and indicate in the caption the meaning of (a)...(f).

We emphasize the separation between the left hand column and right hand PDF by dividing the figure box into dry and wet sections. It is now much clearer which vertical axis belongs to which part of the figure. We used 'top row' and 'lower row' and 'left-right' in the figure caption because we think it improves readability above the alternative: 'for JJA (a-c) and DJF (d-f) for 1- (a,d), 10- (b,e) and 20- (c,f) day sums.' This information is given anyhow in the figure axis labels, so there should be no confusion.

Other modifications

We realised that the statement on P9050, L19-21, 'relative changes between the control and future period ... would be unaltered when the same bias correction is applied to control and future periods', did not hold in the case of an additive error. We have adapted the text accordingly.

C5211

The acknowledgements have been extended to include people behind the ESSENCE project and the CHR dataset.

A number of small grammatical edits were made.

References

Déqué, M., Rowell, D. P., Lüthi, D., Giogi, 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., Scholl, R., Fukutome, S., Schmidli, J., and Vidali, P. L.: Future change of precipitation extremes in Europe: Intercomparison of scenarios from regional climate models, *J. Geophys. Res.*, 111, D06105, doi:10.1029/2005JD005965, 2006.

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, 2007.

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, 2010.

Please also note the supplement to this comment:

<http://www.hydrol-earth-syst-sci-discuss.net/7/C5201/2011/hessd-7-C5201-2011-supplement.pdf>

Interactive comment on *Hydrol. Earth Syst. Sci. Discuss.*, 7, 9043, 2010.

C5212