Royal Netherlands
Meteorological Institute
Ministry of Infrastructure and the Environment

Rainfall generator for the Rhine basin: Description of simulations using gridded precipitation datasets and uncertainty analysis

Maurice J. Schmeits, Erwin L.A. Wolters, Jules J. Beersma, and T. Adri Buishand

De Bilt, 2014 | KNMI publication 186-VII

Rainfall generator for the Rhine basin:

Description of simulations using gridded precipitation datasets and uncertainty analysis

Maurice J. Schmeits, Erwin L.A. Wolters, Jules J. Beersma, and T. Adri Buishand

KNMI publication 186-VII

Work performed under commission of the Ministry of Infrastructure and the Environment: Rijkswaterstaat Water, Transport and Environment (RWS WVL).

CONTENTS

1. INTRODUCTION

The rainfall generator has been used to generate long synthetic series of daily precipitation and temperature for the Rhine basin [see e.g. Wójcik et al. (2000), Buishand and Brandsma (2001), and Beersma (2002)] using the nearest-neighbour resampling (NNR) technique. These simulations were driven by daily precipitation and temperature data for the period 1961 – 1995 from 34 stations across the Rhine basin. For hydrological applications, the simulated point precipitation and temperature data were converted to representative values of 134 HBV[1] sub-basins, using a dataset which is nowadays known as CHR-OBS data [see Görgen et al. (2010)]. Recently, two additional gridded datasets (HYRAS and E-OBS) have become available. These two datasets fulfil the need for extended temperature and precipitation record lengths.

This report is set up as follows. First, the three datasets that were used in this research are highlighted in the next subsections. An intercomparison between the three precipitation datasets is presented in section 2, followed by a description of the nearest neighbour resampling technique in section 3. Results for the various simulation types are shown in section 4, after which results from an uncertainty analysis are presented in section 5. Finally conclusions are drawn in section 6.

1.1 CHR-OBS precipitation and temperature dataset

The Commission for the Hydrology of the Rhine basin (CHR) was officially established in 1970, following advice by UNESCO[2] to promote closer co-operation in international river basins. The major task of the CHR is to perform research on the hydrology of the entire Rhine basin and to exchange this information between the countries situated in the Rhine basin. During the last decade, special attention has been given to the effects of future climate change on the hydrology of the Rhine basin (see for example Görgen et al., 2010) and to the quality of discharge data (Steinrücke et al., 2012).

Some relevant elements in the construction of the CHR-OBS dataset are outlined below. Daily precipitation values for each of the HBV sub-basins are obtained from gridded data. For the German part of the Rhine basin the Regionalisierung der Niederschlagshöhen (REGNIE) dataset (at a 1×1 km^2 resolution) is utilized, except for the Moselle basin. REGNIE uses background climatology fields of monthly mean precipitation to calculate ratio anomalies. Through these background fields elevation is taken into account. First, the value of a weather station is assigned to a grid cell, which is then divided by the background field value at that grid cell. These ratios are then interpolated using inverse squared distance weighting. Subsequently, the result is multiplied by the background field to obtain precipitation values (Steiner, 2009).

For the Moselle basin gridded data (7×7 km^2) generated by the University of Trier are used (de Wit and Buishand, 2007) and for the Swiss basin, the precipitation data are gridded at a 2×2 km^2 resolution (Dällenbach, 2000).

The daily temperature values at 49 stations have been transformed to areal values for the sub-basins using the HBV modelling software (Eberle et al., 2005). For each

[1] HBV: Hydrologiska Byråns Vattenbalansavdelning, a commonly-used hydrological model, described in detail by Lindström *et al.* (1997).
[2] UNESCO: United Nations Educational, Scientific, and Cultural Organization

sub-basin there were user defined input stations and station weights as well as an altitude correction of 6 ˚C km^{-1} to the mean elevation of the sub-basin as derived from the elevation zones in the HBV model.

1.2 HYRAS precipitation dataset

The HYRAS precipitation dataset has been developed by the German Weather Service (DWD) and at present comprises about 6000 rain gauge stations in The Netherlands, Germany, Luxembourg, north-eastern France, northern Switzerland, Austria, and the Czech Republic. Gridded daily rainfall observations are available on a daily basis, with a spatial resolution of 5×5 km^2. The station data are interpolated using the REGNIE method, which has a nominal resolution of about 1x1 km^2 and is subsequently aggregated to 5×5 km^2.

Figure 1.1 shows the geographical distribution of the used rain gauge stations for HYRAS in 1951 and 1991 (versions 1 and 2 of the dataset). During the period 1951 – 1970 the number of rain gauge stations increased from almost 4000 to almost 6000 (Willems and Stricker, 2011). From then the number of stations remained more or less constant until 2000 after which there is a decline to about 4500 stations in 2006 (Figure 1 in Rauthe et al, 2013). Currently, precipitation data is available for the period January 1951 – December 2006. At the time of writing, it is unclear to what extent the dataset will be updated with more recent observations.

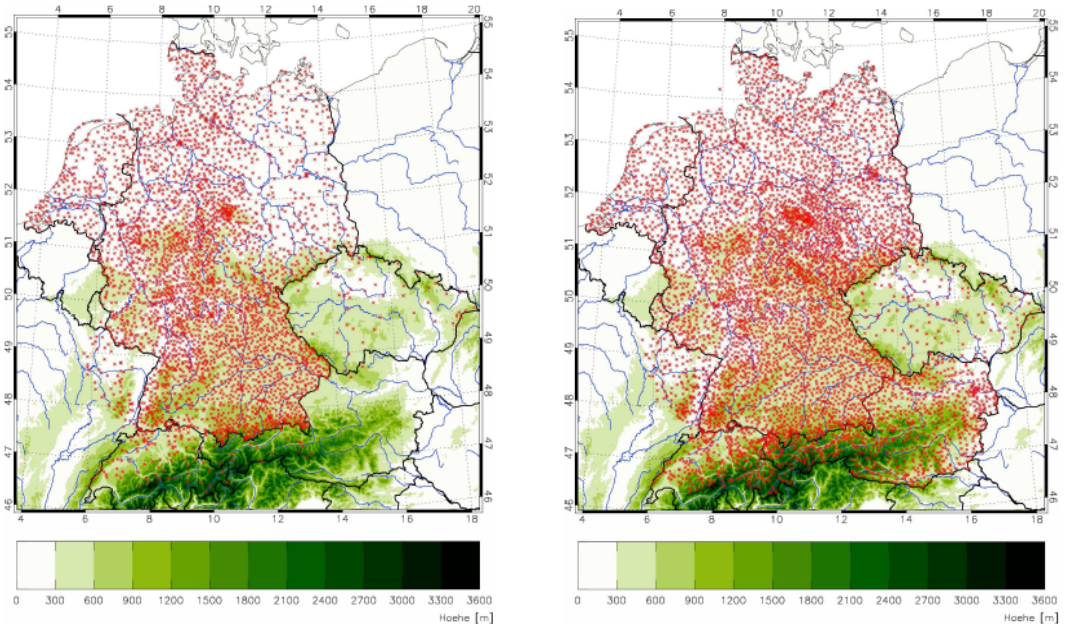


Figure 1.1: *Geographical distribution of HYRAS rain gauge stations in 1951 (left panel) and 1991 (right panel).*

1.3 E-OBS dataset

The European Daily High-Resolution Observational Gridded Dataset (E-OBS) of rain gauge and pluviograph observations is available on various regular spatial grids (in this research a 0.22°×0.22° rotated pole grid was used) and on a daily resolution. The daily observations at point locations are taken from the European Climate Assessment and Dataset (ECA&D; http://www.ecad.eu). At present, the number of

precipitation stations in the Rhine basin upstream of Lobith, where the Rhine enters the Netherlands, is 467 in the E-OBS data set, of which 421 are located in the German part of the basin and the remainder in the Luxembourgian, French, and Swiss parts. The station density in the French and Swiss parts of the Rhine basin is low. E-OBS provides land-only information on precipitation amounts and minimum, maximum, and mean surface temperatures over Europe for the period 1950–2012. Geographical maps with the number of precipitation stations in 1951 and 1991 are shown in Figure 1.2. The dataset improves on previous products for Europe in its spatial resolution, number of contributing stations, and length of records. Research has been done to find the most appropriate method for spatial interpolation of daily climate observations (Haylock et al. 2008).

The dataset has been designed to provide estimates of grid box averages rather than point values to enable direct comparisons with regional climate model output, in particular for the distribution of daily precipitation. Therefore the observations were first interpolated to a 0.1°×0.1° rotated master pole grid and subsequently averaged over the final grid cells. The interpolation process is employed in three steps. First, the monthly precipitation totals and monthly mean temperatures are interpolated using trivariate thin-plate splines, with latitude, longitude, and elevation as the independent variables. Second, the daily anomalies (the difference between the daily observations and the monthly mean and the ratio of the daily sums to the monthly total for temperature and precipitation, respectively) are interpolated using indicator and universal kriging for precipitation and kriging with an external drift for temperature. Finally, the monthly and daily interpolations are combined. Interpolation uncertainty is quantified by the provision of daily standard errors for every grid square (Haylock et al, 2008).

The E-OBS dataset is updated and incremented in a semi-yearly cycle, with version 7 being available from September 2012 onwards. Note that for the analyses described in section 2 of this report both version 5 (September 2011) and 7 were used, as the latter dataset was not yet available at the start of this study. Version 5 consists of daily precipitation and temperature data from January 1950 to June 2011, and in version 7 this period is extended to June 2012. For the Luxembourgian sub-basins, it appeared that for a number of stations in E-OBS (version 5) the precipitation at day *t* was erroneously attributed to day *t*-1 during the period 1954 – 2003. These errors have been corrected for in the last update of E-OBS (version 7), available since September 2012. Therefore, all simulations in section 4 are based on version 7.

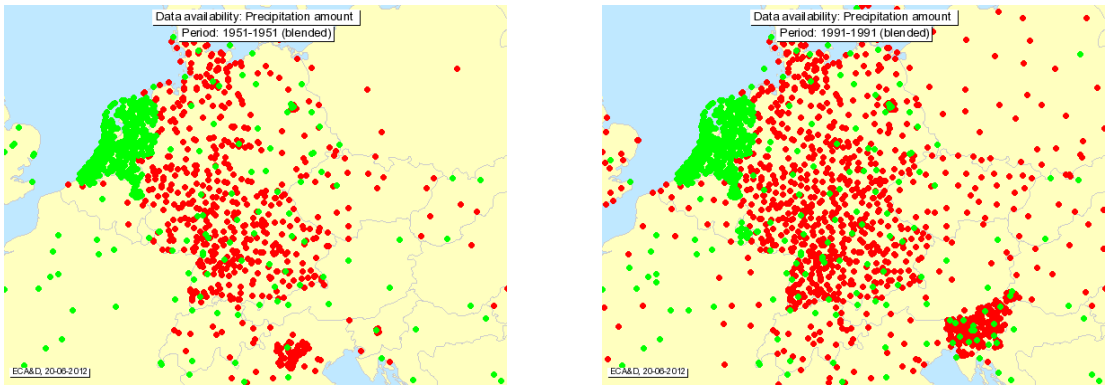


Figure 1.2: *Geographical distribution of E-OBS precipitation stations in 1951 (left panel) and 1991 (right panel). The green colour-coded stations indicate the stations of which the data are publicly available, whereas the data of those in red are not.*

2 COMPARISONS BETWEEN E-OBS, HYRAS, AND CHR-OBS PRECIPITATION DATA

In order to assess the usefulness of the three rainfall datasets for NNR, a comparison was performed. Extreme Rhine discharges in the Netherlands are mainly influenced by precipitation that accumulates over several days rather than by extreme single-day precipitation. In addition to results of the daily means, results are therefore also shown for the maximum 10-day precipitation. The comparison between the three datasets is presented for their common periods, i.e., 1961 – 1995 for the comparison of E-OBS and HYRAS with CHR-OBS and 1951 – 2006 for the comparison of E-OBS with HYRAS. In addition, the daily gridded precipitation data were converted to daily average precipitation amounts for the 134 HBV sub-basins[3]. These sub-basin data were also used in the feature vector calculations, which will be discussed in more detail in section 3.

Figure 2.1 shows the daily averaged and maximum 10-day precipitation over the entire Rhine basin for the winter (October – March) and summer (April – September) halves of the year. The average relative differences are also given in the figure panels. The figure clearly shows that the three datasets are in close agreement, with relative differences between E-OBS (version 5), HYRAS, and CHR-OBS generally within ±5%. It is noted that all three datasets show an increasing trend for the winter precipitation, while for the summer half-year no distinct trend is seen. The increase in winter precipitation is in accordance with the increase in winter precipitation during the 20th century throughout Europe (Moberg et al., 2006) and for parts of the Rhine basin [see e.g., Quirmbach et al. (2012), Hundecha and Bárdossy (2005), and Schmidli and Frei (2005)].

Figure 2.2 presents the same precipitation characteristics as Figure 2.1, but exclusively for the Swiss part of the Rhine basin. Agreement is not as close as for the entire Rhine basin, with differences between the three datasets of 5-10% in the winter half-year. There is no longer a clear positive trend in the HYRAS data, in particular for the 10-day maxima. In the summer half-year, HYRAS shows about 7% larger values than E-OBS (version 5), while CHR-OBS values are about 15% larger than E-OBS. An additional analysis (as shown in Figure 2.3) revealed that the larger values of CHR-OBS compared to HYRAS and E-OBS (version 5) are mainly explained by a much larger precipitation over high-elevation sub-basins (>1000 m mean elevation), with differences of 17-19% between CHR-OBS and HYRAS, and up to 27-29% between CHR-OBS and E-OBS during the summer half-year. Further assessment of the possible causes of these differences over the elevated areas is beyond the scope of this report, but these differences could be partly related to differences in the smoothness of the monthly background fields used for interpolation of the daily values. In the HYRAS dataset, for instance, this background field was obtained from a regression of monthly mean precipitation on 5 geographical variables (Rauthe et al, 2013). The results from the regression were interpolated between stations, which may result in too smooth fields, in particular in mountainous areas. The values for the high-elevation sub-basins from the E-OBS data may be seriously biased because of the low density of the stations in Switzerland used for this dataset.

[3] The gridded daily average temperature data from E-OBS was also converted to daily average temperatures for the 134 HBV sub-basins for use in the feature vector calculations in Section 3.

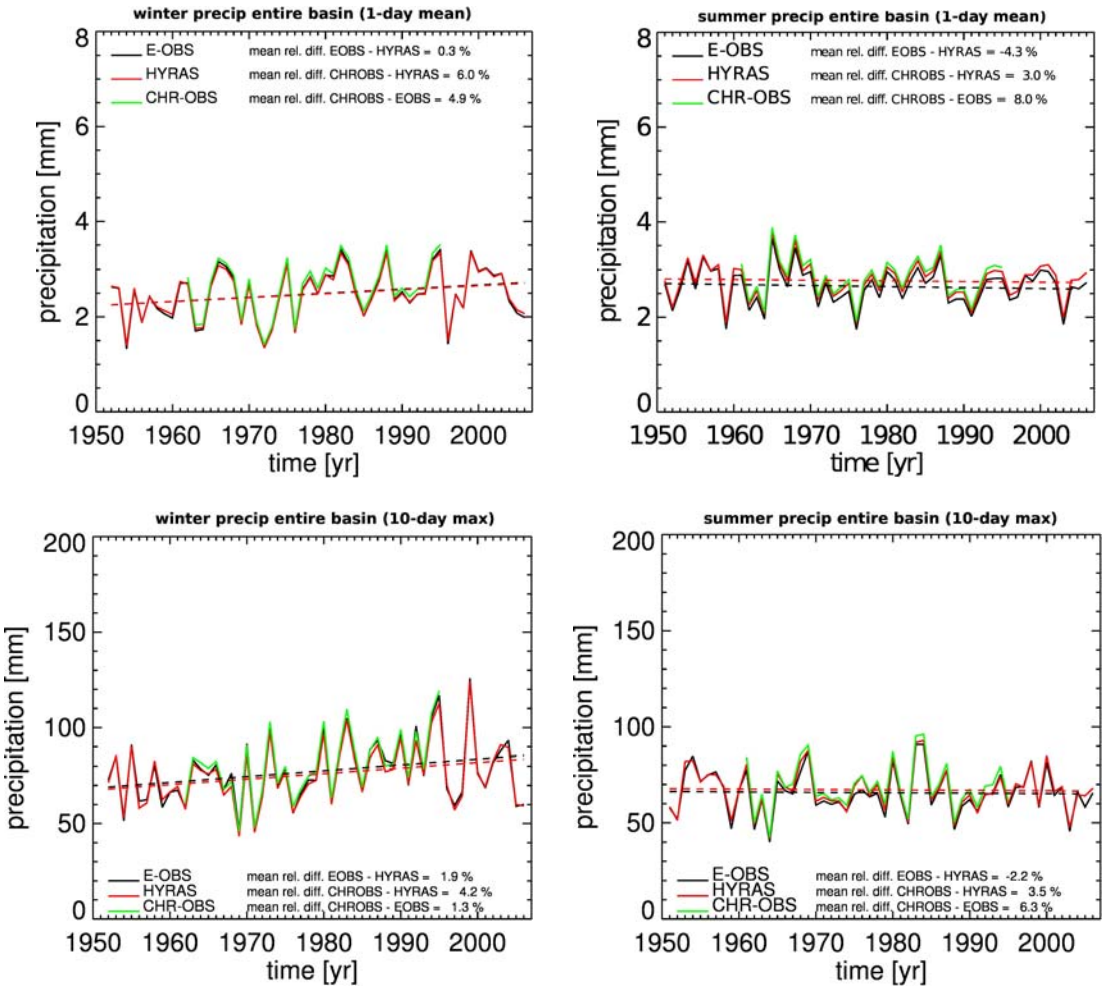


Figure 2.1: *Comparison between E-OBS (version 5), HYRAS, and CHR-OBS precipitation data for the winter (left panels) and summer (right panels) halves of the year over the entire Rhine basin. Results are presented for averaged 1-day precipitation (upper panels) and maximum 10-day precipitation (lower panels). Dashed lines represent linear trends based on E-OBS (black) and HYRAS (red) for the period 1951 – 2006. Note that CHR-OBS data are only available for 1961 – 1995, therefore relative differences involving CHR-OBS were calculated over this period. For the E-OBS – HYRAS comparison, the period 1951 – 2006 was considered.*

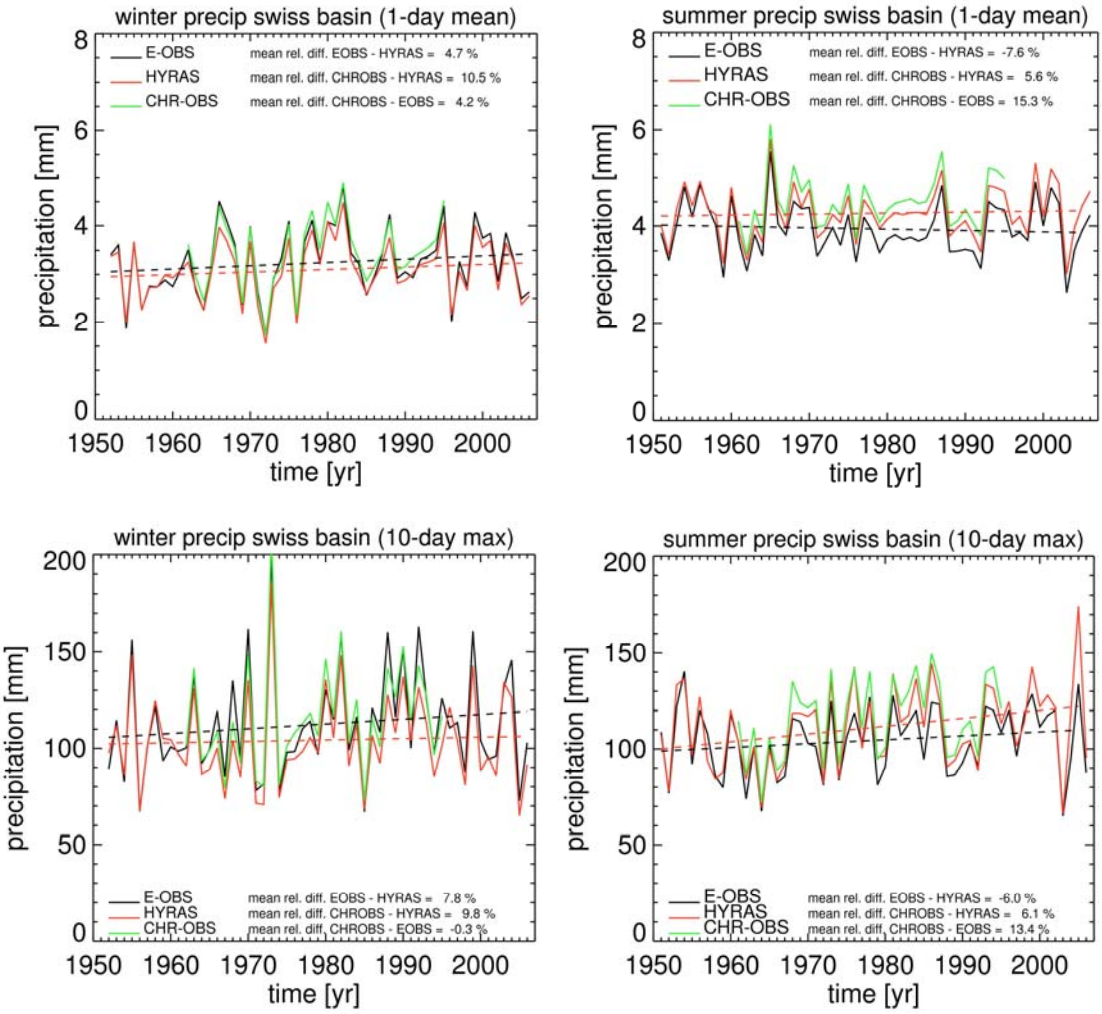


Figure 2.2: *Same as Figure 2.1, but only for the Swiss part of the Rhine basin.*

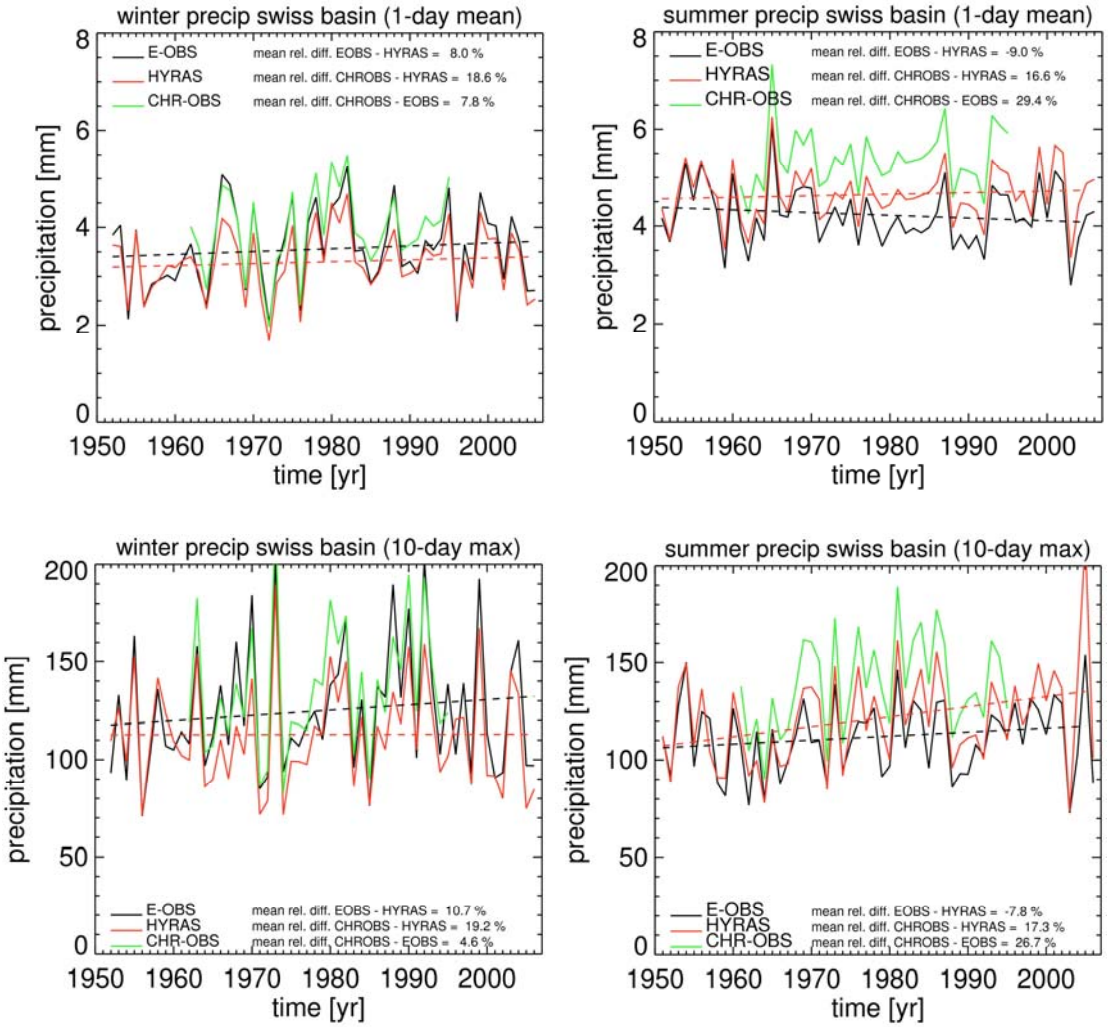


Figure 2.3: *Same as Figure 2.1, but only for the Swiss sub-basins with an average elevation > 1000 m.*

In order to further investigate the three precipitation datasets at smaller spatial scales, the correlation coefficients between their daily precipitation amounts were calculated for the period 1961 – 1995. The left panel of Figure 2.4 shows for each sub-basin and each combination of two data sets the average correlation coefficients. Note that the correlation coefficients were calculated over the whole period without making a distinction between seasons. It can easily be seen in the figure that for almost all sub-basins the correlation between the three datasets is high (>0.9). However, for some sub-basins (50, 51, 55, 107, 108, 109, and 112) correlation coefficients as low as 0.6 – 0.7 are found. These sub-basins are located in Luxembourg (50, 51, 55) and in the south-western part of the Upper Rhine area (107, 108, 109, 112). The problematic areas in the Upper Rhine sub-basins were already identified within the RheinBlick2050 project, with differences between HYRAS and CHR-OBS precipitation of 20 – 30% for the period 1961 – 1990. These differences have been related to the low station density in this area in the CHR-OBS data (Görgen et al., 2010). The low correlation coefficients for the Luxembourgian sub-basins can be explained by the incorrect attribution of E-OBS (version 5) precipitation amounts to the previous day, as mentioned already in section 1.3. These errors have been corrected for in the last update of E-OBS (version 7). Indeed, the right panel of Figure 2.4 shows that the mean correlation of daily precipitation in E-OBS (version 7) and HYRAS is greater than 0.8 for all sub-basins.

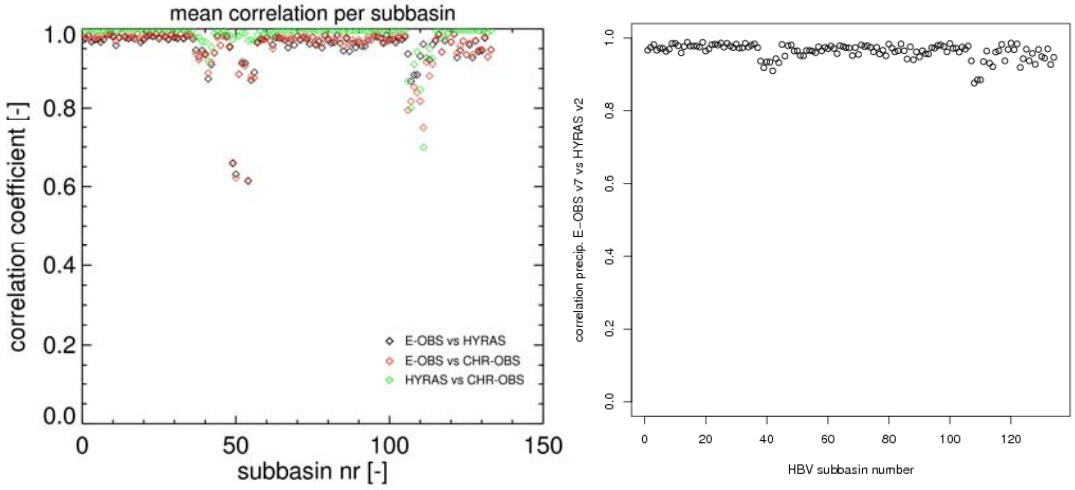


Figure 2.4: *Left panel: Mean correlation coefficients of daily precipitation in E-OBS (version 5) and HYRAS (black), E-OBS (version 5) and CHR-OBS (red), and HYRAS and CHR-OBS (green) per HBV sub-basin for 1961 – 1995. Right panel: Mean correlation coefficients of daily precipitation in E-OBS (version 7) and HYRAS per HBV sub-basin for 1951 – 2006.*

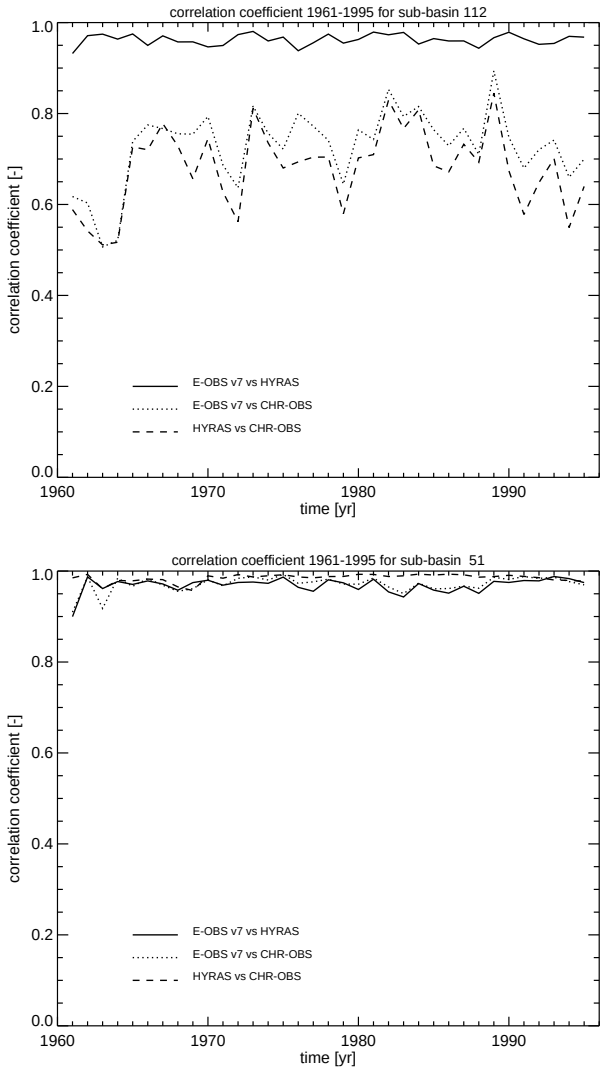


Figure 2.5: *Annual correlation coefficients for the period 1961 – 1995 for HBV sub-basins 112 (Kanal) and 51 (Sûre).*

Examples of the year-to-year variation of the correlation coefficients during the period 1961 – 1995 are shown in Figure 2.5 for HBV sub-basins 112 (Kanal, HBV district Upper Rhine area) and 51 (Sûre, HBV district Moselle). In sub-basin 112, the correlation between E-OBS (version 7) and HYRAS is high (>0.9), while the correlation coefficients in which CHR-OBS is involved fluctuate between 0.5 and 0.8 during the investigated period. This result confirms the limitations of the CHR-OBS data in this area as expressed in Görgen et al. (2010) and suggests that the lack of representativity of the CHR-OBS precipitation data not only relates to the mean amounts. For sub-basin 51 all annual correlation coefficients are high (> 0.9) over the entire period, including those for E-OBS (version 7), because of the correction for the incorrect attribution of E-OBS (version 5) precipitation amounts to the previous day in Luxembourg.

3. SIMULATION SETUP

3.1 General

With the NNR technique, weather variables are resampled simultaneously from the historical data. To incorporate autocorrelation, the resampling depends on the values of the previous resampled day. Therefore the days in the historical record that are most similar to those of the previously simulated day in terms of precipitation and temperature are pre-selected. One of these *k* nearest neighbours is randomly selected and the observed values for the day subsequent to that nearest neighbour are adopted as the simulated values for the next day. In the random selection from the *k* nearest neighbours, a decreasing kernel is used to give more weight to the closest neighbours. In line with Beersma (2002), *k* is set to 10.

A feature vector is used to find the nearest neighbours in the historical record. In the rainfall generator for the Rhine basin the feature vector was composed of standardized daily temperature and precipitation data (see next paragraph) that was averaged over the 134 HBV sub-basins. In addition, the areal precipitation fraction, which is the fraction of sub-basins with daily precipitation larger than 0.3 mm, is included. The frequency distributions of the areal precipitation fraction for the E-OBS and HYRAS data are shown in Figure 3.1, left panel. In both E-OBS and HYRAS, a clear bimodal distribution is seen, with values in the lowest (highest) bins of 24% (19%) and 22% (21%) for E-OBS and HYRAS, respectively. Closer examination revealed that for both E-OBS and HYRAS about 12% of the days have an areal precipitation fraction equal to one, i.e., once every 8-9 days the entire Rhine basin has daily precipitation amount larger than 0.3 mm.

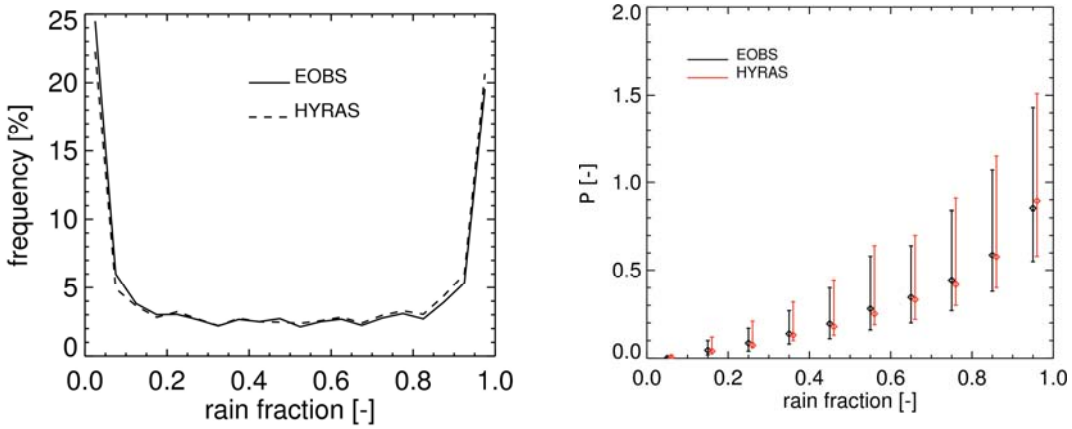


Figure 3.1: *Left panel: Frequency distribution of areal precipitation fraction (i.e., the fraction of the HBV sub-basins with precipitation > 0.3 mm) for E-OBS (solid) and HYRAS (dashed). Right panel: Average standardized precipitation P as a function of areal precipitation fraction for E-OBS (black) and HYRAS (red). Error bars span the range between the 17th and 84th percentiles, median values are denoted by the diamond symbols.*

The effect of seasonal variation is reduced by restricting the search for nearest neighbours to days within a moving window of 61 days centered at the last simulated day (Beersma, 2002; Wójcik et al., 2000). The daily temperatures are standardized

by subtracting the calendar-day mean and dividing by the calendar-day standard deviation before resampling. Daily precipitation is standardized by dividing by the mean wet-day precipitation amount for that calendar day[4]. The standardized precipitation averaged over the entire Rhine basin as a function of the areal precipitation fraction for E-OBS and HYRAS is shown in the right panel of Figure 3.1. Obviously, both the median standardized precipitation and the lower and higher extreme percentiles increase with increasing areal precipitation fraction. Note that temperature data is not yet available in the HYRAS dataset; hence the standardized temperature in the feature vector is currently based on E-OBS data. At the end of the simulation procedure, the resampled standardized variables are transformed back to their original scale.

In the pre-selection of the *k* nearest neighbours the feature vector elements are weighted inversely proportional to their variance. This variance was globally calculated, that is, one value was calculated for the entire record, rather than one value for each calendar day, month or season separately (hence the indication global variance and *"gvar"* in the simulation name).

3.2 Types of resampling

Two types of simulation were carried out. Regarding precipitation, one is solely based on HYRAS precipitation data, i.e., both the feature vector/resampling step and the back transformation/de-standardization step are based on the HYRAS precipitation data, while for temperature E-OBS data are used. This type is denoted as HYRAS simulation. The second type is a so-called passive-HYRAS simulation in which E-OBS (version 7) precipitation data are used in the feature vector/resampling step and HYRAS precipitation data in the back transformation/de-standardization step of the simulation. The two types of resampling are designated *"hyras"* and *"pass_hyras"*, respectively, and are presented schematically in Figure 3.2.

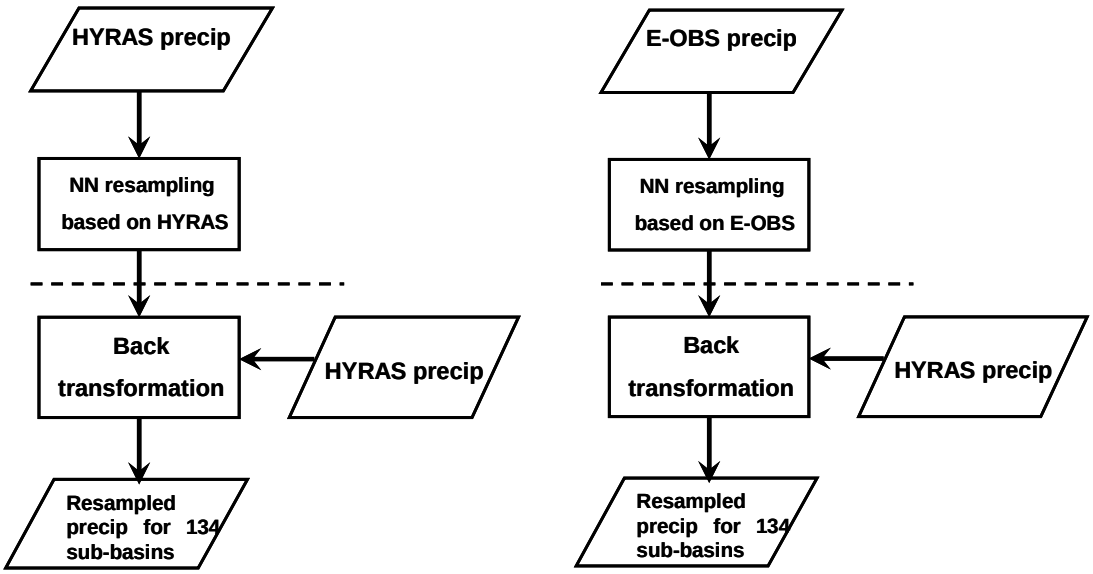


Figure 3.2: *Schematic representation of the different steps in the NNR technique based entirely on HYRAS precipitation data (left) and on E-OBS (version 7) with HYRAS precipitation data used in the back transformation (passive-HYRAS simulation, right). The dashed line indicates the point above which the two types of simulation differ.*

[4] In contrast to earlier simulations for the Rhine basin the wet-day threshold was set here to 0.3 mm.

Indirect simulation using precipitation data from a coarser grid as in the passive-HYRAS simulation may be necessary if one wishes to include historical data from 2007 onwards in the resampling procedure. The passive simulation type is straightforward because both data sets (i.e., the one used for the feature vector/resampling step and the one used for the back transformation/de-standardization step) cover the same period. If not, and this is expected to be the case when in the future E-OBS is regularly updated and HYRAS is not (see section 1.2), a second resampling step is needed, such as introduced by Leander and Buishand (2004) for the Meuse simulation.

In order to improve the reproduction of the autocorrelation of daily precipitation and the standard deviation of the monthly totals, Leander and Buishand (2004) and Leander et al. (2005) included a 4-day memory term in the rainfall generator for the Meuse basin. Beersma (2011) conducted a number of simulations for the Rhine basin with different memory terms based on the CHR-OBS data. It was recommended to repeat a number of these simulations using a different historical dataset. Therefore two other passive-HYRAS simulations were produced, in which the areal precipitation fraction term in the feature vector was replaced with a memory term based on the accumulated precipitation of the previous 4 or 10 days, respectively. In the naming convention, these simulations are labelled *"mem4d"* and *"mem10d"*, while simulations without a memory term included are labelled *"nomem"*.

To summarize, the various simulations can be distinguished using the following convention:
<memory tag>_<variance calculation tag>_<simulation length>_<simulation type tag>
In this report, 4 types of simulations are highlighted (see Table 3.1). Table 3.2 presents for these simulations the feature vector elements and their weighting coefficients.

Table 3.1: *Naming convention and explanation of the four simulation types that were investigated.*

Naming	Meaning
nomem_gvar_50000_hyras	No memory term, globally calculated weights, 50,000 yr, HYRAS simulation
nomem_gvar_50000_pass_hyras	No memory term, globally calculated weights, 50,000 yr, passive-HYRAS simulation
mem4d_gvar_50000_pass_hyras	4-day memory term, globally calculated weights, 50,000 yr, passive-HYRAS simulation
mem10d_gvar_50000_pass_hyras	10-day memory term, globally calculated weights, 50,000 yr, passive-HYRAS simulation

Table 3.2: *Weighting coefficients of the feature vector elements used in the HYRAS and passive-HYRAS simulations. The weighting coefficients for the memory term represent the values for a 4-day and 10-day (between parentheses) memory for precipitation. Note that in simulations without a memory term ('nomem'), the weight for the memory term is set to 0, while in the simulations with a memory term the weight for the areal precipitation fraction term is set to 0.*

Feature vector element	HYRAS	Passive-HYRAS (E-OBS)
Precipitation	2.13	2.33
Temperature	1.10	1.10
Areal precipitation fraction	6.62	6.38
Memory	0.27 (0.08)	0.31 (0.09)

4. SIMULATION RESULTS

4.1 Results for simulations without a memory term

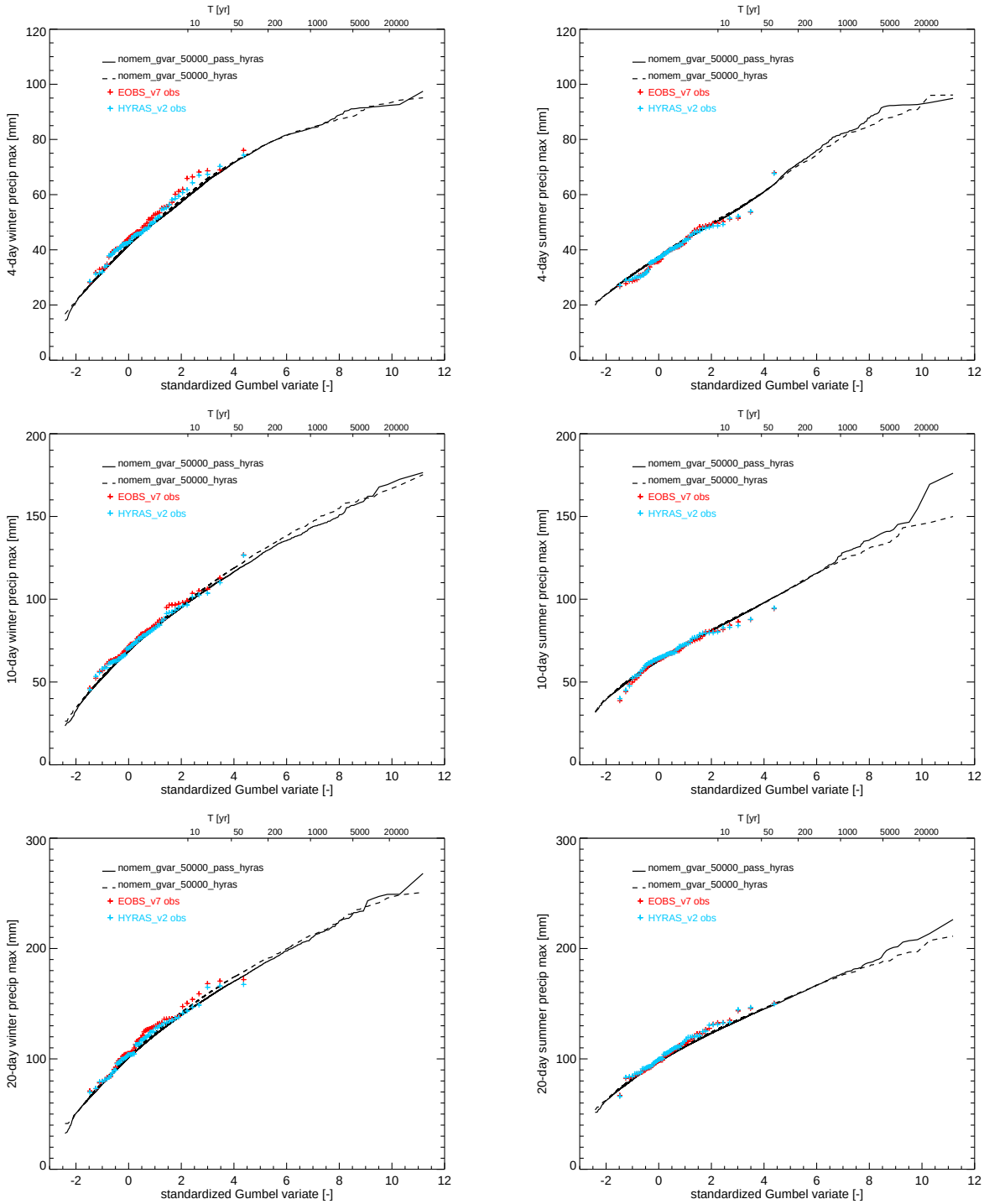


Figure 4.1: *Gumbel plots of 4-, 10-, and 20-day winter (left panels) and summer half-year (right panels) precipitation maxima for the HYRAS (dashed line) and passive-HYRAS (solid line) simulations. The corresponding values from the historical period 1951 – 2006 are denoted by the coloured pluses. T denotes the return period.*

Figure 4.1 presents the results for the 50,000-year HYRAS and passive-HYRAS simulations. Note that the winter precipitation maxima were calculated for the period October – March, while the summer maxima were computed for April – September. Differences between the HYRAS and passive-HYRAS simulations are generally small, even in the upper tail. The only substantial difference occurs in the upper tail of the 10-day summer precipitation maxima. The results indicate that the passive-HYRAS simulation (which uses E-OBS (version 7) precipitation in the feature vector) gives similar results as the HYRAS simulation.

4.2 Results for simulations including a memory term

The results of simulations with a memory term are shown in Figure 4.2. These simulations (with either a 4- or 10-day precipitation memory) have been carried out using passive simulation. For comparison, the results from the simulation without a memory term are included as well. For the winter half-year, the two simulations with a memory term differ somewhat from the simulation without a memory term, in particular for the 10- and 20-day precipitation maxima in the simulation with a 10-day memory. The nature of the differences is quite similar to that in analogous 50,000-year simulations by Beersma (2011) based on the CHR-OBS station data, although in the latter the differences are more distinct for the 10-day precipitation maxima. As in Beersma (2011), the most extreme return levels of the 10- and 20-day maximum precipitation in the simulations with a memory term are lower than those in the simulations without a memory term. The right panels in Figure 4.2 show that the return levels in the summer half-year are also relatively low in the *mem4d* simulation but not in the *mem10d* simulation. This was found in the simulations by Beersma (2011) as well.

In simulations for the Meuse basin with a 4-day memory term, it was observed that certain historical days are often simulated in the most extreme 10-day maxima, in particular some days of January 1995 (Leander and Buishand, 2008). To investigate whether there is a similar selection effect in the simulations for the Rhine basin, the historical days in the most extreme multi-day events in the winter half-year in the simulated series were examined. The relative frequencies of these historical days per winter are shown in Figure 4.3a, b and c for the 250 most extreme 4-day events, the 100 most extreme 10-day events, and the 50 most extreme 20-day events, respectively. For the most extreme 4-day events, days from the winter of 1987 were selected considerably more often in the simulations without a memory term and with a 4-day memory term than historical days from other winters, with a relative frequency of almost 30%. Note that this large relative frequency involves only a few days from October 1986. Days from the winter of 1999 account for almost 24% of the days forming the 100 most extreme 10-day events in the simulation with a 4-day memory term. Both for the most extreme 10-day events and the most extreme 20-day events all frequencies are below 10% in the *nomem* simulation. For the simulations with a 10-day memory term days from the winter of 1994 account for the largest fraction of the resampled days in all panels of Figure 4.3. Days from that winter were also selected relatively often in the *nomem* simulation, albeit with a much smaller relative frequency.

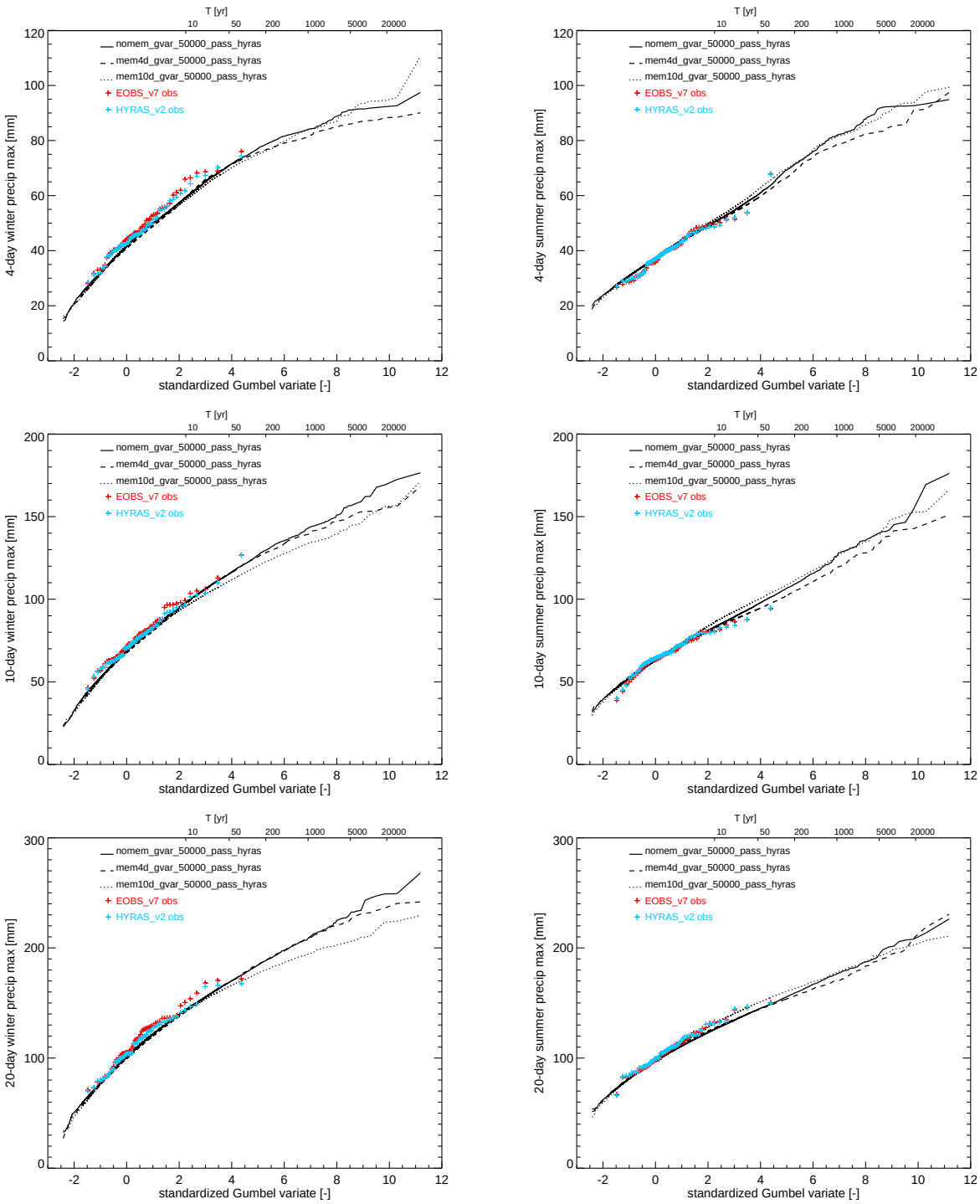
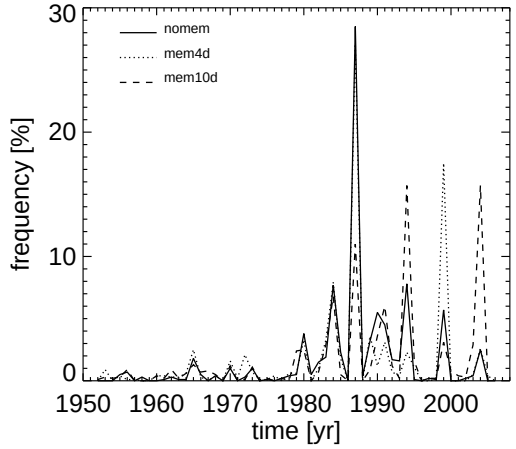
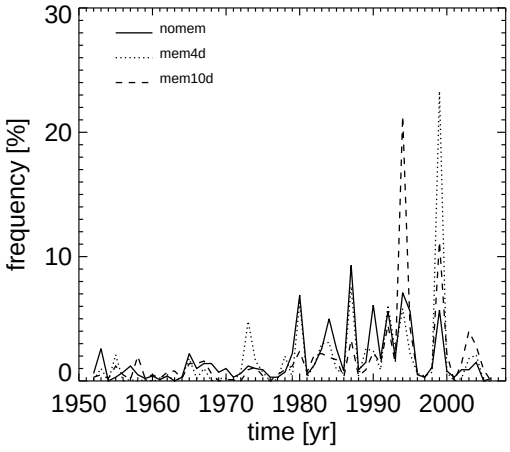


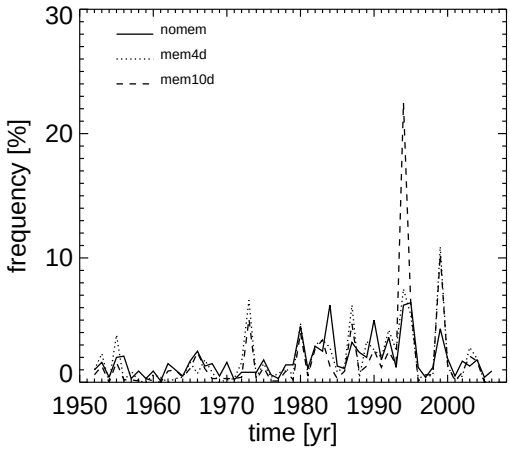
Figure 4.2: *As in Figure 4.1, but now for simulations with a 4-day (dashed) and 10-day (dotted) memory term using the passive-HYRAS simulation. The passive-HYRAS simulation without a memory term is included for reference and is indicated by the solid line.*



(a)



(b)



(c)

Figure 4.3: *Frequency of simulated days from historical years for (a) the 250 most extreme 4-day precipitation winter maxima, (b) the 100 most extreme 10-day precipitation winter maxima, and (c) the 50 most extreme 20-day precipitation winter maxima in the passive-HYRAS simulations without a memory term, and with a 4- and 10-day memory term. The frequencies plotted at each year y were computed using data from the last months of year y-1 through the first months of year y.*

The results as shown in Figure 4.3 give rise to suspect the resampling technique having a preference to specific periods with extremely high multi-day precipitation amounts. This strong preference to certain historical days in the maximum multi-day precipitation amounts, especially from simulations with a memory term, probably leads to an increase of the standard error of the extreme quantiles of these multi-day precipitation amounts compared to simulations without a memory term. This is investigated for the Meuse basin in Schmeits et al. (2014). In contrast to the Meuse basin, simulations with a memory term for the Rhine basin do not really improve the reproduction of the autocorrelation of the daily precipitation (Beersma, 2011). Therefore, for the Rhine basin there is no need to include a memory term in the feature vector of the resampling procedure.

5. UNCERTAINTY ANALYSIS

In this section the uncertainty analysis of the rainfall generator for the Rhine basin (without a memory term in the feature vector) is described. First, we followed the same approach as in Leander and Buishand (2008). So, we also selected randomly about half of the series, i.e., 27 (winter half-)years without replacement in this case, and repeated this procedure, so that an ensemble of 50,000 subsets was created. Figure 5.1 shows the mean daily winter amount of precipitation *R* as a function of the fraction *f* of winter days with ≥ 10 mm of precipitation for each subset of 27 years. As in Leander and Buishand (2008) an ellipse emerges.

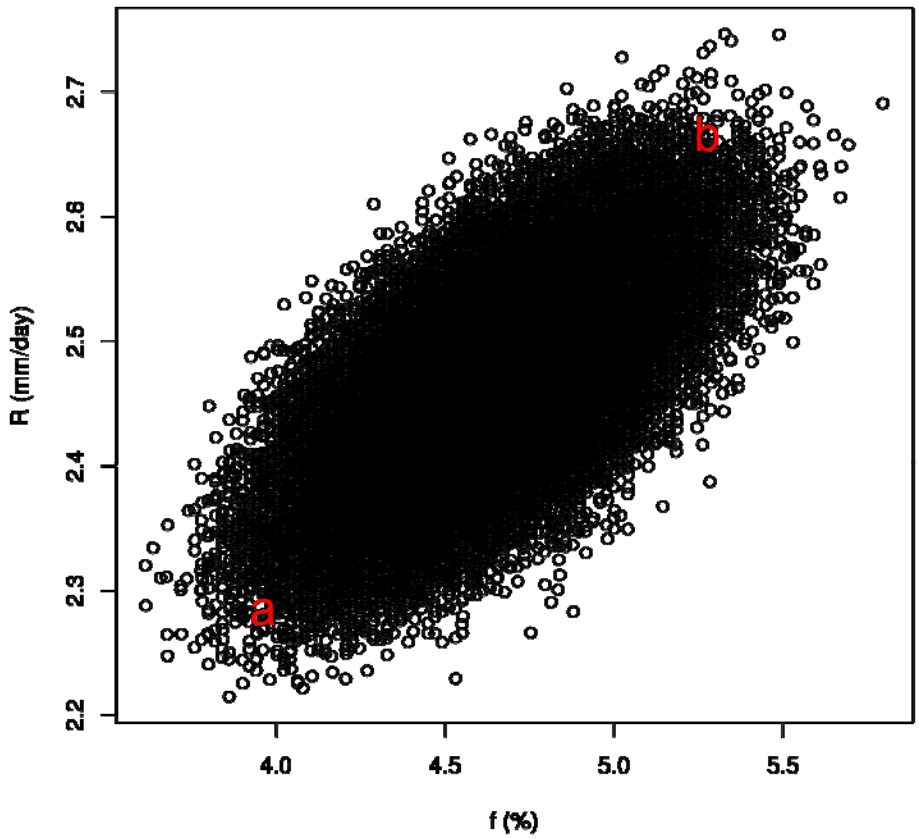


Figure 5.1: *Mean daily winter precipitation amount R versus fraction f of winter days with ≥ 10 mm of precipitation for all 50,000 subsets in the ensemble (circles).*

Because of the conclusion from Figure 4.3 that days from the winters of 1987 and 1994 were selected often in extreme situations, subsets including and excluding both 1987 and 1994 have been selected and for each also a dry and a wet subset (point a and point b in Figure 5.1, respectively). This yields 4 subsets and for each a 20,000-year passive-HYRAS simulation has been performed. The 50,000-year passive-HYRAS simulation based on the entire historical period, which was shown in the previous section, serves as a reference. In Figure 5.2 the Gumbel plots of the 4-, 10- and 20-day precipitation maxima in the winter half-year for these simulations are shown. It can be seen that the return levels are generally higher for the simulations that include 1987 and 1994 (i.e., the dotted curves lie higher than the dashed curves). In addition, for the 10- and 20-day maxima the return levels for the simulations based on the wet subset are also higher than those based on the dry subset (i.e., the green curves lie higher than the black curves).

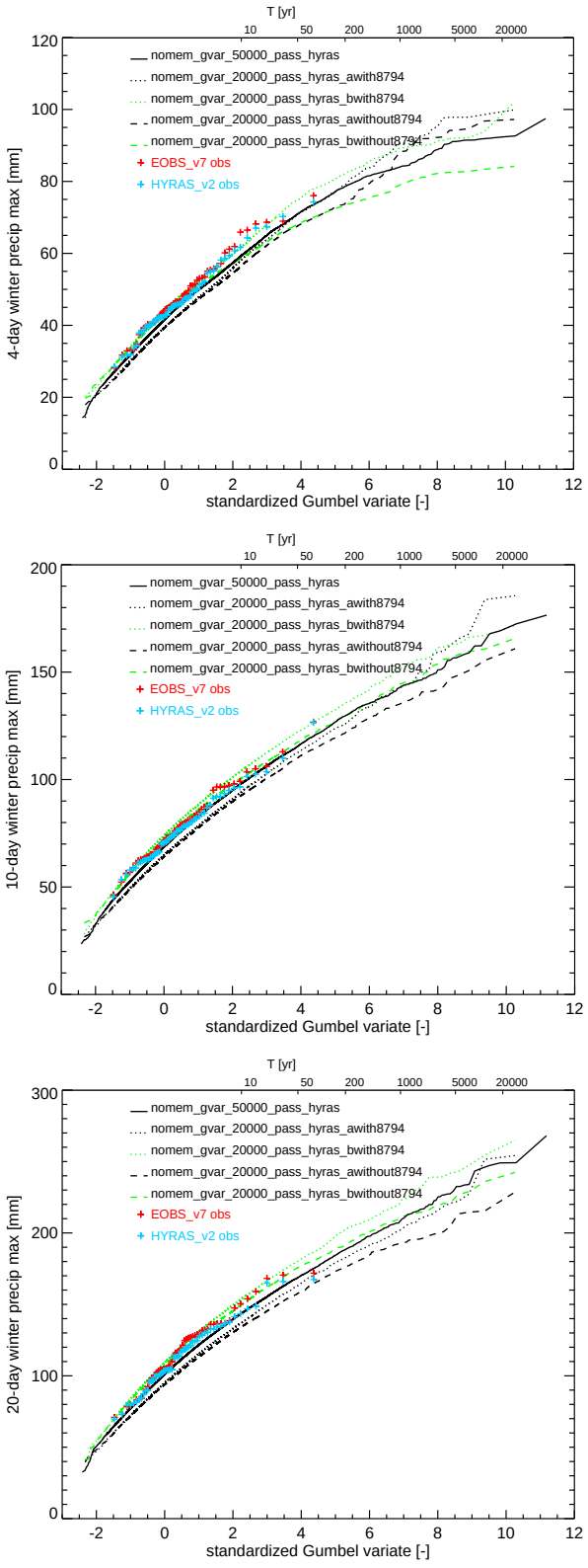


Figure 5.2: *Gumbel plots of 4-, 10-, and 20-day winter half-year precipitation maxima for the 50,000-year passive-HYRAS simulation (solid) and on the four 20,000-year passive-HYRAS simulations with 1987 and 1994 (dotted) and without 1987 and 1994 (dashed), with the letters a and b corresponding to dry and wet winters, respectively (see Figure 5.1) and denoted by black and green, respectively. The corresponding values from the historical period 1951 – 2006 are denoted by the coloured pluses. T denotes the return period.*

However, this approach has the following disadvantages. First, for the relatively short 27-year series (about half of the original series), the influence of two anomalous years can be relatively large. Second, it is not trivial how the standard deviation of an estimated return level can be obtained from these simulations. The computation of this standard deviation as in Kramer and Schroevers (2008) may lead to an overestimation, because the subsets were selected from the two extreme ends of the 'ellipse'. Therefore, we have adopted a different approach, based on the jackknife. Eleven jackknife series of 50 years were formed by leaving out subsequent non-overlapping 5-year blocks from the original series of 55 years (with the first block being 1952-56[5]). For each jackknife series a 20,000-year passive HYRAS simulation was conducted (the names of the NetCDF-files of these simulations and the reference simulation are given in the Appendix). Let $\hat{x}_{T(i)}$ be the estimated *T*-year return level from the *i*-th jackknife series. Then the jackknife standard deviation of the estimated *T*-year return level is given by:

$$s_{jack} = \left\{ \frac{n-1}{n} \sum_{i=1}^{n} \left[\hat{x}_{T(i)} - \hat{x}_{T(\bullet)} \right]^2 \right\}^{1/2}, \qquad (1)$$

where

$$\hat{x}_{T(\bullet)} = \sum_{i=1}^{n} \hat{x}_{T(i)} / n$$

and n (= 11) is the number of jackknife series. The accuracy of s_{jack} decreases with increasing block size. A small block size requires, however, more computations. Because of computational restraints on the hydrological part and because the length of the original series was a multiple of 5 years, a block size of 5 years was chosen. Figure 5.3 shows the Gumbel plots of the 20,000-year passive-HYRAS simulations based on the 11 jackknife series, again with the 50,000-year passive-HYRAS simulation as a reference. The spread between the different 20,000-year simulations is smaller than that between the simulations from the 27-year series in Figure 5.2, because the jackknife series are longer and also more similar (each pair of jackknife series has 45 years in common). Because of the latter, the sum of the squared deviations in Eq. (1) is multiplied by (n-1)/n rather than 1/n.

Table 5.1 compares two estimates of the 1250- and 4000-year return levels of the 4-, 10- and 20-day precipitation maxima in the winter half-year. The empirical estimate is computed as the 16th and 5th largest value in each of the 20,000-year simulations for T = 1250 and 4000 years, respectively, while the Weissman estimate is based on the joint distribution of the r largest values $x_{[1]} \geq x_{[2]} \geq \ldots \geq x_{[r]}$ (Weissman, 1978). In the Weissman method, the T-year return level x_T is estimated as

$$\hat{x}_T = x_{[r]} + \hat{\sigma} \ln(rT / N)$$

with N (= 20,000) the length of the simulation, and

$$\hat{\sigma} = \bar{x}_r - x_{[r]}$$

where $\bar{x}_r$ is the average of the r largest values. A value of r = 100 gave the best

[5] or, more precisely 1 October 1951 – 30 September 1956. The start of a block was set at 1 October to avoid splitting the winter half-years.

results. The estimated return levels in Table 5.1 are the averages $\hat{x}_{T(\bullet)}$ from the 11 jackknife series and their standard deviations are based on Eq. (1). There is not much difference between the empirical and Weissman estimates. The latter has a somewhat smaller standard deviation. Though the estimated 4000-year return level has a larger standard deviation than the 1250-year return level, their relative standard deviations do not differ much. These relative standard deviations vary between 5 and 8%.

A second jackknife set was formed to investigate the sensitivity of the estimation of the standard deviation to the choice of the jackknife-blocks. For this second set also a block size of 5 years was used, which again resulted in 11 jackknife series. The difference however is that the blocks in the second set consisted of different historical years. Ten blocks of consecutive 5-year periods were considered: 1955-1959, …, 2000-2004, and a remaining block containing the years 1952-54 and 2005-06. Table 5.2 shows the estimated 1250-year and 4000-year return levels of the multi-day precipitation maxima and their standard deviations for this second jackknife set. For the return levels of the 4-day and 10-day maxima, the differences between the standard deviations from the first jackknife set are small, but for the return levels of the 20-day maxima the standard deviations for the Weissman estimates are much smaller in the second jackknife set. This indicates that the jackknife standard deviations are quite uncertain. To reduce the uncertainty in the estimation of the standard deviation, the jackknife block size can be reduced, but this requires a new set with a larger number of jackknife simulations. Alternatively, one could combine the two different sets of 11 jackknife simulations and estimate the standard deviation of a return level as the average of the standard deviations of both sets. In the same way the uncertainty in the standard deviation of a return level of river discharge as obtained with GRADE can be reduced.

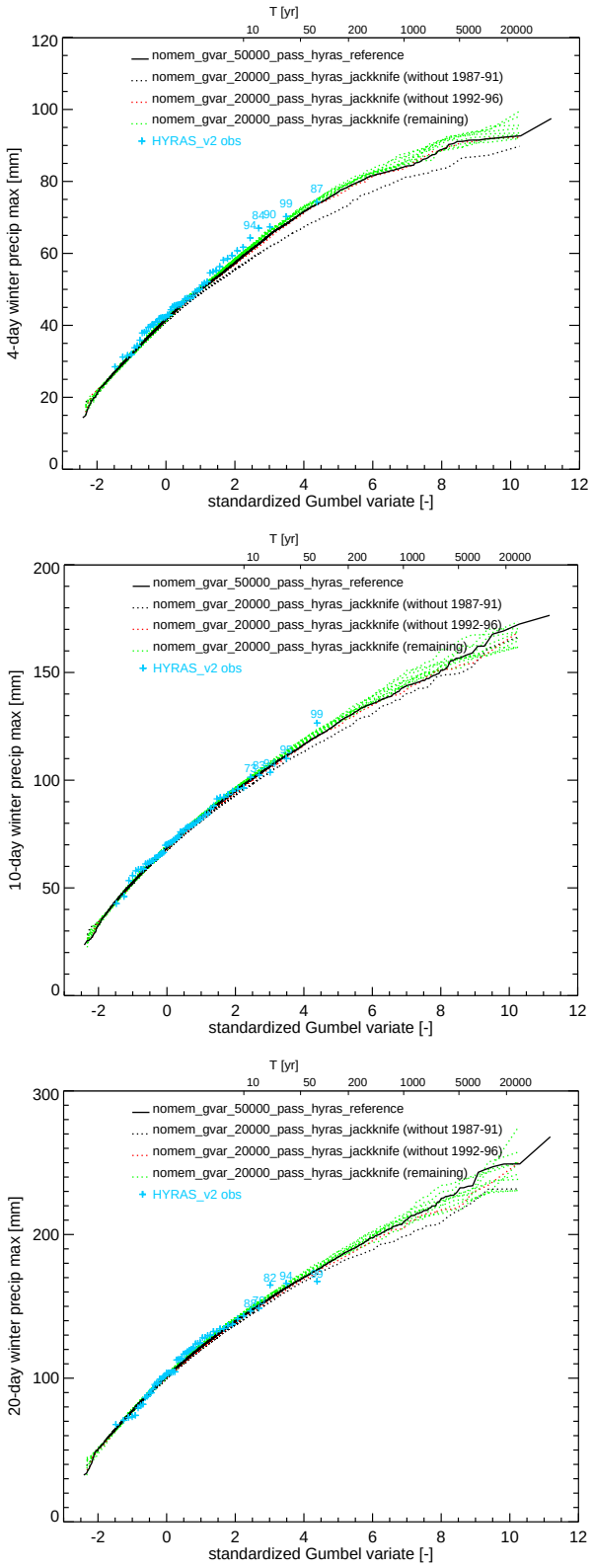


Figure 5.3: *As in Figure 5.2, but now for 20,000-year passive-HYRAS simulations based on 11 jackknife series of 50 years (black, red and green dotted; see legend), again with the 50,000-year passive-HYRAS simulation as a reference (black solid). The plusses indicate the ordered observed maxima for the period 1951-2006 (and for the top 5 the year minus 1900 is added, e.g. 87 indicates the winter half-year October 1986 – March 1987).*

Table 5.1: *Estimated return levels of 4-, 10- and 20-day winter half-year precipitation maxima and their standard deviations for the 20,000-year passive-HYRAS simulations based on 11 jackknife series of 50 historical years and for return periods T of 1250 and 4000 years; both empirically determined and using the Weissman method (r = 100), see text.*

T (yr)	Return level (mm)		Standard deviation	
	Empirical	Weissman	Empirical	Weissman
		4-day precipitation		
1250	86	86	6 mm (7%)	5 mm (6%)
4000	91	91	7 mm (8%)	5 mm (5%)
		10-day precipitation		
1250	147	146	9 mm (6%)	8 mm (5%)
4000	155	156	11 mm (7%)	10 mm (6%)
		20-day precipitation		
1250	212	211	15 mm (7%)	14 mm (7%)
4000	224	226	19 mm (8%)	18 mm (8%)

Table 5.2: *As Table 5.1, but now for an alternative set of 11 jackknife series of 50 historical years (see text).*

T (yr)	Return level (mm)		Standard deviation	
	Empirical	Weissman	Empirical	Weissman
		4-day precipitation		
1250	86	86	7 mm (8%)	5 mm (6%)
4000	91	91	6 mm (7%)	6 mm (7%)
		10-day precipitation		
1250	146	146	11 mm (8%)	9 mm (6%)
4000	157	156	12 mm (8%)	12 mm (8%)
		20-day precipitation		
1250	212	212	12 mm (7%)	9 mm (4%)
4000	227	227	19 mm (8%)	12 mm (5%)

6. CONCLUSIONS

Since recently new and considerably longer gridded daily precipitation and temperature datasets (E-OBS and HYRAS) have been available for usage in nearest neighbour resampling for the Rhine basin. In order to assess the suitability of the datasets, the 1-day mean and 10-day maximum precipitation of these datasets were compared with the CHR-OBS dataset. The HYRAS and E-OBS gridded datasets compare well with CHR-OBS, with relative differences mostly within 5% for the entire Rhine basin. For the Swiss part of the Rhine basin, the differences between E-OBS, HYRAS, and CHR-OBS are larger; CHR-OBS being 25-30% larger than E-OBS and 15-20% larger than HYRAS over the elevated sub-basins (> 1000 m) of Switzerland.

A new series of 50,000-yr simulations was performed. Differences between the simulation entirely based on HYRAS precipitation data and the passive HYRAS simulation are generally small. In the latter only E-OBS data were included in the feature vector, which is useful for future simulation efforts, as the E-OBS precipitation and temperature data are updated semi-yearly.

The simulation with a 4-day memory term shows lower return levels compared to the simulation without a memory term, while the simulation with a 10-day memory term generally shows similar results as the *nomem* simulation, apart from the 10- and 20-day winter precipitation maxima which are considerably lower at long return periods. The nature of the differences between simulations with and without the memory terms is similar to findings from analogous simulations of Beersma (2011) based entirely on the CHR-OBS data. Further, it was shown that especially in the simulations with a memory term wet days from specific wet periods were selected more frequently in situations of extreme multi-day precipitation. It is however not fully clear what causes such selection effects. All these results indicate that inclusion of a memory term has no added value for the Rhine basin, which is in contrast with the simulation results for the Meuse basin of Leander et al. (2005).

Because it is not trivial to estimate the standard deviation of the estimated return levels using the approach in Leander and Buishand (2008), a different approach has been followed in the uncertainty analysis in this study. Jackknife series were formed by leaving out subsequent non-overlapping 5-year blocks from the original 55-year series and for each of these series the desired return level was estimated by resampling a synthetic sequence of 20,000 years. The relative standard deviation of the 1250-year and 4000-year return levels of the 4-, 10- and 20-day maxima in the winter half-year varies between 4 and 8%.

APPENDIX

The names of the NetCDF files start with the string Rhine_2012_v01.1_reference_50K for the reference simulation (with nomem_gvar_50000_pass_hyras[_reference] as the corresponding name in this report). The 25 NetCDF files each contain 2000 years of data and are called:
Rhine_2012_v01.1_reference_50K_part01.nc, …,
Rhine_2012_v01.1_reference_50K_part25.nc.

For the 11 jackknife simulations the names of the NetCDF files start with the string
Rhine_2012_v01.1_jackknife_20K_subseries01,…,
Rhine_2012_v01.1_jackknife_20K_subseries11
(with nomem_gvar_20000_pass_hyras_jackknife as the corresponding name in this report). In each subseries a different 5-year block is deleted; details of which 5-year block is deleted are given in the metadata of the NetCDF files. The 10 NetCDF files for each of the 11 jackknife simulations each contain 2000 years of data and are called:
Rhine_2012_v01.1_jackknife_20K_subseries01_part01.nc,…..,
Rhine_2012_v01.1_jackknife_20K_subseries01_part10.nc
to
Rhine_2012_v01.1_jackknife_20K_subseries11_part01.nc,…..,
Rhine_2012_v01.1_jackknife_20K_subseries11_part10.nc.

The NetCDF files contain both the precipitation and temperature data for each of the 134 HBV_Rhine sub basins and use the Gregorian calendar. Each simulation starts in the year 2001 to avoid the Gregorian correction in the year 1582.

Note that a version “Rhine_2012_v01_” (Rhine_2012_v01_reference_50K_... and Rhine_2012_v01_jackknife_20K_...) was provided earlier to Deltares that used the Julian calendar[6] and which turned out to be incompatible with the Gregorian calendar used in FEWS[7]. Version “Rhine_2012_v01_” should therefore not be used anymore.

[6] In this earlier version each NetCDF file contains 1000 years of simulated data rather than 2000 years.
[7] FEWS: Flood Early Warning System (by Deltares).

REFERENCES

Beersma, J.J., 2002: Rainfall generator for the Rhine, description of 1000-yr simulations, KNMI Publication 186-V, KNMI, De Bilt, Netherlands, 18 pp.

Beersma, J.J., 2011: Rainfall generator for the Rhine Basin, first results of simulations including a memory term in the feature vector, KNMI Memorandum, KNMI, De Bilt, Netherlands, 8 pp.

Buishand, T.A. and T. Brandsma, 2001: Multisite simulation of daily precipitation and temperature in the Rhine Basin by nearest-neighbour resampling, *Water Resour. Res.*, **37,** 2761 **–** 2776**,** doi:10.1029/2001WR000291.

Dällenbach, F., 2000: Gebietsniederschlag Schweiz – Interpolation und Berechnung der Niederschlagsdaten, METEOTEST report for the Landeshydrologie und –geologie (LHG)*,* Bern, Switzerland.

Eberle, M., H. Buiteveld, P. Krahe and K. Wilke, 2005: Hydrological modelling in the river Rhine basin, Part III: Daily HBV model for the Rhine basin. Report No. 1451, Bundesanstalt für Gewässerkunde (BfG), Koblenz, Germany.

Görgen, K., J. Beersma, G. Brahmer, H. Buiteveld, M. Carambia, O. de Keizer, P. Krahe, E. Nilson, R. Lammersen, C. Perrin, and D. Volken, 2010: Assessment of climate change impacts on discharge in the Rhine River Basin: results of the RheinBlick2050 project, CHR report No. I-23, CHR, Lelystad, Netherlands, 211 pp.

Haylock, M.R., N. Hofstra, A.M.G. Klein Tank, E.J. Klok, P.D. Jones, and M. New, 2008: A European daily high-resolution gridded dataset of surface temperature and precipitation for 1950 – 2006, *J. Geophys. Res.*, **113**, D20119, doi: 10.1029/2008JD010201.

Hundecha, Y. and A. Bárdossy, 2005: Trends in daily precipitation and temperature extremes across western Germany in the second half of the 20th century, *Int. J. Climatol.*, **25**, 1189 – 1202, doi:10.1002/joc.1182.

Kramer, N. and R. Schroevers, 2008: Generator of Rainfall and Discharge Extremes (GRADE), Part F. Deltares Report Q4424.

Leander, R. and T.A. Buishand, 2004: Rainfall generator for the Meuse basin, development of a multi-site extension for the entire drainage area. Publication 196-III, KNMI, De Bilt, Netherlands, 40 pp.

Leander, R., T.A. Buishand, P. Aalders, and M.J.M. de Wit, 2005: Estimation of extreme floods of the river Meuse using a stochastic weather generator and rainfall-runoff model, *Hydrolog. Sci. J.*, **50**, 1089 – 1103, doi: 10.1623/hysj.2005.50.6.1089.

Leander, R. and T.A. Buishand, 2008: Rainfall generator for the Meuse basin, description of 20,000-year simulations, KNMI Publication 196-IV, KNMI, De Bilt, Netherlands, 15 pp.

Lindström, G., B. Johansson, N. Persson, M. Gardelin, M. Bergström, 1997: Development and test of the distributed HBV-96 hydrological model, *J. Hydrol.*, **201**, 272 – 288.

Moberg, A. and Coauthors, 2006: Indices for daily temperature and precipitation extremes in Europe analyzed for the period 1901 – 2000, *J. Geophys. Res.*, **111**, D22106, doi: 10.1029/2006JD007103.

Quirmbach, M., T. Einfalt, and G. Langstädtler, 2012: Climate change analysis of precipitation data for North Rhine-Westphalia, *Atmos. Res.,* **109-110**, 1 – 13, doi:10.1016/j.atmosres.2011.10.014.

Rauthe, M., H. Steiner, U. Riediger, A. Mazurkiewicz, and A. Gratzki, 2013: A central European precipitation climatology – Part I: Generation and validation of a high-resolution gridded daily data set (HYRAS), *Meteor. Z.*, **22**, 235-256.

Schmeits, M.J., J.J. Beersma, and T.A. Buishand, 2014: Rainfall generator for the Meuse basin: Description of simulations with and without a memory term and uncertainty analysis. KNMI Publication 196-VI, KNMI, De Bilt.

Schmidli, J. and C. Frei, 2005: Trends of heavy precipitation and wet and dry spells in Switzerland during the 20th century, *Int. J. Climatol.*, **25**, 753 – 771, doi: 10.1002/joc.1179.

Steiner, H., 2009: HYRAS report, chapter 5, Interpolationsverfahren Deutscher Wetterdienst.

Steinrücke J., B. Fröhlings and R. Weißhaupt, 2012: HYMOG - Hydrologische Modellierungsgrundlagen im Rheingebiet, CHR report no I-24, Lelystad, Netherlands, 107 pp.

Weissman, I., 1978: Estimation of parameters and large quantiles based on the *k* largest observations. *Journal of the American Statistical Association*, **73**, 812-815.

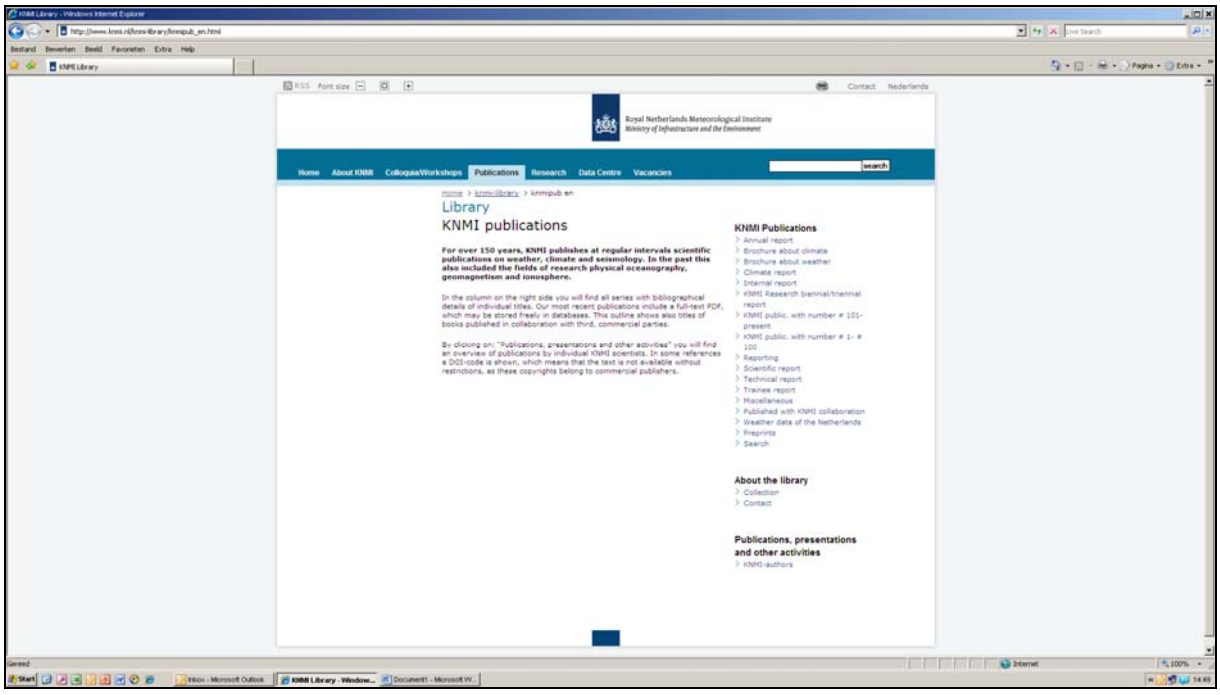
Willems, W. and K. Stricker, 2011: Vergleichende Analyse ausgewählter meteorologischer Observationsproducte (MOBS). Unpublished report, Ingenieurhydrologie Angewandte Wasserwirtschaft und Geoinformatik (IAWG), Ottobrunn, Germany, 111 pp.

Wit, M.J. de and T.A. Buishand, 2007: Generator of Rainfall And Discharge Extremes (GRADE) for the Rhine and Meuse basins. Rijkswaterstaat RIZA report 2007.027/KNMI publication 218, Lelystad/De Bilt, Netherlands, 77 pp.

Wójcik, R., J.J. Beersma, and T.A. Buishand, 2000: Rainfall generator for the Rhine basin – multi-site generation of weather variables for the entire drainage area, KNMI Publication 186-IV, KNMI, De Bilt, Netherlands, 38 pp.

A complete list of all KNMI-publications (1854 – present) can be found on our website

www.knmi.nl/knmi-library/knmipub_en.html



The most recent reports are available as a PDF on this site.