Koninklijk Nederlands
Meteorologisch Instituut
Ministerie van Infrastructuur en Milieu

Sensitivity to the composition of the feature vector and passive simulations

Rainfall Generator for the Rhine Basin

Jules J. Beersma

De Bilt, 2011 | KNMI publication 186-VI

Rainfall Generator for the Rhine Basin

Sensitivity to the composition of the feature vector and passive simulations

Jules J. Beersma

KNMI publication 186-VI

Work performed under contract “Zaaknummer 31007187.0002” to Ministry of Transport, Public Works and Water Management, Rijkswaterstaat Centre for Water Management, Lelystad, The Netherlands.

Contents

1. Introduction

Since the mid 1990s rainfall generators for the Rhine and Meuse basins have been developed. These rainfall generators form part of the GRADE instrument for the Generation of Rainfall and Discharge Extremes (de Wit and Buishand, 2007). This report focuses on some developments of the Rainfall generator for the Rhine basin during the period June 2009 to November 2010 as part of the KNMI contribution to the Waterdienst-Deltares-KNMI collaboration regarding GRADE. Three topics are considered: i) sensitivity of the temporal pattern correlation for the composition of the feature vector, ii) Rainfall simulations for the Rhine basin based on catchment average data rather than station data and, iii) passive simulation of precipitation and temperature in the Meuse basin based on simulations for the Rhine basin. This last topic may seem a bit of an off-topic in a report on the Rainfall generator for the Rhine basin. However, this work gives a first indication of the possibility to develop a unified Rhine-Meuse rainfall generator.

2. Sensitivity of the temporal pattern correlation for the composition of the feature vector

2.1 Introduction

The pattern correlation is a measure of the dependence between the spatial fields of two variables or the spatial fields of one variable at two different times (temporal pattern correlation). This section investigates the sensitivity of the temporal pattern correlations of the simulated precipitation and temperature fields for the composition of the feature vector. The earlier simulations which are used in GRADE and which are described in Beersma (2002) serve as a reference. These simulations were based on daily precipitation and temperature series of 34 stations for the period 1961–1995. The average (standardized) precipitation of these stations $\widetilde{P}^*$, the fraction of stations with precipitation F^*, and the average (standardized) temperature $\widetilde{T}^*$, were used in the feature vector. In earlier simulations by Brandsma and Buishand (1999), that were restricted to the German part of the Rhine basin,

the average precipitation for 5 different regions was used, instead of $\tilde{P}^*$ and F^*. Based on this concept 6 regions are introduced here for the whole Rhine basin upstream of the Netherlands, including the stations in Luxembourg, France and Switzerland (see Figure 2.1). Note that the first 5 regions closely resemble those in Brandsma and Buishand (1999). In the simulations presented here the average precipitation for each of these 6 regions is used in the feature vector, combined with or without the 34-station-average temperature $\tilde{T}^*$. In addition two simulations are performed in which the precipitation of all 34 stations is included in the feature vector (again with and without the 34-station-average temperature).

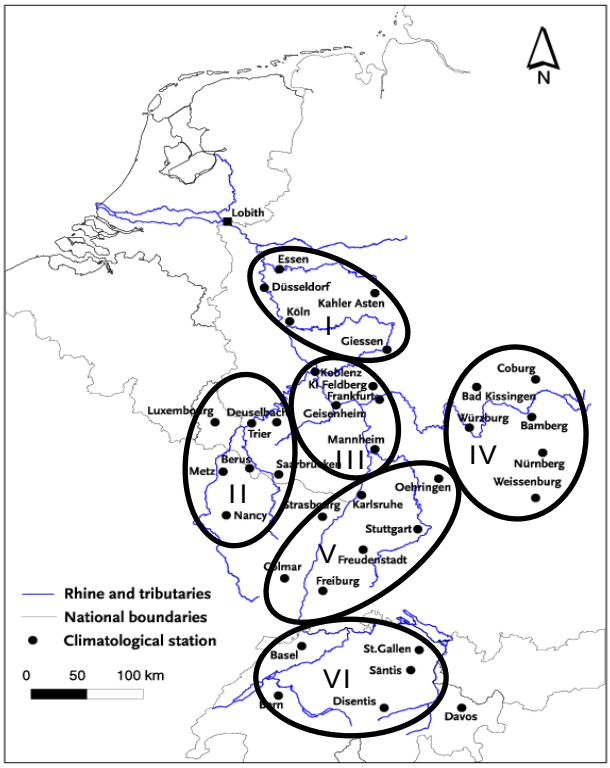


Figure 2.1. Location of 36 stations in the Rhine basin and the subdivision into 6 regions. Note that the stations Davos and Säntis are not used.

2.2 Results

All simulations are performed with the same settings as the 1000-yr reference simulation ue241_k=10_1000_ran1.1_leapyear_chck.log except for the composition of the feature vector. In these simulations 'u' stands for unconditional (i.e. *not* conditional on the atmospheric circulation indices), 'e' stands for Euclidean distance (i.e. the metric used to determine and order the nearest neighbours), '241' denote the weights for the feature vector elements $\tilde{P}^*, F^*, \tilde{T}^*$, 'k=10' refers to the number of nearest neighbours and 'ran1.1' refers to the type of the random number generator and the random seed used. Table 2.1 presents the simulation names and the corresponding feature vector composition. Note that $\tilde{\boldsymbol{P}}_6^* = (\tilde{P}_I^*, \tilde{P}_{II}^*, \tilde{P}_{III}^*, \tilde{P}_{IV}^*, \tilde{P}_V^*, \tilde{P}_{VI}^*)$ is a vector containing the average standardized precipitation for each of the 6 regions and similarly $\tilde{\boldsymbol{P}}_{34}^* = (\tilde{P}_1^*, \tilde{P}_2^*, ..., \tilde{P}_{34}^*)$ is a vector containing the standardized precipitation for each of the 34 stations.

Table 2.1. Composition of the feature vector and corresponding weights for different simulations. Note that a weight **1** refers to a vector of ones (with 6 or 34 elements).

Simulation	Feature vector	Weights
ue241	$\widetilde{P}^*, F^*, \widetilde{T}^*$	2, 4, 1
ue001R_6	$\widetilde{T}^*, \widetilde{\boldsymbol{P}}_6^*$	1, **1**
ue000R_6	$\widetilde{\boldsymbol{P}}_6^*$	**1**
ue001	$\widetilde{T}^*$	1
ue001P_{34}	$\widetilde{T}^*, \widetilde{\boldsymbol{P}}_{34}^*$	17, **1**
ue000P_{34}	$\widetilde{\boldsymbol{P}}_{34}^*$	**1**

The resampling technique used preserves the spatial patterns of the daily precipitation and temperature fields, but it does not necessarily reproduce the dependence between the patterns of successive days. The temporal dependence between two spatial patterns that lie *l* days apart can be characterized by a pattern correlation coefficient. The lag *l* pattern correlation (centred statistic) correlates the spatial patterns, and is averaged over all pairs of patterns that lie *l* days apart (for precipitation, days with no rainfall were excluded).

Tables 2.2 – 2.5 present the lag 1 and lag 2 pattern correlations for the historical data (1961 – 1995) together with the standard statistics presented earlier (Beersma, 2002) and the biases in these quantities for the various simulations. The pattern correlation is calculated in two different ways: i) from the patterns defined by the 34 stations and, ii) from the patterns defined by the 6 regions. Separate tables are given for precipitation and temperature and for winter and summer. In these tables the following colour-coding is adopted: a red value is unacceptably different[1] from the historical value; an orange value is relatively the worst value in a column (but if a column contains a red value it cannot contain an orange value) and a green value is relatively the best value in a column (but no green value possible in a column if all values in the column differ significantly from the historical value – i.e. differ more than twice their standard error *se* from the historical estimate).

The biases in the pattern correlations (derived from the patterns of the 34 stations) for precipitation and temperature for the 'ue241' reference simulation are of comparable size as those for simulations *conditional* on the atmospheric circulation presented in Beersma (2007; see Chapter 3, Table 3.3). In those simulations the lag 1 and 2 pattern correlation biases for winter precipitation were respectively about -0.09 and -0.12 and for winter temperature respectively -0.04 and -0.07.

Incorporation of the average precipitation for the 6 regions, or all 34 precipitation stations in the feature vector leads to a better reproduction of the lag 1 and 2 pattern correlations of precipitation (see Tables 2.2 and 2.3) but the other statistics typically suffer from this improvement. The largest improvement in pattern correlation is found for the simulations with all 34 precipitation stations in the feature vector (the 'P_{34}' simulations) but in these simulations significant biases are found in the mean precipitation and in the daily and monthly standard deviations of precipitation. In the 'R_6' simulations the reduction of the bias in the precipitation pattern correlation is smaller but some additional bias in the daily standard

[1] This is clearly a subjective judgment but first of all the difference should be larger than 2 × *se* from the historical estimate (i.e. statistically significant, roughly at about the 5%-level) and in addition, the difference should be considerably larger than in the alternative simulations in the table.

deviation is found (both in winter and in summer), but at the same time the bias in the lag 1 autocorrelation is reduced (both in winter and in summer). A similar reduction of bias in the the lag 1 autocorrelation for precipitation was already identified by Brandsma and Buishand (1999) with the use of precipitation of 5 regions in the feature vector.

Table 2.2. Differences between statistical properties of the simulated time series and the historical records (1961–1995) for **winter precipitation** (October–March). For the lag 1 and 2 pattern correlation coefficients (r_p), mean precipitation (monthly totals), and the mean lag 1 and 2 autocorrelation coefficients ($\bar{r}$) the absolute differences are given, and for the mean standard deviations of monthly and daily values ($\bar{s}_M$ and $\bar{s}_D$) the percentage differences. Values between ⟨⟩ denote averages over the 34 stations (details in Beersma, 2002). Bottom lines: average historical estimates (mean and standard deviations in mm) and their standard error *se* (standard errors for the mean in mm, for standard deviations in % and for the autocorrelation coefficients dimensionless). Values in bold refer to differences more than 2 × *se* from the historical estimate, where *se* is calculated as in Beersma and Buishand (1999).

	34 stations		6 regions						
Simulation	$r_p(1)$	$r_p(2)$	$r_p(1)$	$r_p(2)$	Mean	$\langle\Delta\bar{s}_M\rangle$	$\langle\Delta\bar{s}_D\rangle$	$\langle\Delta\bar{r}(1)\rangle$	$\langle\Delta\bar{r}(2)\rangle$
ue241	**-0.086**	**-0.044**	**-0.096**	**-0.049**	-0.8	-3.4	-0.7	**-0.036**	-0.009
ue001R$_6$	**-0.055**	**-0.034**	**-0.033**	**-0.030**	-3.1	-7.3	-2.7	-0.015	-0.010
ue000R$_6$	**-0.040**	**-0.030**	-0.001	**-0.017**	-4.1	-5.0	-2.9	-0.011	-0.003
ue001	**-0.096**	**-0.053**	**-0.099**	**-0.053**	-0.8	**-19.6**	-0.7	**-0.156**	**-0.099**
ue001P$_{34}$	**-0.043**	**-0.029**	**-0.034**	**-0.028**	**-7.2**	-6.6	-4.5	-0.003	0.001
ue000P$_{34}$	-0.005	**-0.013**	0.009	-0.010	**-7.5**	-7.3	**-5.6**	0.003	0.004
Historical	0.253	0.161	0.234	0.133	64.1	35.8	4.2	0.285	0.144
Se	0.005	0.005	0.008	0.008	2.47	4.53	2.46	0.008	0.009

Table 2.3. As Table 2.2 but for the **summer precipitation** (April—September).

	34 stations		6 regions						
Simulation	$r_p(1)$	$r_p(2)$	$r_p(1)$	$r_p(2)$	Mean	$\langle\Delta\bar{s}_M\rangle$	$\langle\Delta\bar{s}_D\rangle$	$\langle\Delta\bar{r}(1)\rangle$	$\langle\Delta\bar{r}(2)\rangle$
ue241	**-0.061**	**-0.014**	**-0.083**	**-0.022**	-0.7	**-8.9**	-1.3	**-0.029**	0.008
ue001R$_6$	**-0.032**	-0.005	**-0.017**	-0.002	-3.2	**-9.1**	-3.4	-0.015	-0.008
ue000R$_6$	**-0.024**	-0.002	0.004	0.004	-5.0	-6.6	**-5.3**	-0.013	0.017
ue001	**-0.059**	**-0.014**	**-0.078**	**-0.019**	0.3	**-16.7**	-0.3	**-0.111**	**-0.037**
ue001P$_{34}$	**-0.020**	0.000	-0.016	-0.002	**-8.9**	**-12.4**	**-8.0**	-0.009	-0.014
ue000P$_{34}$	0.004	0.009	0.009	0.005	**-10.6**	**-11.3**	**-11.3**	0.004	**0.025**
Historical	0.158	0.073	0.193	0.084	73.9	36.7	5.3	0.178	0.044
Se	0.005	0.005	0.008	0.008	2.53	3.91	1.92	0.009	0.010

The simulation with only temperature in the feature vector ('ue001') leads to undesirable large biases in relevant precipitation statistics (name in red in Tables 2.2 and 2.3) while the simulations with only precipitation in the feature vector ('ue000R$_6$' and 'ue000P$_{34}$') lead to undesirable biases in relevant temperature statistics (names in red in Tables 2.4 and 2.5). In these three simulations a severe underestimation of the autocorrelation structure leads to a severe underestimation of the monthly standard deviation. This again demonstrates the

importance of including both precipitation and temperature characteristics in the feature vector.

As expected, the influence of the composition of the feature vector on the reproduction of the pattern correlation of temperature is relatively small (see Tables 2.4 and 2.5). Also here some additional bias in the daily and monthly standard deviations is introduced in the 'R_6' simulations, but in contrast to precipitation the bias in the autocorrelation is not reduced but enhanced.

Table 2.4. Differences between statistical properties of the simulated time series and the historical records (1961–1995) for **winter temperature** (October–March). For the lag 1 and 2 pattern correlation coefficients (r_p), mean temperature, and the mean lag 1 and 2 autocorrelation coefficients ($\bar{r}$) the absolute differences are given, and for the mean standard deviations of monthly and daily values ($\bar{s}_M$ and $\bar{s}_D$) the percentage differences. Values between ⟨⟩ denote averages over the 34 stations (details in Beersma, 2002). Bottom lines: average historical estimates (mean and standard deviations in °C) and their standard error *se* (standard errors for the mean in °C, for standard deviations in % and for the autocorrelation coefficients dimensionless). Values in bold refer to differences more than 2 × *se* from the historical estimate, where *se* is calculated as in Beersma and Buishand (1999).

	34 stations		6 regions						
Simulation	$r_p(1)$	$r_p(2)$	$r_p(1)$	$r_p(2)$	Mean	$\langle\Delta\bar{s}_M\rangle$	$\langle\Delta\bar{s}_D\rangle$	$\langle\Delta\bar{r}(1)\rangle$	$\langle\Delta\bar{r}(2)\rangle$
ue241	**-0.129**	**-0.088**	**-0.182**	**-0.122**	0.04	-5.8	-2.4	**-0.045**	-0.006
ue001R_6	**-0.129**	**-0.094**	**-0.165**	**-0.114**	0.20	**-13.5**	-4.0	**-0.064**	**-0.044**
ue000R_6	**-0.131**	**-0.098**	**-0.165**	**-0.111**	0.20	**-43.1**	-3.6	**-0.360**	**-0.400**
ue001	**-0.153**	**-0.110**	**-0.208**	**-0.146**	0.00	-2.3	-1.3	**-0.044**	-0.002
ue001P_{34}	**-0.143**	**-0.114**	**-0.157**	**-0.106**	0.51	-11.2	-4.1	**-0.055**	-0.030
ue000P_{34}	**-0.131**	**-0.107**	**-0.146**	**-0.100**	0.16	**-41.3**	4.4	**-0.334**	**-0.376**
Historical	0.742	0.620	0.584	0.413	3.6	2.1	4.2	0.826	0.639
se	0.004	0.007	0.006	0.009	0.17	6.16	2.49	0.007	0.015

Table 2.5. As Table 2.4 but for the **summer temperature** (April—September).

	34 stations		6 regions						
Simulation	$r_p(1)$	$r_p(2)$	$r_p(1)$	$r_p(2)$	Mean	$\langle\Delta\bar{s}_M\rangle$	$\langle\Delta\bar{s}_D\rangle$	$\langle\Delta\bar{r}(1)\rangle$	$\langle\Delta\bar{r}(2)\rangle$
ue241	**-0.076**	**-0.056**	**-0.170**	**-0.122**	0.11	-3.2	0.1	**-0.025**	0.009
ue001R_6	**-0.074**	**-0.058**	**-0.157**	**-0.116**	**0.25**	-2.7	1.3	**-0.035**	-0.013
ue000R_6	**-0.077**	**-0.060**	**-0.165**	**-0.121**	0.12	**-33.3**	1.6	**-0.339**	**-0.272**
ue001	**-0.086**	**-0.061**	**-0.189**	**-0.133**	-0.03	5.6	-0.5	**-0.029**	**0.027**
ue001P_{34}	**-0.083**	**-0.071**	**-0.156**	**-0.119**	**0.65**	2.8	**4.0**	**-0.027**	-0.003
ue000P_{34}	**-0.077**	**-0.062**	**-0.163**	**-0.119**	0.15	**-30.3**	**3.0**	**-0.318**	**-0.321**
Historical	0.785	0.713	0.577	0.418	14.3	1.5	3.6	0.771	0.533
se	0.003	0.004	0.011	0.013	0.12	4.34	1.20	0.006	0.011

2.3 Conclusions

In this section the influence of the composition of the feature vector on the reproduction of the pattern correlation is assessed. Compared to the reference simulation (ue241), both the 'R_6' and 'P_{34}' type simulations reduce the biases in the precipitation pattern correlation coefficients but they do generally enhance the biases in the daily standard deviations, both for precipitation and temperature. Three simulations: ue000R_6, ue001 and ue000P_{34} largely fail on either one or more relevant statistics. That leaves two candidates that are serious alternatives for the reference simulation: ue001R_6 and ue001P_{34}. Of these two simulations, ue001R_6 performs best overall, i.e. in terms of the relevant statistics for precipitation and temperature and for winter and summer. With ue001R_6 therefore a 10 000-year simulation is performed for comparison with the earlier 10 times 1000-year simulation with ue241. Figure 2.2 presents the maxima of basin-average 10-day precipitation in the historical 1961 - 1995 series and those in the 10 000-year simulated series for the hydrological winter. It can be seen that the maxima in the ue001R_6 simulation are slightly smaller than in the ue241 simulation. It would be interesting to investigate the effect of the ue001R_6 simulation (with a smaller bias in the precipitation pattern correlations) on the (extreme) discharge simulation with GRADE.

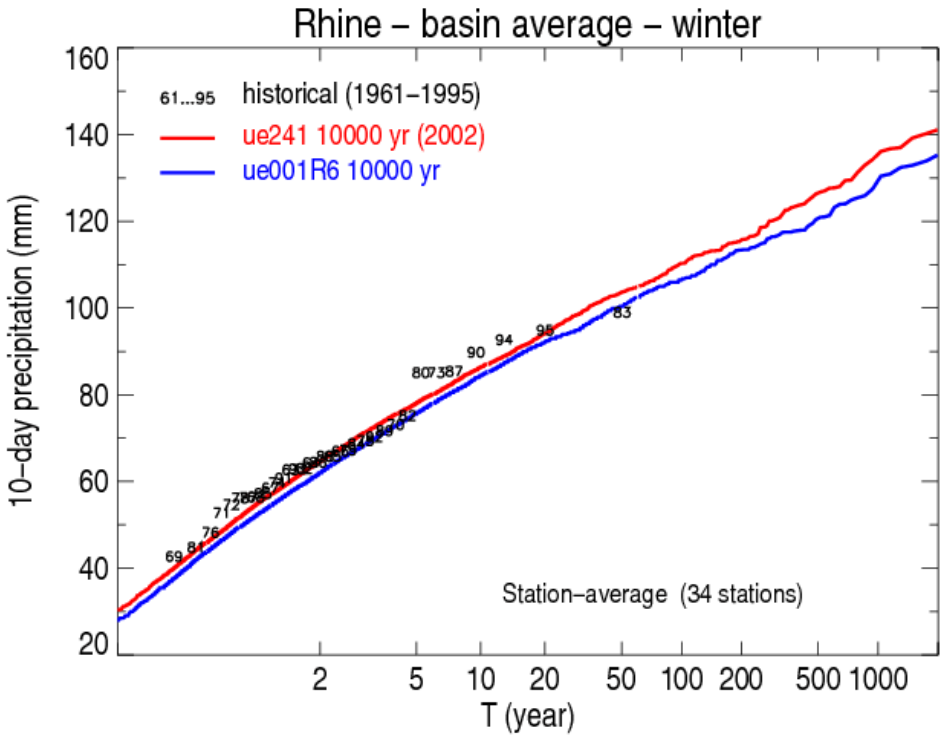


Figure 2.2. Gumbel plot of the maxima of basin-average 10-day precipitation in the historical 1961 - 1995 series and those in two 10 000-year simulated series for the winter (October to March).

For comparison Figure 2.3 shows a similar plot as Figure 2.2 but where the basin average is not obtained by averaging the 34 stations but as the area weighted average of the passively simulated 134 HBV sub-basins (see Section 3 for details). The basin-average obtained from the 134 sub-basins is systematically larger than that from the 34 stations, but again the ue001R6 simulation gives somewhat smaller basin-averages than the ue241 simulation.

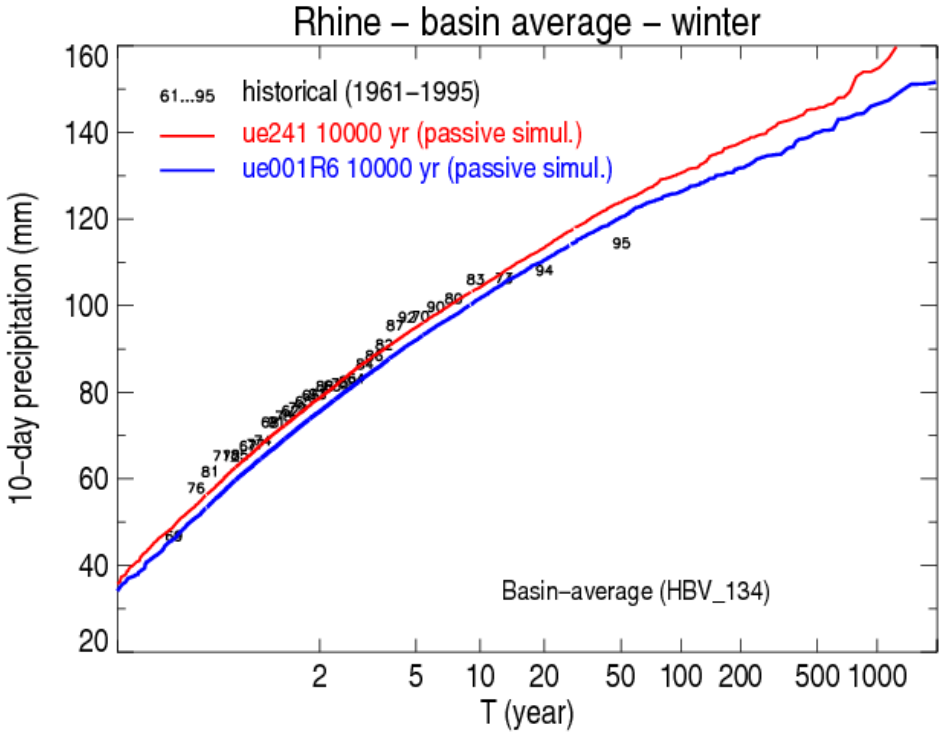


Figure 2.3. As figure 2.2 but basin average determined as the area weighted average of the passively simulated 134 HBV sub-basins (see Section 3 for details).

3. Rainfall simulations for the Rhine basin based on catchment average data rather than station data

In the simulations with the Rainfall generator for the Rhine basin described in Beersma (2002) and in the previous section the resampling procedure is 'driven' by the precipitation and temperature data of the 34 stations that are used in the feature vector. However for coupling with the hydrological model in GRADE simulated time series of precipitation and temperature for 134 HBV sub-basins rather than 34 stations are required. These time series for the 134 HBV sub-basins can easily be obtained from the simulation based on the 34 stations. In the simulated series each simulated day corresponds with a historical date. So, if for every historical date in the simulation also the precipitation and temperature data for the 134 sub-basins are available, the 34 station data can simply be replaced by the 134 HBV sub-basin data. This type of simulation is referred to as passive simulation. Passive refers to the fact that the 134 sub-basins are not directly represented in the feature vector and therefore do not directly or actively 'drive' the simulation. This distinction between active simulation of the station data and passive simulation of the sub-basin data has a historical background. At the start of the development of the Rainfall generator for the Rhine basin no sub-basin data were available at all, only station data. For the development of the first multi-site version of the rainfall generator use was made of 25 stations in German part of the Rhine basin (Brandsma and Buishand, 1999). Later, the area was extended to the whole upstream area of Lobith using the 34 stations (Wójcik et al.,2000, Beersma et al., 2001) but still no sub-basin data were included. In June 2002 the BfG made the HBV_134 sub-basin data for precipitation and temperature covering the period 1961 to 1995 available. These data became later known as the CHR-OBS data.

Up to now the performance of the resampling method for the Rhine basin has been assessed by comparing statistics from the historical series of station data with those from simulated series of station data. However no comparisons were performed between the statistics from the historical series of sub-basin data and those from *passively* simulated series of sub-basin data, despite the fact that the time series of sub-basin data are most relevant for coupling with the hydrological part in GRADE. In this paragraph this comparison is made. In Figure 3.1 the winter and summer maxima of basin-average 10-day precipitation in the historical 1961 – 1995 series and those in the 10 000 (10 x 1000) year series simulated with the rainfall generator are presented. For ease of comparison, the basin-average precipitation is determined in two different ways: i) as the average of the 34 stations (in red) and ii) as the area weighted average of the 134 sub-basins (in black). The colored numbers represent the historical year of the hydrological seasons (minus 1900) in the 1961 – 1995 period. The basin-average 10-day precipitation determined as the area weighted average of the 134 sub-basins (black numbers) is systematically larger than that determined as the average of the 34 stations (red numbers). It is expected that the sub-basin-average better represents the real basin-average than the station-average simply because the sub-basin data makes use of a much larger number of stations. This is also confirmed by the two most extreme 10-day events in the hydrological winters of 1994 and 1995 (which correspond to the December 1993 and January 1995 floods respectively) and which are known to be the largest in the past 50 years. More importantly from Figure 3.1 is that (in terms of reproducing the distribution of the seasonal maxima of the basin-average 10-day precipitation) the *passive* simulation of sub-basin data (black-red dashed line) performs as well as the active simulation of station data (solid red line)[2].

Now the sub-basin data are available it is also possible to drive the rainfall generator actively with the sub-basin data (by using the sub-basin data in the feature vector rather than the station data). This is exactly the configuration of the rainfall generator that has been used when the rainfall generator methodology is applied to (regional) climate model data, because gridded data rather than station data are available. In the RheinBlick2050 project (Görgen et al., 2010) the rainfall generator methodology was applied to a number of RCM simulations where the RCM grid-data was interpolated to the 134 HBV sub-basins first and these sub-basin data were used in the feature vector. As a reference also a 3000-yr simulation was performed using the 1961 – 1995 CHR-OBS sub-basin data in the feature vector of the rainfall generator, i.e. a simulation in which the sub-basin data is *actively* simulated. The composition of the feature vector for this simulation was 'as similar as possible' to the simulations described in Beersma (2002) and the ue241 simulation in the previous section, i.e. the average (standardized) precipitation of the 134 sub-basins, the fraction of the sub-basins with precipitation and the average (standardized) temperature of the sub-basins[3]. The results of this active simulation of sub-basin data, again in terms of the seasonal maxima of basin-average 10-day precipitation, are presented in Figure 3.1 as the solid black lines. Compared with the standard *passive* simulation of the sub-basin data (black-red dashed lines) the active simulation of the sub-basin data looks very similar. Only for the hydrological summer and for return periods longer than about 200 years the two types of simulations start to deviate

[2] Note that in the top panel of Figure 3.1 the solid red line is identical to the solid red line in Figure 2.2 and that the black-red dashed line is identical to the red line in Figure 2.3.

[3] In the ue241 simulations the weights for the feature vector elements were set manually to $W(P,F,T) = (2, 4, 1)$ while in this simulation they are set automatically, and are inversely proportional to the variance of the feature vector element which resulted in $W(P,F,T) \approx (2, 6.5, 1)$, i.e. a slightly larger weight for the fraction of wet sub-basins. Note that since the fraction F is determined from a much larger number when using sub-basins (rather than only 34 stations), it might be expected that its variance becomes smaller, which may at least partly justify the larger weight.

systematically. Apart from this deviation, this is another indication that the passive simulation of the sub-basin data works satisfactorily.

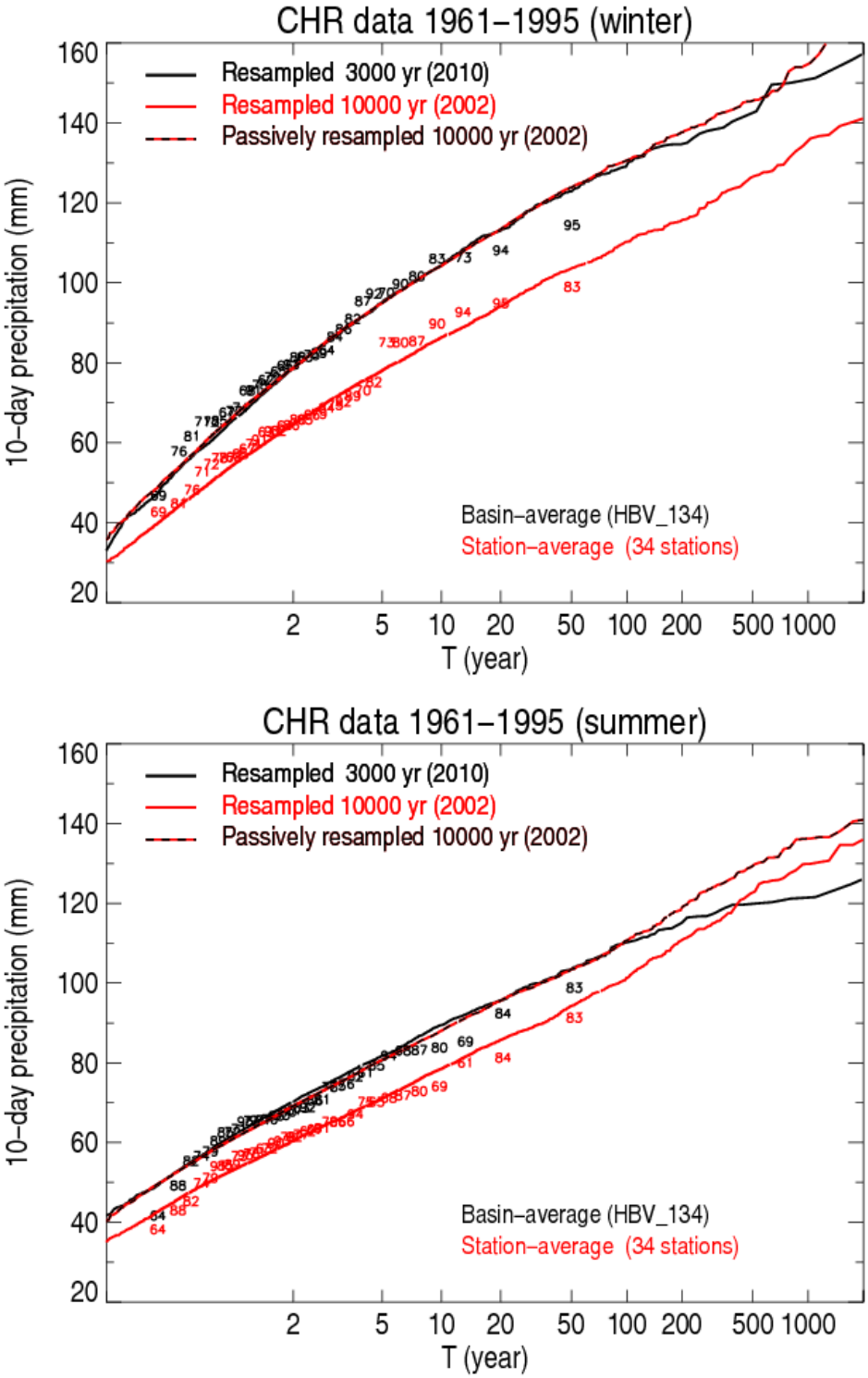


Figure 3.1. Gumbel plots of winter and summer maxima of the 10-day Rhine-basin-average precipitation amount for actively and passively resampled series (respectively solid and dashed lines) and for averaging over 34 precipitation stations (station-average) and averaging over 134 sub-basins (basin-average). The colored numbers refer to the seasonal maxima in the historical data for the period 1961 – 1995; i.e. year minus 1900. In red: 34 station average (station-average); in black: 134 sub-basin average (basin-average). Winter (top panel) refers to the hydrological winter (October to March) and summer to the hydrological summer (April to September).

Table 3.1 compares for winter and summer precipitation, in a similar way as in Tables 2.2 and 2.3, the reproduction of the pattern correlation and the standard statistics for the active simulation of the sub-basin data, which is denoted as ueHBV134, with the reproduction of these quantities for the reference ue241 simulation. In terms of these statistics the results for both types of simulations are quite similar.

Tabel 3.1. Differences between statistical properties of the simulated time series and the historical records (1961–1995) for winter (October–March) and summer (April–September) **precipitation**. For the lag 1 and 2 pattern correlation coefficients (r_p), mean precipitation (monthly totals), and the mean lag 1 and 2 autocorrelation coefficients ($\bar{r}$) the absolute differences are given, and for the mean standard deviations of monthly and daily values ($\bar{s}_M$ and $\bar{s}_D$) the percentage differences. Values between ⟨⟩ denote averages over the 34 stations (details in Beersma, 2002). 'Historical' and '*se*': average historical estimates (mean and standard deviations in mm) and their standard error *se* (standard errors for the mean in mm, for standard deviations in % and for the autocorrelation coefficients dimensionless). Values in bold refer to differences more than 2 × *se* from the historical estimate, where *se* is calculated as in Beersma and Buishand (1999).

	34 stations		6 regions						
Simulation	$r_p(1)$	$r_p(2)$	$r_p(1)$	$r_p(2)$	Mean	$\langle\Delta\bar{s}_M\rangle$	$\langle\Delta\bar{s}_D\rangle$	$\langle\Delta\bar{r}(1)\rangle$	$\langle\Delta\bar{r}(2)\rangle$
Hydrological winter									
ue241	**-0.086**	**-0.044**	**-0.096**	**-0.049**	-0.8	-3.4	-0.7	**-0.036**	-0.009
ueHBV134	**-0.093**	**-0.050**	**-0.103**	**-0.051**	-1.6	-4.6	-1.1	**-0.035**	-0.010
Historical	0.253	0.161	0.234	0.133	64.1	35.8	4.2	0.285	0.144
se	0.005	0.005	0.008	0.008	2.47	4.53	2.46	0.008	0.009
Hydrological summer									
ue241	**-0.061**	**-0.014**	**-0.083**	**-0.022**	-0.7	**-8.9**	-1.3	**-0.029**	0.008
ueHBV134	**-0.064**	**-0.016**	**-0.089**	**-0.026**	-0.6	-6.7	-0.7	-0.014	0.016
Historical	0.158	0.073	0.193	0.084	73.9	36.7	5.3	0.178	0.044
se	0.005	0.005	0.008	0.008	2.53	3.91	1.92	0.009	0.010

4. Passive simulation of precipitation and temperature in the Meuse basin based on simulations for the Rhine basin

So far two different rainfall generators were developed, one for the Rhine basin and one for the Meuse basin. Both rainfall generators are based on a resampling technique known as Nearest-Neighbour resampling. In both cases the rainfall generators are 'fitted' to the basin of interest, i.e. the Rainfall generator for the Rhine basin only uses data in the Rhine basin in the feature vector and similarly the Rainfall generator for the Meuse uses data in the Meuse basin or close to the Meuse basin in the feature vector. Methodologically, both rainfall generators are the same but there are a few differences of which the most important are: 1) the composition of the feature vector is somewhat different, i.e. the Rainfall generator for the Meuse basin uses a 4-day memory term in the feature vector and the Rainfall generator for the Rhine does not, 2) the Rainfall generator for the Rhine basin uses station data from the historical period 1961 – 1995 to drive the simulations while the Rainfall generator for the Meuse basin uses station data from either the historical period 1961 – 1998 (denoted as

sim61) or the period 1930 – 1998 (denoted as sim30)[4], and 3) for the Rainfall generator for the Rhine basin the 134 sub-basin data are available for exactly the same period as the driving station data so there is a one-to-one correspondence between actively simulated station data and the passively simulated sub-basin data. However, for the Rainfall generator for the Meuse basin the 15 sub-basin data were available for precipitation for 1961 – 1998 and for temperature only for 1967 – 1998. Leander and Buishand (2004) applied a second nearest-neighbour search to cope with incomplete sequences for the base period.

Using simulations with the Rainfall generator for the Rhine basin, it is possible to passively simulate the 15 sub-basin data for the *Meuse* basin (in the same way that the 134 sub-basin data for the Rhine basin are passively simulated). It is expected that such a simulation makes sense since the Meuse basin is similar in size as the Mosel basin (which is part of the Rhine basin) and that the Meuse basin is located just west of the Mosel basin. The only extra thing that is needed is an additional nearest-neighbour search to simulate the sub-basin temperature for simulated dates in the 1961 – 1968 period[5]. An advantage of such a simulation is that the generated time series for the Rhine and Meuse basins are spatially and temporally coherent. This coherence is not possible when the Rainfall generators for the Rhine and Meuse basins are run separately. For some applications, e.g. when the simultaneous occurrence of floods in both the Rhine and Meuse basin is relevant this spatial-temporal coherence is an important characteristic.

In this section the performance of the passive simulation of the Meuse sub-basin data based on the ten 1000-year simulations for the Rhine basin is compared with the passive simulation of the Mosel sub-basin data (which is part of the Rhine basin) based on the same ten 1000-year simulations. For this purpose Figure 4.1 presents the hydrological winter and summer maxima of the 4-day, 10-day and 20-day basin-average precipitation for the Mosel basin for the historical data (1961 – 1995) and the ten 1000-year simulations. Figure 4.2 presents similar results for the Meuse basin. Overall the performance for the Meuse basin is not much worse than for the Mosel basin. For the 10- and 20-day sums in the Meuse basin in the hydrological winter there seems to be a discrepancy between the most extreme historical events (December 1993, in the figure presented as '94', and January 1995) and the simulated maxima. A similar difference was also found in the simulations with the Rainfall generator for the Meuse basin regarding the 10- and 30 day sums (Leander and Buishand, 2004). However, the 10-day winter amounts for return periods between 2 and 20 years in those simulations correspond better with the historical data than those in Figure 4.2. For the Meuse basin the maxima of the 20-day amounts in the hydrological summer seem to be systematically underestimated in these passive simulations, a discrepancy that is not found in the 30-day amounts presented in Leander and Buishand (2004). This result may be related to the absence of a memory term in the feature vector of the Rainfall generator for the Rhine basin or alternatively to a lower spatial dependence of summer precipitation. Further research is needed to assess this.

[4] Note, that since early 2011 also Rainfall simulations for the Meuse basin using station data from the period 1930 – 2008 (denoted as sim30-08) are available (Buishand and Leander, 2011).

[5] In this case the additional nearest-neighbour search was based on the four temperature stations (Aachen, Langres, Reims and Uccle) for which the data were made available for the full base period of the Rainfall generator for the Rhine basin (1961 – 1995). Note that this additional nearest-neighbour search somewhat differs from that applied by Leander and Buishand (2004) and that the sub-basin temperature data for 1967 and 1968 were not used in the present study.

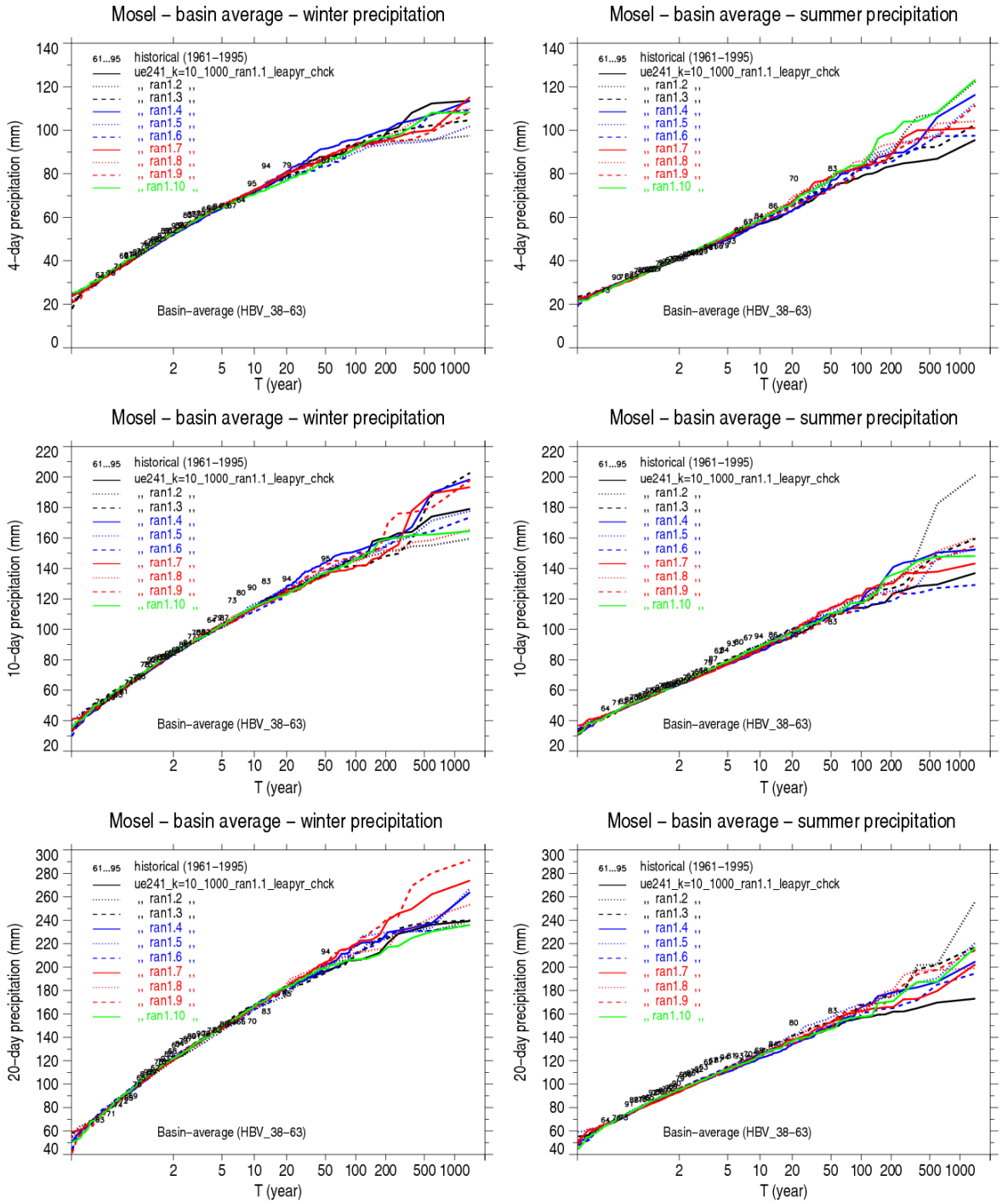


Figure 4.1. Gumbel plots of maxima of 4-day, 10-day and 20-day basin-average precipitation for the Mosel basin for the historical data (1961 – 1995) and for passive simulations based on the ten 1000-year simulations for the Rhine basin, both for the hydrological winter (October to March) and the hydrological summer (April to September). Note, the Mosel-basin-average corresponds with the area weighted average of the corresponding 25 HBV-Rhine sub-basins.

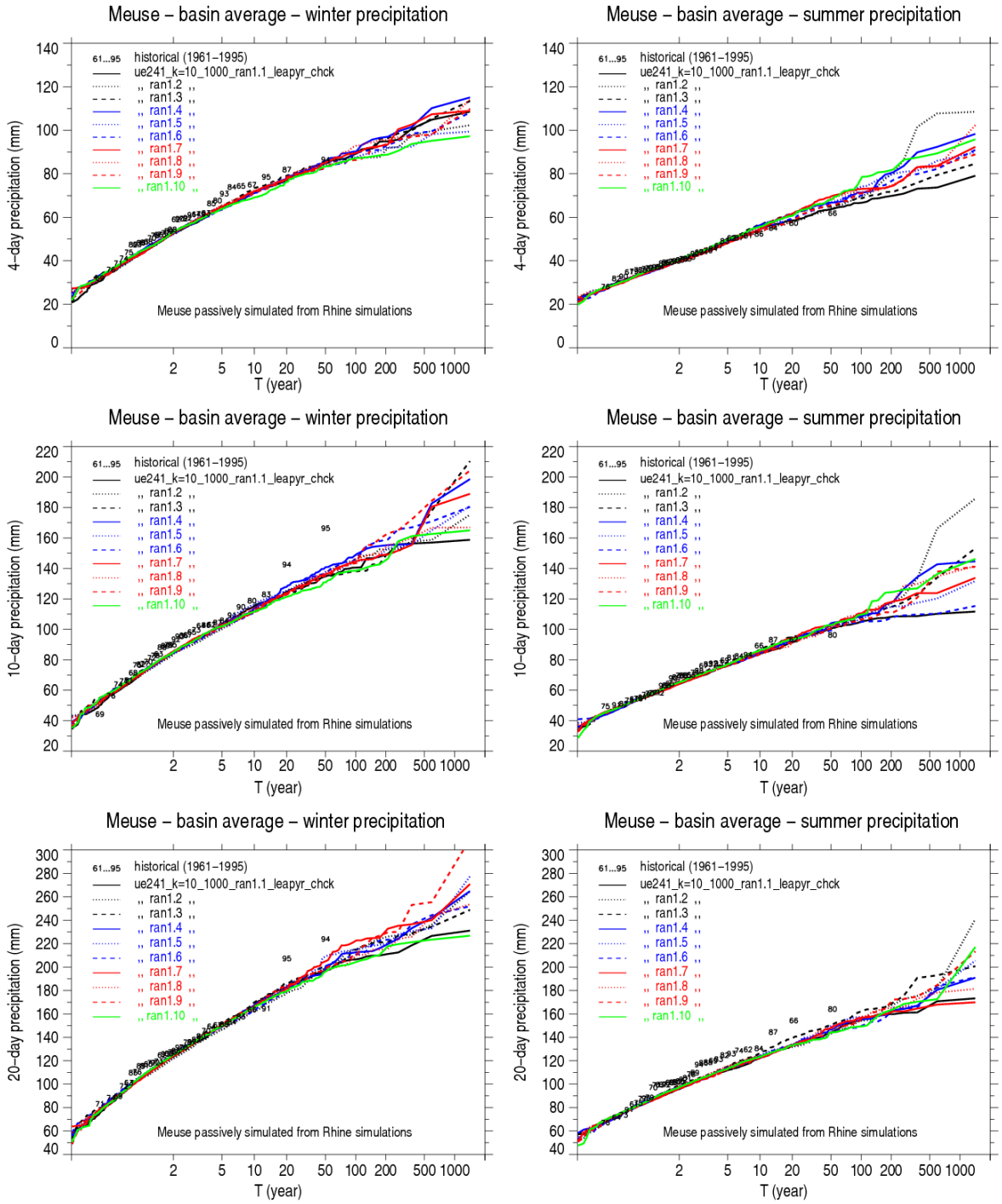


Figure 4.2. As Figure 4.1 but for passive simulations for the Meuse basin based on the ten 1000-year simulations for the Rhine basin. Note, the Meuse-basin-average corresponds with the area weighted average of the 15 HBV-Meuse sub-basins.

Figures 4.3 and 4.4 show for the Mosel and Meuse basins the basin-average autocorrelation coefficients (i.e. an area-weighted average of the coefficients for each individual sub-basin) in the winter (October to March) for the historical (1961 – 1995) data and the passive simulations based on the ten 1000-year simulations for the Rhine basin. Though for both river basins the *N*-day winter maxima (in Figures 4.1 and 4.2) are realistically simulated, the basin-average lag 1 autocorrelation coefficients are underestimated. For the Meuse basin the

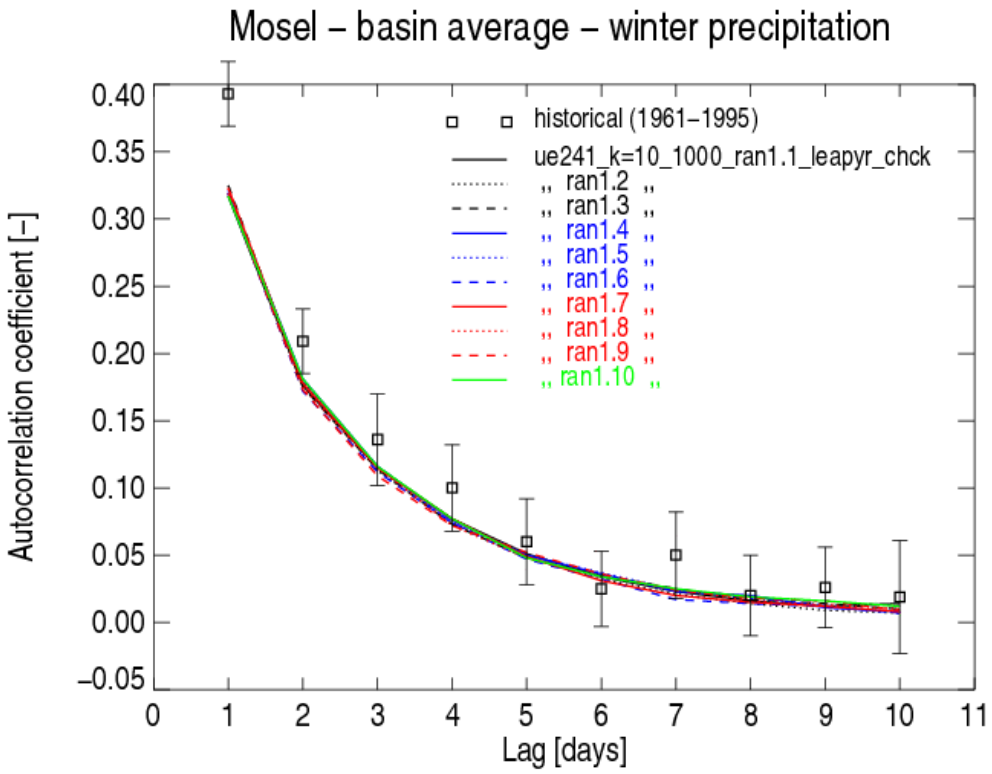


Figure 4.3. Basin-average autocorrelation coefficients for the Mosel basin for the hydrological winter (October to march) for ten 1000-year simulations compared with those for the historical (1961 – 1995) data. The error bars correspond with 2 × *se*-intervals. The standards errors se were calculated using a jackknife technique (Buishand and Beersma, 1993).

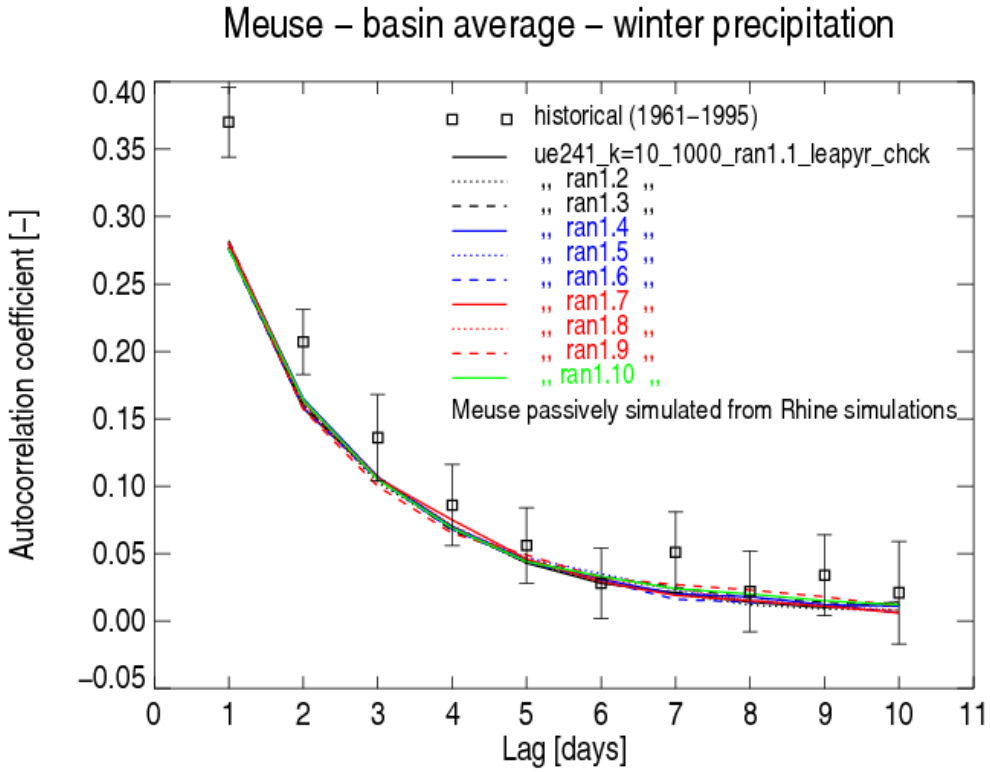


Figure 4.4. As Figure 4.3 but for passive simulations for the Meuse basin based on the ten 1000-year simulations for the Rhine basin. Note, the Meuse-basin-average corresponds with the area weighted average of the 15 HBV-Meuse sub-basins.

underestimation is somewhat larger than for the Mosel basin. This underestimation is also worse in comparison to the (slight) underestimation of the basin-average lag-1 autocorrelation in simulations with the Rainfall generator for the Meuse basin (see Figure 2.2 in Leander, 2009). Leander (2009) presents in his figure two types of simulations one with and one without a 4-day memory term in the feature vector. For higher lags (5-8 days) the results presented here are somewhat better than the results for the simulation without memory term in Leander (2009) but not as good as the results for the simulation with the memory term. This indicates that for the winter the memory term is a relevant element which is by definition 'omitted' in the passive simulations for the Meuse basin based on the simulations for the Rhine basin.

For the Mosel basin, finally, the basin-average autocorrelation coefficients were also calculated for the ue001R_6 simulation introduced in Section 2. For this simulation the underestimation of the lag-1 autocorrelation is about half that for the ue241 simulations presented in Figure 4.3 (not shown).

The results of the passive simulations for the Meuse basin are satisfactory, in particular for extreme discharges in winter. It may therefore be worthwhile to explore a unified Rhine-Meuse rainfall generator further.

References

Beersma, J.J., 2002. Rainfall generator for the Rhine basin: Description of 1000-year simulations. Publication 186-V, KNMI, De Bilt.

Beersma, J.J., 2007. Extreme hydro-meteorological events and their probabilities. PhD-thesis, ISBN 90-8504-615-7.

Beersma, J.J. and T.A. Buishand, 1999. Rainfall generator for the Rhine basin: Nearest-neighbour resampling of daily circulation indices and conditional generation of weather variables. Publication 186-III, KNMI, De Bilt.

Beersma, J.J., T.A. Buishand and R. Wójcik, 2001. Rainfall generator for the Rhine basin: multi-site simulation of daily weather variables by nearest-neighbour resampling. In: Generation of hydrometeorological reference conditions for the assessment of flood hazard in large river basins, P. Krahe and D. Herpertz (Eds.), pp. 69-77, CHR-Report No. I-20, CHR, Lelystad, The Netherlands, ISBN 90-36954-18-5.

Brandsma, T. and T.A. Buishand, 1999. Rainfall generator for the Rhine basin: Multi-site generation of weather variables by nearest-neighbour resampling. Publication 186-II, KNMI, De Bilt.

Buishand, T.A. and J.J. Beersma, 1993. Jackknife tests for differences in autocorrelation between climate time series. Journal of Climate, 6, 2490-2495.

Buishand, T.A. and R. Leander, 2011. Rainfall generator for the Meuse basin: Extension of the base period with the years 1999 – 2008. Publication 196-V, KNMI, De Bilt.

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, The Netherlands, ISBN 90-70980-35-1.

Leander, R., 2009. Simulation of precipitation and discharge extremes of the river Meuse in current en future climate. PhD-thesis, ISBN 978-90-393-5045-4.

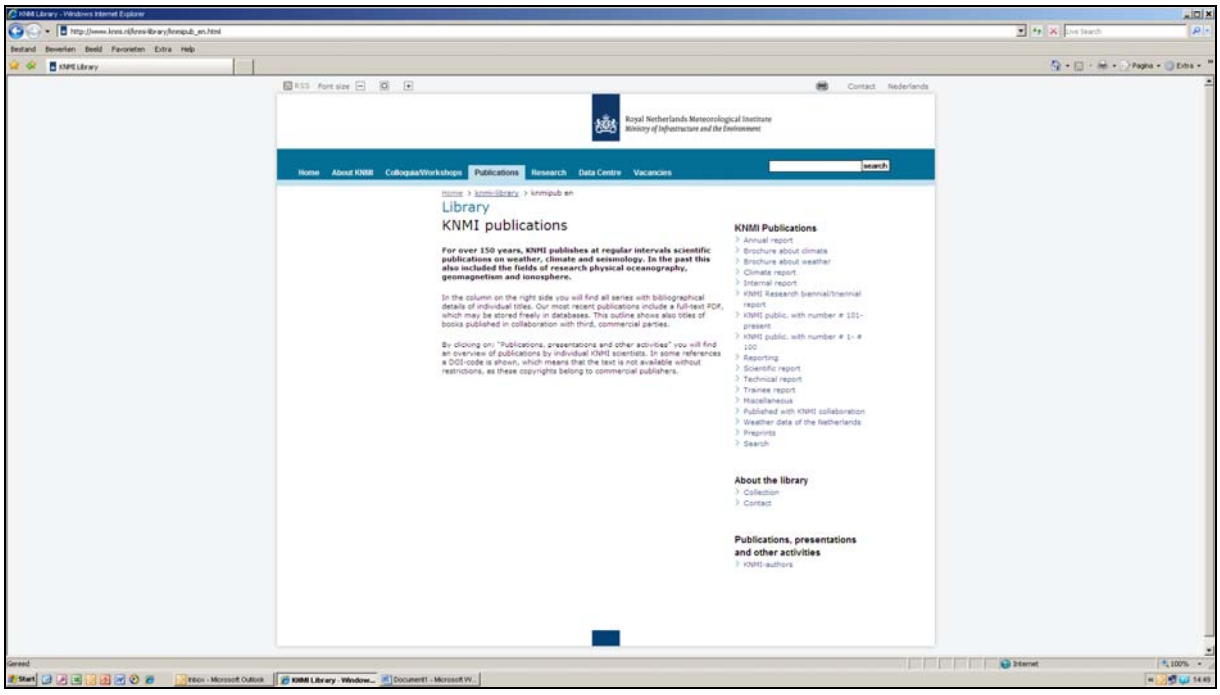
Leander, R. and T.A. Buishand, 2004. Rainfall generator for the Meuse basin: Development of a multi-site extension for the entire drainage area. Publication 196-III, KNMI, De Bilt.

Wit, M.J.M. 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, The Netherlands

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. Publication 186-IV, KNMI, De Bilt.

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.