# Rainfall Generator for the Rhine Basin

## Multi-site generation of weather variables for the entire drainage area

Rafał Wójcik

Jules J. Beersma

T. Adri Buishand

KNMI publication 186–IV

Work performed under contract RI-2041 to Ministry of Transport, Public Works and Water Management, Institute for Inland Water Management and Waste Water Treatment RIZA, P.O. Box 17, 8200 AA Lelystad (The Netherlands) Telephone: +31.320.298411; Telefax: +31.320.249218

# Contents

# Summary

This is the final report of a project on the development of a rainfall generator for the Rhine basin. The request for this generator arose from the need to study the likelihood of extreme river discharges in the Netherlands, using a hydrological/hydraulic model. Long-duration, multi-site simulations of daily precipitation and temperature time series for the entire Rhine basin are needed for this purpose. Temperature is included to determine snow accumulation and snowmelt. Daily precipitation and temperature data from 36 stations in Germany, Luxemburg, France and Switzerland for the period 1961-1995 are considered. Studies in earlier reports dealt with the German part of the basin only.

Time series simulation is done by nearest-neighbour resampling. The method does not make restrictive assumptions about the underlying joint distribution of the multi-site precipitation and temperature data. In order to generate weather variables for day $t$, a feature vector $\mathbf{D}_t$ is formed to find the nearest neighbours of this day, or the previous day, in the historical data. For unconditional simulations $\mathbf{D}_t$ contains variables that characterize the weather on the previous day $t-1$ (first order model) or a number of previous days (higher order model). Circulation indices for day $t$ are included in $\mathbf{D}_t$ in the case of conditional simulation on the atmospheric circulation. A finite number $k$ of nearest neighbours in terms of a weighted Euclidean or the Mahalanobis distance is selected from the historical record. One of these $k$ nearest neighbours is finally "resampled" using a discrete probability kernel.

The criteria used to assess the performance of different simulation methods are the ability to reproduce the second-order moment statistics of daily and monthly values of precipitation and temperature and the distribution of multi-day winter (October-March) precipitation amounts. For the high-elevation stations in Germany and Switzerland the snowmelt simulation is also evaluated. First-order conditional and unconditional models for the generation of daily precipitation and temperature are considered. Conditional simulations were done with simulated circulation indices produced by a separate second-order resampling model.

With respect to the reproduction of the above mentioned precipitation and temperature statistics unconditional simulations perform better than conditional simulations. Inclusion of circulation indices in the feature vector for unconditional simulations worsens the reproduction of the temperature statistics. Due to minor modifications in the resampling model the second-order moment statistics of precipitation were somewhat better preserved in the conditional simulations than in those for the German part of the Rhine basin in earlier reports. As a result the reproduction of multi-day winter maximum precipitation also compares favourably with the conditional simulations in earlier studies.

For the unconditional models the reproduction of multi-day maximum snowmelt was satisfactory despite the slight underprediction of the temperature autocorrelation. The conditional simulations showed a significant underestimation (up to 20-30%) of the median and the upper quintile mean of multi-day snowmelt amounts at four of the six high-elevation stations.

Realistic multi-day precipitation amounts much larger than the largest historical precipitation amounts were generated in simulation runs of 1000 years. The largest events in such runs are subject to large sampling variability. This variability can be reduced by averaging over several simulation runs.

Unlike the weighted Euclidean metric the use of the Mahalanobis metric does not require the specification of scaling weights and thus reduces the number of parameters involved in the resampling algorithm. The results obtained in this report show that the simulations with the Mahalanobis metric compare well to those with the weighted Euclidean metric.

# 1. Introduction

## 1.1. Background

The Rhine is the most important river in the Netherlands. The river originates in the Swiss Alps. Large parts of its catchment area are situated in Switzerland, Germany, France and the Netherlands. Small parts of Austria, Belgium and almost the whole country of Luxemburg also drain to the river. Protection against flooding is a point of continuous concern. According to safety standards, laid down in the Flood Protection Act, measures against flooding in the non-tidal part of the Rhine in the Netherlands have

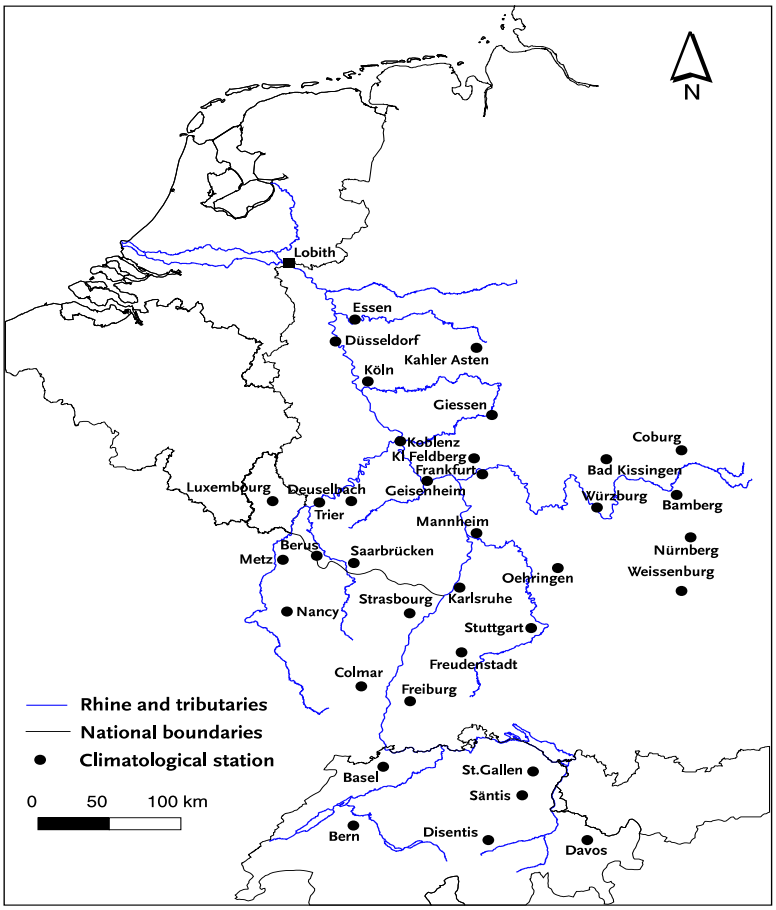


Figure 1.1: Location of Lobith in the Netherlands and the 36 stations in the drainage area of the river Rhine used in this study.

to withstand a discharge that is exceeded on average once in 1250 years. Traditionally this design discharge has been obtained from a statistical analysis of peak discharges (data from 1901 onwards) at Lobith, where the river enters the country. Several probability distributions have been fitted to the discharge maxima of that record. The long return period requires an extrapolation far beyond the length of the observed discharge record. Different distributions then usually lead to quite different design discharges. The fact that the parameters of these distributions have to be estimated from a finite sample introduces yet another uncertainty.

In the most recent re-evaluation of the design discharge at Lobith, there was a strong feeling that the uncertainties of extrapolation could be reduced by taking the physical behaviour of the river basin into account [Delft Hydraulics and EAC-RAND, 1993]. For this purpose, it was suggested to develop a hydrological/hydraulic model for the whole basin. With such a model, it would also be possible to quantify the effects of changes in the catchment and the river geometry and to predict the potential impact of climate change. The Institute for Inland Water Management and Waste Water Treatment (RIZA) adopted this idea in a research plan for a new methodology to determine the design discharge [Bennekom and Parmet, 1998]. Besides the hydrological/hydraulic model, the development of a stochastic rainfall generator was also planned in order to produce long-duration rainfall series over the basin. Unprecedented extreme rainfall events are expected if the simulation run is considerably longer (300-1000 years) than the observed rainfall record. Such rainfall events in turn, may lead to more extreme river discharges at Lobith than those experienced in the past century. The use of synthetic rainfall series in combination with a hydrological/hydraulic model does not only provide the peak discharges but also the durations of these extreme events. This may lead to a better insight into the profile of the design flood.

## 1.2. Previous research

Before the project started a feasibility study had been carried out by KNMI [Buishand and Brandsma, 1996]. Several techniques for the simulation of daily rainfall time series were considered with respect to multi-site applications in large river basins. One possible option was to describe the sequences of daily rainfall by a truncated multivariate first-order autoregressive process. Truncation is necessary to avoid the generation of negative rainfall amounts. Further, a power transformation is usually applied to achieve that the normal distribution can be used [Bárdossy and Plate, 1992]. These modifications complicate parameter estimation for the underlying multivariate autoregressive process. It is also not clear whether the generation of this process from a multivariate normal distribution is capable to reproduce the dependence of extreme values. The alternative to parametric time-series modelling is to use a non-parametric resampling technique. Non-parametric techniques are much more computationally intensive. Multi-site extensions are, however, rather straightforward. Partly because of the good results of the non-parametric nearest-neighbour resampling technique for generating weather variables of a single site in a draft of the paper of Rajagopalan and Lall [1999], it was decided to give priority to that approach.

In a first attempt towards the practical application of nearest-neighbour resampling

for the Rhine basin, the joint simulation of daily precipitation and temperature at single sites situated in the German part of the basin was investigated [Brandsma and Buishand, 1997, 1998]. The simulation of temperature is neccesary to account for the effect of snow and frozen soils on large river discharges. Much attention was given to the conditional simulation of the weather variables on atmospheric flow indices. This kind of simulation is particularly relevant to study the impact of past or future changes in the atmospheric circulation on precipitation and river discharges. The performance of nearest-neighbour resampling was evaluated with respect to the reproduction of temperature and rainfall autocorrelation coefficients and the $N$-day maximum rainfall and snowmelt distributions. The results were encouraging enough to proceed with a multi-site extension [Brandsma and Buishand, 1999]. Despite the emphasis on large-scale features in that extension, the reproduction of autocorrelation coefficients and extremes at single sites remained satisfactory. For the multi-site version it further turned out that unconditional simulation of precipitation, temperature and circulation indices performed somewhat better than conditional simulation of precipitation and temperature on circulation indices.

Both in the single-site and multi-site studies, a number of long-duration unconditional simulations were performed to show that much higher multi-day precipitation amounts can be generated than ever observed in the past. For example, for one particular area in the south of Germany a 10-day precipitation maximum almost 70% larger than the historical maximum was found in a multi-site 1000-year simulation [Brandsma and Buishand, 1999].

Initially, simulation of daily precipitation and temperature conditional on circulation indices was restricted to the length of the historic record of circulation indices. Several separate stochastic models for generating these indices were then developed [Beersma and Buishand, 1999a] to make long-duration conditional simulations possible. It was shown that the persistence of the atmospheric circulation was quite well preserved by the simulated indices. The performance of the precipitation and temperature simulations conditional on these indices fell, however, behind that for the unconditional ones. At best an underestimation of extreme value properties of precipitation of 10% could be achieved. Attempts to surpass this shortcoming remained without success.

## 1.3. Scope and objectives

The present report is the final report on the development of the rainfall generator for the river Rhine. The report deals with multi-site simulations of daily precipitation and temperature time series for the whole Rhine basin. Besides the data from Germany used in earlier studies, data from Luxemburgian, French and Swiss stations in the basin are included. The main objectives of this study are to:

- perform simulations for the whole Rhine basin and to compare the results with those for the German part
- evaluate snowmelt simulations for high Alpine stations situated in Switzerland

Three resampling models are considered: two for unconditional simulation and one for conditional simulation. The reproduction of second-order moment statistics of tem-

perature and precipitation and properties of extreme winter precipitation and snowmelt is examined. Additionally, an alternative metric is introduced in the nearest-neighbour resampling technique.

# 2. Data description

Daily temperature and precipitation data from 36 stations were used. The stations are distributed all over the Rhine basin: 25 stations in Germany, 1 station in Luxemburg, 4 stations in France and 6 stations in Switzerland (see Figure 1.1). For the 35-year study period (1961-1995) the data were provided, via the "International Commission for the Hydrology of the Rhine Basin", by the following institutions:

- Deutscher Wetterdienst
- Service de la météorologie et de l'hydrologie de Luxembourg
- Météo France
- Swiss Meteorological Institute

Some relevant characteristics of the above mentioned data set are displayed in Table 2.1. Most stations in Germany, Luxemburg and France are lowland stations with mean annual rainfall ranging from 500 to 900 mm. There are, however, two stations in Germany, Kahler Asten and Freudenstadt, with a much larger mean annual rainfall ($\approx$1500 mm). This is due to orographic enhancement. The Swiss station Säntis lying at an altitude of 2500 m is the only station which has a higher mean annual rainfall than Kahler Asten and Freudenstadt. It is further obvious from Table 2.1 that high-elevation stations have a relatively low mean annual temperature.

Figure 2.1 shows the annual cycle of precipitation and temperature at a number of stations. At lowland stations the monthly mean temperature is always greater than $0^oC$. Mean monthly rainfall is more or less evenly distributed over the year for stations in the northern part of the basin. Exceptions are Kahler Asten and Freudenstadt which have much precipitation in winter. For stations in the southern part of the basin mean monthly rainfall in summer exceeds that in winter, which is mainly due to convection. More details on climate characteristics of the Rhine basin can be found in CHR [1978].

Because precipitation $P$ and temperature $T$ depend on the atmospheric flow, three daily circulation indices are also considered: 1) relative vorticity $Z$; 2) strength of the westerly flow $W$ and 3) strength of the southerly flow $S$. These circulation indices were computed from daily mean sea-level pressure data on a regular $5^o$ latitude and $10^o$ longitude grid. The derivation of the circulation indices is similar to that in Jones et al. [1993], except that the grid was centered at the Rhine basin instead of the British Isles.

Table 2.1: Characteristics of the stations that have been used in this study ( annual mean values are for the period 1961-1995)

| Nr | Station | Altitude [meters above m.s.l] | Mean annual temperature [$^{o}$C] | Mean annual precipitation [mm] |
|---|---|---|---|---|
| | GERMANY | | | |
| 1 | Stuttgart (Echterdingen) | 373 | 9.0 | 713 |
| 2 | Frankfurt | 112 | 9.9 | 645 |
| 3 | Kahler Asten | 839 | 5.0 | 1474 |
| 4 | Trier (Petrisberg) | 265 | 9.2 | 783 |
| 5 | Essen (Bredeney) | 152 | 9.7 | 928 |
| 6 | Bamberg | 239 | 8.6 | 632 |
| 7 | Freudenstadt | 797 | 6.8 | 1691 |
| 8 | Düsseldorf | 37 | 10.4 | 759 |
| 9 | Saarbrücken | 319 | 9.0 | 867 |
| 10 | Berus | 363 | 8.9 | 835 |
| 11 | Köln (Wahn) | 92 | 9.9 | 807 |
| 12 | Geisenheim | 118 | 10.0 | 542 |
| 13 | Koblenz (Horchheim) | 85 | 10.6 | 670 |
| 14 | Deuselbach | 480 | 8.0 | 808 |
| 15 | Freiburg | 269 | 10.9 | 944 |
| 16 | Giessen (Liebigshöhe) | 186 | 9.2 | 655 |
| 17 | Kl.Feldberg | 805 | 5.7 | 998 |
| 18 | Würzburg | 268 | 9.2 | 601 |
| 19 | Oehringen | 276 | 9.2 | 833 |
| 20 | Mannheim | 96 | 10.4 | 664 |
| 21 | Karlsruhe | 112 | 10.4 | 771 |
| 22 | Coburg | 322 | 8.2 | 738 |
| 23 | Bad Kissingen | 262 | 8.7 | 735 |
| 24 | Nürnberg | 310 | 8.9 | 640 |
| 25 | Weissenburg | 422 | 8.3 | 664 |
| | LUXEMBURG | | | |
| 26 | Luxembourg (Findel) | 380 | 8.5 | 862 |
| | FRANCE | | | |
| 27 | Colmar (Meyenheim) | 220 | 10.0 | 575 |
| 28 | Metz (Augny) | 190 | 9.8 | 756 |
| 29 | Nancy (Tomblaine, Essey) | 212 | 9.6 | 752 |
| 30 | Strasbourg (Entzheim) | 150 | 10.0 | 612 |
| | SWITZERLAND | | | |
| 31 | Basel (Binningen) | 316 | 9.8 | 787 |
| 32 | Bern (Liebefeld) | 572 | 8.8 | 1030 |
| 33 | Disentis | 1190 | 6.4 | 1116 |
| 34 | St. Gallen | 779 | 7.8 | 1283 |
| 35 | Davos (Dorf) | 1590 | 3.2 | 1022 |
| 36 | Säntis | 2490 | -1.7 | 2385 |

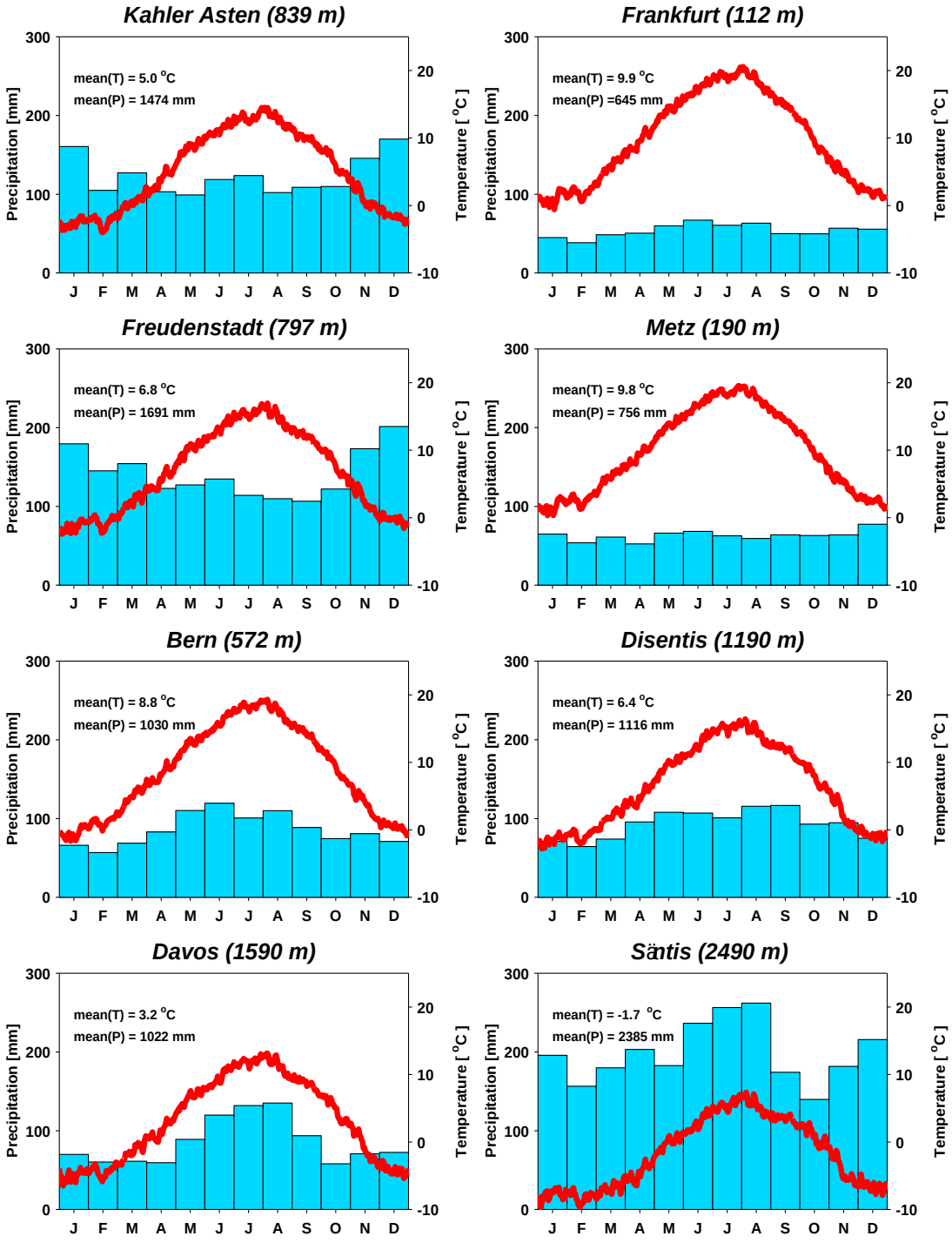


Figure 2.1: Precipitation and temperature characteristics of eight stations (1961-1990) together with their elevation above mean sea level. The bars and lines give the mean monthly precipitation and mean calendar-day temperature, respectively. Additionally, the annual mean temperature and precipitation are given.

# 3. Methods

## 3.1. Nearest-neighbour resampling

### 3.1.1. A historical note

Nearest-neighbour resampling was originally proposed by Young [1994] to simulate daily minimum and maximum temperatures and precipitation. Independently, Lall and Sharma [1996] discussed a nearest-neighbour bootstrap to generate hydrological time series. An application of their algorithm to daily precipitation and five other weather variables was presented in Rajagopalan and Lall [1999]. The ability of nearest-neighbour resampling to reproduce several sample statistics, precipitation spell structure and amount, was demonstrated. The method used for generating daily precipitation and temperature in the Rhine basin is basically the same as that in Rajagopalan and Lall [1999]. Especially for multi-site simulations summary statistics are needed to avoid problems with the high dimensional data space [Buishand and Brandsma, 2000].

### 3.1.2. The nearest-neighbour technique

In the nearest-neighbour method weather variables like precipitation and temperature are sampled simultaneously with replacement from the historical data. To generate weather variables for a new day $t$, one first abstracts the days from the historical record with similar characteristics as those simulated for the previous day. One of these nearest neighbours is randomly selected and the observed values for the day subsequent to that nearest neighbour are adopted as the simulated values for day $t$. A feature (or state) vector $\mathbf{D}_t$ is used to find the nearest neighbours in the historical record. In Rajagopalan and Lall [1999] $\mathbf{D}_t$ was formed out of the standardized weather variables generated for day $t-1$. The $k$ nearest neighbours of $\mathbf{D}_t$ were selected in terms of a weighted Euclidean distance. For two $q$-dimensional vectors $\mathbf{D}_t$ and $\mathbf{D}_u$ the latter is defined as:

$$\delta(\mathbf{D}_t, \mathbf{D}_u) = \left( \sum_{j=1}^{q} w_j \left( v_{tj} - v_{uj} \right)^2 \right)^{\frac{1}{2}} \tag{3.1}$$

where $v_{tj}$ and $v_{uj}$ are the $j$th components of $\mathbf{D}_t$ and $\mathbf{D}_u$ respectively and the $w_j$'s are scaling weights. In matrix notation, equation (3.1) can be rewritten as :

$$\delta(\mathbf{D}_t, \mathbf{D}_u) = \left( \left( \mathbf{D}_t - \mathbf{D}_u \right)^T \mathbf{W} \left( \mathbf{D}_t - \mathbf{D}_u \right) \right)^{\frac{1}{2}} \tag{3.2}$$

where $\mathbf{W}$ is a $q \times q$ diagonal matrix with the weights $w_j$ on the main diagonal, and $T$ stands for the transpose of a vector or matrix.

A discrete probability distribution or kernel is required for resampling from the $k$ nearest neighbours. Lall and Sharma [1996] recommended a kernel that gives higher weight to the closer neighbours. For this decreasing kernel the probability $p_n$ that the $n$th closest neighbour is resampled is given by:

$$p_n = \frac{1/n}{\sum_{i=1}^{k} 1/i}, \quad n = 1, ..., k \tag{3.3}$$

From the above description it is clear that apart from creating a feature vector (for a detailed discussion see Section 3.3), the user has to set the values of the number $k$ of nearest neighbours and the weights $w_j$. A sensitivity analysis in Brandsma and Buishand [1999] showed that $k = 5$ usually works well. In this study we also use this value of $k$. A more difficult issue is the optimal choice of the weights $w_j$. A sensitivity analysis with some theoretical hints (see next section) is again a way to select those parameters. It might be very time consuming to make this selection, however, especially if the dimension of the feature vector is high. An alternative approach is to avoid specification of the weights at all (and thus reduce the number of free parameters involved in the resampling algorithm) by introducing the *Mahalanobis distance* function [see, e.g., Kendall et al., 1983, p.290]. This metric incorporates the weights automatically in a data-adaptive manner.

### 3.1.3. Mahalanobis distance

In searching for a particular set of weights one should take the following into account :

- proper scaling in each direction of the feature space so that the feature vector components have comparable variances
- intercorrelations between feature vector components (it is pointless to give equal weights to variables which are highly correlated as to those which are uncorrelated)

The first point can be met by the weighted Euclidean distance by choosing the weights $w_j$ inversly proportional to the variances of the feature vector components $v_{tj}$ :

$$w_j = \frac{1}{Var(v_{tj})}\ , \quad j = 1, ..., q \tag{3.4}$$

To meet the second point one should consider the covariance matrix of the feature vector. A commonly used metric that directly incorporates these two points is the Mahalanobis metric **:**

$$\delta_{Mh}(\mathbf{D}_t, \mathbf{D}_u) = \left((\mathbf{D}_t - \mathbf{D}_u)^T \mathbf{B}^{-1} (\mathbf{D}_t - \mathbf{D}_u)\right)^{\frac{1}{2}} \tag{3.5}$$

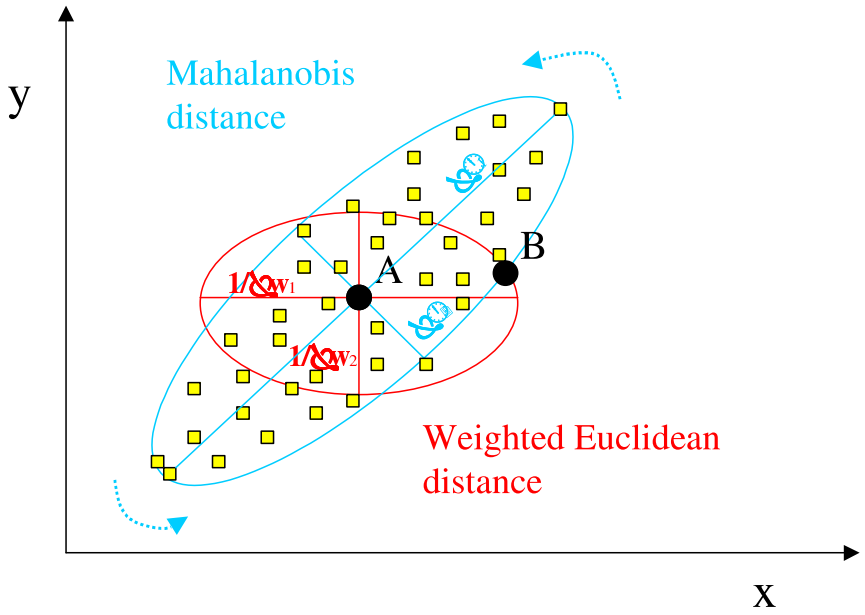


Figure 3.1: An intuitive approach to weighted Euclidean and Mahalanobis distance

where $\mathbf{B}$ is the covariance matrix of the feature vector $\mathbf{D}_t$. The elements of this matrix are the covariances between the components of $\mathbf{D}_t$:

$$B_{ij} = Cov(v_{ti}, v_{tj}) \ , \quad i, j = 1, ..., q \tag{3.6}$$

To build an intuition of what the difference is between the Mahalanobis metric and the weighted Euclidean metric let us consider a simple example. Figure 3.1 shows points in a two-dimensional plane. Points for which the distance to point A equals $\delta(A, B)$ in the weighted Euclidean metric, lie on the red ellipse. The axes of this ellipse are aligned parallel to the coordinate axes and their lengths are inversly proportional to the square root of the weights $w_1$ and $w_2$. The situation is different for the Mahalanobis distance. Points for which the distance to point A equals $\delta_{Mh}(A, B)$ in the Mahalanobis metric lie on the blue ellipse. The axes of the blue ellipse are not necceserily parallel to the coordinate axes and their lengths are proportional to the square root of the eigenvalues $\lambda_1$ and $\lambda_2$ of the matrix $\mathbf{B}$. So in comparison to the weighted Euclidean metric one has an extra possibility of rotating the "equidistance" ellipse. The angle of this rotation is determined by the correlation coefficient between the variables (yellow squares represent the values of $x$ and $y$) in the feature space and the ratio of their standard deviations.

## 3.2. Standardization procedure

Before resampling the data were deseasonalized through standardization. The daily temperatures and circulation indices were standardized by subtracting an estimate $m_d$ of the mean and dividing by an estimate $s_d$ of the standard deviation for the calendar day $d$ of interest:

$$\widetilde{x}_t = (x_t - m_d)/s_d, \quad t = 1, ..., 365J; \quad d = (t-1) \bmod 365 + 1 \tag{3.7}$$

where $x_t$ and $\widetilde{x}_t$ are the original and standardized variables, respectively, for day $t$, and $J$ is the total number of years in the record. The estimates $m_d$ and $s_d$ were obtained by smoothing the sample mean and standard deviation of the successive calendar days using the Nadaraya-Watson smoother [for more details see, e.g., Hastie and Tibshirani, 1990, p.19]. The smoothed statistic $g(d)$ for day $d$ is given by:

$$g(d) = \frac{\sum_{\tau=d-\sigma}^{d+\sigma} \kappa(\frac{d-\tau}{\sigma}) z_\tau}{\sum_{\tau=d-\sigma}^{d+\sigma} \kappa(\frac{d-\tau}{\sigma})}, \quad d = 1, ..., 365 \tag{3.8}$$

where $z_\tau$ is the raw value of the statistic for calendar day $\tau$, $\kappa(\cdot)$ is the kernel function and $\sigma$ is the bandwidth[1]. In this study the Epanechnikov kernel was applied:

$$\kappa(a) = \begin{cases} \frac{3}{4}(1-a^2), & \text{for} \quad |a| \leq 1 \\ 0 & \text{otherwise} \end{cases} \tag{3.9}$$

with a bandwidth $\sigma = 30$ days for temperature and circulation indices and $\sigma = 45$ days for precipitation.

Daily precipitation was standardized by dividing by a smooth estimate $m_{d,wet}$ of the mean wet-day precipitation amount:

$$\widetilde{x}_t = x_t / m_{d,wet}, \quad t = 1, ..., 365J; \quad d = (t-1) \bmod 365 + 1 \tag{3.10}$$

A wet day was defined here as a day with $P \geq 0.1$ mm.

Figure 3.2 shows examples of seasonal variation. Values of $m_d$ and $s_d$ for $Z, W, S$ and $T$ and $m_{d,wet}$ for $P$ are presented together with their smoothed approximations ($P$ and $T$ refer to Bern). Due to sampling effects there are large day-to-day fluctuations in the statistics presented in Fig. 3.2. The largest standard deviations of the flow indices are found in winter. The mean westerly flow is also relatively large in that season. The largest mean wet-day precipitation amounts are found in summer, which is due to the influence of convection (summer showers). The annual cycle is here stronger than that of the monthly mean precipitation amounts in Fig. 2.1.

To reduce the effect of seasonal variation further, the search for nearest neighbours was restricted to days within a moving window, centered on the calendar day of interest. The width of this window ($W_{\text{mw}}$) was 61 days as in Brandsma and Buishand [1999] for rainfall and temperature simulations and 121 days as in Beersma and Buishand [1999a] for simulations of circulation indices.

[1]To apply equation (3.8) the values $z_\tau$ for $\tau = 365 - \sigma, ..., 365$ were inserted for $\tau = 1 - \sigma, ..., 0$ and the values $z_\tau$ for $\tau = 1, ..., \sigma$ were inserted for $\tau = 366, ..., 365 + \sigma$.

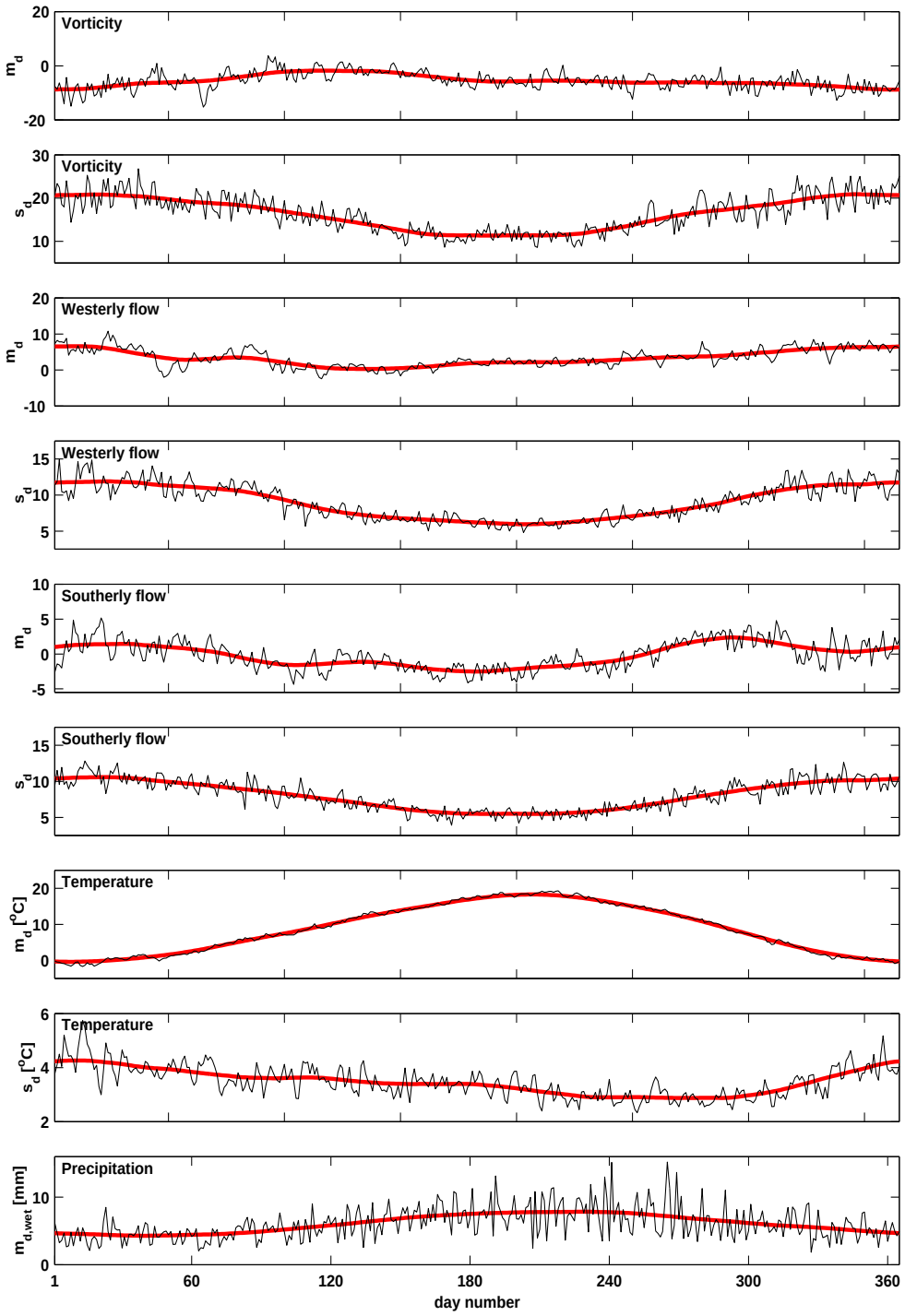


Figure 3.2: Values of $m_d$ and $s_d$ for vorticity ($Z$), westerly flow ($W$), southerly flow ($S$), temperature ($T$) and $m_{d,wet}$ for precipitation ($P$) together with their smoothed approximations ($P$ and $T$ refer to Bern) as a function of calendar day $d$ for the period 1961-1995. Vorticity and flow units are geostrophic, expressed as hPa per 10° lattitude at 50°$N$ (1 unit is equivalent to $0.73 \times 10^{-6}\text{s}^{-1}$ and 0.65 $\text{ms}^{-1}$ for vorticity and flow, respectively). The smooth curves are computed using the Nadaraya-Watson estimator with the Epanechnikov kernel.

## 3.3. Model identification

### 3.3.1. The feature vector

Daily $P$ and $T$ observations were available for the 36 stations listed in Table 2.1. Because of their rather extreme weather characteristics, the two Swiss mountain stations Davos and Säntis are not included in the feature vector. It is, however, still possible to simulate values for these stations passively (see Section 3.4).

To keep the dimension of the feature vector low, a small number of summary statistics was calculated for the remaining 34 stations. Both for $P$ and $T$ the arithmetic mean of the standardized daily values was used. In addition, the fraction $F$ of stations with $P \geq 0.1$ mm was considered. $F$ helps to distinguish between large-scale and convective precipitation. To keep the notation compact, the above components of the feature vector will be referred to as a sub-vector $\mathbf{V} = [\widetilde{P}, F, \widetilde{T}]^T$ where the tilde indicates standardized values. In some cases, the feature vector also contains the standardized circulation indices $\widetilde{\mathbf{C}} = [\widetilde{Z}, \widetilde{W}, \widetilde{S}]^T$.

### 3.3.2. The test cases

Basically two different kinds of simulations can be distinguished: *unconditional* simulations and *conditional* simulations on the atmospheric flow indices. In the unconditional simulation of Young [1994] and Rajagopalan and Lall [1999] the feature vector $\mathbf{D}_t$ comprises generated variables for the previous day as shown in Fig. 3.3 (cases 1 and 2).

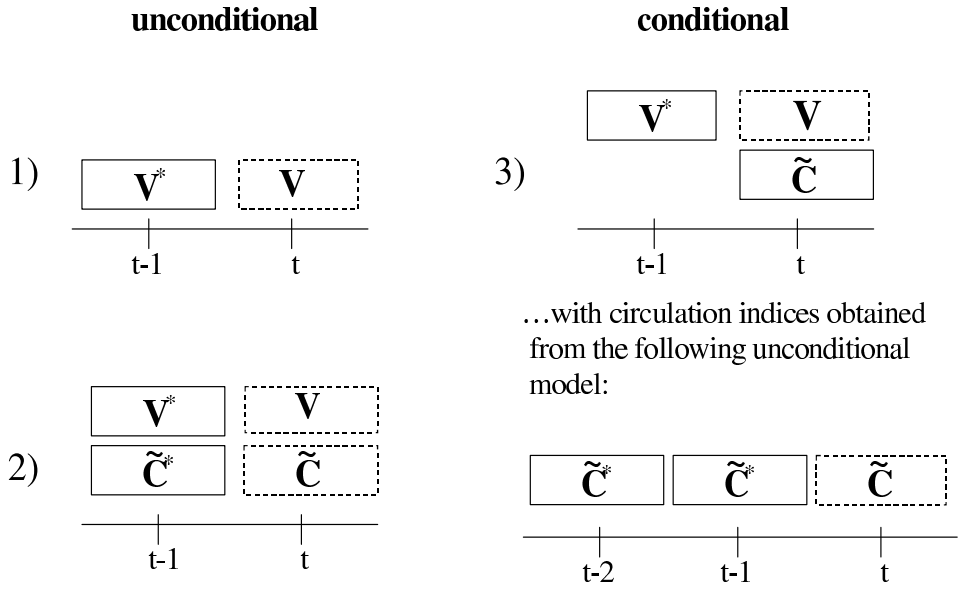


Figure 3.3: Components of the feature vector (solid boxes) for unconditional simulations 1), 2) and conditional simulation 3). The dashed boxes relate to variables to be resampled. The asterisks indicate that the corresponding variables are resampled values of the previous time steps.

Table 3.1: Definition of models for unconditional and conditional simulation The weights for the circulation (printed in bold) apply to all three components of $\widetilde{\mathbf{C}}$. $\widetilde{P}$ and $\widetilde{T}$ denote respectively the standardized precipitation and standardized temperature averaged over 34 stations, and $F$ denotes the fraction of these stations with $P \geq 0.1$ mm. An asterisk indicates that a value was resampled in a previous time step.

| **Case** | **elements of $\mathbf{D}_t$** | **weights** |
|---|---|---|
| **unconditional** | | |
| UE | $\widetilde{P}^*_{t-1}, F^*_{t-1}, \widetilde{T}^*_{t-1}$ | $2, 4, 1$ |
| UM | as case UE but with the Mahalanobis distance | – |
| UEc | $\widetilde{\mathbf{C}}^*_{t-1}, \widetilde{P}^*_{t-1}, F^*_{t-1}, \widetilde{T}^*_{t-1}$ | $\mathbf{1}$,3, 5, 2 |
| UMc | as case UEc but with the Mahalanobis distance | – |
| **conditional** | | |
| CE | $\widetilde{\mathbf{C}}_t, \widetilde{P}^*_{t-1}, F^*_{t-1}, \widetilde{T}^*_{t-1}$ | $\mathbf{1}$,3, 5, 2 |
| CM | as case CE but with the Mahalanobis distance | – |
| CMs | as case CM but with the Mahalanobis distance computed using the splitted covariance matrix | – |

For the simulation of circulation indices Beersma and Buishand [1999a] also studied the inclusion of the simulated indices for two or three previous days in the feature vector. These resampling schemes are indicated as second-order and third-order models, respectively (see lower panel of case 3 in Fig. 3.3). Conditional simulation on the atmospheric flow requires that the circulation indices for day $t$ are included in the feature vector as schematically represented in the upper panel of case 3 in Fig. 3.3. Nearest-neighbour resampling is then closely related to the analogue method used in climate change studies [Zorita and von Storch, 1999].

The resampling models for precipitation and temperature in this paper resemble those in Brandsma and Buishand [1999] for the German part of the basin. The differences in the model architecture originate from the composition of the feature vector, the choice of the scaling weights and the distance function used. Additionally, all conditional models are based on *simulated* circulation indices obtained with the second-order model (circ2.5) described in Beersma and Buishand [1999a]. The details of various cases are given in Table 3.1. In simulations which incorporated the Mahalanobis distance, the covariance matrix $\mathbf{B}$ was estimated locally, i.e., using only the values of the standardized weather variables lying within the moving data window. So, instead of having one global $\mathbf{B}$ matrix we computed 365 matrices to take the seasonal variation in the covariances into account. Moreover, in the conditional model CMs the Mahalanobis distance was applied with a new twist. Rather than using the full covariance matrix we only considered its two minors:

$$\mathbf{B}_{(1)} = Cov[\widetilde{\mathbf{C}}] \tag{3.11}$$

$$\mathbf{B}_{(2)} = Cov[\mathbf{V}] \tag{3.12}$$

Observing that for model CMs $\mathbf{D}_t = [\widetilde{\mathbf{C}}_t, \mathbf{V}^*_{t-1}]^T$ and $\mathbf{D}_u = [\widetilde{\mathbf{C}}_u, \mathbf{V}_{u-1}]^T$, the modified

Mahalanobis distance is given by the following expression:

$$\delta_{Mh}(\mathbf{D}_t, \mathbf{D}_u) = \left( (\widetilde{\mathbf{C}}_t - \widetilde{\mathbf{C}}_u)^T \mathbf{B}_{(1)}^{-1} (\widetilde{\mathbf{C}}_t - \widetilde{\mathbf{C}}_u) + (\mathbf{V}_{t-1}^* - \mathbf{V}_{u-1})^T \mathbf{B}_{(2)}^{-1} (\mathbf{V}_{t-1}^* - \mathbf{V}_{u-1}) \right)^{\frac{1}{2}} \tag{3.13}$$

The reason for this is that in conditional resampling the circulation indices are fed into the model from an external source (in this study the second-order model for the indices). On the other hand, the matrix $\mathbf{B}$ describes the covariance between $\widetilde{\mathbf{C}}$ and $\mathbf{V}$ within the historical record. The use of this covariance in the Mahalanobis distance is questionable, since the feature vector contains combinations of $\widetilde{\mathbf{C}}$ and $\mathbf{V}$ that are generally not found in the historical record. Equation (3.13) therefore neglects the covariance between $\widetilde{\mathbf{C}}$ and $\mathbf{V}$.

## 3.4. Resampling algorithm

To complete the discussion about nearest-neighbour resampling, the scheme for unconditional simulation of a multi-site precipitation and temperature record of $J^*$ years is listed below:

1. Calculate raw values of $m_d, s_d$ and $m_{d,wet}$ statistics and smooth them using (3.8).
2. Standardize $P$ and $T$ data using (3.7) and (3.10) respectively.
3. Calculate the fraction $F$ of "wet" stations and spatial averages of standardized $P$ and $T$ data.

4a. Choose one of the following distance functions:

- weighted Euclidean metric (3.2)
- Mahalanobis metric (3.5)

4b. Choose $k$ and $W_{\mathrm{mw}}$.

5. Generate data for $t = 1$, e.g. by randomly sampling a day within the window for the 1st of January.
6. Repeat steps 7-9 for $t = 2, 3, ..., 365J^*$.
7. Create a feature vector $\mathbf{D}_t$ from :

- the fraction $F$ of wet stations and spatial averages of standardized $P$ and $T$ data generated for day $t-1$ (cases UE and UM in Table 3.1), or
- the fraction $F$ of wet stations, spatial averages of standardized $P$ and $T$ data and circulation indices generated for day $t-1$ (cases UEc and UMc in Table 3.1).

8. Determine the $k$ nearest neighbours of $\mathbf{D}_t$ within the window for day $t-1$, using the distance function from step 4a.

9. Sample one of the $k$ nearest neighbours using the decreasing kernel (3.3) and deliver the values of the historical successor to the selected nearest neighbour as the values for day $t$.

10. Backtransform the resampled standardized variables to their original scale using the inverse of (3.7) and (3.10).

The algorithm above describes only unconditional simulation. For the conditional simulation on the atmospheric flow the following adjustments apply:

- In step 4a the Mahalanobis metric with the splitted covariance matrix may be considered as an additional option.

- The feature vector $\mathbf{D}_t$ in step 7 is created from the fraction $F$ of wet stations and spatial averages of standardized $P$ and $T$ data generated for day $t-1$ and the observed or simulated (by another model) circulation indices for day $t$ (cases CE, CM and CMs in Table 3.1).

- Sampling in step 5 should be restricted to days with circulation similar to that for $t = 1$.

- The nearest neighbour selected in step 9 straightforwardly provides (no successors needed) the $P$ and $T$ values for day $t$.

It should finally be noted that the values delivered in step 9 may include data from stations that were not used in the feature vector (Davos and Säntis) or area-average precipitation over subcatchments for the selected day. Henceforth the simulation of such additional data is designated as *passive* simulation. Daily area-average precipitation data for 230 subcatchments in Germany were passively generated in Brandsma and Buishand [1999] to compare the space-time patterns of simulated and historical extreme 10-day precipitation events.

# 4. Results

## 4.1. Reproduction of standard deviations and autocorrelation

Extreme river discharges in the lower part of the Rhine basin are mostly caused by prolonged heavy rainfall in winter. The reproduction of the standard deviation and autocorrelation coefficients was therefore only studied for the winter half-year (October - March). To reduce the influence of the annual cycle these second-order moment statistics were first calculated for each calendar month separately. The estimates were then averaged over the six calendar months October,...,March:

$$\overline{s_{\mathrm{D}i}} = \frac{1}{6}\sum_{k=1}^{6} \mathrm{S}_{\mathrm{D}i,k} \ , \quad i = 1, ..., 34 \tag{4.1}$$

$$\overline{r}_i(l) = \frac{1}{6}\sum_{k=1}^{6} r_{i,k}(l) \ , \qquad l = 1, 2; \ i = 1, ..., 34 \tag{4.2}$$

where $s_{\mathrm{D}i,k}$ and $r_{i,k}(l)$ are the sample standard deviation of the daily values and lag $l$ autocorrelation coefficient of the daily values respectively, calculated for the $i$th station and the $k$th calendar month ($k = 1$ corresponds to October). Besides, the reproduction of the standard deviations of the monthly values was considered as an additional performance measure. These were also averaged over the winter half-year:

$$\overline{s_{\mathrm{M}i}} = \frac{1}{6}\sum_{k=1}^{6} s_{\mathrm{M}i,k} \ , \quad i = 1, ..., 34 \tag{4.3}$$

where $s_{\mathrm{M}i,k}$ is the sample standard deviation of the monthly values for the $i$th station and the $k$th calendar month. The reproduction of the monthly standard deviations depends on that of the daily standard deviations and the autocorrelation coefficients. In particular, the monthly standard deviation tends to be too small if the autocorrelation coefficients decrease too rapidly with increasing lag. The statistics in (4.1),(4.2) and (4.3) were further averaged over all stations (except Davos and Säntis) to obtain the values $\langle \overline{s_{\mathrm{D}}} \rangle$ , $\langle \overline{r}(l) \rangle$ and $\langle \overline{s_{\mathrm{M}}} \rangle$ respectively.

Twenty-eight runs of 35 years were generated to investigate the performance of the resampling procedure. The standard deviations and autocorrelation coefficients were first estimated for each simulation run separately and then averaged over the 28 runs.

Table 4.1: Percentage differences between the mean standard deviations of monthly and daily values, and absolute differences between the mean lag 1,2 autocorrelation coefficients of daily values in the simulated time series (twenty-eight runs of 35 years for each case) and the historical records (1961-1995), averaged over 34 stations. In the lower part of the table the estimates for the historical data are listed (standard deviations of precipitation are in mm and those of temperature in $^o$C) . Values in bold refer to differences more than $2 \times se$ from the estimate for the historical data. Note that cases 2, 18 and 4.1(circ2.5) relate to the German part of the Rhine basin only.

| | $\langle \Delta \overline{s_M} \rangle$ | | $\langle \Delta \overline{s_D} \rangle$ | | $\langle \Delta \overline{r}(1) \rangle$ | | $\langle \Delta \overline{r}(2) \rangle$ | |
|---|---|---|---|---|---|---|---|---|
| Case | $P$ | $T$ | $P$ | $T$ | $P$ | $T$ | $P$ | $T$ |
| UE | 0.3 | -1.1 | 0.2 | 0.2 | **-0.019** | **-0.032** | -0.001 | 0.006 |
| 2[1] | -2.3 | -4.2 | -0.3 | -0.5 | -0.017 | **-0.030** | -0.007 | 0.002 |
| UM | -0.9 | -2.5 | 0.2 | -0.5 | **-0.020** | **-0.035** | -0.005 | 0.000 |
| UEc | -1.7 | -8.2 | -1.2 | -1.9 | **-0.018** | **-0.036** | 0.001 | -0.020 |
| 18[1] | -3.6 | -11.4 | -1.2 | -2.8 | -0.017 | **-0.035** | 0.004 | -0.020 |
| UMc | -0.6 | **-13.8** | -0.6 | -3.5 | -0.016 | **-0.054** | -0.003 | **-0.045** |
| CE | -6.4 | **-18.8** | -2.3 | **-7.0** | **-0.052** | **-0.087** | **-0.022** | **-0.050** |
| 4.1(circ 2.5)[2] | **-10.9** | **-25.4** | -3.8 | **-7.7** | **-0.060** | **-0.096** | **-0.038** | **-0.069** |
| CM | -9.1 | **-27.7** | -2.1 | **-9.2** | **-0.068** | **-0.134** | **-0.037** | **-0.115** |
| CMs | -8.9 | **-24.2** | -2.4 | **-8.1** | **-0.063** | **-0.110** | **-0.031** | **-0.086** |
| Historical | 35.7 | 2.1 | 4.2 | 4.2 | 0.283 | 0.826 | 0.144 | 0.639 |
| *se* | 4.53 | 6.16 | 2.45 | 2.49 | 0.008 | 0.007 | 0.009 | 0.015 |

[1] simulations from Buishand and Brandsma (2000)

[2] from simulations in Beersma and Buishand (1999a)

Afterwards, these average estimates $\overline{s_{D}}_i^*$, $\overline{s_{M}}_i^*$, $\overline{r}_i^*(l)$ were compared with the estimates $\overline{s_{D}}_i$, $\overline{s_{M}}_i$, $\overline{r}_i(l)$ for the historical data using:

$$\langle \Delta \overline{s_D} \rangle = \frac{1}{34} \sum_{i=1}^{34} \frac{(\overline{s_{D}}_i^* - \overline{s_{D}}_i)}{\overline{s_{D}}_i} \ 100\% \tag{4.4}$$

$$\langle \Delta \overline{s_M} \rangle = \frac{1}{34} \sum_{i=1}^{34} \frac{(\overline{s_{M}}_i^* - \overline{s_{M}}_i)}{\overline{s_{M}}_i} \ 100\% \tag{4.5}$$

$$\langle \Delta \overline{r}(l) \rangle = \frac{1}{34} \sum_{i=1}^{34} (\overline{r}_i^*(l) - \overline{r}_i(l)) \tag{4.6}$$

In order to evaluate the statistical significance of $\langle \Delta \overline{s_D} \rangle$ , $\langle \Delta \overline{s_M} \rangle$ and $\langle \Delta \overline{r}(l) \rangle$ standard errors $se$ were calculated for the historical record. The standard error of the mean lag $l$

autocorrelation estimate $\langle \overline{r}(l) \rangle$ was obtained by the jackknife method in Buishand and Beersma [1993]. Likewise, the jackknife procedures in Buishand and Beersma [1996] and Beersma and Buishand [1999b] were used to compute the standard errors of $\langle \overline{s_{\mathrm{D}}} \rangle$ and $\langle \overline{s_{\mathrm{M}}} \rangle$ respectively. A criterion of $2 \times se$ was used to indicate significant differences between the historical and simulated values. This corresponds to a two-sided test at 5% level [Brandsma and Buishand, 1998].

Table 4.1 presents $\langle \Delta \overline{s_{\mathrm{M}}} \rangle$, $\langle \Delta \overline{s_{\mathrm{D}}} \rangle$ and $\langle \Delta \overline{r}(l) \rangle$ for the models defined in Table 3.1. To make a comparison with the earlier unconditional simulations for the German part of the Rhine basin, the results for the cases 2 and 18 in Buishand and Brandsma [2000] are also listed in Table 4.1. These cases are similar[1] to UE and UEc respectively. Analogously, conditional simulations performed in the present report can be compared[2] with the results of model 4.1(circ2.5) for the German part of the Rhine basin in Beersma and Buishand [1999a].

For the unconditional models which incorporate only the large-scale features of the $P$ and $T$ fields Table 4.1 shows that the precipitation statistics are well reproduced. A slight, but statistically significant, bias is present in the lag 1 autocorrelation coefficient. The performance of model UE is more or less the same as that of model 2 in Buishand and Brandsma [2000]. The largest differences are found for the standard deviations of monthly temperatures. Model UM with Mahalanobis metric performs equally well as its twinning case UE. Incorporation of the circulation indices into the feature vector (cases UEc, UMc and 18) generally worsens the reproduction of daily temperature statistics in particular if the Mahalanobis metric (UMc) is used. The results for precipitation are, however, similar to those obtained in the unconditional models without circulation indices. This insensitivity of the quality of the reproduction of precipitation statistics to the inclusion of the circulation indices is in line with the results in Buishand and Brandsma [2000].

Conditional resampling of $P$ and $T$ on simulated circulation indices (cases CE, CM, CMs) lags behind. In Beersma and Buishand [1999a] it was also shown that this occurs for conditional simulations on historical circulation indices. All statistics of daily and monthly temperatures and the lag 1 and 2 autocorrelation coefficients for precipitation are significantly underestimated. In general, the quality of the reproduction of $\langle \overline{s_{\mathrm{M}}} \rangle, \langle \overline{r}(1) \rangle, \langle \overline{r}(2) \rangle$ for $T$ is similar to that obtained from the conditional simulation 4.1(circ2.5) in Beersma and Buishand [1999a]. Model CM with the "classical" Mahalanobis distance shows the strongest underestimation of the standard deviations of monthly and daily temperatures. The reproduction of these statistics is improved using the splitted covariance matrix (see Section 3.3.2) in the Mahalanobis distance. The resulting model CMs performs still poorer than model CE with the weighted Euclidean distance. The second-order moment statistics of daily precipitation are better repro-

[1]The similarity refers to the feature vector composition and the choice of the number $k$ of nearest neighbours. In Buishand and Brandsma (2000) twenty-five runs of 35-years were conducted instead of 28 runs. Moreover, the fraction $F$ of wet stations in the feature vector was based on a wet-day threshold of 0.3 mm instead of 0.1 mm.

[2]Note that in Beersma and Buishand [1999a] ten runs of 35-years were performed, and the fraction $F$ of wet stations in the feature vector had a relatively low weight. Moreover, $F$ was based on a wet-day threshold of 0.3 mm instead of 0.1 mm.

duced by the conditional model CE than by model 4.1(circ2.5) for the German part of the Rhine basin in Beersma and Buishand [1999a]. The improvement is due to the use of a higher weight of the fraction $F$ of wet stations (this reduces the bias in the lag 2 autocorrelation coefficient) and the use of a lower threshold for the definition of wet days (this reduces the bias in the daily standard deviation) in the present study.

## 4.2. Reproduction of $N$-day winter maximum precipitation amounts

For the 34 stations, the $N$-day ($N = 1, 4, 10, 20$) winter (October-March) precipitation amounts were abstracted from the historical record and all simulated cases. Like in previous studies for the German part of the basin, the following three quantities are considered to verify the reproduction of the $N$-day winter maxima distributions:

1. The maximum $MAX$ of the $N$-day winter maxima (highest $N$-day precipitation amount in the record).
2. The upper quintile mean $QM5$ of the $N$-day winter maxima.
3. The median $M$ of the $N$-day winter maxima.

$QM5$ refers to the mean of the data beyond the highest quintile (upper 20%). Because taking 20% of the 34 winters in our 35-year record does not result in a whole number, we

Table 4.2: Percentage differences between the maxima ($MAX$), upper quintile means ($QM5$) and medians ($M$) of the $N$-day winter (October-March) precipitation maxima in the simulated data (twenty-eight 35-year runs for each case) and the historical records (1961-1995), averaged over 34 stations. In the lower part of the table the estimates for the historical data are listed. Note that cases 2, 18 and 4.1(circ2.5) are related to the German part of the Rhine basin only.

| | *MAX* (%) | | | | *QM5* (%) | | | | *M* (%) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Case | $N$=1 | $N$=4 | $N$=10 | $N$=20 | $N$=1 | $N$=4 | $N$=10 | $N$=20 | $N$=1 | $N$=4 | $N$=10 | $N$=20 |
| UE | -3.4 | -1.8 | -1.4 | -0.4 | 0.6 | -1.0 | -0.6 | 0.2 | -1.3 | -2.1 | -0.2 | -1.7 |
| 2[1] | -4.9 | -1.5 | 0.1 | -2.9 | -1.0 | -1.3 | -0.6 | -0.1 | -3.0 | -3.1 | -1.3 | -1.4 |
| UM | -4.0 | -0.8 | -0.9 | -0.8 | -0.2 | -1.0 | -1.0 | -0.4 | -1.8 | -2.7 | -0.7 | -1.5 |
| UEc | -5.6 | -3.1 | -0.9 | -0.7 | -2.3 | -3.7 | -2.0 | -0.6 | -2.8 | -3.3 | -1.9 | -2.6 |
| 18[1] | -7.8 | -5.1 | -3.3 | -4.5 | -3.8 | -3.9 | -2.8 | -1.9 | -4.3 | -3.5 | -1.5 | -2.0 |
| UMc | -5.6 | -2.1 | -0.4 | 0.2 | -1.5 | -2.0 | -1.4 | 0.1 | -2.8 | -2.6 | -1.3 | -2.2 |
| CE | -7.4 | -5.1 | -5.5 | -6.6 | -3.5 | -5.7 | -5.3 | -5.0 | -4.0 | -6.4 | -4.5 | -5.4 |
| 4.1(circ 2.5)[2] | -9.1 | -7.1 | -6.9 | -8.4 | -5.5 | -7.2 | -7.9 | -8.2 | -5.9 | -8.0 | -7.2 | -7.7 |
| CM | -7.3 | -6.4 | -8.2 | -6.3 | -3.4 | -6.6 | -7.3 | -5.8 | -3.5 | -7.5 | -6.9 | -6.9 |
| CMs | -6.4 | -5.9 | -6.7 | -5.6 | -2.9 | -6.5 | -6.9 | -5.4 | -3.5 | -6.8 | -6.1 | -6.4 |
| Historical [mm] | 56.6 | 95.7 | 138.5 | 189.4 | 42.7 | 76.7 | 111.1 | 152.6 | 27.2 | 51.1 | 75.2 | 106.9 |

[1] from simulations in Buishand and Brandsma (2000)

[2] from simulations in Beersma and Buishand (1999a)

obtained $QM5$ as the average of the mean of the 7 largest winter maxima (with weight 0.8) and the mean of the 6 largest winter maxima (with weight 0.2). This procedure gives almost identical results as that followed in the UK Flood Studies Report [NERC, 1975] to derive the quartile means of annual maxima as summary statistics.

Analogous to equations (4.4) and (4.5), we calculated for each of the three quantities the percentage differences between the values for the simulated and historical data averaged over all stations. Table 4.2 presents the results for the cases discussed in the previous section. The unconditional model UE, which did not use circulation indices in the search for the nearest neighbours, reproduces the extreme-value statistics virtually as well as the corresponding model 2 for the German part of the Rhine basin. Model UM with Mahalanobis metric performs equally well. Furthermore, in line with the results from the previous section, the quality of the reproduction of the extreme-value properties of precipitation remains more or less the same if circulation indices are included in the feature vector. Especially, cases UEc and UMc give results comparable to cases UE and UM. Model 18 from Buishand and Brandsma [2000] shows, however, a somewhat larger underprediction of the maximum and the upper quintile mean.

Conditioning the resampling procedure on circulation indices (cases CE, CM ,CMs), results in a bit larger underestimation of the extreme-value statistics than in the unconditional cases. This is in agreement with the poorer reproduction of second-order moment statistics for conditional simulations as observed in Table 4.1. It may further be noted that the results obtained with model CE are slightly better than those for the two conditional models CM and CMs using the Mahalanobis distance. Model CE also performs better than model 4.1(circ2.5) for the German part of the Rhine basin, which is in line with the results for the second-order moment statistics in the previous section.

## 4.3. Reproduction of $N$-day maximum snowmelt amounts

Snowmelt generally, contributes to extreme river discharges in the lower part of the Rhine basin. It is, however, only for the highest stations Kahler Asten, Freudenstadt, Kl. Feldberg , Disentis, Davos and Säntis in this study that a considerable part of the winter precipitation falls in the form of snow. For these six stations the reproduction of extreme-value properties of $N$-day snowmelt has been analyzed. The results for Kahler Asten, Freudenstadt and Kl. Feldberg are compared with those in Brandsma and Buishand [1999][3] and Buishand and Brandsma [2000].

Historical estimates and simulated values of snowmelt were derived from the historical and generated daily precipitation and temperature[4], respectively. It was assumed that precipitation accumulates on the ground as snow if $T < 0\,^oC$ and that the melt on days with $T > 0\,^oC$ is proportional to $T$ as long as there is snow on the ground. The constant of proportionality (degree-day-factor) was set equal to 4 mm/$^o$C which is an average value from the literature [Linsey et al., 1988, Gray and Prowse, 1993]. The $N$-day winter maxima ($N$=1, 4, 10 or 20) were abstracted from the calculated snowmelt

[3]Only the simulations for Kahler Asten and Freudenstadt are compared with the simulations in this report.

[4]For Davos and Säntis precipitation and temperature time series were simulated passively (see Section 3.4).

Table 4.3: Percentage differences between the maxima ($MAX$), upper quintile means ($QM5$) and medians ($M$) of the $N$-day snowmelt extremes for the simulated data (28 runs of 35 years for each case) and the historical records (1961-1995) for six stations in the Rhine basin. In the lower part of the table the values of $MAX$, $QM5$ and $M$ for the historical data are listed.

| | *MAX* (%) | | | | *QM5* (%) | | | | *M* (%) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Station | $N$=1 | $N$=4 | $N$=10 | $N$=20 | $N$=1 | $N$=4 | $N$=10 | $N$=20 | $N$=1 | $N$=4 | $N$=10 | $N$=20 |
| UE | | | | | | | | | | | | |
| Kahler Asten | -15.5 | -15.7 | -12.1 | 3.1 | 6.4 | 8.6 | 2.6 | 5.6 | 15.9 | 12.3 | 20.7 | 3.6 |
| Freudenstadt | 5.1 | 1.5 | -10.0 | 5.4 | 4.8 | -2.8 | -11.2 | -5.8 | 8.7 | -0.6 | -6.7 | -6.6 |
| Kl. Feldberg | -1.1 | -0.4 | 5.6 | -10.5 | 6.1 | -3.0 | -1.7 | -12.5 | 11.2 | 6.6 | -6.5 | -3.4 |
| Disentis | 17.2 | 10.9 | -13.6 | -1.8 | 13.8 | -4.7 | -7.0 | -6.0 | 5.9 | -1.6 | -7.6 | -10.7 |
| Davos | 24.5 | 5.9 | -6.7 | -18.2 | 20.0 | 4.2 | -4.8 | 1.1 | 14.3 | 13.3 | 11.0 | 0.9 |
| Säntis | -0.8 | 4.4 | 1.1 | -8.7 | 2.1 | -3.2 | -13.6 | -12.0 | 6.1 | -9.7 | -0.5 | 5.2 |
| UEc | | | | | | | | | | | | |
| Kahler Asten | -13.9 | -14.5 | -12.8 | 0.4 | 6.9 | 8.3 | 2.8 | 3.8 | 15.1 | 14.3 | 20.3 | 5.2 |
| Freudenstadt | 7.3 | 6.0 | -6.4 | 14.3 | 4.1 | -2.1 | -10.5 | -4.0 | 5.6 | -2.3 | -8.8 | -9.0 |
| Kl. Feldberg | 1.4 | 4.1 | 2.2 | -10.8 | 8.6 | -2.5 | -6.0 | -15.0 | 8.9 | 3.7 | -11.0 | -7.4 |
| Disentis | 19.1 | 14.7 | -3.4 | 6.9 | 12.7 | -1.6 | -3.3 | -1.8 | 3.3 | -1.2 | -8.6 | -10.1 |
| Davos | 22.0 | 8.1 | -4.9 | -17.6 | 18.5 | 7.4 | -3.7 | 2.0 | 11.4 | 12.7 | 14.5 | 2.4 |
| Säntis | -0.4 | 7.8 | -5.8 | -24.1 | 1.8 | -4.1 | -16.2 | -18.2 | 4.3 | -9.6 | -1.9 | 0.9 |
| CE | | | | | | | | | | | | |
| Kahler Asten | -16.4 | -20.9 | -20.9 | -8.5 | 3.4 | 0.0 | -7.3 | -6.5 | 12.9 | 3.9 | 9.2 | -2.8 |
| Freudenstadt | 1.6 | -11.3 | -31.6 | -18.5 | -1.4 | -18.7 | -30.2 | -24.6 | 0.2 | -15.8 | -22.5 | -21.0 |
| Kl. Feldberg | -4.0 | -14.0 | -18.4 | -34.1 | 0.9 | -17.0 | -22.3 | -30.8 | 4.2 | -11.1 | -23.0 | -18.3 |
| Disentis | 13.1 | -8.1 | -34.2 | -24.9 | 7.7 | -19.0 | -25.2 | -23.6 | -2.4 | -17.7 | -23.0 | -24.0 |
| Davos | 20.9 | 1.7 | -15.1 | -24.7 | 16.0 | -1.9 | -12.8 | -6.9 | 13.3 | 8.5 | 8.7 | 0.0 |
| Säntis | -6.6 | -11.2 | -22.0 | -36.2 | -3.3 | -17.9 | -29.2 | -30.1 | 3.7 | -19.5 | -11.7 | -5.8 |
| Historical [mm] | | | | | | | | | | | | |
| Kahler Asten | 51.2 | 164.0 | 287.2 | 314.0 | 35.1 | 104.8 | 184.6 | 236.2 | 22.9 | 62.0 | 86.6 | 134.6 |
| Freudenstadt | 42.4 | 126.8 | 234.7 | 243.1 | 36.4 | 108.2 | 180.4 | 212.3 | 24.8 | 64.0 | 93.7 | 118.7 |
| Kl. Feldberg | 38.0 | 105.9 | 151.6 | 212.2 | 29.8 | 85.7 | 121.7 | 165.1 | 20.0 | 46.6 | 71.3 | 86.1 |
| Disentis | 29.2 | 87.0 | 171.8 | 191.4 | 26.3 | 79.3 | 119.1 | 149.0 | 20.4 | 47.4 | 67.8 | 87.6 |
| Davos | 26.9 | 85.6 | 176.4 | 285.9 | 22.7 | 71.6 | 137.8 | 180.8 | 16.6 | 41.0 | 63.6 | 98.6 |
| Säntis | 40.0 | 109.6 | 198.0 | 326.0 | 32.8 | 90.9 | 157.0 | 211.5 | 20.6 | 50.9 | 63.9 | 79.2 |

amounts. As in the previous section, the statistics $MAX, QM5$ and $M$ of these extremes were used to assess the reproduction of $N$-day maximum snowmelt in the simulated data.

Analogous to Table 4.2, Table 4.3 presents the average percentage differences between the values of $MAX, QM5$ and $M$ for 28 simulation runs and the values of these statistics for the historical data of the six stations of interest. For Kahler Asten, Freudenstadt and Kl. Feldberg the results obtained with model UE are very similar to those of model 2 for the German part of the Rhine basin in Buishand and Brandsma [2000]. The extremes are satisfactorily reproduced by model UE. The largest discrepancies are found for Kahler Asten and Davos (overestimation of the median of the 1, 4 and 10-day maxima).

On comparing the results of the unconditional experiment UEc with case 4.4 in

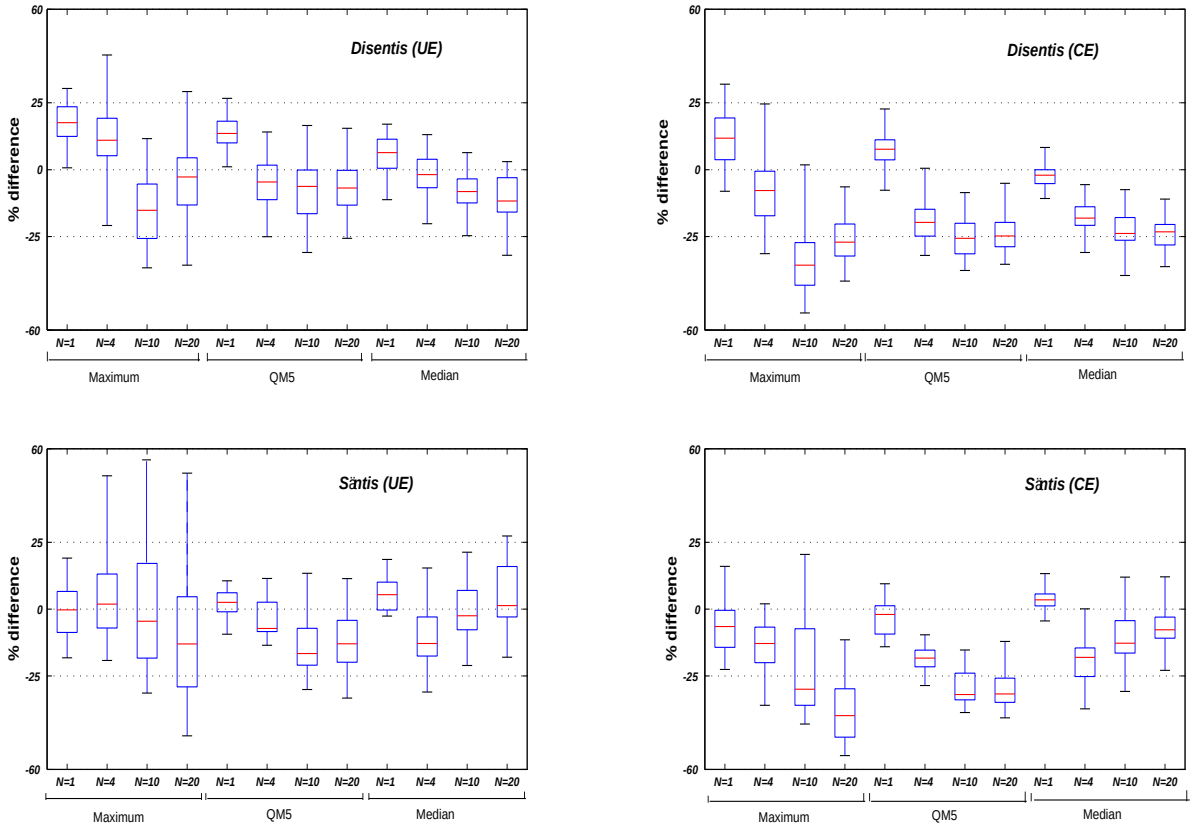


Figure 4.1: Boxplots of the percentage differences between the maxima, upper quintile means and medians of the $N$-day winter (October-March) snowmelt maxima for the simulated data and the historical records (1961-1995), for two mountain stations in Switzerland. A box represents a sample of twenty-eight runs of 35 years. The lower and upper lines of the box identify the 25th and 75th percentiles of the sample. The line in the middle of the box is the sample median. The "whiskers" at either end extend to the extreme values, i.e., the minimum and the maximum of the sample.

Brandsma and Buishand [1999] the reproduction of the median and upper quintile mean of the calculated multi-day maximum snowmelt for Freudenstadt is much improved. As in case UE the medians of the 1,4 and 10-day maxima remain overestimated for Kahler Asten and Davos.

Conditional simulation (case CE) exhibits, for most stations, a relatively large underprediction of the extreme-value properties of 10 and 20-day snowmelt. This phenomenon can partly be explained by the considerable negative bias in the daily temperature autocorrelation coefficients, which reduces the likelihood that snow accumulates over long periods and thus the probability of extreme multi-day snowmelt.

Figure 4.1 shows boxplots of the percentage differences for Disentis and Säntis. The extreme-value characteristics are rather well reproduced in experiment UE. The reproduction of these characteristics turns out to be less successful in the conditional case CE. In almost all simulations in that experiment, the upper quintile mean $QM5$ of the calculated multi-day maximum snowmelt is lower than the values derived from the historical data (for Disentis this holds also for the median $M$).

The historical winter snowmelt maxima at the Swiss stations are not higher than those at Kahler Asten and Freudenstadt. In particular for Säntis there is, however, a lot of snowmelt outside the winter period. For example, the maximum ($MAX$) 10-day snowmelt amount calculated for the whole year at this station is as high as 444.8 mm, while for the winter period it is only 198.0 mm.

In comparison to the models with the weighted Euclidean metric discussed above, the quality of the reproduction of the statistics of $N$-day snowmelt extremes turns out to be more or less the same for the models with the Mahalanobis metric. This is a somewhat surprising result, taking into account that for conditional simulations with the Mahalanobis metric relatively large negative biases in the second-order moment statistics for temperature were found (Table 4.1).

## 4.4. Long-duration simulations

A number of 1000-year conditional and unconditional simulations were performed. For three of these simulations Figure 4.2 shows Gumbel plots of the 10-day winter precipitation maxima for five stations and the average of the 34 stations used in the feature vector. The reason why the 10-day winter precipitation maxima for observed and simulated data were chosen here is that large river discharges at Lobith in the past were often the result of extensive precipitation over a period of about ten days. There is a reasonable correspondence between the historical and simulated distributions. The figure clearly shows the underestimation of the extreme-value properties for the conditional model CE, discussed in Section 4.2. Furthermore, in all cases displayed in Fig. 4.2, a large part of the curve for the conditional 1000-year simulation lies below the curves for the unconditional 1000-year simulations. Realistic multi-day precipitation much larger than the largest historical precipitation amounts are generated in all simulation experiments shown in Fig. 4.2.

Table 4.4 presents some properties of the $N$-day winter maxima. The results for models with the Mahalanobis metric are comparable to those with the weighted Euclidean metric. For both models the extreme-value statistics in the conditional simulations tend

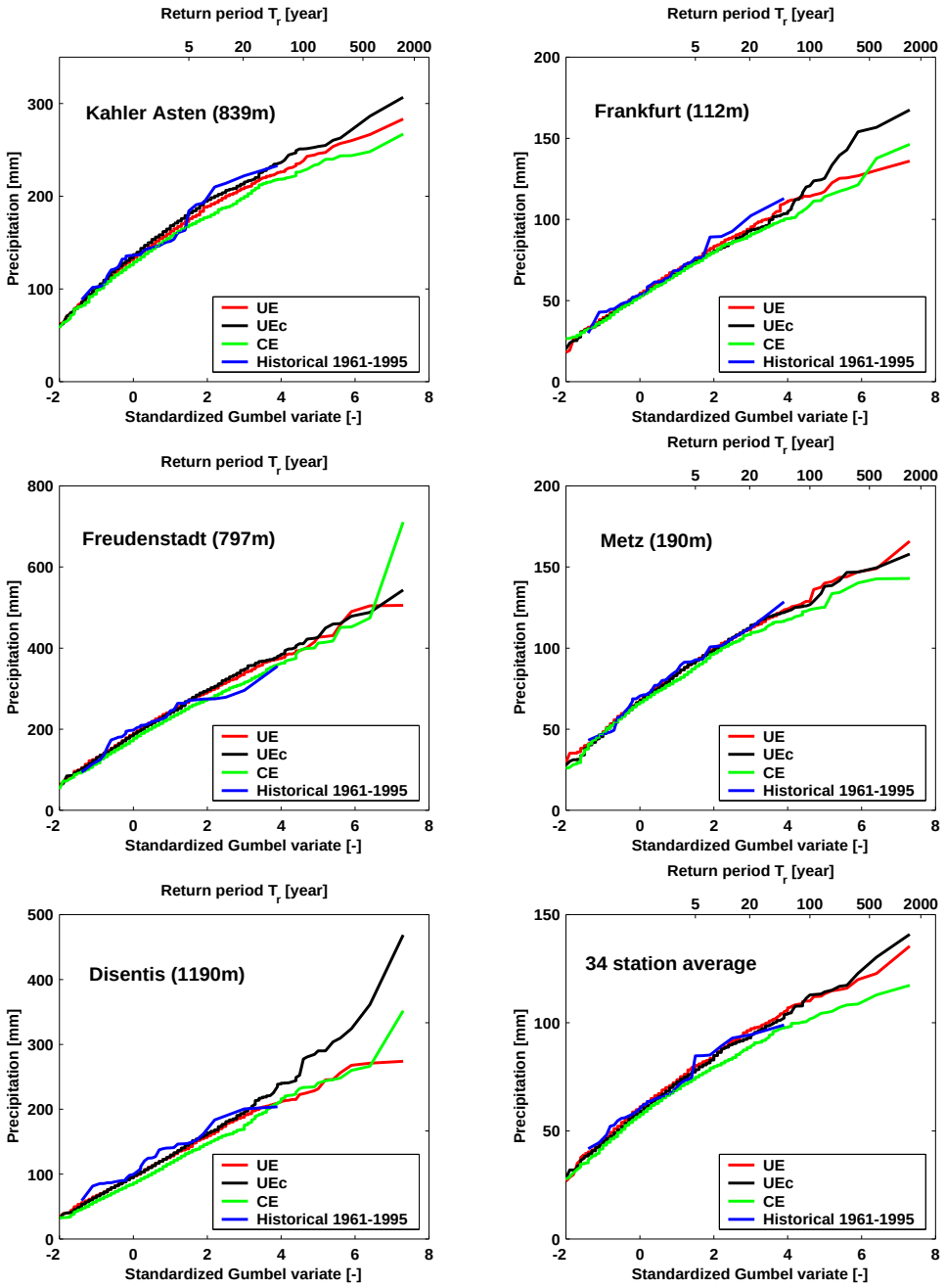


Figure 4.2: Gumbel plots of 10-day winter precipitation maxima for observed and simulated data (runs of 1000 years)

Table 4.4: The largest value ($MAX$), the 200-year and 50-year events[1] of $N$-day winter (October-March) precipitation [mm] in 1000-year simulations, averaged over 34 stations. The bottom part of the table gives the relative differences between some of the conditional and unconditional simulations, averaged over 34 stations

| | $MAX$ | | | 200-year event | | | 50-year event | | |
|---|---|---|---|---|---|---|---|---|---|
| Case | $N$=4 | $N$=10 | $N$=20 | $N$=4 | $N$=10 | $N$=20 | $N$=4 | $N$=10 | $N$=20 |
| UE | 121.6 | 180.7 | 241.5 | 106.1 | 153.2 | 210.0 | 92.6 | 132.4 | 181.0 |
| UM | 129.1 | 184.0 | 242.2 | 107.3 | 156.4 | 211.5 | 92.8 | 133.4 | 182.1 |
| UEc | 122.2 | 191.1 | 252.8 | 105.1 | 156.8 | 210.2 | 90.7 | 132.1 | 181.1 |
| UMc | 124.0 | 187.2 | 249.0 | 106.2 | 153.6 | 207.1 | 92.1 | 133.1 | 182.7 |
| CE | 122.7 | 177.4 | 225.7 | 102.0 | 145.1 | 192.5 | 87.6 | 125.8 | 170.0 |
| CM | 120.5 | 166.9 | 233.3 | 101.8 | 142.4 | 194.8 | 87.0 | 123.0 | 170.5 |
| CMs | 130.4 | 169.3 | 235.2 | 102.2 | 144.4 | 196.7 | 87.3 | 124.4 | 171.1 |
| Avg. relative differences [%] | | | | | | | | | |
| CE vs. UEc | 0.9 | -10.6 | -11.0 | -2.4 | -6.7 | -7.5 | -2.6 | -4.1 | -5.6 |
| CM vs. UMc | -2.1 | -10.3 | -5.6 | -3.7 | -6.7 | -6.2 | -4.6 | -6.9 | -6.4 |

[1] For the 200-year and 50-year events, respectively the 5th and the 20th highest values out of 999 simulated winters were used.

to be smaller than those in the unconditional simulations. The average relative differences given in the bottom part of Table 4.4 are, however, quite uncertain because they are based on only one simulation run. To demonstrate this uncertainty, five additional 1000-year simulations were performed with the conditional model CE and the unconditional model UEc. The results of these simulations are displayed in Table 4.5. Especially for $MAX$ they are quite different from those in the bottom part of Table 4.4. The uncertainty can be reduced by taking the average (or the median) of the relative differences from the five simulation runs. The relative differences between conditional

Table 4.5: The relative differences between the conditional simulation CE and unconditional simulation UEc, averaged over 34 stations (results of five runs of 1000 years each).

| | $MAX$ | | | 200-year event | | | 50-year event | | |
|---|---|---|---|---|---|---|---|---|---|
| | $N$=4 | $N$=10 | $N$=20 | $N$=4 | $N$=10 | $N$=20 | $N$=4 | $N$=10 | $N$=20 |
| CE vs. UEc | Avg. relative differences [%] | | | | | | | | |
| run nr 1 | 4.1 | 1.4 | -1.2 | 1.0 | -2.4 | -1.5 | -2.0 | -1.3 | -2.0 |
| run nr 2 | 5.4 | -2.4 | -4.1 | 1.0 | -3.0 | -4.0 | -1.8 | -3.0 | -3.8 |
| run nr 3 | 3.4 | 0.7 | -2.5 | 1.1 | -2.4 | -2.1 | -2.6 | -2.3 | -1.0 |
| run nr 4 | 3.2 | 0.6 | -6.2 | 1.4 | -2.1 | -5.3 | -1.3 | -2.5 | –3.8 |
| run nr 5 | -2.1 | -4.4 | -1.6 | -1.3 | -2.7 | -2.4 | -3.1 | -3.3 | -3.4 |
| Average | 2.8 | -0.8 | -3.1 | 0.6 | -2.5 | -3.1 | -2.1 | -2.5 | -2.8 |
| Median | 3.4 | 0.6 | -2.5 | 1.0 | -2.4 | -2.4 | -2.0 | -2.5 | -3.4 |

and unconditional simulations do not increase with increasing return period in Table 4.5. Another noticeable point in that table is that the 4-day precipitation amounts at long return periods for the conditional model exceed those for the unconditional one.

# 5. Conclusions

In this report the multi-site simulation of daily precipitation and temperature for the entire Rhine basin was explored. Besides the data from Germany used in earlier studies, data from Luxemburgian, French and Swiss stations in the basin were included. Both unconditional and conditional simulations were performed. To reduce the number of parameters involved in these simulations, the Mahalanobis metric was applied as an alternative to the weighted Euclidean metric.

The unconditional simulations preserved the second-order moment statistics of daily and monthly precipitation and $N$-day maximum precipitation well. The lag 1 autocorrelation coefficient for daily temperature was, however, significantly underestimated. The reproduction of the second order moments of temperature became worse in simulations where atmospheric circulation indices were added to the feature vector. Despite this deficiency the reproduction of $N$-day maximum snowmelt was satisfactory.

Multi-site simulations of $P$ and $T$ conditional on simulated atmospheric circulation indices performed somewhat poorer than the unconditional simulations. Especially for temperature the reproduction of second-order moment statistics became worse. As a result a significant underestimation (up to 20-30%) of the median and the upper quintile mean of multi-day snowmelt amounts was observed for four high-elevation stations (Freudenstadt, Kl. Feldberg, Disentis, Säntis). For daily precipitation the second-order moment statistics were somewhat better preserved than in similar conditional simulations for the German part of the Rhine basin in Beersma and Buishand [1999a] due to minor modifications in the resampling model. As a consequence, the results for the extreme $N$-day precipitation amounts for the whole basin compare favourably to those for the German part in earlier studies.

The ability of both unconditional and conditional models to generate realistic unprecedented multi-day rainfall events was demonstrated with simulation runs of 1000 years. Especially those extreme events may cause large peak discharges of the river Rhine in the Netherlands. A single simulation run of 1000 years does not provide, however, an accurate estimate of a 1000-year event. More simulations are needed for that purpose and even then a considerable uncertainty remains due to the use of a relatively short 35-year historical record for resampling.

The unconditional models with the Mahalanobis metric generated results comparable to those with the weighted Euclidean metric. For the conditional ones a large bias was found in the second-order moment statistics of daily temperature. This bias could be somewhat reduced by using a splitted covariance matrix in the Mahalanobis distance computation. On the other hand, the above shortcoming had no effect on the quality of reproduction of $N$-day snowmelt maxima. Furthermore, the magnitude of extreme rainfall events in long-duration simulations was not influenced by the metric used.

# 6. Recommendations

A number of resampling models for multi-site generation of daily precipitation and temperature for the entire drainage area of the Rhine basin have been developed. The statistical properties of observed precipitation and temperature were best preserved by a resampling model with a 3-dimensional feature vector containing only the area-averages of the standardized daily temperature and precipitation values and the fraction of stations with precipitation. These models should therefore be considered first for design discharge estimation under present-day climate conditions.

Future work on the application of a resampling model for design discharge estimation should include a further validation using extreme-value characteristics of river discharges rather than those of precipitation and calculated snowmelt. In order to discriminate deficiencies in the resampling model from those in hydraulic and hydrological modelling, such a validation should comprise a comparison of observed discharges at Lobith over the period 1961-1995 with,

- computed discharges at Lobith from observed precipitation and temperature over the period 1961-1995, and
- computed discharges at Lobith from 35-year simulations of precipitation and temperature.

Several 35-year simulations should be done to obtain accurate estimates of the extreme-value characteristics for the resampling model used and to get some idea of the uncertainty of these estimates. The required daily average precipitation over subcatchments can be passively simulated.

The final application to the design discharge requires long-duration simulations and a method to estimate the 1250-year discharge event from these simulations. The estimated design discharge is subject to random errors resulting from the finite length of the simulation runs and the use of a relatively short historical record for resampling. Especially the latter uncertainty is difficult to quantify. One possibility is to repeat the long-duration simulations with subsamples of the 1961-1995 period.

Conditional simulations show a relatively large bias in extreme-value properties of $N$-day precipitation and snowmelt. Regarding applications for future climate conditions, it is expedient to investigate how far these biases influence the distribution of extreme river discharges. Such an investigation requires an extension of the validation procedure above with a number of 35-year simulations from a conditional model.

# Acknowledgements

We thank W. van de Langemheen, W.E. van Vuuren, H.C. van Twuiver and C. Geerse (RIZA) for comments on the earlier drafts of the report. The UK Meteorological Office gridded MSLP data were kindly provided by P.D. Jones (Climatic Research Unit, University of East Anglia, Norwich). The daily precipitation and temperature data for the German, Luxemburgian, French and Swiss stations were made available by the following institutions:

- Deutscher Wetterdienst
- Service de la météorologie et de l 'hydrologie de Luxembourg
- Météo France
- Swiss Meteorological Institute

via the International Commission for the Hydrology of the Rhine Basin (CHR).

# References


A. Bárdossy and E.J. Plate. Space-time model for daily rainfall using atmospheric circulation patterns. *Water Resour. Res.*, 28:1247–1259, 1992.

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

J.J. Beersma and T.A. Buishand. A simple test for equality of monthly variances in climate time series. *J. Climate*, 12:1770–1779, 1999b.

A.R. van Bennekom and B.W.A.H. Parmet. *Bemessungsabfluß in den Niederlanden; Menschliche Einflüsse und Andere Unsicherheiten*, pages 125–131. In: Zukunft der Hydrologie in Deutschland, BfG Mitteilung 16. Bundesanstalt für Gewässerkunde, Koblenz, 1998.

T. Brandsma and T.A. Buishand. Rainfall generator for the Rhine basin: Single-site generation of weather variables by nearest-neighbour resampling. KNMI-publication 186-I, KNMI, De Bilt, 1997.

T. Brandsma and T.A. Buishand. Simulation of extreme precipitation in the Rhine basin by nearest-neighbour resampling. *Hydrol. Earth Syst. Sci.*, 2:195–209, 1998.

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

T.A. Buishand and J.J. Beersma. Jacknife tests for differences in autocorrelation between climate time series. *J. Climate*, 6:2490–2495, 1993.

T.A. Buishand and J.J. Beersma. Statistical tests for comparison of of daily variability in observed and simulated climates. *J. Climate*, 9:2538–2550, 1996.

T.A. Buishand and T. Brandsma. Rainfall generator for the Rhine catchment: A feasibility study. Technical Report TR-183, KNMI, De Bilt, 1996.

T.A. Buishand and T. Brandsma. Multi-site simulation of daily precipitation and temperature in the Rhine basin by nearest-neighbour resampling. Preprint 2000-14, KNMI, De Bilt (Submitted to Water Resour. Res.), 2000.

CHR. Le bassin du Rhin, Monographie Hydrologique, Tome A: Textes. Comission Internationale de l' Hydrologie du Basin du Rhin (CHR), Den Haag, 1978.

Delft Hydraulics and EAC-RAND. Toetsing uitsgangpunten rivierdijkversterkingen, Deelrapport 2: Maatgevende belastingen. Delft Hydraulics, Emmeloord, and European American center for Policy Analysis (EAC-RAND), Delft, the Netherlands, 1993.

D.M. Gray and T.D. Prowse. *Snow and Floating Ice*. Chapter 7 in Handbook of Hydrology (D.R. Maidment, ed.), McGraw-Hill, New York, 1993.

T.J. Hastie and R.J. Tibshirani. *Generalized Additive Models*. Chapman and Hall, London, 1990.

P.D. Jones, M. Hulme, and K.R. Briffa. A comparison of Lamb circulation types with an objective classification scheme. *Int. J. Climatol.*, 13:655–663, 1993.

M. Kendall, A. Stuart, and J.K. Ord. *The Advanced Theory of Statistics*, volume 3: *Design and Analysis, and Time Series*. Charles Griffin, London & High Wycombe, 1983.

U. Lall and A. Sharma. A nearest neighbor bootstrap for resampling hydrologic time series. *Water Resour. Res.*, 32:679–693, 1996.

R.K. Linsey, M.A. Kohler, and J.L.H. Paulhus. *Hydrology for Engineers*. McGraw-Hill, London, 1988.

NERC. *Flood Studies Report, Volume II: Meteorological Studies*. Natural Environment Research Council (NERC), London, 1975.

B. Rajagopalan and U. Lall. A k-nearest-neighbor simulator for daily precipitation and other variables. *Water Resour. Res.*, 35:3089–3101, 1999.

K. C. Young. A multivariate chain model for simulating climatic parameters from daily data. *J. Appl. Meteorol.*, 33:661–671, 1994.

E. Zorita and H. von Storch. The analog method as a simple statistical downscaling technique: Comparison with more complicated models. *J. Climate*, 12:2474–2488, 1999.

# List of publications on the rainfall generator for the Rhine Basin

## Reports

T.A. Buishand and T. Brandsma. Rainfall generator for the Rhine catchment: A feasibility study. Technical Report TR-183, KNMI, De Bilt, 1996

T. Brandsma and T.A. Buishand. Rainfall generator for the Rhine basin: Single-site generation of weather variables by nearest-neighbour resampling. KNMI-publication 186-I, KNMI, De Bilt, 1997

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

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

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

## Journal articles

T. Brandsma and T.A. Buishand. Simulation of extreme precipitation in the Rhine basin by nearest-neighbour resampling. *Hydrol.Earth.Syst.Sci.,* 2:195-209, 1998

T.A. Buishand and T. Brandsma. Multi-site simulation of daily precipitation and temperature in the Rhine basin by nearest-neighbour resampling. Preprint 2000-14, KNMI, De Bilt, 2000 (Submitted to Water Resour. Res.)