

Rainfall Generator for the Rhine Basin

**Multi-site generation of weather variables
by nearest-neighbour resampling**

*Theo Brandsma
T. Adri Buishand*

KNMI publication 186-II

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

Table of contents

TABLE OF CONTENTS	3
SUMMARY	4
1. INTRODUCTION	7
1.1 BACKGROUND.....	7
1.2 PREVIOUS RESEARCH	7
1.3 SCOPE AND OBJECTIVES	10
1.4 OUTLINE.....	10
2. METHODOLOGY	11
2.1 NEAREST-NEIGHBOUR RESAMPLING	11
2.2 DATA DESCRIPTION	14
3. CONSTRUCTION OF D_T AND INFLUENCE OF K	17
3.1 SUMMARY STATISTICS FOR P AND T FIELDS.....	17
3.2 CONSTRUCTION OF D_T	18
3.2.1 Daily autocorrelation and variability of monthly values	20
3.2.2 N -day winter maximum precipitation amounts	23
3.2.3 Correlation with circulation indices	25
3.3 INFLUENCE OF K	26
3.3.1 The need for further study	26
3.3.2 Daily autocorrelation and variability of monthly values	26
3.3.3 N -day winter maximum precipitation amounts	27
3.3.4 Shape parameter of the GEV distribution	28
3.3.5 Correlation with circulation indices	29
3.3.6 Run lengths in unconditional simulations	30
4. FURTHER ANALYSIS OF THE DISTRIBUTION OF EXTREMES	31
4.1 N -DAY WINTER MAXIMUM PRECIPITATION AMOUNTS	31
4.2 N -DAY MAXIMUM SNOWMELT AMOUNTS	33
5. SPATIAL ASSOCIATION OF N-DAY PRECIPITATION MAXIMA	35
5.1 THEORETICAL BACKGROUND.....	35
5.2 APPLICATION TO THE RHINE BASIN.....	37
6. LONG-DURATION SIMULATIONS	41
6.1 DISTRIBUTION OF THE N -DAY WINTER MAXIMA.....	41
6.2 SPACE-TIME PATTERN OF SELECTED 10-DAY EVENTS	43
7. DISCUSSION AND CONCLUSIONS	49
ACKNOWLEDGEMENTS	51
REFERENCES	52
APPENDICES	54
A. STATISTICAL PROPERTIES OF RUN LENGTHS	54
B. RELATIVE VARIABILITY OF STATISTICS USED TO COMPARE EXTREME-VALUE PROPERTIES	55

Summary

This report presents the second phase 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. The first phase dealt with the single-site generation of weather variables by nearest-neighbour resampling for seven stations in the German part of the Rhine basin. The second phase is concerned with the multi-site extension of the first phase, using daily precipitation and temperature data for twenty-five stations (1961–1995) in the German part of the Rhine basin. Joint simulation of daily precipitation and daily temperature is necessary to account for the effect of snow and frozen soils on large river discharges.

The nearest-neighbour resampling technique is a method of simulation that can easily handle multi-site daily precipitation and temperature data without making restrictive assumptions concerning the joint distribution of those data. The essence of this technique is that the variables for a new day are sampled with replacement from a selected set of historical data (the nearest neighbours or analogues). In order to generate weather variables for day t , the method needs a feature (or state) vector \mathbf{D}_t to find nearest neighbours in the historical data. For unconditional simulation, \mathbf{D}_t contains variables that characterise the state of the weather on day $t-1$, like summary statistics of the daily precipitation and temperature fields and atmospheric circulation indices. Conditional simulation on the atmospheric flow is also possible with the nearest-neighbour resampling technique. Circulation indices for day t are then included in \mathbf{D}_t . A (weighted) Euclidean distance measure is used to determine the k nearest neighbours of \mathbf{D}_t . From the k nearest neighbours, one day is sampled with a predefined probability.

Different compositions of the feature vector are investigated using $k = 20$ nearest neighbours in the resampling procedure. The influence of the magnitude of k is further studied both for a selected case of unconditional simulation and for one of conditional simulation. The most important criteria used to judge the performance of a particular resampling technique are the ability to reproduce the autocorrelation structure, the variability of monthly values, the distribution of multi-day winter (October-March) maximum precipitation amounts and the spatial dependence of these amounts.

With respect to the reproduction of the autocorrelation structure and the variability of monthly values, there are small, but often statistically significant, departures from the historical values for all choices of \mathbf{D}_t . The unconditional method performs somewhat better than the conditional method, which may partly be ascribed to the fact that the former often selects successive days in the historical record. For unconditional multi-site simulation, it is advantageous to omit the circulation indices in \mathbf{D}_t . Especially the reproduction of the temperature autocorrelation benefits from this.

The success of reproducing the distributions of the multi-day winter maximum precipitation amounts is strongly related to that of the autocorrelation properties. For all choices of \mathbf{D}_t , there is a small systematic underprediction of the median and upper quantiles of these distributions. As for the reproduction of the autocor-

relation structure and the variability of monthly values, the unconditional method performs somewhat better than the conditional method.

For most conditional simulations the correlation coefficients between the circulation indices and the weather variables are also considered. The reproduction of these coefficients is important for applications to simulated atmospheric-flow indices (e.g. from a general circulation model) or to observed circulation indices outside the period 1961–1995. For the winter half-year both precipitation and temperature have the strongest correlation with the westerly flow index, which is somewhat too low in the simulated cases. This underestimation is also found for the correlation between precipitation and vorticity, whereas the correlation between temperature and the southerly flow index is slightly overestimated.

The reproduction of the autocorrelation structure and the variability of monthly values clearly improves at small k ($k \leq 5$). This is reflected in the reproduction of the distribution of the multi-day winter maximum precipitation amounts. In the conditional simulations, the correlation coefficients between the circulation indices and the weather variables are also better preserved for small k . The Generalised Extreme Value distribution has been fitted to the N -day winter maximum precipitation amounts to study the reproduction of its shape parameter. For all values of k this parameter is adequately preserved.

For the $k = 5$ simulations, the distributions of multi-day winter maximum precipitation amounts and multi-day maximum snowmelt are analysed in more detail for seven stations and five large subareas of the basin. The selected stations correspond to those used in the single-site simulations in the first phase of the project. For both the unconditional and conditional case, the results for the seven stations are somewhat better than those reported for the single-site simulations with $k = 20$. Especially, the reproduction of the median of the multi-day winter maximum precipitation amounts improves. The results for the maxima of area-average rainfall are comparable with those for the selected stations.

The daily temperatures are used to determine snow accumulation and melt. Despite the systematic underprediction of the temperature autocorrelation, the reproduction of the distribution of the multi-day maximum snowmelt is satisfactory.

The spatial dependence of the multi-day winter precipitation maxima is analysed by counting the years that a given threshold is not exceeded at pairs of stations. Choosing different heights of the threshold gives a more complete description of the degree of spatial association than just a correlation coefficient. The dependence is somewhat stronger at the median than in the upper tail of the distribution. The spatial dependence of the two $k = 5$ simulations is compared with that of the historical data. The results for the unconditional simulation are in good agreement with the historical data, while for the conditional simulation the spatial dependence is stronger, though not statistically significant, than that of the historical data.

The simulation of unprecedented extreme rainfall events is illustrated with two unconditional simulations of 1000 years, one for $k = 5$ and one for $k = 20$. Three historical and six simulated events with extreme 10-day precipitation over the total area are selected from the $k = 5$ simulation and their space-time patterns are compared. There is often a clear maximum over the Black Forest in these events,

in agreement with that in the long-term mean rainfall over the southern part of Germany. Sometimes, however, the maximum 10-day amount occurs in the western part of Germany over the Mosel and the Nahe subbasins. The temporal patterns of the daily amounts over large areas differ widely between the selected extreme events.

We conclude that with nearest-neighbour resampling a reasonable reproduction of autocorrelation coefficients and properties of extreme rainfall and snowmelt can be achieved, in particular in the unconditional simulations. In the two 1000-year simulations, multi-day precipitation amounts were much larger than the observed maxima. The space-time patterns of daily rainfall in the selected simulated extreme events also differed from those in historical extreme events. An important future development is to generate circulation indices separately, which can then be used to obtain long-duration synthetic sequences of daily precipitation and temperature by conditional simulation.

1. Introduction

1.1 Background

The Rhine is the most important river in the Netherlands. The river flows through several countries (Figure 1). Large parts of its drainage basin are situated in Switzerland, Germany, France and the Netherlands. 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 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 large river 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 record. Different distributions then lead to quite different design discharges. The fact that the parameters of these distributions have to be estimated from a finite record introduces 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 bed and to predict the potential impacts of climate change. The Institute of 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 a hydrological/hydraulic model, the development of a stochastic rainfall generator was also planned in order to produce long-duration daily rainfall series over the basin. Different spatial and temporal patterns of heavy precipitation in the generated series may lead to more extreme discharges at Lobith than those experienced in the past century. The use of such synthetic data in combination with a hydrological/hydraulic model does not only provide the peak discharges but also the duration of extreme river discharges, which may lead to a better insight into the shape of the design flood.

1.2 Previous research

At the request of RIZA, KNMI carried out a feasibility study into the possibilities of a rainfall generator (Buishand and Brandsma, 1996). In that study, the statistical techniques for generating daily rainfall sequences were reviewed. For a multi-site application in a large catchment like the Rhine basin, two, quite different, alternatives were discussed: (1) parametric time series modelling of the observed daily precipitation using a transformed multivariate AR(1) process, and (2) non-parametric resampling from historical data. Although some promising results have been reported for the two methods, there is a serious lack of knowledge about the reproduction of properties of extreme rainfall. Some of these properties, like the extreme-value distributions of multi-day amounts and the spatial association of large amounts during winter, are important for the peak discharges of the Rhine in the Netherlands.

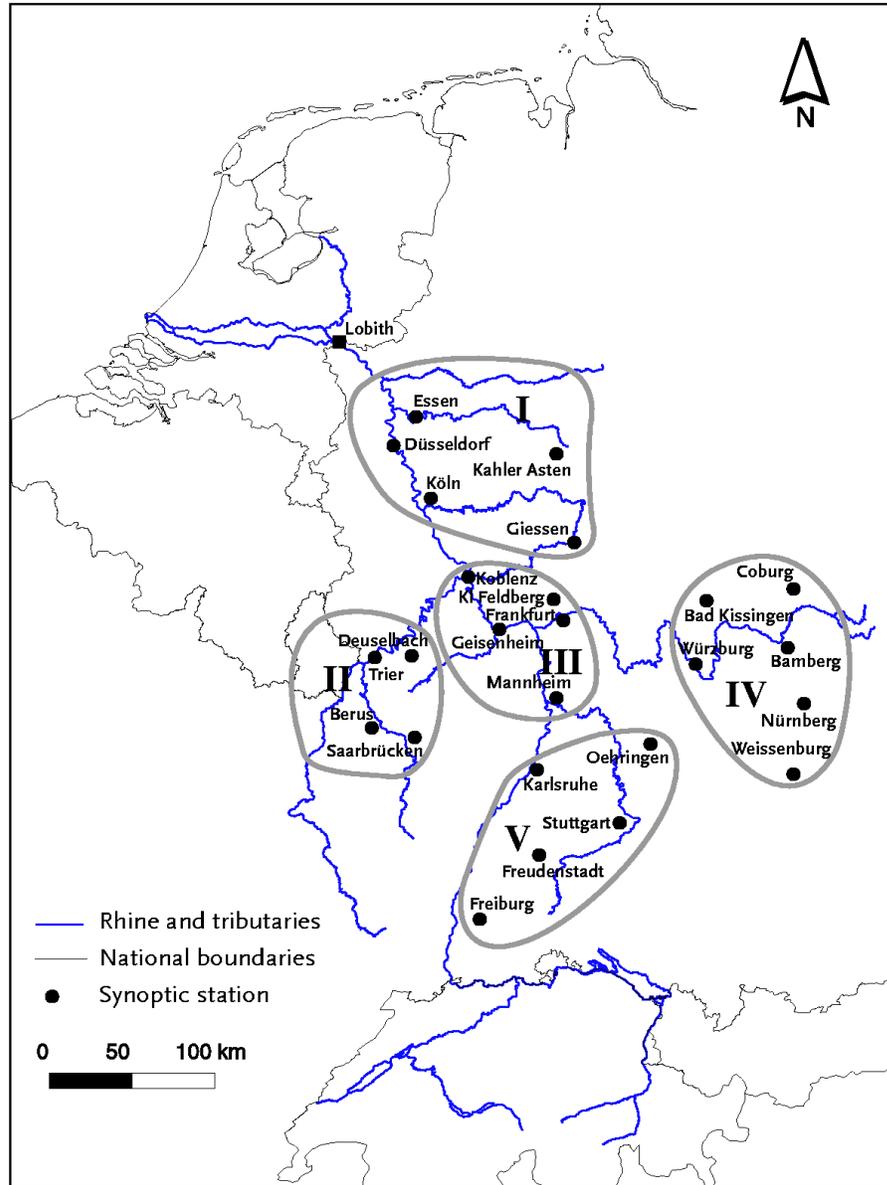


Figure 1: Location of Lobith in the Netherlands and the twenty-five German stations used in this study in the drainage basin of the river Rhine. A subdivision of the stations into five groups is also shown (see Section 3).

The results of the feasibility study were promising enough to start the development of a rainfall generator for the Rhine basin. For the intended application, it is important to extend the generation of daily precipitation with that of daily temperature in order to account for the effect of snow and frozen soils on large river discharges. The dependence of precipitation on the atmospheric circulation has also been considered in the development of the rainfall generator. This linkage has been used by others to improve the reproduction of the persistence of daily rainfall (Katz and Parlage, 1993) and to assess the effects of systematic changes in the atmospheric circulation, e.g. resulting from increased atmospheric greenhouse gas concentrations (Bárdossy and Plate, 1992; Wilby and Wigley, 1997).

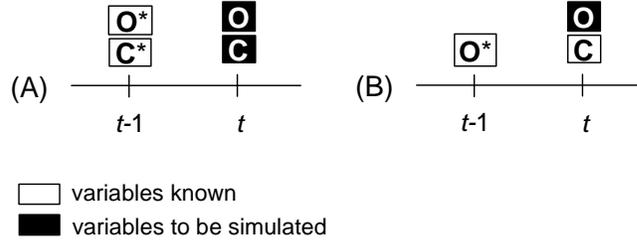


Figure 2: Two methods for the generation of new variables (solid squares) using different sets of known variables (open squares). O refers to weather variables like precipitation and temperature and C is a vector of circulation indices. The asterisks indicate that the corresponding variables are simulated values of the previous time steps.

The present report is the second in a series of reports on the development of a rainfall generator. In the first report (Brandsma and Buishand, 1997; further denoted as BB97), we studied the nonparametric nearest-neighbour method for single-site generation of daily precipitation and temperature. This method has been introduced in the hydrological literature by Rajagopalan and Lall (1995) and Lall and Sharma (1996). It is strongly related to resampling from analogues in climate change studies (Zorita *et al.*, 1995). In BB97 encouraging results were obtained for seven stations in the German part of the Rhine basin. These results are discussed further in Brandsma and Buishand (1998). The main features of nearest-neighbour resampling are summarised here. Further details are presented in Section 2.1.

The essence of the nearest-neighbour resampling technique is that the variables for a new day are sampled with replacement from a selected set of historical data (the nearest neighbours or analogues). In order to generate weather variables for day t , the method needs a feature (or state) vector \mathbf{D}_t to find nearest neighbours in the historical data. Figure 2 shows compositions of \mathbf{D}_t for an unconditional simulation (A) and a conditional simulation (B). For unconditional simulation, \mathbf{D}_t contains variables that characterise the state of the weather on day $t-1$, e.g. daily precipitation, temperature and a number of atmospheric circulation indices. Circulation indices for day t are included in \mathbf{D}_t in case of conditional simulation. A (weighted) Euclidean distance measure is used to determine the k nearest neighbours of \mathbf{D}_t . From the k nearest neighbours, one day is sampled with a pre-defined probability.

In BB97, we studied the influence of k , the composition of \mathbf{D}_t , and some other options in the resampling procedure. The emphasis was on the reproduction of autocorrelation coefficients and extreme N -day precipitation amounts in the winter half-year (October-March). With respect to the reproduction of these properties, unconditional simulation of precipitation, temperature and circulation indices turned out to be at least as good as conditional simulation of precipitation and temperature on circulation indices. Furthermore, we concluded that both the weather variables and circulation indices must be considered in \mathbf{D}_t . A point of some concern is the systematic underprediction of the median of the N -day precipitation maxima.

An important objective of the rainfall generator is the simulation of unprecedented extreme rainfall situations over the Rhine basin. With a 300-year simulation we showed in BB97 that more extreme multi-day precipitation amounts could be generated than the largest observed values. Unconditional simulation

was considered only, because conditional simulation would require a separate model to extend the observed record of circulation indices. The latter is, however, a point of future research.

1.3 Scope and objectives

The main objective of this report is the multi-site extension of BB97. The results in BB97 justify the development of a such an extension. At first, this extension can be restricted to the German part of the Rhine basin. For this part of the river basin we consider daily precipitation and temperature data for 25 stations (1961–1995). The work can be regarded as the first detailed investigation and application of the works of Rajagopalan and Lall (1995) and Lall and Sharma (1996) for a large river basin. Because BB97 showed no clear preference to either the conditional or unconditional method, we continue to study both methods further.

Multi-site generation by nearest-neighbour resampling is, in fact, rather straightforward. The composition of the feature vector \mathbf{D}_i needs, however, special attention. A considerable growth of the dimension of \mathbf{D}_i may not be desirable. Therefore, we have to use summary statistics of the daily precipitation and temperature fields, just as we used three air-flow indices to characterise the mean sea-level pressure (MSLP) field.

Although a resampling technique preserves the spatial dependence of the 1-day amounts, this is not necessarily true for the multi-day amounts. The reproduction of the spatial association of large multi-day amounts should therefore be tested.

1.4 Outline

The report is organised as follows. Section 2 provides the necessary background of the nearest-neighbour method and describes the data. Section 3 deals with the construction of \mathbf{D}_i and the influence of k on the autocorrelation coefficients and the distribution of N -day maximum winter precipitation. We then proceed with one conditional and one unconditional case that gave interesting results in Section 3. In Section 4, we present results for the distributions of N -day winter maximum precipitation and N -day maximum snowmelt. The spatial association of the precipitation maxima is dealt with in Section 5. In Section 6, we show some results for two 1000-year simulations. The methodology is evaluated in Section 7.

2. Methodology

In the present section we first introduce the nearest-neighbour resampling method followed by a description of the data. The presented material is taken from BB97 and only slightly adapted for the purpose of this report.

2.1 Nearest-neighbour resampling

The principle of the nearest-neighbour resampling method is simple. For the generation of precipitation and temperature on day t first a feature vector \mathbf{D}_t is formed to find analogue situations in the historical data. In the method of Rajagopalan and Lall (1995) for generating multivariate daily weather data at a single site, \mathbf{D}_t contains the values of the weather variables generated for day $t-1$. The k nearest neighbours (k -NN) of \mathbf{D}_t , in terms of Euclidean distance, are abstracted from the historical record. Let $t(j)$, $j=1, \dots, k$ be the times associated with these nearest neighbours, such that the distance of $\mathbf{D}_{t(j)}$ to \mathbf{D}_t increases with increasing j . The vector of weather variables following $\mathbf{D}_{t(j)}$, the successor to $\mathbf{D}_{t(j)}$, is denoted as $\mathbf{x}_{t(j)}$.

One of the successors of the k -NN is sampled using a discrete probability distribution or kernel $\{p_j\}$. For the uniform kernel, p_j is given by:

$$p_j = 1/k, \quad j = 1, \dots, k \quad (1)$$

In Lall and Sharma (1996) the following decreasing kernel was recommended:

$$p_j = \frac{1/j}{\sum_{i=1}^k 1/i}, \quad j = 1, \dots, k \quad (2)$$

Figure 3 shows the two kernels for $k=5$ and $k=20$. The decreasing kernel gives relatively high probability mass to the closest neighbours, whereas the uniform kernel assigns the same probability to all neighbours. In BB97 comparable results were obtained for both kernels, if k for the uniform kernel was taken relatively small ($k \approx 5$) compared to the value for the decreasing kernel ($k \approx 20$). Besides

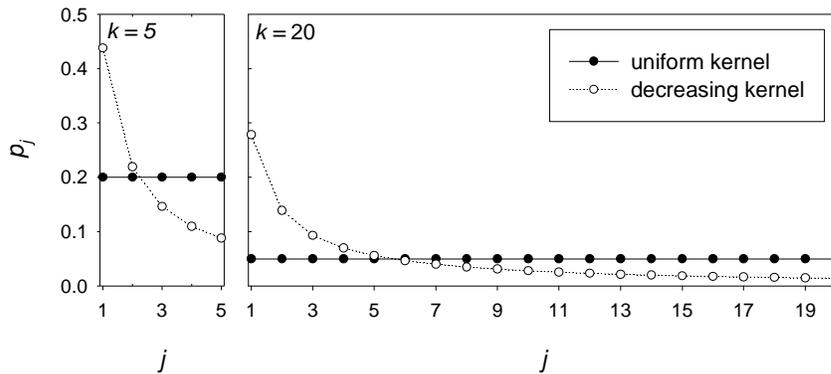


Figure 3: Resampling probability p_j as a function of the index j of the sorted Euclidean distances for $k=5$ and $k=20$ for both the uniform and decreasing kernels.

the kernel type, the choice of k depends on the number q of variables in \mathbf{D}_t and the record length. Brandsma and Buishand (1998) resampled from historical records of 30 years using a decreasing kernel with $k = 20$, which is somewhat lower than that recommended in Lall and Sharma (1996). In this report we only consider the decreasing kernel and the sensitivity to k is further explored.

We distinguish between unconditional and conditional simulation. Method A in Figure 2 is an example of unconditional simulation, where \mathbf{D}_t contains the weather variables and circulation indices generated for the previous day. Resampling occurs from the successors to the k nearest neighbours. Method B, is an example of conditional simulation of weather variables given the atmospheric circulation. \mathbf{D}_t then contains circulation indices on day t as well as the values of weather variables that were generated for the previous day. Resampling occurs from the observed precipitation and temperatures on the days $t(j)$ in the nearest neighbourhood. In BB97 we compared six possible compositions of \mathbf{D}_t . Here we consider eight compositions of \mathbf{D}_t , adapted for multi-site generation of weather variables (see Section 3).

To account for the systematic annual cycle in the various weather variables, the search for the k nearest neighbours of the feature vector is restricted to days in a specified moving window of width W_{mw} days, centred at the day of interest (see Figure 4). The use of a moving window, instead of fixed seasons, prevents sharp transitions between seasons. For $W_{mw} = 61$ days and a historical time series of 35 years, as considered in this report, the Euclidean distances for a specific day are calculated for $61 \times 35 = 2135$ days. In BB97, we compared several values of W_{mw} and obtained satisfactory results for $W_{mw} = 61$ days. Consequently, only this value is used in the present report.

A further reduction of seasonal variation can be achieved by working with standardised variables. In Rajagopalan and Lall (1995) standardisation was done by subtracting the calendar day's mean m_d and dividing by the calendar day's sample standard deviation s_d :

$$\tilde{x}_t = (x_t - m_d) / s_d, \quad t = 1, \dots, 365J; d = (t - 1) \bmod 365 + 1 \quad (3)$$

where x_t and \tilde{x}_t are the original and standardised variable, respectively, for day t , and J is the total number of years in the time series. For variables with a normal

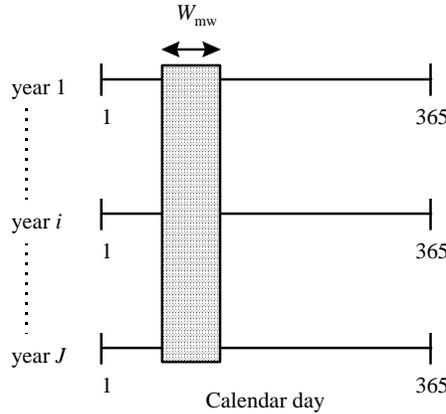


Figure 4: Moving window

or almost normal distribution, \tilde{x}_t usually takes values between -3 and $+3$. However, for daily precipitation the range of \tilde{x}_t is quite different. For a dry day, $\tilde{x}_t = -m_d / s_d \approx -0.5$ at a lowland station in the Rhine basin, whereas for days with heavy precipitation $\tilde{x}_t \approx 10$.

For daily precipitation, resampling of negative standardised values from adjacent days may result in negative precipitation amounts. Thus equation (3) does not give the most appropriate standardisation for that variable. In hydrology, division by the mean is a popular method to standardise non-negative variables. For daily precipitation, we therefore considered:

$$\tilde{x}_t = x_t / m_{d,\text{wet}} \quad (4)$$

where $m_{d,\text{wet}}$ is the calendar day's mean precipitation for wet days. For dry days $\tilde{x}_t = 0$ and for the most extreme wet days \tilde{x}_t is about 10. In BB97 we also considered some alternatives to equation (4) which, however, did not lead to an improvement of the results.

To reduce the effect of sampling variability, we used smooth approximations of m_d , $m_{d,\text{wet}}$ and s_d instead of the raw values (see further Section 2.2).

Through the standardisation, the q elements of the feature vector \mathbf{D}_t are dimensionless quantities. The weighted Euclidean distance between two vectors \mathbf{D}_t and \mathbf{D}_u is given by:

$$\delta_{tu} = \sqrt{\sum_{i=1}^q w_i (v_{ti} - v_{ui})^2} \quad (5)$$

with v_{ti} and v_{ui} the elements of the vectors \mathbf{D}_t and \mathbf{D}_u , and w_i the weight associated with the i th variable. In BB97 we mostly set w_i equal to 1 for all elements of \mathbf{D}_t . In this report, the influence of the weights w_i is further explored.

The final simulated value $x_{t,\text{sim}}$ for day t is obtained from the standardised resampled value $\tilde{x}_{t(j)}$ by inverting equation (3) or (4):

$$x_{t,\text{sim}} = m_d + s_d \tilde{x}_{t(j)} \quad (6)$$

$$x_{t,\text{sim}} = m_{d,\text{wet}} \tilde{x}_{t(j)} \quad (7)$$

Because $\tilde{x}_t \geq 0$ in equation (4), the simulated value $x_{t,\text{sim}}$ in equation (7) cannot be negative, as required for precipitation. This is clearly not true for equation (6).

Figure 5 presents a flow diagram of the full resampling procedure for the unconditional method A of Figure 2. The starting day in this figure is generated by random sampling a day within the window for 1 January. For method B, the starting day is generated by resampling a day from the 50 nearest neighbours of the observed circulation on the first day, using a uniform kernel.

From the presentation above, it is clear that there are various options in the nearest-neighbour method. In BB97 we showed that the most important of these is the construction of the feature vector \mathbf{D}_t . Here we study the composition of \mathbf{D}_t for multi-site resampling. As mentioned earlier, the effects of varying k and the weights w_i are also explored.

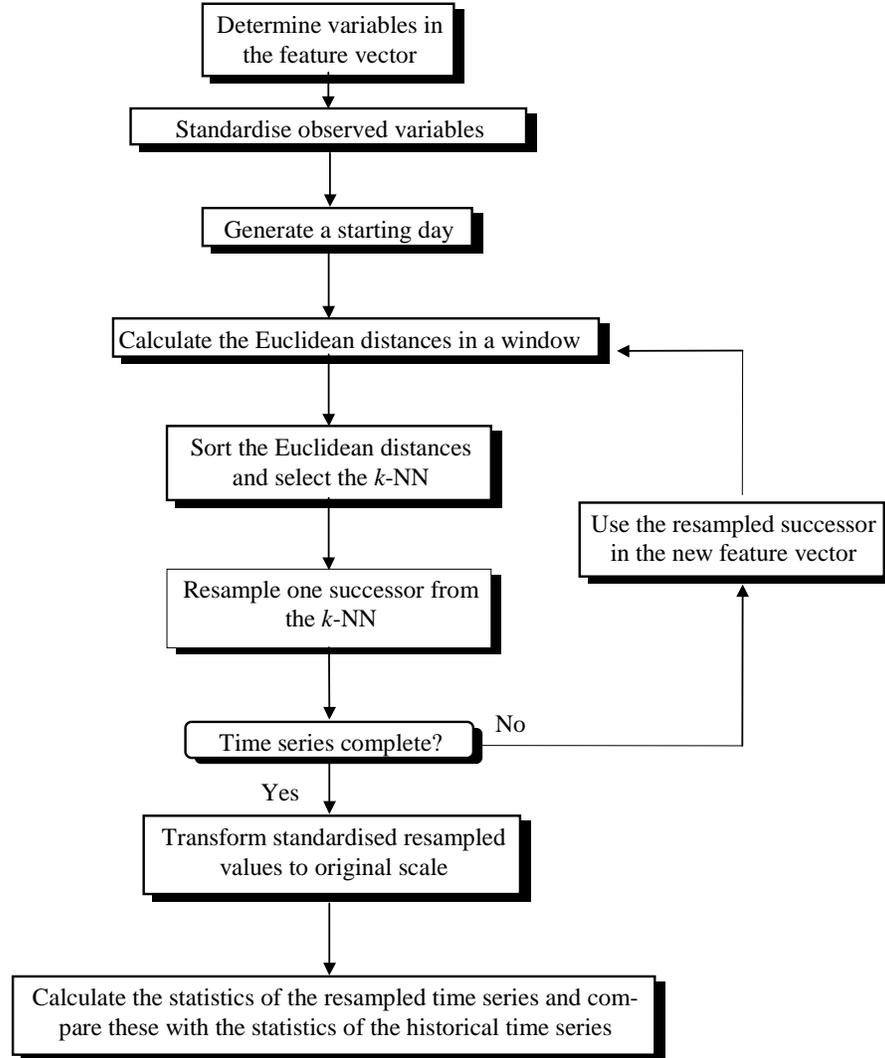


Figure 5: Flow diagram for method A (see Figure 2).

2.2 Data description

For the research described in this report, precipitation and temperature data for 25 stations were analysed for the period 1961–1995, which is 5 years more than in BB97. The stations are situated in the German part of the Rhine basin (Figure 1). The data were made available by the Deutscher Wetterdienst via the ‘International Commission for the Hydrology of the Rhine Basin’ (CHR/KHR). Table 1 presents the mean annual temperature and precipitation of these stations, together with the station elevation. There are three stations at an altitude of about 800 m: Kahler Asten (in Sauerland in the northern part of the basin), Freudensstadt (in the Black Forrest in the southern part of Germany) and Kl. Feldberg. The annual mean temperature is relatively low at these stations. The other sta-

No.	Station	Altitude (m above m.s.l.)	Mean annual temperature (°C)	Mean annual pre- cipitation (mm)
1	Stuttgart	373	9.0	713
2	Frankfurt	112	9.9	645
3	Kahler Asten	839	5.0	1474
4	Trier	265	9.2	783
5	Essen	152	9.7	928
6	Bamberg	239	8.6	632
7	Freudenstadt	797	6.7	1691
8	Düsseldorf	37	10.4	759
9	Saarbrücken	319	9.0	867
10	Berun	363	8.9	835
11	Köln	92	9.9	807
12	Geisenheim	118	10.0	542
13	Koblenz	85	10.6	670
14	Deuselbach	480	8.0	808
15	Freiburg	269	10.9	944
16	Giessen	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

Table 1: Characteristics of the stations that have been used in the study (mean annual values for the period 1961–1995).

tions are below 500 m. The highest mean annual precipitation is found at Kahler Asten and Freudenstadt, which is caused by orographic enhancement.

In a later stage of the project, we obtained daily precipitation data for 230 areas in Germany for the period 1961–1995, also made available by the Deutscher Wetterdienst. These data were derived from interpolated daily values on a 30'' latitude by 60'' longitude grid (B. Dietzer, pers. comm., 1998, 1999). The interpolation procedure used is similar to that described in Müller-Westermeier (1995), except that relative daily precipitation (relative with respect to the 1961–1990 monthly mean) was considered to account for topographic influences. The map of the mean precipitation in the winter half-year in Figure 6 is derived from these area-average data, using ArcView GIS. It shows marked variations in mean winter precipitation over the German part of the Rhine basin. The largest precipitation amounts (> 600 mm) are found in the Black Forest in the south and Sauerland in the north. The data for the 230 areas are only used further in Section 6 to present the space-time pattern of extreme precipitation events.

To incorporate atmospheric flow characteristics, we considered daily MSLP data from the UK Meteorological Office on a 5° latitude by 10° longitude grid. These data extend back to December 1880. For a grid centred at the Rhine basin (see Figure 7), we calculated three daily air-flow indices: (1) total shear vorticity Z ; (2) strength of the westerly flow W ; and (3) strength of the southerly flow S (see also Jones *et al.*, 1993). These three indices form the elements of the vector \mathbf{C} in Figure 2.

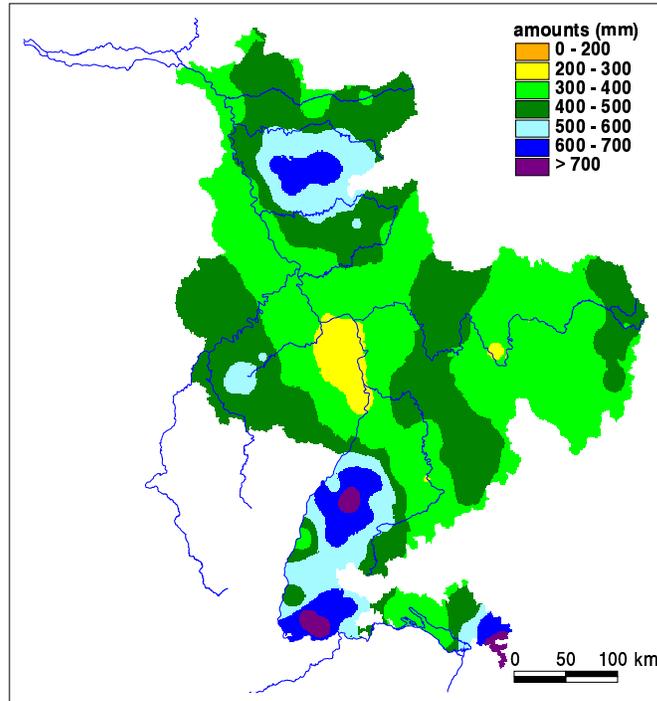


Figure 6: Spatial distribution of the mean winter (October–March) precipitation for the period 1961–1995 over the German part of the Rhine basin.

Before resampling, the data have been standardised using the smoothed values of the calendar day's mean, m_d or $m_{d,wet}$, and standard deviation, s_d , as described in Section 2.1. The smoothed values are based on Friedman's supersmoother (Härdle, 1990). Before calculating the smooths, the values for $d = 336, \dots, 365$ were inserted for $d < 1$ and the values for $d = 1, \dots, 30$ for $d > 365$ to harmonize the smoothed values at the beginning and end of the year. The largest mean wet-

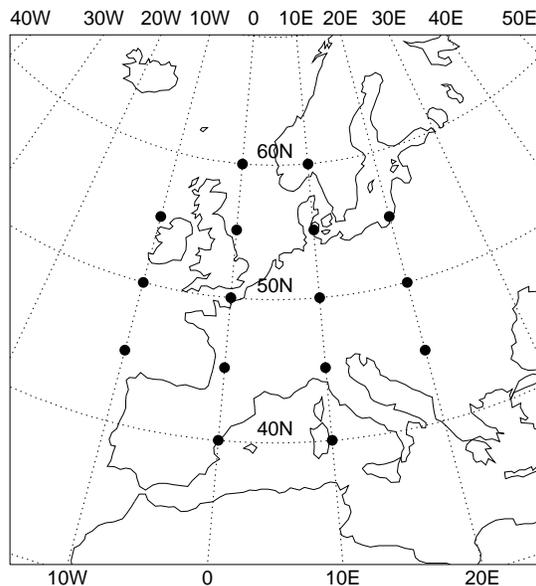


Figure 7: Grid points of mean sea-level pressure used for the calculation of the air-flow indices over the Rhine basin.

day precipitation amounts are found in summer (except for the high-elevation stations Kahler Asten and Freudenstadt), which is due to the influence of convection (summer showers). The mean westerly flow is relatively large in winter. The largest standard deviations of the circulation indices (vorticity, strength of the flow) and temperature are also found in winter (BB97).

3. Construction of \mathbf{D}_t and influence of k

In the present section we study the construction of the feature vector \mathbf{D}_t and the number of nearest neighbours k . In BB97 we used the daily precipitation (P) and temperature (T) observations for a single station in \mathbf{D}_t , in most cases supplemented with three atmospheric circulation indices. Here we have to deal with 25 stations simultaneously. It is not feasible to incorporate the individual daily P and T observations from these stations in \mathbf{D}_t . Considerable differences between the k nearest neighbours may then occur because of the large dimension of \mathbf{D}_t and there may be too much emphasis on the differences between local precipitation amounts in the selection of nearest neighbours. In order to keep the dimension of \mathbf{D}_t within reasonable limits, the P and T fields are described by a small number of statistics, just as we used the air-flow indices to characterise the MSLP field. These summary statistics are discussed first. Several test cases are then defined using different sets of the summary statistics for the MSLP, P and T fields in \mathbf{D}_t , and taking a range of values for the weights w_i (equation (5)) for each combination of the summary statistics. For all test cases, autocorrelation coefficients and properties of extreme N -day winter precipitation are compared with those in the historical record and the reproduction of the correlation between weather variables and circulation indices is explored. Thereafter, the influence of k on these quantities is studied for just two test cases.

3.1 Summary statistics for P and T fields

For each day we have daily P and T fields consisting of 25 stations for each field. An important summary statistic is the mean of these fields. Here we use the arithmetic means \tilde{P} and \tilde{T} of the standardised values for the 25 stations

$$\tilde{P} = \sum_{i=1}^{25} \tilde{P}_i / 25 \quad (8)$$

$$\tilde{T} = \sum_{i=1}^{25} \tilde{T}_i / 25 \quad (9)$$

where \tilde{P}_i and \tilde{T}_i are the standardised P and T values, respectively, for the i th station. The averaging reduces the variances. For an individual station $\text{var}(\tilde{P}_i) \approx 1.1$ and $\text{var}(\tilde{T}_i) = 1$, whereas $\text{var}(\tilde{P}) \approx 0.60$ and $\text{var}(\tilde{T}) \approx 0.89$ for the area-averages.

The daily field averages \tilde{T} explain about 90% of the daily temperature variance. For precipitation, however, \tilde{P} explains only 53% of the daily variance due to the relatively large spatial variation of this element. There is therefore some need for a more complete summary of the precipitation field than just \tilde{P} . One possibility is to use the daily averages $\tilde{P}_I, \dots, \tilde{P}_V$ over the five areas in Figure 1. The vector $\tilde{\mathbf{P}} = (\tilde{P}_I, \tilde{P}_{II}, \tilde{P}_{III}, \tilde{P}_{IV}, \tilde{P}_V)$ explains about 73% of the daily precipitation variance. This is almost the same as that achieved by the five leading principal components ($\approx 74\%$) obtained from the sample covariance matrix of the \tilde{P}_i 's. In contrast to principal component scores, the area-averages have about the same variance (≈ 0.83) and they show a rather strong correlation (ranging between 0.53 in summer

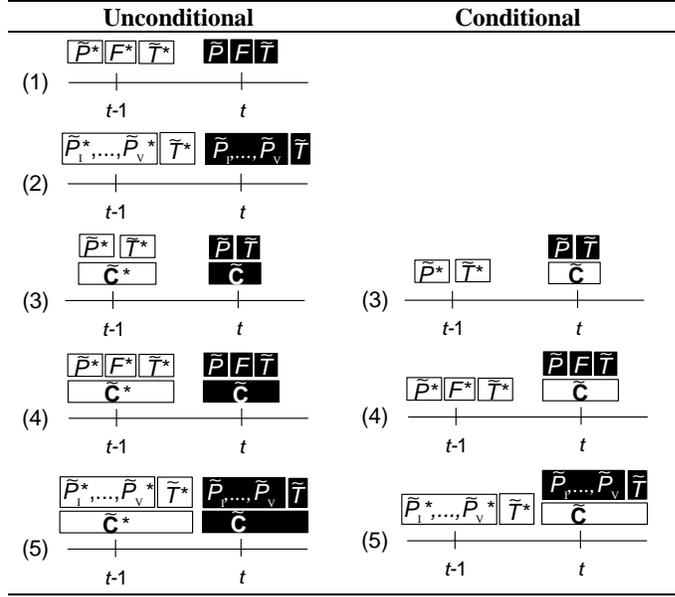


Table 2: Eight feature vectors for the generation of new variables (solid squares) using different sets of known variables (open squares). P refers to precipitation, T to temperature, F to the fraction of stations with $P > 0.2$ mm, \mathbf{C} is a vector of circulation indices. The asterisks indicate that the corresponding variables are resampled values of the previous time step and the tilde refers to standardised values.

and 0.75 in winter). An advantage of $\tilde{\mathbf{P}}$ is that the area-averages have an unambiguous interpretation. Principal component analysis is not considered further in this report.

As an alternative summary measure of P , we used the fraction F of stations with precipitation ($P > 0.2$ mm) in combination with \tilde{P} . The variable F helps to distinguish between large-scale and convective precipitation. The variance of F equals 0.15 and is thus much smaller than that of the other variables. It should further be noted that F and \tilde{P} are correlated. The correlation coefficient ranges between 0.69 in winter and 0.77 in summer, which is comparable to that between the area-averages in $\tilde{\mathbf{P}}$.

Summarising, we now have the following variables that can be used in \mathbf{D}_t : (1) the standardised circulation indices $\tilde{\mathbf{C}} = (\tilde{Z}, \tilde{W}, \tilde{S})$; (2) the mean \tilde{T} of the 25 standardised T values; (3) the mean \tilde{P} of the 25 standardised P values; (4) the means $\tilde{\mathbf{P}} = (\tilde{P}_I, \tilde{P}_{II}, \tilde{P}_{III}, \tilde{P}_{IV}, \tilde{P}_V)$ of the standardised P values for the five areas; and (5) the fraction F of stations with precipitation.

3.2 Construction of \mathbf{D}_t

Table 2 shows eight feature vectors \mathbf{D}_t that have been used in this study. A first division is made between unconditional and conditional methods. A second division is made within each category according to the weather variables included in \mathbf{D}_t . Note that the unconditional methods also consider two feature vectors without atmospheric circulation.

Case	Unconditional		Conditional	
	\mathbf{D}_t	w_i	\mathbf{D}_t	w_i
0.1				
1.1	$(\tilde{P}_{t-1}^*, F_{t-1}^*, \tilde{T}_{t-1}^*)$	(1,2,1)		
1.2	$(\tilde{P}_{t-1}^*, F_{t-1}^*, \tilde{T}_{t-1}^*)$	(2,2,2)		
2.1	$(\tilde{\mathbf{P}}_{t-1}^*, \tilde{T}_{t-1}^*)$	(1,5)		
2.2	$(\tilde{\mathbf{P}}_{t-1}^*, \tilde{T}_{t-1}^*)$	(1,4)		
2.3	$(\tilde{\mathbf{P}}_{t-1}^*, \tilde{T}_{t-1}^*)$	(1,3)		
2.4	$(\tilde{\mathbf{P}}_{t-1}^*, \tilde{T}_{t-1}^*)$	(1,2)		
2.5	$(\tilde{\mathbf{P}}_{t-1}^*, \tilde{T}_{t-1}^*)$	(1,1)		
3.1	$(\tilde{\mathbf{C}}_{t-1}^*, \tilde{P}_{t-1}^*, \tilde{T}_{t-1}^*)$	(1,1,1)	$(\tilde{\mathbf{C}}_t, \tilde{P}_{t-1}^*, \tilde{T}_{t-1}^*)$	(1,1,1)
3.2	$(\tilde{\mathbf{C}}_{t-1}^*, \tilde{P}_{t-1}^*, \tilde{T}_{t-1}^*)$	(1,2,2)	$(\tilde{\mathbf{C}}_t, \tilde{P}_{t-1}^*, \tilde{T}_{t-1}^*)$	(1,2,2)
3.3	$(\tilde{\mathbf{C}}_{t-1}^*, \tilde{P}_{t-1}^*, \tilde{T}_{t-1}^*)$	(1,3,3)	$(\tilde{\mathbf{C}}_t, \tilde{P}_{t-1}^*, \tilde{T}_{t-1}^*)$	(1,3,3)
4.1	$(\tilde{\mathbf{C}}_{t-1}^*, \tilde{P}_{t-1}^*, F_{t-1}^*, \tilde{T}_{t-1}^*)$	(1,1,1,1)	$(\tilde{\mathbf{C}}_t, \tilde{P}_{t-1}^*, F_{t-1}^*, \tilde{T}_{t-1}^*)$	(1,1,1,1)
4.2	$(\tilde{\mathbf{C}}_{t-1}^*, \tilde{P}_{t-1}^*, F_{t-1}^*, \tilde{T}_{t-1}^*)$	(1,1,2,1)	$(\tilde{\mathbf{C}}_t, \tilde{P}_{t-1}^*, F_{t-1}^*, \tilde{T}_{t-1}^*)$	(1,1,2,1)
4.3	$(\tilde{\mathbf{C}}_{t-1}^*, \tilde{P}_{t-1}^*, F_{t-1}^*, \tilde{T}_{t-1}^*)$	(1,1.5,2.5,1)	$(\tilde{\mathbf{C}}_t, \tilde{P}_{t-1}^*, F_{t-1}^*, \tilde{T}_{t-1}^*)$	(1,1.5,2.5,1)
4.4	$(\tilde{\mathbf{C}}_{t-1}^*, \tilde{P}_{t-1}^*, F_{t-1}^*, \tilde{T}_{t-1}^*)$	(1,2,2,2)	$(\tilde{\mathbf{C}}_t, \tilde{P}_{t-1}^*, F_{t-1}^*, \tilde{T}_{t-1}^*)$	(1,2,2,2)
4.5	$(\tilde{\mathbf{C}}_{t-1}^*, \tilde{P}_{t-1}^*, F_{t-1}^*, \tilde{T}_{t-1}^*)$	(1,2,4,2)	$(\tilde{\mathbf{C}}_t, \tilde{P}_{t-1}^*, F_{t-1}^*, \tilde{T}_{t-1}^*)$	(1,2,4,2)
4.6	$(\tilde{\mathbf{C}}_{t-1}^*, \tilde{P}_{t-1}^*, F_{t-1}^*, \tilde{T}_{t-1}^*)$	(1,3,5,2)	$(\tilde{\mathbf{C}}_t, \tilde{P}_{t-1}^*, F_{t-1}^*, \tilde{T}_{t-1}^*)$	(1,3,5,2)
4.7	$(\tilde{\mathbf{C}}_{t-1}^*, \tilde{P}_{t-1}^*, F_{t-1}^*, \tilde{T}_{t-1}^*)$	(1,3,3,3)	$(\tilde{\mathbf{C}}_t, \tilde{P}_{t-1}^*, F_{t-1}^*, \tilde{T}_{t-1}^*)$	(1,3,3,3)
4.8	$(\tilde{\mathbf{C}}_{t-1}^*, \tilde{P}_{t-1}^*, F_{t-1}^*, \tilde{T}_{t-1}^*)$	(1,3,6,3)	$(\tilde{\mathbf{C}}_t, \tilde{P}_{t-1}^*, F_{t-1}^*, \tilde{T}_{t-1}^*)$	(1,3,6,3)
5.1	$(\tilde{\mathbf{C}}_{t-1}^*, \tilde{\mathbf{P}}_{t-1}^*, \tilde{T}_{t-1}^*)$	(1,1/5,1)	$(\tilde{\mathbf{C}}_t, \tilde{\mathbf{P}}_{t-1}^*, \tilde{T}_{t-1}^*)$	(1,1/5,1)
5.2	$(\tilde{\mathbf{C}}_{t-1}^*, \tilde{\mathbf{P}}_{t-1}^*, \tilde{T}_{t-1}^*)$	(1,2/5,2)	$(\tilde{\mathbf{C}}_t, \tilde{\mathbf{P}}_{t-1}^*, \tilde{T}_{t-1}^*)$	(1,2/5,2)
5.3	$(\tilde{\mathbf{C}}_{t-1}^*, \tilde{\mathbf{P}}_{t-1}^*, \tilde{T}_{t-1}^*)$	(1,3/5,3)	$(\tilde{\mathbf{C}}_t, \tilde{\mathbf{P}}_{t-1}^*, \tilde{T}_{t-1}^*)$	(1,3/5,3)

Table 3: Definition of test cases for unconditional and conditional simulation. The w_i values for the circulation apply to all three components \tilde{Z} , \tilde{W} and \tilde{S} of the vector $\tilde{\mathbf{C}}$, likewise the w_i values for $\tilde{\mathbf{P}}_{t-1}^*$ refer to its five components. The asterisks indicate that the corresponding variables are resampled values of the previous time step and the tilde refers to standardised values.

For each of the eight feature vectors Table 3 defines a number of test cases. The test cases are obtained by varying the weights w_i (equation (5)) for the variables in \mathbf{D}_t . In this section $k = 20$ for all cases. Case 0.1 is a reference case, in which new days are resampled randomly from all days in the window (35x61 days).

For each case in Table 3, ten 35-year simulations were carried out. In case of unconditional simulation, these 35-year runs are independent through the use of different random number seeds. For the conditional methods, there is some correlation because each run is related to the observed circulation. Unless specified otherwise, the results in this report are obtained for each simulation separately and then averaged over all ten simulations.

Although the weights w_i for \tilde{P} , F , and \tilde{T} were taken to be the same in a number of cases, alternative choices of w_i were also considered because of the differences between the variances of these summary statistics. These differences suggest that F should have the largest weight and \tilde{T} the smallest. Alternative combinations of weights for these statistics are considered in cases 1.1, 4.2, 4.3, 4.5, 4.6 and 4.8.

3.2.1 Daily autocorrelation and variability of monthly values

The reproduction of the occurrence of extreme multi-day precipitation amounts requires that not only the lag 1 autocorrelation coefficient is preserved, but also the higher order autocorrelation coefficients. Here the autocorrelation coefficients were estimated for each calendar month separately, using the jackknife method of Buishand and Beersma (1993), in order to account for the annual cycles in the first and second order moments. Both for the historical and simulated data we calculated the mean lag 1, 2 and 3 autocorrelation coefficients for each station

$$\bar{r}_i(l) = \sum_{m=1}^{12} r_{i,m}(l)/12, \quad l = 1,2,3; \quad i = 1, \dots, 25 \quad (10)$$

with $r_{i,m}(l)$ the estimated lag l autocorrelation coefficient for the i th station and the m th month. We then computed the differences $\Delta\bar{r}(l)$ between the autocorrelation coefficients of the simulated and historical data, averaged over all stations

$$\Delta\bar{r}(l) = \frac{1}{25} \sum_{i=1}^{25} \bar{r}_i^*(l) - \bar{r}_i(l), \quad l = 1,2,3 \quad (11)$$

where \bar{r}_i^* and \bar{r}_i refer to the simulated and historical data, respectively.

The variability of monthly values is summarised here by the mean standard deviation of the monthly precipitation totals and monthly mean temperatures

$$\bar{s}_i = \sum_{m=1}^{12} s_{i,m}/12, \quad i = 1, \dots, 25 \quad (12)$$

From these values we calculated the percentage difference $\Delta\bar{s}$ between the standard deviations of the simulated and historical monthly values, averaged over all stations

$$\Delta\bar{s} = \frac{1}{25} \sum_{i=1}^{25} \frac{\bar{s}_i^* - \bar{s}_i}{\bar{s}_i} 100 \quad (13)$$

where \bar{s}_i^* and \bar{s}_i refer to the simulated and historical data, respectively.

For the simulated data, the monthly standard deviations tend to be too small if their autocorrelation is not adequately reproduced. In contrast to the comparisons for each individual lag, a test on the monthly standard deviation considers all lags simultaneously. This test is useful to discover a too rapid decay of the autocorrelation coefficients with increasing lag in the simulated data. Such a departure has often been observed with unconditional simulation of daily rainfall using simple parametric models (Buishand, 1978; Katz and Parlange, 1996).

In order to judge the statistical significance of the differences between the auto-correlation coefficients and standard deviations in the historical and simulated data, standard errors se were calculated for the 35-year historical record. The standard error of the mean lag l autocorrelation estimate $\bar{r}(l)$ was obtained by the jackknife method in Buishand and Beersma (1993). The jackknife procedure in Beersma and Buishand (1999) was used to obtain the standard deviation of $\sum_{i=1}^{25} \ln(\bar{s}_i) / 25$, which is approximately equal to the standard deviation of $\frac{1}{25} \sum_{i=1}^{25} [\bar{s}_i - E(\bar{s}_i)] / E(\bar{s}_i)$, using a first-order Taylor expansion of $\ln(\bar{s}_i)$ about the expected value of \bar{s}_i (Stuart and Ord, 1986, p.324). The latter gives an average relative deviation over all stations comparable with that in equation (13). A criterion of $2 \times se$ is used to indicate significant differences between historical and simulated values.

Table 4 presents $\Delta \bar{r}(l)$ and $\Delta \bar{s}$ for the cases defined in Table 3. The bottom line of the table gives $\bar{r}(l)$ and the mean standard deviation \bar{s} for the historical monthly values, averaged over all stations. From Table 4 the following can be noted:

- The search for nearest neighbours is necessary. Random sampling from all days in the window (case 0.1) leads to large systematic errors. Because there is no correlation in that case, the underprediction of $\bar{r}(l)$ is almost the same as $\bar{r}(l)$ itself. As a consequence \bar{s} is also strongly underpredicted (27.9% for P and 61.7% for T).
- There is always an underprediction of the lag 1 autocorrelation coefficient and the standard deviation of the monthly values, which is for most cases statistically significant. The unconditional cases perform somewhat better than the corresponding conditional cases. In the latter, the lag 2 and lag 3 autocorrelation coefficients are always slightly underpredicted.
- The omission of circulation indices in the feature vector of the unconditional cases 1.x and 2.x leads for temperature to a better reproduction of $\bar{r}(l)$ and \bar{s} . This can be seen by comparing cases with corresponding weights: 1.1 with 4.2, 1.2 with 4.4, and 2.1 with 5.1. Furthermore, the underprediction of the standard deviation of the monthly temperatures is no longer statistically significant in cases 1.x and 2.x. For precipitation, there is little difference between cases with and without circulation indices.
- The performance of cases 5.x and 2.x, with detailed information on the spatial distribution of precipitation, is not better than that of the corresponding cases with less detailed information on precipitation (compare cases 5.x with 4.x and 3.x and cases 2.x with 1.x).
- The effect of the variation of weights for \tilde{P} , F and \tilde{T} is not clear. Comparison of cases 4.1, 4.2 and 4.3 shows, e.g., that the use of different weights for F has little effect. This may be a result of the rather large correlation between \tilde{P} and F .

The better performance of the unconditional method can partly be ascribed to the fact that this method often selects successive days in the historical data set. This occurs because the latest resampled day is also part of the k nearest neighbours. If that day is excluded in the search for nearest neighbours, the unconditional method still performs slightly better than the conditional method. In Section 3.3 we will further deal with selection of successive days in the historical record.

Case	Precipitation				Temperature			
	$\Delta\bar{r}(1)$	$\Delta\bar{r}(2)$	$\Delta\bar{r}(3)$	$\Delta\bar{s}$ (%)	$\Delta\bar{r}(1)$	$\Delta\bar{r}(2)$	$\Delta\bar{r}(3)$	$\Delta\bar{s}$ (%)
Unconditional								
0.1	-0.229	-0.101	-0.057	-27.9	-0.726	-0.523	-0.389	-61.7
1.1	-0.019	-0.002	-0.006	-6.7	-0.034	0.007	0.013	-4.9
1.2	-0.020	0.001	-0.005	-7.3	-0.034	0.007	0.014	-4.4
2.1	-0.010	0.000	-0.009	-7.7	-0.036	0.003	0.011	-4.0
2.2	-0.013	0.000	-0.009	-9.4	-0.038	-0.001	0.005	-5.1
2.3	-0.016	-0.002	-0.010	-10.7	-0.043	-0.009	-0.006	-7.2
2.4	-0.008	-0.001	-0.009	-8.0	-0.041	-0.007	-0.003	-6.9
2.5	-0.005	-0.001	-0.004	-6.6	-0.030	-0.004	-0.003	-6.2
3.1	-0.031	0.002	-0.005	-7.5	-0.059	-0.051	-0.062	-16.3
3.2	-0.027	-0.001	-0.007	-8.5	-0.046	-0.030	-0.035	-12.7
3.3	-0.018	0.004	-0.005	-6.9	-0.044	-0.026	-0.032	-12.5
4.1	-0.025	0.005	-0.006	-7.3	-0.057	-0.047	-0.057	-15.6
4.2	-0.021	0.006	-0.002	-5.6	-0.056	-0.044	-0.051	-14.9
4.3	-0.019	0.007	-0.004	-6.7	-0.060	-0.054	-0.066	-16.1
4.4	-0.018	0.009	-0.001	-4.5	-0.044	-0.027	-0.035	-11.6
4.5	-0.017	0.005	-0.003	-5.6	-0.047	-0.032	-0.039	-12.9
4.6	-0.016	0.007	-0.002	-5.1	-0.048	-0.033	-0.040	-13.2
4.7	-0.021	0.010	-0.003	-6.1	-0.041	-0.021	-0.028	-11.4
4.8	-0.015	0.013	0.001	-4.7	-0.044	-0.024	-0.030	-12.8
5.1	-0.024	0.001	-0.005	-9.0	-0.058	-0.049	-0.057	-14.8
5.2	-0.020	0.005	-0.004	-7.8	-0.046	-0.030	-0.038	-11.7
5.3	-0.015	0.008	-0.002	-8.2	-0.041	-0.021	-0.028	-10.9
Conditional								
3.1	-0.065	-0.023	-0.014	-10.1	-0.091	-0.059	-0.051	-17.1
3.2	-0.057	-0.016	-0.013	-8.7	-0.073	-0.036	-0.027	-14.4
3.3	-0.051	-0.018	-0.012	-8.0	-0.066	-0.025	-0.015	-13.9
4.1	-0.057	-0.020	-0.015	-8.9	-0.090	-0.061	-0.052	-18.2
4.2	-0.056	-0.016	-0.011	-7.2	-0.091	-0.063	-0.057	-17.8
4.3	-0.054	-0.014	-0.014	-8.1	-0.098	-0.074	-0.067	-19.3
4.4	-0.050	-0.017	-0.012	-8.5	-0.077	-0.042	-0.034	-15.2
4.5	-0.046	-0.015	-0.012	-8.6	-0.078	-0.043	-0.034	-15.9
4.6	-0.048	-0.014	-0.011	-7.2	-0.079	-0.047	-0.040	-16.3
4.7	-0.046	-0.014	-0.011	-7.9	-0.067	-0.029	-0.021	-14.9
4.8	-0.046	-0.012	-0.010	-7.5	-0.069	-0.030	-0.020	-14.9
5.1	-0.054	-0.025	-0.019	-10.6	-0.090	-0.060	-0.052	-17.3
5.2	-0.046	-0.019	-0.011	-10.0	-0.075	-0.038	-0.030	-14.1
5.3	-0.041	-0.019	-0.014	-10.4	-0.066	-0.027	-0.017	-13.4
Historical								
	0.230	0.098	0.059	35.6	0.794	0.582	0.447	1.81

Table 4: Differences between the mean lag 1, 2 and 3 autocorrelation coefficients of daily values and percentage differences between the mean standard deviations of monthly values for the simulated data (ten runs of 35 years for each case) and the historical records (1961–1995), averaged over 25 stations. The bottom line gives the estimates for the historical data. Estimates in italics for the simulated data indicate that they differ more than $2 \times se$ from the corresponding estimate for the historical data.

The better reproduction of $\bar{r}(l)$ and \bar{s} for temperature in the unconditional cases without circulation indices, 1.x and 2.x, is in line with the results of single-site simulations for Stuttgart in BB97. The fact that the results for precipitation are not sensitive to the inclusion of circulation indices is, however, in sharp contrast with the results of those single-site simulations, where the use of circulation indices gave a marked improvement in the reproduction of the autocorrelation properties. Although the inclusion of large-scale features of the weather seems desirable to obtain a satisfactory reproduction of the autocorrelation properties of local precipitation, this needs not to be restricted to the circulation indices as in the single-site simulations. The large-scale features $(\tilde{T}, \tilde{P}, F)$ in the multi-site generation may partly take over the role of the circulation indices. This is supported by the fact that the correlation between the vorticity and westerly flow indices and the area-averages \tilde{T}, \tilde{P} is stronger than that between these indices and P, T of the individual stations.

3.2.2 N -day winter maximum precipitation amounts

For the 25 stations, the N -day ($N = 1, 4, 10, 20$) winter (October-March) maximum precipitation amounts were abstracted from the historical record and all simulated cases. Like in BB97, the following three quantities are considered to verify the reproduction of the N -day winter maxima distributions:

1. The maximum of the N -day winter maxima (highest N -day 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 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 equation (13), we calculated for each of the three quantities the percentage difference between the values for the simulated and historical data, averaged over all stations. Table 5 presents these differences for the cases defined in Table 3. The bottom line of the table gives the absolute values for the historical data. From Table 5 the following can be noted:

- In general, the maximum, the upper quintile mean and the median are all somewhat underestimated.
- In case 0.1 there is for $N > 1$ a relatively large underprediction of the maximum, the upper quintile mean and the median, ranging between 19% and 25%. This is because there is no autocorrelation in the data in this case. For $N = 1$, the underprediction in case 0.1 is for the maximum comparable to that of the other cases. This underprediction is caused by the fact that the largest recorded amount is not always selected. The median of the 1-day amounts is, however, much better reproduced in case 0.1 than in the other cases. Apparently, the introduction of autocorrelation by nearest-neighbour resampling results in the selection of too few extreme days.

- The somewhat better reproduction of the autocorrelation coefficients by the unconditional method in Table 4, is reflected in the reproduction of the distribution of the N -day winter maximum precipitation amounts.
- The performance of cases 1.x and 2.x is comparable to that of the corresponding unconditional cases that include circulation indices in the feature vector (compare e.g. cases 1.1 and 4.2). In fact, there is no clear preference to one of the feature vectors.
- As in Section 3.2.1, the variation of weights for \tilde{P} , F , and \tilde{T} seems to have little effect.

Case	Maximum (%)				Upper quintile mean (%)				Median (%)			
	$N=1$	$N=4$	$N=10$	$N=20$	$N=1$	$N=4$	$N=10$	$N=20$	$N=1$	$N=4$	$N=10$	$N=20$
Unconditional												
0.1	-7.4	-23.9	-24.2	-24.7	-3.6	-23.8	-24.5	-22.2	-2.9	-21.8	-21.3	-18.6
1.1	-6.1	-5.1	-3.8	-4.3	-3.0	-5.8	-4.8	-2.6	-4.9	-5.3	-5.0	-4.8
1.2	-6.0	-5.8	0.1	-5.2	-3.5	-4.8	-2.6	-3.6	-4.2	-4.3	-3.4	-3.9
2.1	-8.2	-7.6	-4.0	-4.4	-4.6	-6.3	-5.3	-3.5	-5.0	-5.3	-4.5	-5.3
2.2	-6.8	-3.7	-2.8	-7.3	-3.7	-5.3	-4.9	-5.7	-6.3	-6.5	-5.7	-7.0
2.3	-9.1	-4.2	-5.0	-7.9	-6.9	-7.4	-7.8	-7.4	-7.9	-8.2	-7.5	-9.5
2.4	-7.5	-0.3	1.2	2.2	-4.0	-3.0	-3.3	-2.6	-4.1	-5.5	-5.3	-6.8
2.5	-5.8	-3.7	-3.0	-0.2	-2.6	-3.3	-3.0	0.0	-4.5	-4.4	-3.9	-4.2
3.1	-6.1	-8.4	-6.4	-6.9	-3.4	-6.7	-6.7	-5.7	-5.8	-6.2	-5.6	-6.4
3.2	-7.8	-9.2	-7.2	-10.3	-5.9	-8.2	-7.2	-7.5	-7.9	-8.2	-6.8	-8.1
3.3	-6.7	-3.9	-0.3	-2.0	-3.5	-3.9	-3.5	-2.5	-5.8	-5.2	-5.1	-6.4
4.1	-6.7	-7.3	-9.0	-8.5	-2.5	-6.0	-7.3	-6.0	-4.6	-5.6	-4.4	-6.6
4.2	-6.5	-2.7	-1.9	-5.7	-3.2	-3.1	-3.2	-2.5	-6.0	-5.2	-5.8	-5.4
4.3	-7.7	-5.6	-3.8	-3.4	-4.8	-5.5	-4.3	-2.6	-6.3	-6.3	-4.7	-5.6
4.4	-8.0	-3.6	-2.4	-3.2	-5.3	-4.7	-4.2	-2.7	-6.8	-5.6	-4.2	-4.4
4.5	-8.1	-3.8	-2.1	-6.3	-4.1	-4.4	-3.6	-3.8	-5.1	-4.9	-3.0	-3.1
4.6	-5.2	-4.9	-2.1	-2.1	-2.8	-4.7	-3.1	-1.7	-4.8	-4.9	-4.6	-5.4
4.7	-7.6	-6.4	-4.5	-5.6	-4.5	-5.4	-4.1	-2.5	-6.2	-5.6	-4.6	-5.5
4.8	-7.8	-6.2	-5.0	-5.4	-4.2	-5.0	-3.6	-2.6	-6.3	-4.0	-2.7	-3.3
5.1	-10.1	-7.2	-5.5	-8.4	-7.1	-7.7	-6.7	-7.5	-9.1	-8.5	-7.5	-8.1
5.2	-6.6	-6.7	-3.3	-3.8	-4.1	-5.8	-4.3	-2.5	-6.5	-5.4	-5.0	-5.9
5.3	-9.2	-4.3	-1.2	-6.0	-5.0	-3.6	-3.2	-3.8	-5.0	-4.3	-3.8	-6.0
Conditional												
3.1	-10.3	-8.4	-9.2	-8.8	-6.6	-9.3	-9.8	-7.7	-7.6	-10.1	-8.3	-7.1
3.2	-6.3	-4.2	-3.5	-5.0	-4.4	-6.0	-5.7	-4.7	-5.8	-6.6	-5.5	-4.8
3.3	-8.1	-7.9	-4.5	-3.9	-4.5	-6.3	-4.9	-3.0	-5.4	-6.5	-5.2	-5.0
4.1	-6.6	-5.6	-3.5	-6.6	-3.7	-6.6	-6.2	-5.1	-6.0	-7.6	-6.4	-5.3
4.2	-7.2	-8.5	-5.6	-6.1	-4.3	-8.0	-6.6	-4.3	-5.4	-7.9	-5.4	-4.4
4.3	-9.8	-9.2	-8.6	-11.5	-5.2	-8.4	-8.3	-7.4	-5.6	-7.9	-6.1	-5.4
4.4	-8.7	-9.1	-9.7	-9.1	-6.0	-8.2	-8.2	-6.4	-6.3	-8.6	-6.1	-5.7
4.5	-8.3	-7.0	-6.5	-7.4	-5.8	-8.1	-7.5	-5.4	-6.9	-9.0	-7.0	-5.8
4.6	-9.0	-5.4	-3.6	-5.3	-4.1	-7.0	-5.6	-4.2	-4.2	-6.8	-4.9	-4.3
4.7	-8.1	-7.3	-6.8	-8.4	-5.8	-7.6	-8.0	-5.9	-6.3	-7.3	-6.8	-5.7
4.8	-7.9	-6.0	-2.8	-3.6	-5.6	-7.0	-5.2	-4.2	-5.7	-7.3	-5.8	-5.0
5.1	-9.3	-7.9	-7.6	-10.0	-6.1	-8.4	-8.5	-7.3	-5.6	-8.0	-7.0	-6.3
5.2	-9.3	-8.5	-6.9	-9.3	-6.3	-8.7	-7.1	-6.3	-6.8	-8.6	-6.9	-6.0
5.3	-9.3	-9.5	-7.8	-8.9	-5.9	-9.4	-7.8	-7.2	-5.2	-7.2	-5.9	-6.9
Historical (mm)												
	56.5	95.9	136.8	189.8	42.3	76.2	110.8	152.3	26.9	50.8	75.1	107.0

Table 5: Percentage differences between the maxima, upper quintile means and medians of the N -day winter (October-March) precipitation maxima for the simulated data (ten runs of 35 years for each case) and the historical records (1961–1995), averaged over 25 stations. The bottom line gives the estimates for the historical data.

The systematic underprediction of the N -day winter precipitation maxima may have different causes. The first is the systematic underprediction of the precipitation autocorrelation coefficients as shown in Table 4. A second cause is the inability of the resampling procedure to simulate sufficient extreme 1-day precipitation amounts.

3.2.3 Correlation with circulation indices

While the cross-correlation coefficients between the circulation indices and the weather variables are more or less automatically preserved in the unconditional simulations, this is not necessarily so for the conditional simulations. The reproduction of these cross-correlation coefficients is, however, important for studies concerning the effects of long-term variations and potential future systematic changes in the atmospheric circulation. Therefore, we calculated the cross-correlation coefficients between the standardised daily values for most conditional cases in Table 3 and compared these with the historical values for the winter half-year (October - March). The daily average standardised values \tilde{P} , \tilde{T} , $\tilde{P}_1, \dots, \tilde{P}_V$ were considered here rather than the standardised values for the individual stations.

Table 6 summarises the results. The coefficients $\bar{r}(\tilde{Z}, \tilde{\mathbf{P}})$, $\bar{r}(\tilde{W}, \tilde{\mathbf{P}})$ and $\bar{r}(\tilde{S}, \tilde{\mathbf{P}})$ refer here to the correlation coefficients for the elements of $\tilde{\mathbf{P}}$, averaged over the five areas. From the table it is seen that both precipitation and temperature show the strongest correlation with the westerly flow index, which is somewhat too low in the simulated cases. This underestimation is also found for the correlation between precipitation and vorticity, whereas the correlation between temperature and the southerly flow index is slightly overestimated. For the cases considered in Table 6, the reproduction of cross-correlation coefficients improves with increasing relative weight of the circulation indices. The best results are generally obtained if the total weight of the circulation indices is about the same as that of the weather variables (cases 3.1, 4.1 and 5.1).

It should be noted that the cross-correlation coefficient only measures the strength of linear dependence. Especially the relationships between temperature and vorticity and between precipitation and the southerly flow index are non-linear here. The non-linearity is such that the correlation coefficient is almost zero for these cases.

3.3 Influence of k

3.3.1 The need for further study

It was already noted in Section 2.1 that the value of k depends on the type of kernel, the number n of daily values used in the search for nearest neighbours, and the number q of elements in \mathbf{D}_t . For the decreasing kernel in equation (2), Lall and Sharma (1996) recommended to use $k = n^{1/2}$ provided that $1 \leq q \leq 6$ and $n \geq 100$. This recommendation was based partly on experience and partly on an asymptotic result for the mean square error of probability density estimates. In our case, with $n = 2135$ days in the moving window, it yields $k = 46$.

Lall and Sharma (1996) also suggested to use cross-validation to determine the best value of k . Cross-validation considers, however, prediction errors, rather than time series properties as in Section 3.2. It would be a suitable technique when the nearest-neighbour method is used for forecasting. Nevertheless, in BB97 we studied cross-validation, using a uniform kernel. Both for P and T a value of k of about 25 turned out to be optimal in terms of the cross-validation score, but autocorrelation coefficients were better preserved for smaller values of k , e.g. $k = 5$. For the decreasing kernel, BB97 only presents some comparisons between $k = 10, 20$ and 40 . The results for the uniform kernel, however, give rise to consider a broader range of k , in particular values of k less than 10.

3.3.2 Daily autocorrelation and variability of monthly values

Table 7 presents $\Delta\bar{r}(l)$ and $\Delta\bar{s}$ for the cases 4.4 (unconditional) and 4.1 (conditional). For temperature, the table shows a clear improvement of the reproduction

Case	$\Delta\bar{r}(\tilde{Z}, \tilde{P})$	$\Delta r(\tilde{Z}, \tilde{P})$	$\Delta r(\tilde{Z}, \tilde{T})$	$\Delta\bar{r}(\tilde{W}, \tilde{P})$	$\Delta r(\tilde{W}, \tilde{P})$	$\Delta r(\tilde{W}, \tilde{T})$	$\Delta\bar{r}(\tilde{S}, \tilde{P})$	$\Delta r(\tilde{S}, \tilde{P})$	$\Delta r(\tilde{S}, \tilde{T})$
3.1	-0.024	-0.026	0.007	-0.023	-0.022	-0.051	0.016	0.019	0.018
3.2	-0.031	-0.035	0.007	-0.033	-0.035	-0.067	0.018	0.021	0.013
3.3	-0.040	-0.044	0.010	-0.043	-0.046	-0.078	0.020	0.022	0.015
4.1	-0.033	-0.038	0.009	-0.027	-0.029	-0.048	0.023	0.027	0.021
4.4	-0.041	-0.046	0.010	-0.045	-0.048	-0.066	0.028	0.032	0.014
4.7	-0.049	-0.054	0.012	-0.053	-0.056	-0.082	0.026	0.030	0.008
5.1	-0.028	-0.032	0.000	-0.029	-0.031	-0.053	0.009	0.011	0.016
5.2	-0.041	-0.046	0.000	-0.036	-0.039	-0.071	0.017	0.020	0.014
5.3	-0.048	-0.054	0.004	-0.053	-0.058	-0.085	0.019	0.023	0.002
Hist	0.221	0.247	0.021	0.366	0.410	0.561	-0.044	-0.055	0.241

Table 6: Differences between the cross-correlation coefficients of the standardised daily circulation indices and weather variables for the simulated data (ten conditional simulations of 35 years for each case) and the historical records (1961–1995). The differences apply to the winter half-year (October - March). The bottom line gives the cross-correlation coefficients for the historical data.

of $\bar{r}(l)$ and \bar{s} at small k . This is also the case for precipitation in the conditional simulations. Despite these improvements, the underestimation of the lag 1 autocorrelation coefficient is still statistically significant for $k = 2$ and $k = 5$. The improvement is less marked for precipitation in the unconditional simulations. The

Case	Maximum (%)				Upper quintile mean (%)				Median (%)			
	<i>N</i> =1	<i>N</i> =4	<i>N</i> =10	<i>N</i> =20	<i>N</i> =1	<i>N</i> =4	<i>N</i> =10	<i>N</i> =20	<i>N</i> =1	<i>N</i> =4	<i>N</i> =10	<i>N</i> =20
Unconditional												
4.4 (<i>k</i> = 2)	-5.2	-3.4	-0.5	0.7	-0.6	-2.6	-2.4	-0.9	-3.7	-3.3	-3.0	-2.4
4.4 (<i>k</i> = 5)	-6.4	-2.8	-1.7	-2.5	-2.1	-1.2	-1.4	-0.2	-2.4	-1.1	-0.3	-1.1
4.4 (<i>k</i> = 20)	-8.0	-3.6	-2.4	-3.2	-5.3	-4.7	-4.2	-2.7	-6.8	-5.6	-4.2	-4.4
4.4 (<i>k</i> = 50)	-7.3	-8.9	-4.8	-5.1	-6.0	-8.2	-5.3	-4.2	-8.4	-7.0	-5.3	-5.8
Conditional												
4.1 (<i>k</i> = 2)	-6.4	-3.8	0.8	-4.3	-5.8	-5.6	-3.8	-3.0	-6.7	-6.8	-4.7	-2.6
4.1 (<i>k</i> = 5)	-5.1	-2.9	-0.6	-3.6	-2.7	-3.9	-2.9	-2.2	-4.1	-5.8	-3.9	-3.5
4.1 (<i>k</i> = 20)	-8.7	-9.1	-9.7	-9.1	-6.0	-8.2	-8.2	-6.4	-6.3	-8.6	-6.1	-5.7
4.1 (<i>k</i> = 50)	-8.5	-9.7	-7.1	-11.0	-5.9	-9.4	-8.5	-7.3	-7.0	-8.9	-7.7	-6.0

Table 8: Percentage differences between the maxima, upper quintile means and medians of the *N*-day winter (October–March) precipitation maxima for the simulated data (ten runs of 35 years for each case) and the historical records (1961–1995), averaged over 25 stations.

autocorrelation properties are already rather well preserved for *k* = 20 in that case.

3.3.3 *N*-day winter maximum precipitation amounts

Table 8 presents the results for the *N*-day winter precipitation maxima. For both the conditional and unconditional simulations, the reproduction of extreme value properties improves for *k* decreasing from 50 to 5. This is partly due to the better reproduction of the autocorrelation properties for *k* = 5. Another point is that for *k* = 5 the underprediction of the median of the 1-day winter maxima is small compared with that for larger values of *k*, indicating that more extreme days are resampled if *k* = 5. This does not improve for *k* = 2. The effect of selecting too few extreme days seems to override the slightly better reproduction of the autocorrelation properties for *k* = 2.

For *k* = 5 and *k* = 20, Figure 8 presents boxplots of the relative differences be-

Case	Precipitation				Temperature			
	$\Delta\bar{r}(1)$	$\Delta\bar{r}(2)$	$\Delta\bar{r}(3)$	$\Delta\bar{s}$ (%)	$\Delta\bar{r}(1)$	$\Delta\bar{r}(2)$	$\Delta\bar{r}(3)$	$\Delta\bar{s}$ (%)
Unconditional								
4.4 (<i>k</i> = 2)	-0.006	0.006	0.002	-2.4	<i>-0.017</i>	<i>-0.010</i>	<i>-0.012</i>	<i>-4.7</i>
4.4 (<i>k</i> = 5)	<i>-0.012</i>	0.009	-0.003	-5.3	<i>-0.031</i>	<i>-0.019</i>	<i>-0.027</i>	<i>-10.2</i>
4.4 (<i>k</i> = 20)	<i>-0.018</i>	0.009	-0.001	-4.5	<i>-0.044</i>	<i>-0.027</i>	<i>-0.035</i>	<i>-11.6</i>
4.4 (<i>k</i> = 50)	<i>-0.026</i>	0.004	-0.003	-7.6	<i>-0.057</i>	<i>-0.039</i>	<i>-0.045</i>	<i>-14.3</i>
Conditional								
4.1 (<i>k</i> = 2)	<i>-0.035</i>	-0.006	-0.008	-4.7	<i>-0.052</i>	<i>-0.027</i>	<i>-0.021</i>	<i>-12.1</i>
4.1 (<i>k</i> = 5)	<i>-0.045</i>	-0.009	-0.010	-6.4	<i>-0.071</i>	<i>-0.040</i>	<i>-0.033</i>	<i>-14.0</i>
4.1 (<i>k</i> = 20)	<i>-0.057</i>	<i>-0.020</i>	<i>-0.015</i>	-8.9	<i>-0.090</i>	<i>-0.061</i>	<i>-0.052</i>	<i>-18.2</i>
4.1 (<i>k</i> = 50)	<i>-0.063</i>	<i>-0.020</i>	<i>-0.016</i>	-9.7	<i>-0.108</i>	<i>-0.083</i>	<i>-0.076</i>	<i>-20.9</i>

Table 7: Differences between the mean lag 1, 2 and 3 autocorrelation coefficients of daily values and percentage differences between the mean standard deviations of monthly values for the simulated data (ten runs of 35 years for each case) and the historical records (1961–1995), averaged over 25 stations. Estimates in italics for the simulated data indicate that they differ more than $2 \times se$ from the corresponding estimate for the historical data.

tween the observed and simulated extreme-value properties. Each boxplot represents a sample of relative differences derived from the ten simulation runs, where the dots depict the minimum and maximum value, the whiskers the 10th and 90th

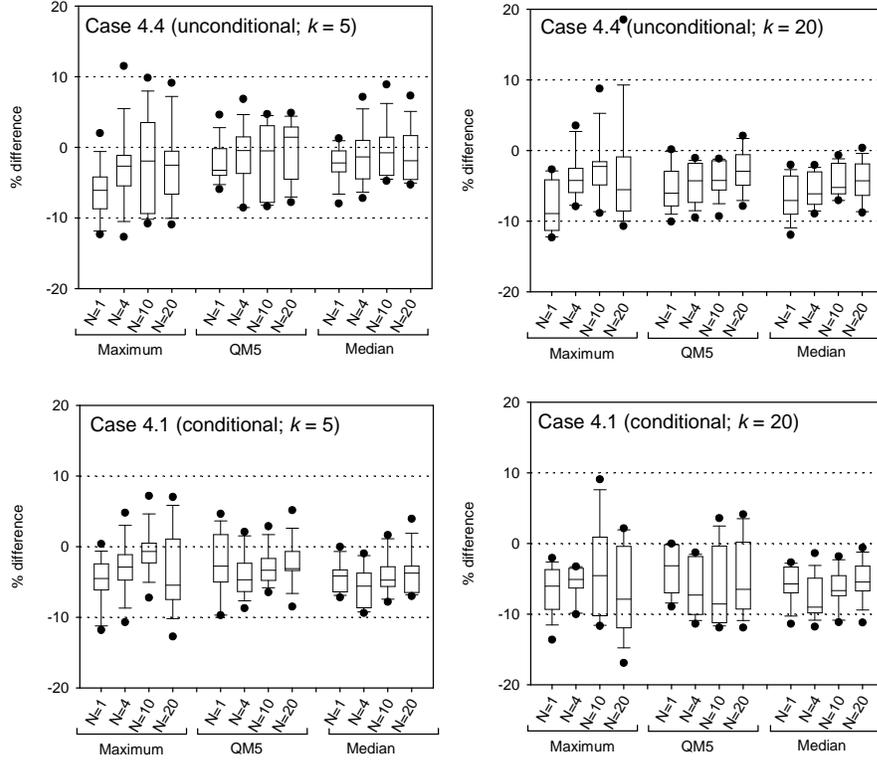


Figure 8: Boxplots of the percentage differences between the maxima, upper quintile means and medians of the N -day winter (October–March) precipitation maxima for the simulated data (each boxplot represents ten runs of 35 years) and the historical records (1961–1995), averaged over 25 stations.

percentiles, the upper and lower boundaries of the box the 25th and 75th percentiles, and the horizontal line in the box the median. The figure gives an impression of the statistical significance of the underprediction of the maximum, QM5 and median of the N -day winter maximum precipitation amounts in Table 8. For instance, the fact that in the $k = 20$ simulation for case 4.4 (unconditional) five of the eight boxplots for QM5 and the median are entirely below the 0% difference line indicates a significant underprediction of these quantities. For $k = 5$ there is no such evidence.

3.3.4 Shape parameter of the GEV distribution

The Generalised Extreme-Value (GEV) distribution is a three parameter distribution that combines into a single form the three possible types of limiting distributions for extreme values:

$$F(x) = \begin{cases} \exp\left\{-\left[1 - \kappa(x - \xi) / \alpha\right]^{1/\kappa}\right\}, & \kappa < 0, \quad x > \xi + \alpha / \kappa \\ \exp\left\{-\exp\left[-(x - \xi) / \alpha\right]\right\}, & \kappa = 0, \\ \exp\left\{-\left[1 - \kappa(x - \xi) / \alpha\right]^{1/\kappa}\right\}, & \kappa > 0, \quad x < \xi + \alpha / \kappa \end{cases} \quad (14)$$

Here $F(x)$ is the distribution function, ξ and α are location and scale parameters, respectively, and the shape parameter κ determines which extreme value distribu-

Case	$N = 1$	$N = 4$	$N = 10$	$N = 20$
Unconditional				
4.4 ($k = 2$)	-0.0415	0.0280	0.0395	0.0646
4.4 ($k = 5$)	-0.0054	0.0468	0.0816	0.1004
4.4 ($k = 20$)	-0.0210	0.0230	0.0663	0.1116
4.4 ($k = 50$)	-0.0464	0.0433	0.0745	0.1035
Conditional				
4.1 ($k = 2$)	-0.0426	0.0151	0.0382	0.1130
4.1 ($k = 5$)	-0.0238	0.0220	0.0592	0.0878
4.1 ($k = 20$)	-0.0387	0.0108	0.0459	0.1070
4.1 ($k = 50$)	-0.0258	0.0331	0.0719	0.1296
Historical	-0.0420	-0.019	0.0391	0.0704
(se)	(0.0391)	(0.0416)	(0.0442)	(0.0557)

Table 9: Shape parameter κ of the GEV distribution of N -day winter maxima for the simulated data (ten runs of 35 years for each case) and the historical records (1961–1995), averaged over 25 stations. The bottom row gives the standard error se of the estimate for the historical data. None of the estimates for the simulated data differs more than $2 \times \text{se}$ from the corresponding estimates for the historical data.

tion is represented: Type I or Gumbel distribution ($\kappa = 0$), Type II ($\kappa < 0$) or Type III ($\kappa > 0$). The Type II distribution has a thicker upper tail than the Gumbel distribution. The values at long return periods are therefore greater than those expected from the Gumbel distribution. The opposite holds for the Type III distribution. It is of interest to see: (1) whether κ of the historical maxima is reproduced by the simulations; and (2) which type of distribution can be fitted to the N -day maxima.

The parameters of the GEV distribution were estimated by the method of probability-weighted moments (PWM), according to Hosking *et al.* (1985). This method requires no iteration and provides for samples of size 34 better estimates than the more computer-intensive maximum likelihood method. Table 9 presents the PWM estimate of κ for the 1, 4, 10 and 20-day winter maxima in the simulated and historical data, averaged over 25 stations. For the historical data, the standard error of this estimate is also given. A bootstrap method, in which years were randomly sampled with replacement, was used to estimate the standard error (see BB97 for further details of this method). The standard errors in Table 9 are based on 500 bootstrap samples. The table shows that κ is adequately reproduced in the simulations and that the quality of the reproduction does not depend on k . The estimates of κ increase with increasing N . This is in agreement with regional analyses of long-duration records of N -day annual maximum precipitation amounts in the United Kingdom and the Low Countries (Dupriez and Demarée, 1988; Dales and Reed, 1989; Buishand, 1991). However, for the 1-day maxima the departures from the Gumbel distribution in these studies are generally larger than those found in Table 9. There is often a strong statistical evidence of a Type II distribution for these maxima. For $N = 10$ and $N = 20$, the values of κ are positive in Table 9, indicating a Type III distribution. The standard error is, however, just as large.

3.3.5 Correlation with circulation indices

For the conditional simulations (case 4.1), we again calculated the cross-correlation coefficients of the standardised daily circulation indices and weather variables. Table 10 shows that the reproduction of these cross-correlation coefficients improves with decreasing k .

3.3.6 Run lengths in unconditional simulations

In Section 3.2 we already noted that the unconditional method often selects successive days in the historical data set. Although this is good for the reproduction of autocorrelation coefficients, there is a risk that large parts of the historical record are duplicated in the simulation run. It is, therefore, of interest to study the influence of k on properties of run lengths of successive historical days in the simulations. The mean run length $E(R)$ is given by (see Appendix A):

$$E(R) = 1/(1 - p_1) \quad (15)$$

where p_1 is the probability that the closest neighbour is selected as defined in equation (2). This probability increases with decreasing k (see also Figure 3) and $E(R)$ will thus be large for small k . For the mean of the longest run R_{\max} in a simulation of J years the following approximation can be derived (see Appendix A):

$$E(R_{\max}) \approx -\frac{\gamma + \ln[365J(1 - p_1)]}{\ln p_1} + 1/2 \quad (16)$$

where $\gamma = 0.5772\dots$ is Euler's constant. Table 11 presents p_1 , $E(R)$ and, for $J = 35$ and 1000 years, $E(R_{\max})$ for several values of k . The table shows that considerable changes in p_1 and run length properties occur at small k . For k as small as 2, runs of consecutive historical days can be quite long for the number of simulated years considered in this report.

The question remains how far it is allowable that large parts of the historical data set are duplicated in a simulation run. This depends, of course, on the desired application of the simulated data. Here it is important that e.g. more extreme 10-day precipitation amounts can be simulated than those observed. There is no indication that a simulation with $k = 2$ performs worse than a simulation with $k = 5$, in this respect.

k	p_1	$E(R)$	$E(R_{\max})$	
			$J = 35$	$J = 1000$
2	0.6667	3.000	22.5	30.8
5	0.4380	1.779	12.0	16.0
20	0.2780	1.358	8.1	10.7
50	0.2223	1.286	7.0	9.2

Table 11: Relation between k and the expected run length $E(R)$ and the expected longest run length $E(R_{\max})$ in simulations of $J = 35$ and 1000 years for the decreasing kernel in equation (2). The variable p_1 denotes the probability that a successive historical day is re-sampled.

k	$\Delta\bar{r}(\tilde{Z}, \tilde{\mathbf{P}})$	$\Delta r(\tilde{Z}, \tilde{P})$	$\Delta r(\tilde{Z}, \tilde{T})$	$\Delta\bar{r}(\tilde{W}, \tilde{\mathbf{P}})$	$\Delta r(\tilde{W}, \tilde{P})$	$\Delta r(\tilde{W}, \tilde{T})$	$\Delta\bar{r}(\tilde{S}, \tilde{\mathbf{P}})$	$\Delta r(\tilde{S}, \tilde{P})$	$\Delta r(\tilde{S}, \tilde{T})$
2	-0.019	-0.021	0.005	-0.013	-0.011	-0.029	0.006	0.006	0.014
5	-0.030	-0.033	0.005	-0.022	-0.023	-0.033	0.018	0.020	0.018
20	-0.033	-0.038	0.009	-0.027	-0.029	-0.048	0.023	0.027	0.021
50	-0.040	-0.044	0.002	-0.034	-0.036	-0.050	0.019	0.022	0.024

Table 10: Differences between the cross-correlation coefficients of the standardised daily circulation indices and weather variables for the simulated data (ten conditional simulations of 35 years for each value of k) and the historical records (1961–1995) for case 4.1. The differences apply to the winter half-year (October - March).

4. Further analysis of the distribution of extremes

In the previous section, we compared the performance of various resampling schemes using the percentage differences between observed and simulated properties of N -day winter maximum precipitation amounts, averaged over the 25 stations. Here we look at these extreme-value properties in more detail for the stations Essen, Kahler Asten, Trier, Frankfurt Bamberg, Freudenstadt and Stuttgart and the five areas shown in Figure 1. The selected stations correspond to those in BB97 for the single-site simulations. As in BB97 we also study the distribution of the snowmelt maxima.

4.1 N -day winter maximum precipitation amounts

Table 12 presents for cases 4.4 (unconditional; $k = 5$) and 4.1 (conditional; $k = 5$) the results for the N -day winter maximum precipitation amounts. It should be noted that the mean in the left column of the table refers to the mean percentage difference for the seven stations between the observed and simulated properties of the single-site N -day maxima. In contrast, for the five areas and the total area the N -day maxima of the mean daily precipitation of the corresponding stations in those areas are considered.

As in Section 3, the differences in Table 12 for the unconditional case are smaller than those for the conditional case. For both cases the results for the seven stations are somewhat better than the corresponding cases reported in BB97 for the single-site simulations with $k = 20$. As could be expected, the mean percentage differences of the seven stations are comparable to those of the 25 stations for the corresponding cases in Table 8, in particular at QM5 and the median. For some stations percentage differences between 10% and 20% are found in Table 12. These large differences can still be explained by natural variability (Appendix B).

In BB97 the underprediction of the median was a point of concern. Compared to BB97, the median for the conditional simulations in Table 12 remains somewhat underpredicted for $N = 1$ and $N = 4$. On the other hand, for $N = 10$ and $N = 20$ the reproduction of the median is improved. For the unconditional simulations, it is noteworthy that the underprediction of the median for $k = 5$ becomes small (on average only 0.8% for $N = 10$ and not more than 4.7% for $N = 1$). In BB97 we found an average systematic underprediction of the median up to about 8% for $N = 20$ using $k = 20$. Especially for Freudenstadt the improvement is striking, ranging between a 5% less underprediction for $N = 1$ and an 11% less underprediction for $N = 20$.

Because the selection of nearest neighbours uses large-scale features of precipitation rather than local precipitation, one may expect that the extreme-value properties of area-average rainfall are better preserved than those of point rainfall. This is, however, not so. The percentage differences for the five areas and the total area are comparable to those for individual stations.

Figure 9 compares for three stations and three areas Gumbel plots of the observed 10-day winter maxima with those of the ten conditional 35-year simulations. For Freudenstadt the large underprediction of the median (9.4% in Table 12) is clearly visible, all ten simulated medians ($T_r = 2$ yr) are below the observed median. The large underprediction of QM5 for area II (10.3% in Table 12) is also apparent.

Station/Area	Maximum (%)				Upper quintile mean (%)				Median (%)			
	N=1	N=4	N=10	N=20	N=1	N=4	N=10	N=20	N=1	N=4	N=10	N=20
Unconditional (4.4; $k = 5$)												
Essen	6.9	-3.3	15.3	10.8	1.9	-1.3	9.1	7.2	1.1	3.4	-1.2	1.2
Kahler Asten	1.3	-6.4	1.7	2.2	-0.5	-1.4	-1.2	0.0	-7.0	-2.1	7.0	-0.1
Trier	-1.6	-4.3	-17.5	-1.2	0.6	-1.6	-6.7	2.5	-6.6	-0.6	-2.2	1.5
Frankfurt	-3.9	-16.8	-5.9	-2.2	-0.9	-1.3	-4.8	-3.9	-0.3	-5.2	-2.8	0.4
Bamberg	-9.4	9.2	13.0	-5.9	-0.4	2.1	3.4	-1.3	-6.1	0.6	-0.7	-1.3
Freudenstadt	7.1	9.0	5.4	-4.4	-0.8	-0.3	8.1	1.3	-9.0	-8.0	-3.2	-6.8
Stuttgart	-6.0	7.3	-0.7	-1.3	-0.3	0.6	5.4	1.4	-4.7	-2.3	-2.7	-4.6
Mean	-0.8	-0.8	1.6	-0.3	-0.1	-0.5	1.9	1.0	-4.7	-2.0	-0.8	-1.4
Area I	-0.4	-14.8	17.7	18.8	-1.0	0.0	8.4	6.5	0.5	2.5	1.4	-3.3
Area II	0.5	-3.2	-12.3	-7.4	0.2	-1.9	-9.3	-1.0	-2.1	-6.8	2.1	2.9
Area III	-13.6	-10.0	-17.1	-14.5	-4.9	-1.7	-2.3	-0.7	-4.4	-5.2	-3.1	2.2
Area IV	-10.4	-1.9	4.3	3.9	0.5	2.9	1.9	2.1	-1.3	-2.4	-0.9	-2.4
Area V	-4.8	3.8	7.9	4.1	-0.5	-0.6	-0.1	1.0	-3.6	-2.2	2.1	-3.2
Total area	-7.0	-7.8	2.1	-1.7	-4.3	-1.5	-3.0	3.7	-4.4	0.0	-1.8	-4.3
Conditional (4.1; $k = 5$)												
Essen	13.2	-6.8	4.1	0.7	4.5	-4.0	2.9	1.1	-2.4	-0.2	-3.6	-1.7
Kahler Asten	4.6	-8.5	-7.3	-6.4	2.5	-3.3	-9.4	-5.7	-5.6	-8.3	-0.9	-3.9
Trier	-13.2	9.1	-7.1	-0.2	-3.7	-2.9	-5.2	1.5	-8.1	-8.5	-4.2	-0.3
Frankfurt	-1.6	-16.6	1.9	3.4	-2.4	-3.9	-2.9	-0.9	-4.4	-7.3	-5.6	-1.3
Bamberg	-5.7	-0.5	8.1	-11.9	-0.3	-7.5	-1.7	-5.3	-9.0	-1.3	-0.3	-4.2
Freudenstadt	4.0	1.0	-2.7	-9.6	-3.9	-6.4	-1.6	-8.3	-11.6	-16.1	-9.4	-12.6
Stuttgart	-0.9	8.2	1.9	1.4	-0.7	4.1	3.7	-2.1	-4.6	-7.8	-7.7	-8.1
Mean	0.1	-2.0	-0.2	-3.2	-0.6	-3.4	-2.0	-2.8	-6.5	-7.1	-4.5	-4.6
Area I	1.9	-17.9	8.8	10.7	1.9	-3.5	2.1	0.5	-2.7	-0.6	-2.6	-5.3
Area II	-6.5	-4.8	-8.7	-12.3	-6.0	-8.7	-10.3	-3.5	-4.5	-12.0	0.2	-0.1
Area III	-11.3	-8.1	-11.8	-10.3	-5.1	-1.4	2.3	1.7	-3.9	-6.8	-7.1	0.1
Area IV	-16.2	-10.0	4.0	-0.5	-7.4	-5.3	-3.4	-2.1	-1.5	-6.5	-4.8	-4.5
Area V	-3.1	9.7	11.9	0.2	-3.4	-0.6	-2.6	-5.1	-5.6	-10	-3.5	-5.2
Total area	-10.6	-5.5	4.5	-4.2	-5.1	-5.7	-6.2	-1.5	-7.8	-5.1	-2.4	-6.2
Historical (mm)												
Essen	38.3	83.8	114.8	180.6	36.5	70.4	102.8	155.2	28.3	50.3	83.2	118.6
Kahler Asten	66.4	147.0	232.9	352.3	59.8	122.7	207.4	304.2	45.3	94.4	142.7	218.0
Trier	51.3	87.6	151.9	176.9	38.9	73.0	113.0	144.4	26.8	49.3	75.3	103.1
Frankfurt	38.4	94.9	112.9	140.5	32.9	63.6	91.9	121.2	23.4	41.6	60.6	81.9
Bamberg	54.0	66.5	87.9	145.9	36.6	59.3	78.7	112.7	21.6	36.2	55.8	79.6
Freudenstadt	112.6	246.4	355.4	547.3	104.4	215.7	289.5	429.7	76.0	155.4	211.4	303.5
Stuttgart	50.8	69.8	97.2	128.4	35.7	57.6	76.2	106.4	22.0	38.4	55.2	77.8
Area I	40.2	87.1	107.6	160.9	34.6	66.3	100.2	150.4	24.0	48.0	78.0	121.1
Area II	49.8	96.5	150.8	207.9	40.9	79.7	124.7	158.5	25.8	54.2	75.9	107.7
Area III	40.2	79.7	119.5	153.9	30.8	56.8	83.4	114.5	19.5	38.0	58.8	80.4
Area IV	47.8	72.3	94.8	136.2	30.9	55.2	81.1	112.3	18.2	37.4	57.0	82.0
Area V	50.3	95.0	137.3	203.7	41.7	84.4	125.6	176.7	30.0	60.8	87.1	127.9
Total area	37.2	75.2	105.9	162.8	29.3	61.3	96.8	131.2	20.9	42.7	67.9	101.6

Table 12: Percentage differences between the maxima, upper quintile means and medians of the N -day winter (October-March) precipitation maxima for the simulated data (ten runs of 35 years for each case) and the historical records (1961–1995) for seven stations in the Rhine basin, for the five areas (Figure 1) and for the total area. The lower part of the table gives the values for the historical data.

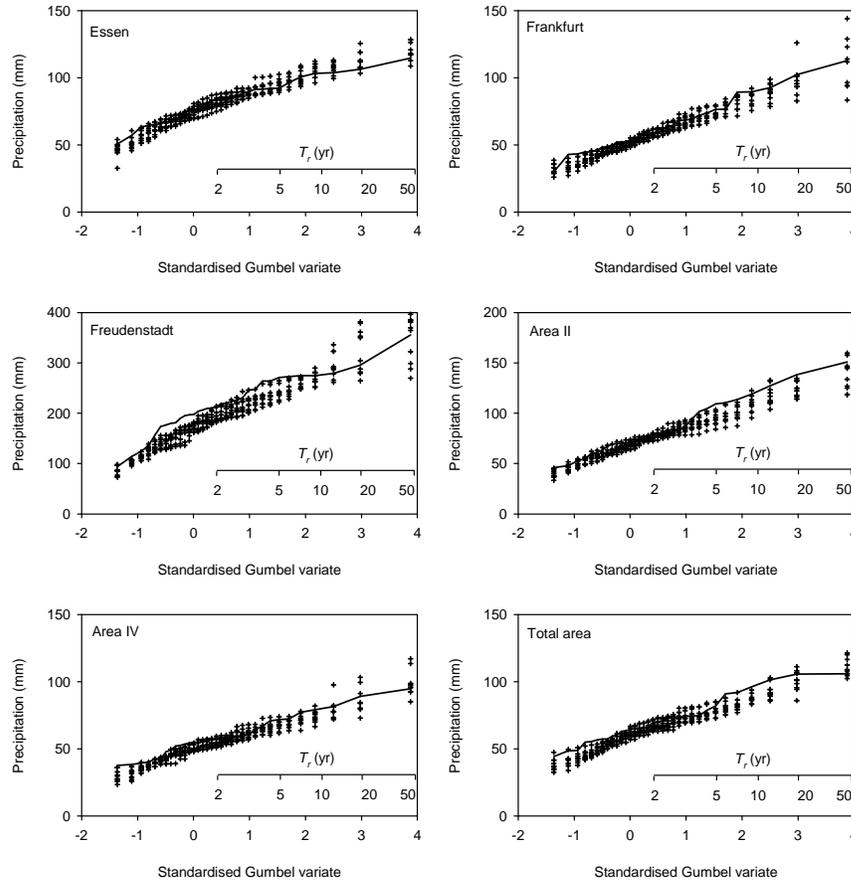


Figure 9: Gumbel plots of the 10-day winter (October–March) maxima for the observed precipitation (1961–1995) at Essen, Frankfurt, Freudenstadt, area II, area IV and the total area (solid lines) and the corresponding simulated data for case 4.1 using $k = 5$ (ten conditional simulations of 35 years, represented by pluses). T_r denotes the return period.

4.2 N -day maximum snowmelt amounts

Large river discharges may be partly caused by snowmelt. Therefore, the reproduction of snowmelt is also considered.

As in BB97, historical estimates and simulated values of snowmelt were derived from daily precipitation and temperature. It was assumed that for $T < 0$ precipitation accumulates on the surface as snow. To calculate the N -day maximum snowmelt, snow was transformed into snowmelt using the degree days method. In that method, the amount of snowmelt on a certain day is proportional to the temperature excess (number of degrees Celsius above freezing point on that day), of course as long as there is solid precipitation stored on the surface. The constant of proportionality is known as the degree days factor ($\text{mm}/^\circ\text{C}$). This factor was set equal to 4, which is an average of the values found in the literature (Linsley *et al.*, 1988; Gray and Prowse, 1993).

Analogous to Table 12, Table 13 shows the results for the N -day maximum snowmelt in the winter half-year. The contrast between the historical values of the two highest stations, Kahler Asten and Freudenstadt, and the other stations is

Station/Area	Maximum (%)				Upper quintile mean (%)				Median (%)			
	<i>N</i> =1	<i>N</i> =4	<i>N</i> =10	<i>N</i> =20	<i>N</i> =1	<i>N</i> =4	<i>N</i> =10	<i>N</i> =20	<i>N</i> =1	<i>N</i> =4	<i>N</i> =10	<i>N</i> =20
Unconditional (4.4; <i>k</i> = 5)												
Essen	3.2	10.9	17.8	8.2	-3.8	0.0	-0.1	5.6	18.8	40.5	49.3	28.0
Kahler Asten	-14.1	-15.7	-14.2	1.4	8.4	8.1	1.3	2.9	17.6	15.3	21.7	6.0
Trier	8.2	20.5	10.1	26.3	11.1	2.1	-6.7	5.1	7.5	22.3	17.7	16.8
Frankfurt	-1.6	13.9	27.2	39.3	16.7	14.3	14.8	19.8	24.9	30.9	29.8	44.9
Bamberg	5.5	9.3	16.1	17.0	7.4	0.7	3.0	0.3	7.9	14.2	7.1	4.6
Freudenstadt	6.2	2.1	-17.4	4.2	4.8	-5.0	-16.1	-11.3	4.7	-5.2	-14.1	-13.5
Stuttgart	0.8	12.1	21.8	11.3	-7.6	2.8	7.2	10.8	6.6	-1.7	-14.7	-11.8
Mean	1.2	7.6	8.8	15.4	5.3	3.3	0.5	4.7	12.6	16.6	13.8	10.7
Conditional (4.1; <i>k</i> = 5)												
Essen	-2.7	-15.6	-0.7	-5.9	-11.5	-12.7	-7.4	-2.8	14.9	28.1	38.0	22.0
Kahler Asten	-9.8	-14.9	-8.6	6.5	7.0	1.1	-5.5	-0.4	14.7	8.2	14.7	5.0
Trier	13.5	16.7	18.1	18.8	10.6	-10.2	-13.0	-3.1	3.4	6.7	3.2	-1.3
Frankfurt	3.9	7.0	17.3	28.2	14.1	-2.6	-2.1	4.8	9.0	9.4	10.9	25.7
Bamberg	-1.1	-7.2	5.4	18.9	0.9	-6.9	-1.9	2.5	-3.3	2.7	-2.7	-4.4
Freudenstadt	2.4	-5.6	-20.3	-7.6	-0.9	-13.1	-20.9	-18.0	6.7	-10.3	-13.8	-12.0
Stuttgart	-5.9	0.3	13.6	12.5	-6.9	-5.7	-1.1	6.6	7.3	-1.1	-8.2	-7.7
Mean	0.0	-2.8	3.5	10.2	1.9	-7.2	-7.4	-1.5	7.5	6.2	6.0	3.9
Historical (mm)												
Essen	20.4	30.6	32.2	38.7	16.5	22.8	25.6	28.6	6.9	7.8	8.2	10.8
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
Trier	21.6	30.3	34.8	37.2	15.2	25.2	30.7	31.8	7.7	8.7	10.5	12.6
Frankfurt	18.7	24.6	24.6	24.6	11.2	17.2	18.7	19.7	5.4	6.3	7.0	7.2
Bamberg	21.4	30.9	30.9	33.9	14.9	23.2	25.1	28.8	7.8	9.7	11.8	14.0
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
Stuttgart	26.4	42.7	42.7	51.8	20.1	32.0	34.6	38.1	9.6	14.4	18.6	20.8

Table 13: Percentage differences between the maxima, upper quintile means and medians of the *N*-day snowmelt maxima for the simulated data (ten runs of 35 years for each case) and the historical records (1961–1995) for seven stations in the Rhine basin. The lower part of the table gives the values for the historical data.

even more striking for snowmelt than for precipitation. For Kahler Asten and Freudenstadt, the maximum of the 10-day snowmelt is for instance of the same order of magnitude as the corresponding maximum precipitation amount in Table 12, whereas for the other stations it is on average a third of that value.

Despite the systematic underprediction of the autocorrelation of daily temperatures, the correspondence between the historical and simulated values is satisfactory in Table 13. The large percentage differences for, e.g., the median in the unconditional simulations for Essen and Frankfurt are not alarming. Because there is little snow at these stations, the percentage differences are rather sensitive to small departures in the simulations. Furthermore, the relative variation of the median is large for such stations (Appendix B). The relatively large underprediction of the extreme-value properties of 10-day snowmelt at Freudenstadt for both the unconditional and conditional simulations in Table 13, is for all cases less than twice the standard error of the estimate from the historical data (Appendix B). The underprediction for Freudenstadt is also smaller than that in the single-site simulations (Brandsma and Buishand, 1998).

5. Spatial association of N -day precipitation maxima

Extreme rainfall over a large area is needed to obtain exceptional river discharges at Lobith. A correct simulation of such events requires that the spatial dependence of multi-day amounts is preserved. We first present some theory about the spatial dependence of extremes. Thereafter, we discuss the dependence structure of the observed N -day winter maximum precipitation amounts and its reproduction in the two simulations used in Section 4.

5.1 Theoretical Background

There are various measures to compare the spatial dependence of large N -day winter precipitation amounts in the observed and simulated data. A natural measure is the cross-correlation coefficient between the N -day winter maxima at different sites. However, a correlation coefficient only measures the strength of linear dependence between two variables. Counting the number of joint exceedances (or non-exceedances) of various thresholds gives a more complete description of spatial association. These counts characterise the structure of the joint distribution of the winter maxima X_{1j}, X_{2j} at two sites:

$$F(x_1, x_2) = \Pr(X_{1j} \leq x_1, X_{2j} \leq x_2) \quad (17)$$

where j is a year index. It should be noted that X_{1j} and X_{2j} need not necessarily relate to the same N -day period in the j th winter. This is a standard situation in the mathematical theory of bivariate extremes (Tiago de Oliveira, 1984; Tawn, 1988). The marginal distributions are denoted as $F_1(x_1) = \Pr(X_{1j} \leq x_1)$ and $F_2(x_2) = \Pr(X_{2j} \leq x_2)$.

In this section, particular attention is given to the case where x_1 and x_2 correspond to the p -quantiles x_{1p} and x_{2p} of the marginal distributions, i.e. $F_1(x_{1p}) = p$ and $F_2(x_{2p}) = p$. Then, if X_{1j} and X_{2j} are independent:

$$F(x_{1p}, x_{2p}) = F_1(x_{1p})F_2(x_{2p}) = p^2 \quad (18)$$

whereas, in case of complete positive dependence:

$$F(x_{1p}, x_{2p}) = F_1(x_{1p}) = p \quad (19)$$

More general, for the joint distribution at the p -quantiles we may write:

$$F(x_{1p}, x_{2p}) = p^{h(p)} \quad (20)$$

where

$$h(p) = \frac{\ln F(x_{1p}, x_{2p})}{\ln p} \quad (21)$$

The quantity $h(p)$ determines the degree of association at the p -quantile. There is little association if $h(p)$ is close to 2. If this holds for large p , then the probability is small that both X_{1j} and X_{2j} are extreme. The extremes tend to occur simultaneously when there is strong association in the upper tail, i.e. when $h(p)$ is close to 1 for large p .

For bivariate normal variables with correlation coefficient < 1 the function $h(p)$ tends to 2 as $p \rightarrow 1$. Theoretical limiting distributions of bivariate maxima have the property that $h(p)$ is constant. For these distributions $h(p)$ is known as the extremal coefficient (Coles, 1993). As an example, we consider the logistic model for bivariate Gumbel variables (Tiago de Oliveira, 1984):

$$\Pr(X_{1j} \leq x_1, X_{2j} \leq x_2) = \exp\left\{-\left[e^{-x_1/(1-\theta)} + e^{-x_2/(1-\theta)}\right]^{(1-\theta)}\right\}, \quad 0 \leq \theta \leq 1 \quad (22)$$

For ease of exposition we have assumed here that the location parameter is 0 at both sites and that the scale parameter is 1. The parameter θ controls the dependence, $\theta = 0$ implies independence and $\theta = 1$ complete dependence. It is easily verified that $h(p) = 2^{(1-\theta)}$ for this bivariate extreme-value distribution. For the case that X_{1j} and X_{2j} have a correlation coefficient of 0.6 ($\theta = 0.3675$), which is found for inter-station distances of about 50 km, Table 14 compares the value of $h(p)$ with those obtained for the bivariate normal distribution. The table shows that around the median the two bivariate distributions have about the same value of $h(p)$. For the bivariate normal distribution $h(p)$ increases with increasing p . However, even at the 0.99-quantile, $h(p)$ still differs considerably from the limiting value 2 for this distribution.

Joe *et al.* (1992) considered as an alternative dependence measure Kendall's τ_b :

$$\tau_b(p) = \frac{F(x_{1p}, x_{2p}) - p^2}{p(1-p)} \quad (23)$$

Like a correlation coefficient $\tau_b(p) = 0$ for independent data and $\tau_b(p) = 1$ in case of complete positive dependence. The quantity $\tau_b(p)$ is, however, not constant for the theoretical limiting distributions of bivariate maxima.

To estimate $F(x_{1p}, x_{2p})$ from a sequence of paired winter maxima $(X_{11}, X_{21}), \dots, (X_{1K}, X_{2K})$, we first replace the unknown p -quantiles by their sample equivalents $\hat{x}_{1p}, \hat{x}_{2p}$ and then count the number K_{joint} of pairs for which both $X_{1j} \leq \hat{x}_{1p}$ and $X_{2j} \leq \hat{x}_{2p}$, giving:

$$\hat{F}(x_{1p}, x_{2p}) = K_{\text{joint}} / K \quad (24)$$

The estimate $\hat{h}(p)$ follows then by substituting $\hat{F}(x_{1p}, x_{2p})$ in the right-hand side of equation (21).

A slight modification is necessary to estimate $h(p)$ at the upper quintile mean QM5, introduced in Section 3. The probability associated with QM5 weakly var-

p	$h(p)$	
	Gumbel	Normal
0.5	1.55	1.50
0.8	1.55	1.60
0.9	1.55	1.67
0.95	1.55	1.72
0.99	1.55	1.82

Table 14: Variation of $h(p)$ with p for the bivariate Gumbel distribution in equation (22) and the bivariate normal distribution, each with correlation coefficient 0.6.

ies with K and depends on the type of distribution as well. We therefore also replace p in equation (21) by an estimate from the paired winter maxima. This estimate is obtained as:

$$\hat{p} = (K_1 + K_2) / (2K) \tag{25}$$

with K_1 the number of winter maxima at gauge 1 for which $X_{1j} \leq x_1$ and K_2 the corresponding number for gauge 2. The final estimate of $h(p)$ then becomes:

$$\hat{h}(p) = \frac{\ln \hat{F}(x_{1p}, x_{2p})}{\ln \hat{p}} \tag{26}$$

Unfortunately, the estimates of $h(p)$ at different quantiles tend to be rather erratic for single gauge pairs. Buishand (1984) therefore recommended to pool the bivariate maxima of all gauge pairs within comparable inter-station distances.

5.2 Application to the Rhine basin

Table 15 presents the distances between the 25 German stations in Table 1. For these stations we defined distance intervals of 30 km. Table 16 shows that the number n_s of gauge pairs in a distance interval then typically ranges between 20 and 40 for inter-station distances between 50 and 350 km and $n_s \approx 10$ outside this range. For each distance interval, equations (24) and (25) were used with $K = n_s J$, where J is the number of years ($J = 34$).

Figure 10 shows the values of $\hat{h}(p)$ for the upper quintile mean ($p = 0.91$) and the median ($p = 0.5$) of the 1, 4, 10 and 20-day winter maximum precipitation amounts. It is seen that $\hat{h}(p)$ increases with increasing distance. This demonstrates that dependence is relatively strong at short distances and relatively weak at long distances. The strength of spatial association further increases with increasing duration and the dependence is also somewhat stronger at the median than at the upper quintile mean. The latter is at variance with Buishand (1984)

No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1	0	159	283	221	343	182	64	331	165	197	285	171	218	199	127	215	180	133	64	105	74	219	180	164	134
2	-	0	126	143	189	167	178	183	141	163	136	47	78	116	234	59	22	103	114	59	115	172	107	189	206
3	-	-	0	204	108	225	304	121	239	248	98	139	113	187	356	68	107	189	230	185	239	203	157	263	299
4	-	-	0	185	306	193	161	67	53	130	95	93	27	213	172	138	238	215	138	147	314	250	320	320	322
5	-	-	-	0	326	344	27	243	238	61	172	127	181	383	152	168	279	303	237	282	308	257	360	389	
6	-	-	-	-	0	242	330	284	314	288	213	243	279	307	176	181	69	125	176	208	45	70	44	96	
7	-	-	-	-	-	0	328	128	157	284	174	217	178	66	238	197	185	118	119	65	276	229	227	198	
8	-	-	-	-	-	0	222	215	47	160	114	160	365	153	162	278	296	226	268	315	260	361	387		
9	-	-	-	-	-	-	0	33	183	104	128	61	146	189	147	215	175	109	94	302	240	289	282		
10	-	-	-	-	-	-	-	0	181	121	135	62	165	205	166	245	207	139	126	329	266	320	314		
11	-	-	-	-	-	-	-	-	0	113	67	123	323	113	116	234	249	180	222	278	219	318	341		
12	-	-	-	-	-	-	-	-	-	0	46	69	220	85	45	147	141	67	110	219	154	232	243		
13	-	-	-	-	-	-	-	-	-	-	0	74	260	83	62	181	185	113	155	241	177	266	283		
14	-	-	-	-	-	-	-	-	-	-	-	0	205	148	112	210	189	112	126	287	222	293	296		
15	-	-	-	-	-	-	-	-	-	-	-	-	0	293	251	251	183	177	121	341	294	290	257		
16	-	-	-	-	-	-	-	-	-	-	-	-	-	0	44	128	162	118	174	165	107	208	238		
17	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	120	135	78	132	180	116	205	226		
18	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	69	106	142	92	48	86	111		
19	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	78	87	158	116	117	108		
20	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	56	194	133	183	185		
21	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	234	179	204	190		
22	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	65	87	140		
23	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	106	146		
24	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	54		
25	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0		

Table 15: Distances (km) between the stations in Table 1.

class	class boundaries (km)	pairs	avdist (km)
1	distance ≤ 50	12	39.5
2	50 < distance ≤ 80	27	65.9
3	80 < distance ≤ 110	23	98.9
4	110 < distance ≤ 140	44	123.5
5	140 < distance ≤ 170	32	155.1
6	170 < distance ≤ 200	45	183.0
7	200 < distance ≤ 230	32	215.1
8	230 < distance ≤ 260	27	242.5
9	260 < distance ≤ 290	20	278.0
10	290 < distance ≤ 320	19	304.5
11	320 < distance ≤ 350	12	331.7
12	distance > 350	7	371.7

Table 16: Definition of class boundaries, number of station pairs and average distance (avdist) for each distance interval.

who found for 1-day winter maximum precipitation in the Netherlands a decrease of $\hat{h}(p)$ with increasing p . The observed differences at the median and the upper quintile mean are also not in agreement with the constancy of $h(p)$ for theoretical limiting distributions of bivariate maxima. The validity of such a limiting distribution is in fact a very strong assumption, which is often not met for observed extreme-value data (Ledford and Tawn, 1996)

For case 4.4 (unconditional) and $k = 5$, Figure 11 compares the simulated and historical values of $\hat{h}(p)$ for the upper quintile mean and the median of the 1 and 10-day winter maximum precipitation amounts. The figure shows that the historical values of $\hat{h}(p)$ are generally well reproduced by the simulations. This could be expected for the 1-day maxima because we resample daily precipitation fields. For the median of the 10-day maxima, the simulations show on average somewhat more dependence (underprediction of $\hat{h}(p)$) than the historical data. There is, however, no statistical evidence that the differences are significant because the observed values of $\hat{h}(p)$ fall well within the boxplots.

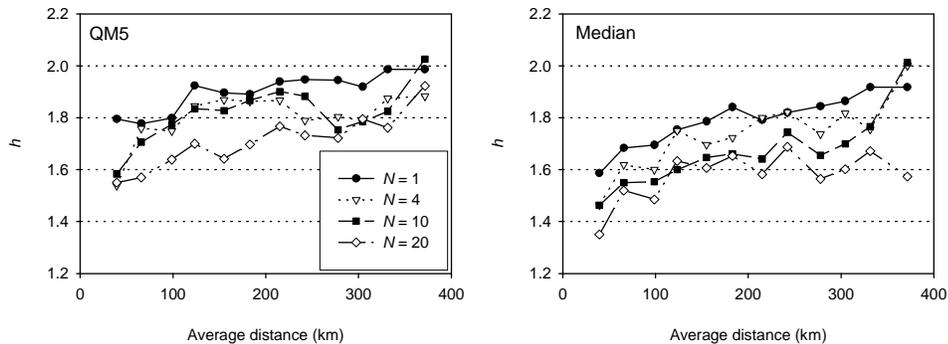


Figure 10: Estimates of h for the upper quintile mean QM5 and the median of the historical (1961–1995) 1, 4, 10 and 20-day winter maximum precipitation amounts as a function of inter-station distance.

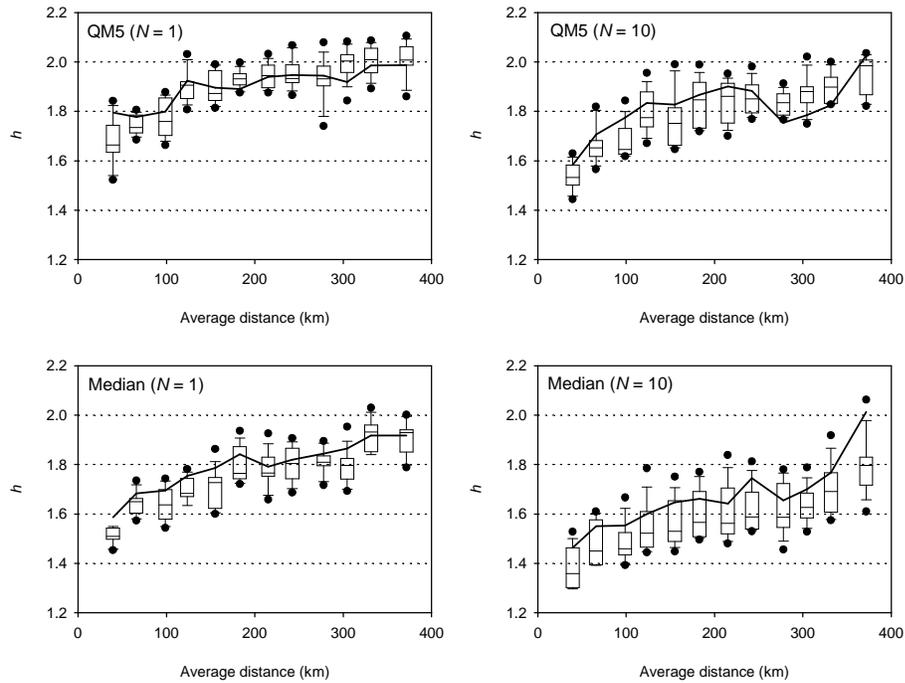


Figure 11: Boxplots of the estimates of h for the upper quintile mean QM5 and the median of the 1 and 10-day winter maximum precipitation amounts for case 4.4 (unconditional; $k = 5$). Each boxplot represents ten runs of 35 years. The estimates for the historical (1961–1995) data are represented by the solid lines.

Figure 12 presents the corresponding estimates of $h(p)$ for case 4.1 (conditional). These results are similar to those for the unconditional simulations in Figure 11, except for $N = 10$ at the upper quintile mean. In the latter case, four out of the twelve boxplots are entirely below the historical values indicating a strong tendency to underpredict $\hat{h}(p)$. To investigate this further, we calculated bootstrap standard errors se of $\hat{h}(p)$ for $N = 10$ using 500 bootstrap samples. Figure 13 shows the $2 \times se$ bands for the historical data together with the mean for the simulations. For inter-station distances ranging between 170 and 260 km the underprediction is larger than $2 \times se$. This is, however, not sufficient to suspect a systematic departure of the simulated data. It is quite natural that a small number of the differences between the observed and simulated values will be declared as statistically significant if one makes several comparisons.

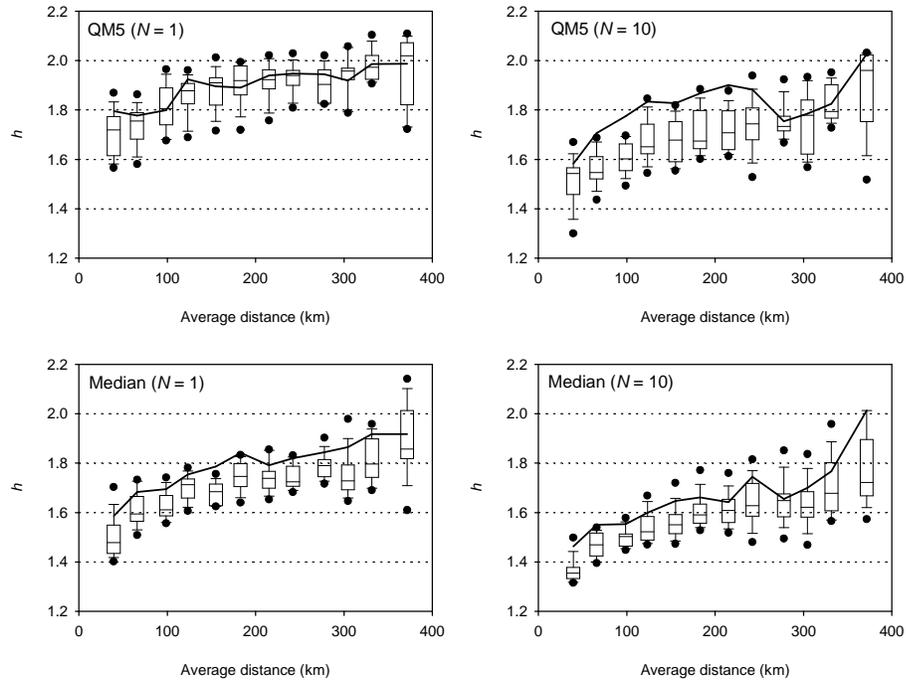


Figure 12: Boxplots of the estimates of h for the upper quintile mean QM5 and the median of the 1 and 10-day winter maximum precipitation amounts for case 4.1 (conditional; $k = 5$). Each box represents ten runs of 35 years. The estimates for the historical (1961–1995) data are represented by the solid lines.

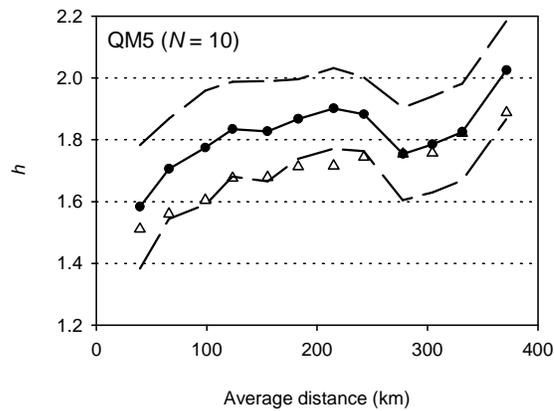


Figure 13: Bootstrap $2 \times \text{se}$ bands (dashed lines) for the estimates of h for the upper quintile mean of the 10-day winter maximum precipitation amounts of the historical (1961–1995) data (solid line through dots). The triangles represent the mean estimate of h for case 4.1 (conditional; $k = 5$).

6. Long-duration simulations

The most important objective of the rainfall generator is the simulation of unprecedented extreme rainfall events. Figures 8 and 9, for instance, already showed that more extreme multi-day precipitation amounts occurred in simulated runs of 35 years than the largest observed values. In the present section we study long-duration simulations with a length of 1000 years. First the distributions of the N -day winter precipitation maxima in two different simulation runs are considered. Thereafter, three historical and six simulated events with extreme 10-day precipitation over the total area are selected from one of these runs and their space-time patterns are discussed. Large river discharges at Lobith in the past were often accompanied by extensive precipitation over a period of about ten days. This section only deals with unconditional simulations, because the conditional methods would need an additional model for generating circulation indices to obtain long-duration simulations.

6.1 Distribution of the N -day winter maxima

For case 4.4 two unconditional 1000-year simulations were performed, one for $k = 5$ and one for $k = 20$. Table 17 compares the N -day winter maximum precipitation amounts in both simulations with the historical N -day winter maximum. The table clearly shows that the objective of generating unprecedented extreme rainfall sequences is met. For area V, the 10-day maximum is in the $k = 20$ simulation even 69.6% larger than the historical maximum. The largest 10 and 20-day amounts in the $k = 5$ simulations are on average greater than those in the $k = 20$ simulation. Furthermore, both simulations show a large variability in the percentage differences for the five areas. This illustrates that there is little dependence between the largest N -day precipitation amounts over the various areas, as expected from the results for the spatial dependence of the maxima at the indi-

	Maximum (%)			
	$N = 1$	$N = 4$	$N = 10$	$N = 20$
Case 4.4 ($k = 5$)				
Mean 25 stations	8.5	24.7	33.0	24.3
Area I	12.2	3.9	39.5	43.9
Area II	12.2	15.6	12.7	8.1
Area III	2.5	40.2	40.5	14.2
Area IV	2.5	12.0	38.7	24.9
Area V	0.2	26.1	48.3	34.1
Total area	2.7	5.9	34.4	13.6
Case 4.4 ($k = 20$)				
Mean 25 stations	8.5	26.2	24.6	18.7
Area I	12.2	11.3	23.2	37.1
Area II	11.8	22.8	6.7	13.6
Area III	2.0	30.0	0.4	6.1
Area IV	2.5	23.4	33.4	16.5
Area V	1.2	22.6	69.6	36.6
Total area	2.4	25.5	24.8	10.0

Table 17: Percentage differences between the largest N -day winter (October-March) precipitation maxima in the 1000-year simulations and the historical records (1961–1995) for point precipitation (mean N -day maximum of 25 stations), for each of the five areas and for the total area.

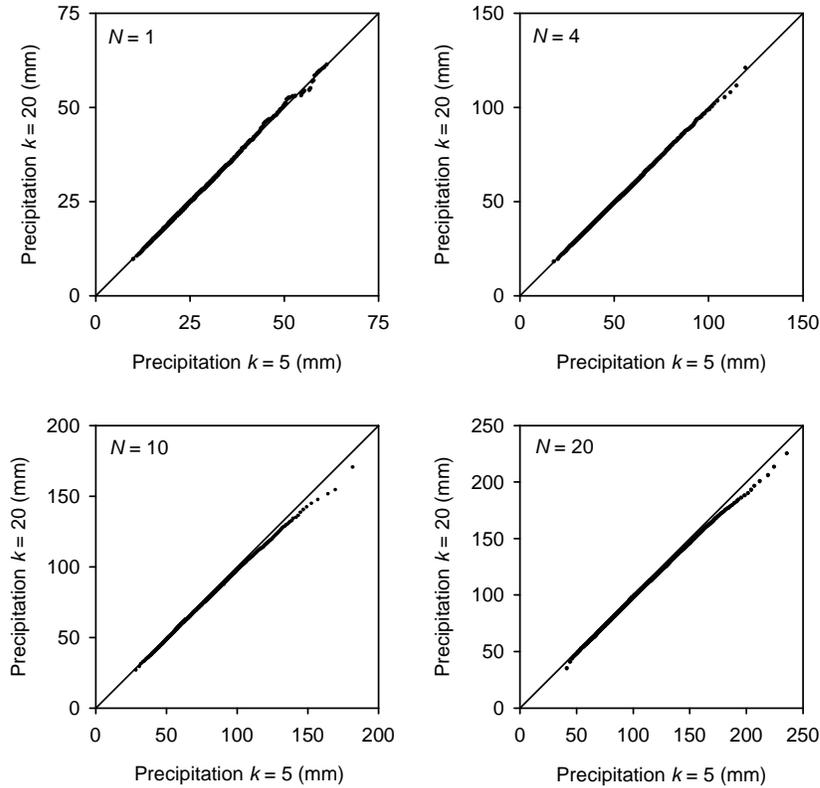


Figure 14: Comparison of the sorted N -day winter precipitation maxima for the two unconditional 1000-year simulations (case 4.4: $k = 5$ and $k = 20$), averaged over 25 stations.

vidual stations in Section 5. It is also noteworthy that the simulated 1-day maxima are greater than the largest historical values. This may be due to: (1) the use of a slowly seasonally varying mean to standardise the observations before resampling; and (2) the use of the moving window, which allows for resampling of days outside the boundaries of the winter half-year.

To study the differences between the two simulations further, the sorted N -day winter maximum precipitation amounts, averaged over the 25 stations, for $k = 20$ are plotted against those for $k = 5$ in Figure 14. The figure clearly shows that the 1 and 4-day maxima are comparable for the two simulations, but for the $k = 5$ simulation the 10 and 20-day maxima are more extreme than those in the $k = 20$ simulation. Although the variability of the most extreme order statistics is large, a systematic effect of the choice of k also may play a role (Appendix B). Because of the results in Section 3.3, the $k = 5$ simulation should be regarded as more realistic than the $k = 20$ simulation.

Figure 15 compares for $k = 5$ the Gumbel plots of the 10-day winter maxima of the historical data with those of the 1000-year simulation for Essen, Frankfurt, Freudenstadt, area II, area IV and the total area. There is a reasonable correspondence between the historical and simulated distributions. The plot for the simulated values shows some tendency to flatten at long return periods, in agreement with the positive value of the GEV shape parameter for the 10-day winter maxima (Table 9).

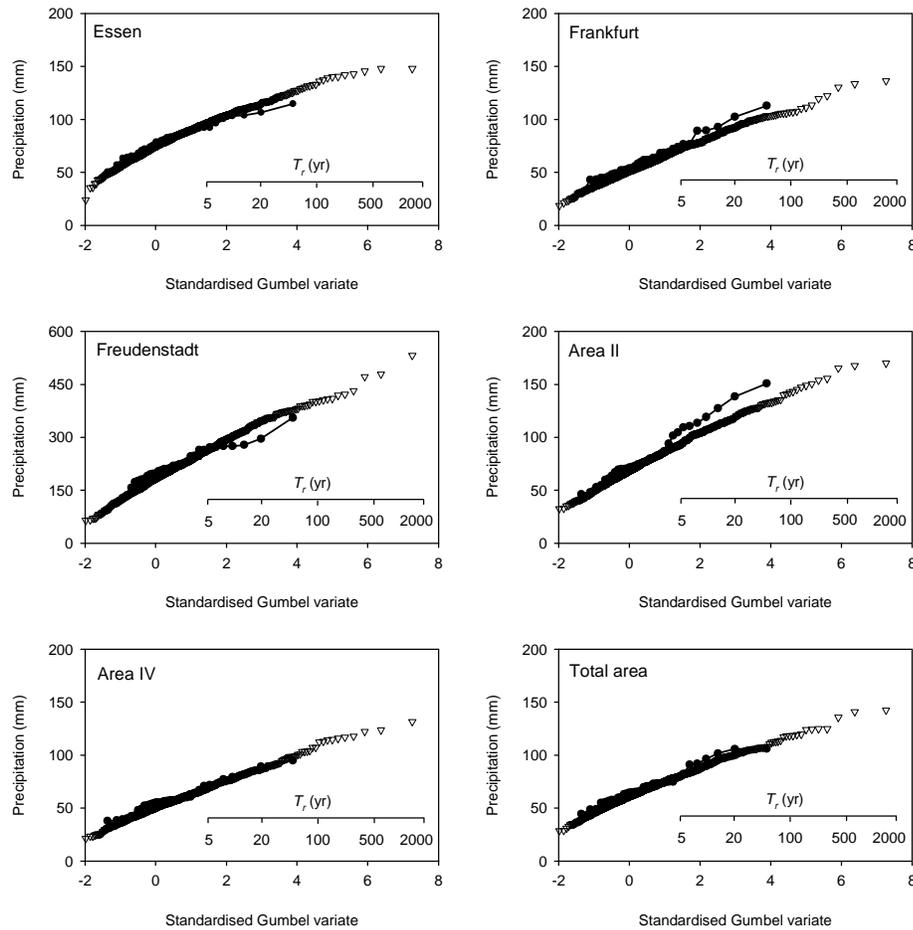


Figure 15: Gumbel plots of the 10-day winter (October–March) maxima for the observed precipitation (1961–1995) at Essen, Frankfurt, Freudenstadt, area II, area IV and the total area (solid dots and lines) and the corresponding simulated data for case 4.4 using $k = 5$ (one unconditional simulation of 1000-years, represented by open triangles). T_r denotes the return period.

6.2 Space-time pattern of selected 10-day events

Together with the antecedent conditions, the spatial and temporal pattern of extreme precipitation events strongly influences the magnitude of floods. Therefore, it may be of interest to compare the space-time pattern of some extreme historical precipitation events with simulated extreme precipitation events.

In this section, extreme precipitation events are defined as events where the winter maximum of the average 10-day precipitation over the total area exceeds 100 mm. For the historical 1961–1995 period, three events are found: (1) 6 October - 15 October 1982; (2) 12 December - 21 December 1993; and (3) 20 January - 29 January 1995. The area-average 10-day amounts of these events equal 105.9, 105.7 and 101.6 mm, respectively. Although the October 1982 event has the largest area-average 10-day amount, it did not result in a large river discharge because at the end of the summer the Rhine basin can normally store much water. The second event led to the Christmas 1993 floods in Germany. Only the January 1995 event caused a threatening situation in the Netherlands, giving rise to the

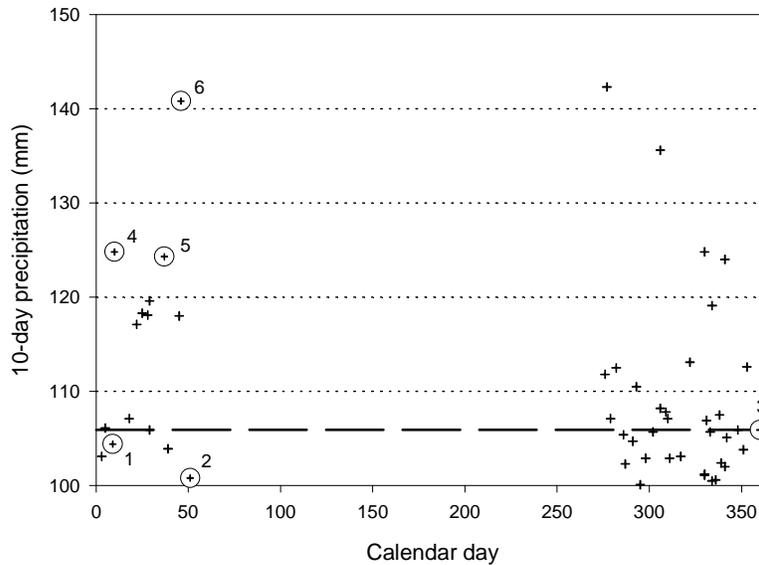


Figure 16: Events with a 10-day winter maximum precipitation (for the total area) > 100 mm in a 1000-year simulation (case 4.4, $k = 5$). The dashed line indicates the largest historical event in the period 1961–1995. The encircled events are selected for further analysis.

evacuation of more than 200 000 people. A special law was issued to complete the most urgent dike reconstructions in 1995 and 1996. The meteorological and hydrological aspects of the January 1995 floods in Germany have been described by Fink *et al.* (1996). Upstream of Köln (Cologne), the peak discharges of the Rhine were less extreme than in the December 1993 event.

For the 1000-year simulation of case 4.4 ($k = 5$), Figure 16 shows all 10-day winter maximum events with area-average precipitation exceeding 100 mm. The horizontal dashed line represents the largest recorded area-average 10-day winter amount of October 1982. The encircled events in Figure 16 are selected for comparison with the historical events of October 1982, December 1993 and January 1995. Table 18 presents some details of these historical and simulated events. In a number of cases the average temperature over the 30-day period preceding the extreme 10-day event is negative or slightly above zero, indicating that snowmelt and frozen soils may significantly contribute to large river discharges at Lobith. The antecedent precipitation of 131.8 mm in the Sim3 case (year 473) is much larger than that of the other selected cases.

Figures 17, 18 and 19 present the spatial and temporal distribution of the selected 10-day events. For each event the spatial distribution of the 10-day precipitation amounts is shown on the left, together with the temporal distribution of the daily precipitation over the total area in the lower left corner. The temporal distribution for the five areas in Figure 1 is presented on the right.

The spatial variation of the long-term mean winter precipitation in Figure 6 is, to a certain extent, reflected in Figures 17, 18 and 19. There is often a maximum over the Black Forest. Two exceptions are the January 1995 event and the simulated 10-day precipitation amounts in year 473, which show a clear maximum over the Eifel and Hunsrück along the Mosel tributary in the western part of Germany. In the 1995 event there is a second maximum over the Sauerland region in the north of the basin. The coincidence of runoff from that area with the flood wave from the south was an important factor contributing to the peak dis-

Case	Month	Year	ΣP (mm)	Antecedent (30d)	
				P (mm)	T (°C)
Hist1	October	1982	105.9	36.4	15.8
Hist2	December	1993	105.7	54.9	0.5
Hist3	January	1995	101.6	85.6	0.1
Sim1	January	861	104.4	84.9	0.3
Sim2	February	864	100.8	45.6	-3.2
Sim3	December	473	105.9	131.8	3.1
Sim4	January	338	124.7	86.7	4.1
Sim5	February	239	124.3	76.4	2.3
Sim6	February	396	140.8	53.8	-2.8

Table 18: Characteristics of the selected historical and simulated 10-day precipitation events (see Figure 16). The values for precipitation and temperature are for the total area.

charges at Lobith during that event (Fink *et al.*, 1996). The average 10-day rainfall over the total area and the antecedent average temperature and precipitation of the 1995 event are comparable with that of the Sim1 case (Table 18), but the spatial and temporal distribution of the precipitation is somewhat less favourable for the 1995 event.

The most extreme selected event (Sim6, year 356) is characterised by excessive precipitation over the northern part (Sauerland), the eastern part (Main tributary) and the southern part (Black Forest) of the area. The 30-day period preceding that event is rather cold (Table 18).

The examples in Figures 17, 18 and 19 show that excessive precipitation over a large part of the area is generally needed to obtain extreme area-average precipitation. These examples may therefore wrongly suggest that the degree of dependence between the extremes over the subareas is rather strong. In fact, in this 1000-year simulation, the largest 10-day precipitation amounts over the five areas are found in different years, which is in line with earlier conclusions about spatial dependence between extremes.

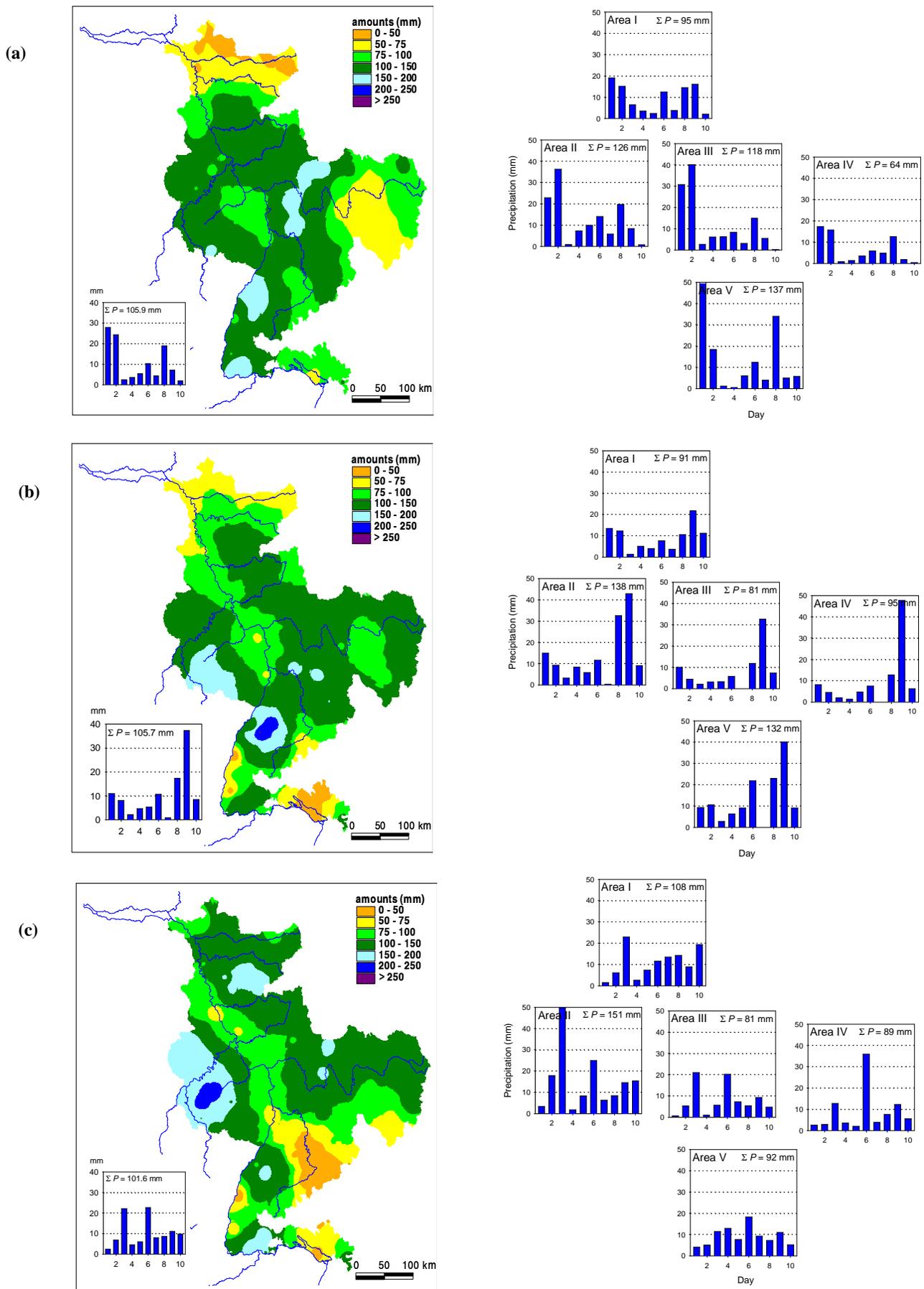


Figure 17: Spatial and temporal distribution of historical (1961–1995) 10-day maximum precipitation events: (a) 6–15 October 1982; (b) 12–21 December 1993; and (c) 20–29 January 1995.

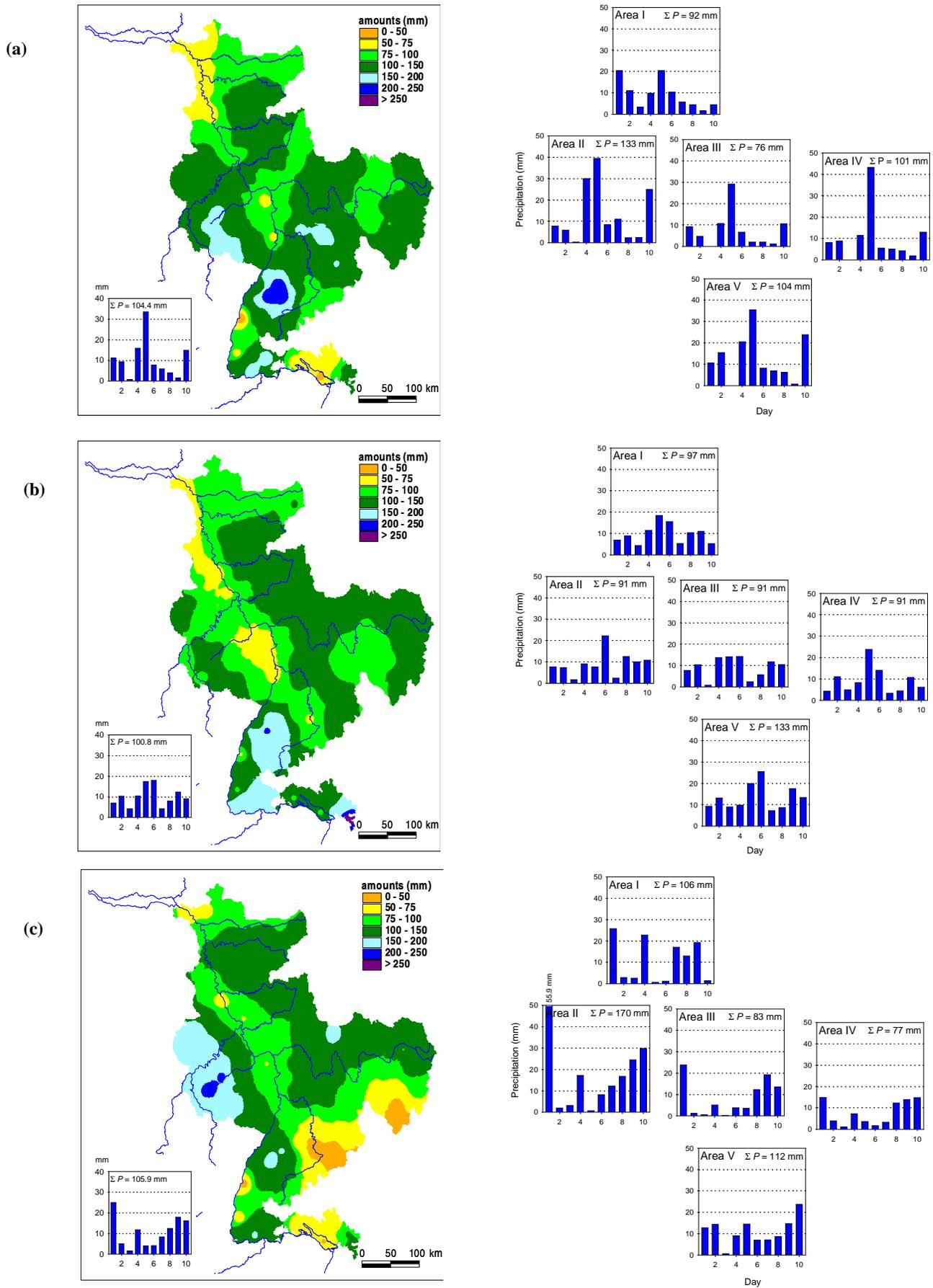


Figure 18: Spatial and temporal distribution of 10-day maximum precipitation events in a 1000-year simulation: (a) 5–14 January 861; (b) 16–25 February 864; and (c) 22–31 December 473.

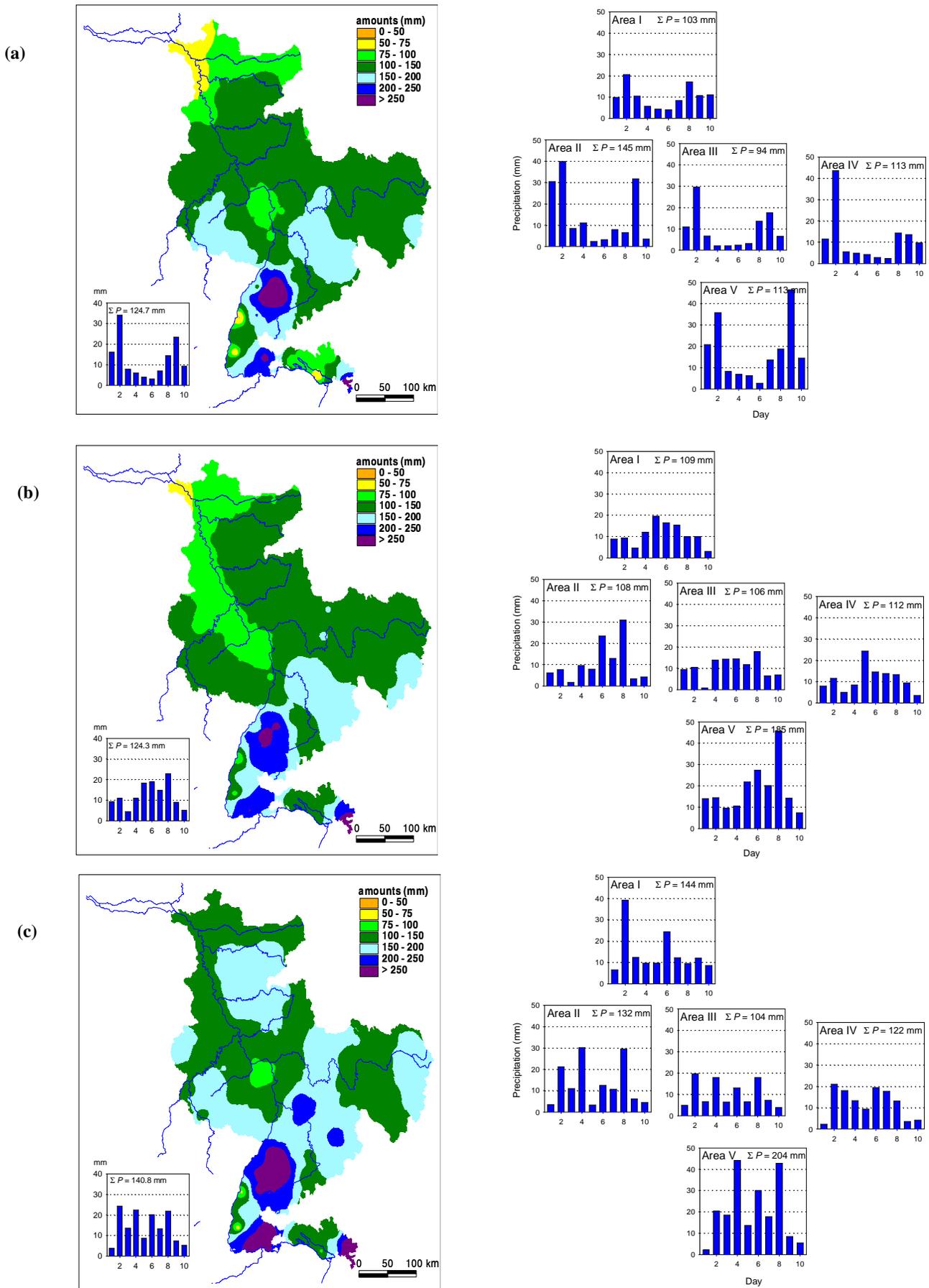


Figure 19: Spatial and temporal distribution of 10-day maximum precipitation events in a 1000-year simulation: (a) 6–15 January 338; (b) 2–11 February 239; and (c) 11–20 February 396.

7. Discussion and conclusions

In this report the multi-site simulation of daily precipitation and temperature was explored. Twenty-five stations with daily precipitation and temperature data in the German part of the Rhine basin were considered.

For multi-site simulations, the dimension of \mathbf{D}_t should be kept within reasonable limits. Therefore, we used summary statistics of the P and T fields in \mathbf{D}_t , in most cases supplemented with circulation indices. It is noteworthy that several combinations of these summary statistics produced comparable results. Also the use of weights in the calculation of the Euclidean distance has only a marginal influence.

The choice of k proved to be rather crucial. Autocorrelation coefficients and distributions of N -day winter maximum precipitation amounts were better reproduced by taking k as small as 2 or 5 instead of $k = 20$ as in the single-site simulations in Brandsma and Buishand (1998). For $k = 2$, the simulated data will contain runs of more than 20 consecutive historical days. This may, however, not be detrimental for the present application. The method e.g. still produces more extreme 10-day precipitation amounts than those observed in the past. The choices of $k = 2$ and $k = 5$ need further attention in subsequent phases of the project.

It is striking that the reproduction of the distribution of multi-day winter maximum precipitation amounts in Section 4 is at least as good as in BB97 for the single-site simulations. The large underprediction of the median in BB97 disappeared in the present study. With respect to snowmelt, the results are comparable to those reported for the single-site simulations in Brandsma and Buishand (1998). Although the autocorrelation of the daily temperatures is not fully reproduced, the distributions of the N -day maximum snowmelt derived from the simulated temperature and precipitation data were close to those derived from the observed data.

In order to generate realistic events causing large river discharges at Lobith, it is necessary that the spatial association of large multi-day precipitation amounts is preserved. Although the conditional $k = 5$ simulation showed a somewhat stronger dependence in the N -day winter maximum precipitation amounts than the historical data, it can be concluded that both the unconditional and conditional method are able to reproduce the spatial association of large multi-day amounts.

It turned out in this study that unconditional simulation of precipitation, temperature and circulation indices performed somewhat better than conditional simulation of precipitation and temperature on circulation indices. For unconditional simulation of daily precipitation and temperature the need to incorporate circulation indices in \mathbf{D}_t is, however, questionable. In contrast with the single-site simulations in BB97, the use of these indices does not lead to a better reproduction of the autocorrelation properties of daily precipitation. Generating daily precipitation and temperature conditional on circulation indices remains, however, an important topic. A separate stochastic model for generating circulation indices is now being developed at KNMI to make it possible to produce simulation runs that can exceed the length of the MSLP data set (about 120 years). Such a model

may better reproduce the autocorrelation properties of these indices than in the unconditional simulations presented here. Furthermore, conditional simulation has potential applications for downscaling the output of general circulation models.

A difficult question is, how far the length of the observed record may put limitations on the length of a simulation run. For $N = 4, 10$ and 20 a repetition of large N -day precipitation amounts was not found in the 1000-year unconditional simulations in Section 6. The most extreme simulated values were generally found to be well above the observed winter maxima.

Despite some shortcomings, the multi-site extension of the nearest-neighbour method looks promising. The selected extreme events in the 1000-year $k = 5$ simulation have different antecedent conditions and show a large variation in the temporal and spatial patterns of the daily amounts.

Acknowledgements

We thank J.J. Beersma (KNMI), B.W.A.H. Parmet and H.C. van Twuiver (RIZA) for comments on earlier drafts of the report and F.J.M. van der Wel (KNMI) for his assistance with using GIS. 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 stations were made available by the Deutscher Wetterdienst via the 'International Commission for the Hydrology of the Rhine Basin' (CHR).

References

- Bárdossy, A., and E.J. Plate, Space-time model for daily rainfall using atmospheric circulation patterns, *Water Resour. Res.*, **28**: 1247–1259, 1992.
- Beersma, J.J., and T.A. Buishand, A simple test for equality of monthly variances in climate time series, *J. Climate*, **12**: 1770–1779, 1999.
- Bennekom, A.R. van, and B.W.A.H. Parmet, *Bemessungsabfluß in den Niederlanden; menschliche Einflüsse und andere Unsicherheiten*, In: *Zukunft der Hydrologie in Deutschland*, BfG Mitteilung 16, 125–131, Bundesanstalt für Gewässerkunde, Koblenz, 1998 (in German).
- Brandsma, T., 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, 47 pp., 1997.
- Brandsma, T., and T.A. Buishand, Simulation of extreme precipitation in the Rhine basin by nearest-neighbour resampling, *Hydrol. Earth Syst. Sci.*, **2**: 195–209, 1998.
- Buishand, T.A., Some remarks on the use of daily rainfall models, *J. Hydrol.*, **36**: 295–308, 1978.
- Buishand, T.A., Bivariate extreme-value data and the station-year method, *J. Hydrol.*, **69**: 77–95, 1984.
- Buishand, T.A., *Het sneeuwdek in Nederland*, Wetenschappelijk rapport WR 86-6, KNMI, De Bilt, 94 pp., 1986 (in Dutch).
- Buishand, T.A., Extreme rainfall estimation by combining data from several sites, *Hydrol. Sci. J.*, **36**: 345–365, 1991.
- Buishand, T.A., and J.J. Beersma, Jackknife tests for differences in autocorrelation between climate time series, *J. Climate*, **6**: 2490–2495, 1993.
- Buishand, T.A., and T. Brandsma, *Rainfall Generator for the Rhine catchment: a feasibility study*, Technical report TR-183, KNMI, De Bilt, 54 pp., 1996.
- Coles, S.G., Regional modelling of extreme storms via max-stable processes, *J. R. Statist. Soc. B*, **55**: 797–816, 1993.
- Cramér, H., *Mathematical Methods of Statistics*, Princeton University Press, 575 pp., 1946.
- Dales, M.Y., and D.W. Reed, *Regional flood and storm hazard assessment*, Report No. 102, Institute of Hydrology, Wallingford, 159 pp., 1989.
- Delft Hydraulics and EAC-RAND, *Toetsing uitgangspunten rivierdijkversterkingen, Deelrapport 2: Maatgevende belastingen*, Delft Hydraulics, Emmeloord, and European American Center for Policy Analysis (EAC-RAND), Delft, the Netherlands 1993 (in Dutch).
- Dupriez, G.L., and G. Demarée, *Totaux Pluviométriques sur des Périodes Continues de 1 à 30 Jours: I. Analyse de 11 Séries Pluviométriques de plus de 80 Ans*, Miscellanea Série A, No. 8, Institut Royal Météorologique de Belgique, Bruxelles, 154 pp., 1988.
- Feller, W., *An Introduction to Probability Theory and Its Applications*, Volume I, 3rd edition, Wiley, New York, 509 pp., 1968.
- Fink, A., U. Ulbrich, and H. Engel, Aspects of the January 1995 flood in Germany, *Weather*, **51**: 34–39, 1996.
- Gordon, L., M.F. Schilling and M.S. Waterman, An extreme value theory for long head runs, *Prob. Th. Rel. Fields*, **72**: 279–287, 1986.
- Gradshteyn, I.S., and I.M. Ryzhik, *Table of Integrals, Series, and Products*, Corrected and enlarged edition, Academic Press, San Diego, Ca, 1160 pp., 1980.
- Gray, D.M., and T.D. Prowse, *Snow and Floating Ice*, Chapter 7 in *Handbook of Hydrology* (D.R. Maidment, ed.), McGraw-Hill, New York, 1993.
- Härdle, W., *Applied Nonparametric Regression*, Cambridge University Press, Cambridge, UK, 333 pp., 1990.
- Hosking, J.R.M., J.R. Wallis and E.F. Wood, Estimation of the generalized extreme-value distribution by the method of probability-weighted moments, *Technometrics*, **27**: 251–261, 1985.

- Joe, H., R.L. Smith, and I. Weissman, Bivariate thresholds methods for extremes, *J. R. Statist. Soc. B*, **54**: 171–183, 1992.
- Jones, P.D., M. Hulme, and K.R. Briffa, A comparison of Lamb circulation types with an objective classification scheme, *Int. J. Climatol.*, **13**: 655–663, 1993.
- Katz, R.W., and M.B. Parlange, Effect of an index of atmospheric circulation on stochastic properties of precipitation, *Water Resour. Res.*, **29**: 2335–2344, 1993.
- Katz, R.W., and M.B. Parlange, Mixtures of stochastic processes: application to statistical downscaling, *Clim. Res.*, **7**: 185–193, 1996.
- Lall, U., and A. Sharma, A nearest neighbor bootstrap for resampling hydrologic time series, *Water Resour. Res.*, **32**: 679–693, 1996.
- Ledford, A.W., and J.A. Tawn, Statistics for near independence in multivariate extreme values, *Biometrika*, **83**: 169–187, 1996.
- Linsley, R.K., M.A. Kohler, and J.L.H. Paulhus, *Hydrology for Engineers*, McGraw-Hill, London, 492 pp., 1988.
- Lynn, P.P., and M.A. Beran, A fast method for generating ordered samples of random numbers from the extreme value distributions, *J. Hydrol.*, **40**: 377–379, 1979.
- Müller-Westermeier, G., *Numerisches Verfahren zu Erstellung klimatologischer Karten*, Berichte des Deutschen Wetterdienstes 193, Deutscher Wetterdienst, Offenbach am Main, 30 pp., 1995 (in German).
- NERC, *Flood Studies Report, Volume II: Meteorological Studies*, Natural Environment Research Council (NERC), London, 81 pp., 1975.
- Rajagopalan, B., and U. Lall, A nearest neighbor bootstrap for resampling daily precipitation and other weather variables, Working paper WP-95-HWR-UL/013, Utah State University, Logan, Utah, 1995.
- Stedinger, J.R., R.M. Vogel, and E. Foufoula-Georgiou, *Frequency analysis of extreme events*, Chapter 18 in *Handbook of Hydrology* (D.R. Maidment, ed.), McGraw-Hill, New York, 1993.
- Stuart, A., and J.K. Ord, *Kendall's Advanced Theory of Statistics, Volume 1: Distribution Theory*, 5th edition, Charles Griffin, London, 604 pp., 1986.
- Tawn, J.A., Bivariate extreme value theory: models and estimation, *Biometrika*, **75**: 397–415, 1988.
- Tiago de Oliveira, J., *Bivariate models for extremes; statistical decision*, In: *Statistical Extremes and Applications* (J. Tiago de Oliveira, ed.), 131–153, D. Reidel Publishing Company, 1984.
- Wang, Q.J., Using partial probability weighted moments to fit the extreme value distributions to censored samples, *Water Resour. Res.*, **32**: 1767–1771, 1996.
- Wilby, R.L., and T.M.L. Wigley, Downscaling general circulation model output: a review of methods and limitations, *Progress Phys. Geography*, **21**: 530–548, 1997.
- Zorita, E., J.P. Hughes, D.P. Lettenmaier, and H. von Storch, Stochastic characterization of regional circulation patterns for climate model diagnosis and estimation of local precipitation, *J. Climate*, **8**: 1023–1042, 1995.

Appendices

A. Statistical properties of run lengths

The method followed here for unconditional simulation of weather variables repeatedly generates successive days in the historical data set. This occurs because the closest neighbour of the most recent generated value (with Euclidean distance 0) is the day in the historical data set that has been resampled. Let this be day τ . Then the probability that day $\tau + 1$ is resampled in the next step equals $p_1 = \left\{ \sum_{i=1}^k 1/i \right\}^{-1}$. This probability does not depend on the actual values of the weather variables.

Let Z_t be the binary variable that takes the value 1 if the successor to the closest neighbour is selected at day t , and 0 if this is not the case. We further specify that the event \mathcal{E} occurs at day t if $Z_t = 0$. The time R between two successive occurrences of \mathcal{E} is known as the recurrence time.

For the distribution of R we obtain:

$$\Pr(R = 1) = \Pr(Z_{t+1} = 0 \mid Z_t = 0) = 1 - p_1 \quad (\text{A1})$$

and for $r > 1$:

$$\Pr(R = r) = \Pr(Z_{t+r} = 0, Z_{t+r-1} = 1, \dots, Z_{t+1} = 1 \mid Z_t = 0) = (1 - p_1) p_1^{r-1} \quad (\text{A2})$$

In the latter case we have a sequence of r consecutive historical days in the simulation run, beginning at day t . $R = 1$ at day t implies that the resampled values for days t and $t + 1$ do not form a pair of consecutive historical days. Equations (A1) and (A2) can be combined into:

$$\Pr(R = r) = q_1 p_1^{r-1}, \quad r = 1, 2, \dots \quad (\text{A3})$$

where $q_1 = 1 - p_1$. This distribution is known as the geometric distribution. The mean of R is given by:

$$E(R) = 1 / q_1 \quad (\text{A4})$$

Let L denote the length of a run of ones in the binary sequence $\{Z_t\}$. A remarkable point is that the geometric distribution in equation (A3) also applies to these run lengths:

$$\Pr(L = k) = \Pr(R = k + 1 \mid R > 1) = \frac{\Pr(R = k + 1)}{\Pr(R > 1)} = \frac{q_1 p_1^k}{p_1} = q_1 p_1^{k-1}, \quad k=1, 2, \dots \quad (\text{A5})$$

Let L_{\max} and R_{\max} be the maximum values of L and R in a sequence of n trials. In general, $R_{\max} = L_{\max} + 1$, except when $Z_t = 0$ for all t or if L_{\max} occurs in the first or the last L_{\max} trials. The probability of these exceptions can, however, be neglected if n is sufficiently large. Therefore:

$$E(R_{\max}) \approx E(L_{\max}) + 1 \quad (\text{A6})$$

An elegant approximation to the distribution of L_{\max} is presented in Feller (1968, pp. 322-326):

$$\Pr(L_{\max} < k) \approx \frac{1 - p_1 x}{q_1 (k + 1 - kx) x^{n+1}} \quad (\text{A7})$$

where

$$x = 1 + q_1 p_1^k + (k+1) q_1^2 p_1^{2k} + (k+1)^2 q_1^3 p_1^{3k} + \dots \quad (\text{A8})$$

Feller demonstrates that this approximation works well, even for very small n . For both k and n sufficiently large, equation (A7) can be simplified as (Cramér, 1946, p. 258):

$$\Pr(L_{\max} < k) \approx \exp(-nq_1 p_1^k) \quad (\text{A9})$$

Equation (A9) can be rewritten as:

$$\Pr(L_{\max} < k) \approx \exp\left\{-\exp\left[-\lambda\left(k - \frac{1}{\lambda} \ln(nq_1)\right)\right]\right\} \quad (\text{A10})$$

where $\lambda = -\ln(p_1)$. The right-hand side of equation (A10) represents the distribution function of a Gumbel variable X with location parameter $\xi = \frac{1}{\lambda} \ln(nq_1)$ and scale parameter $\alpha = 1/\lambda$. In contrast to the Gumbel variate, L_{\max} can only take integer values. In fact, the distribution of L_{\max} converges to that of the integer part of X . The mean of L_{\max} can then be approximated as (Gordon *et al.*, 1986):

$$\mathbb{E}(L_{\max}) \approx \mathbb{E}(X) - 1/2 = \frac{1}{\lambda} [\gamma + \ln(nq_1)] - 1/2 \quad (\text{A11})$$

where γ is Euler's constant. The term $1/2$ is a continuity correction. Substitution in equation (A6) finally gives:

$$\mathbb{E}(R_{\max}) \approx \frac{1}{\lambda} [\gamma + \ln(nq_1)] + 1/2 \quad (\text{A12})$$

B. Relative variability of statistics used to compare extreme-value properties

In order to get an impression of the statistical significance of the percentage differences between simulated and observed extreme-value properties (largest value X_{\max} , upper quintile mean QM5 and median M), we determine the coefficient of variation CV of these properties. We assume that the N -day maximum precipitation amount X in a winter half-year follows the GEV distribution in equation (14) and that the maxima for different winters are stochastically independent. The largest value X_{\max} in a sequence of J years is then also a GEV variable with the same shape parameter κ and scale and location parameters (Dales and Reed, 1989):

$$\alpha_J = \alpha J^{-\kappa} \quad (\text{B1})$$

$$\xi_J = \xi + \frac{\alpha}{\kappa} (1 - J^{-\kappa}) \quad (\text{B2})$$

which for $\kappa = 0$ should be interpreted as the limit $\kappa \rightarrow 0$, that is:

$$\alpha_j = \alpha \quad (\text{B3})$$

$$\xi_j = \xi + \alpha \ln J \quad (\text{B4})$$

The CV of X_{\max} can therefore be obtained from the expression for CV of the GEV variable X .

The mean and variance of the GEV variable X are given by (Stedinger *et al.*, 1993):

$$E(X) = \xi + \frac{\alpha}{\kappa} [1 - \Gamma(1 + \kappa)], \quad \kappa > -1 \text{ and } \kappa \neq 0 \quad (\text{B5})$$

$$\text{var}(X) = \frac{\alpha^2}{\kappa^2} [\Gamma(1 + 2\kappa) - \Gamma^2(1 + \kappa)], \quad \kappa > -0.5 \text{ and } \kappa \neq 0 \quad (\text{B6})$$

where $\Gamma(\cdot)$ denotes the gamma function. For $\kappa = 0$, the mean and variance of X can be derived from the above equations by approximating $\Gamma(1 + \kappa)$ for small κ as (Gradshteyn and Ryzhik, 1980, p. 935):

$$\Gamma(1 + \kappa) \approx 1 - \gamma\kappa + \left(\frac{1}{2}\gamma^2 + \frac{1}{12}\pi^2\right)\kappa^2 \quad (\text{B7})$$

where γ is Euler's constant. This results in:

$$E(X) = \xi + \alpha\gamma \quad (\text{B8})$$

$$\text{var } X = \pi^2 \alpha^2 / 6 \quad (\text{B9})$$

The coefficient of variation then becomes:

$$\text{CV} = \frac{(\alpha / \xi) \sqrt{[\Gamma(1 + 2\kappa) - \Gamma^2(1 + \kappa)] / \kappa^2}}{1 + (\alpha / \xi) [1 - \Gamma(1 + \kappa)] / \kappa} \quad (\text{B10})$$

which for $\kappa = 0$ reduces to:

$$\text{CV} = \frac{(\alpha / \xi) \pi / \sqrt{6}}{1 + \gamma \alpha / \xi} \quad (\text{B11})$$

The value of CV is thus determined by κ and α/ξ . For the N -day winter maximum precipitation amounts in this study, a representative value of α/ξ is 0.25 (the average of α/ξ is 0.26 for the 25 stations and 0.23 for the 5 subareas), which results in $\text{CV} = 0.280$ for the Gumbel variable ($\kappa = 0$). Table B1 shows that CV is somewhat larger for the relatively thick-tailed Type II distribution with $\kappa = -0.1$ and is somewhat smaller if $\kappa = 0.1$. The table also presents CV of X_{\max} for $J = 34$ and $J = 999$. For the values of κ considered in Table B1, CV decreases with increasing J because of the increase of $E(X_{\max})$ with increasing J . The fastest decrease is found for the Type III distribution because for this distribution the variance of X_{\max} decreases with increasing J . The observed percentage differences of

	$\kappa = -0.1$	$\kappa = 0$	$\kappa = 0.1$
X	0.318	0.280	0.255
$X_{\max}, J = 34$	0.231	0.158	0.110
$X_{\max}, J = 999$	0.194	0.112	0.062

Table B1: Coefficient of variation (CV) of the GEV variable X and the maximum X_{\max} of J independent GEV variables with $\alpha/\xi = 0.25$ and $\kappa = -0.1, 0$ or 0.1 .

10% to 20% between the historical N -day winter maximum and the average values of X_{\max} in the 35-year simulation runs in Table 12 are in line with the values of CV in Table B1 for $J = 34$ (CV ≈ 0.15 for κ near 0).

The average difference between the largest 10-day precipitation amounts in the $k = 5$ simulation and the $k = 20$ simulation in Figure 14 is about 7%, whereas the coefficient of variation of X_{\max} for $J = 999$ is about 0.10. Despite the relatively large influence of a few mountain stations, it may be expected that the average X_{\max} of 25 stations has a much smaller CV because of the weak dependence in the upper tail of the bivariate distributions of the 10-day maxima (Section 5). The observed difference between the values of X_{\max} in Figure 14 should therefore be regarded as considerable compared with that expected from pure random variation. It is, however, impossible to specify random and systematic effects further. This would require many repetitions of the $k = 5$ and $k = 20$ long-duration simulations.

The CVs of QM5 and M were determined by simulation using the method presented by Lynn and Beran (1979) to generate ordered samples of random numbers from the GEV distribution. Table B2 shows the results. Each value in the table is based on 10 000 repetitions. Samples of size $J = 34$ were considered to judge the statistical significance of the percentage differences between the simulated and observed values of QM5 and M in Section 4. The shape parameter κ has little effect on the CV of the median. The CV of both QM5 and M strongly depends on α/ξ . For the winter maxima in Section 4, representative values of α/ξ are: 0.25 for precipitation, 0.50 for multi-day snowmelt at high-elevation stations (Kahler Asten, Freudenstadt and Kl. Feldberg), and 0.25 - 1.50 for multi-day snowmelt at the other stations. Thus for the precipitation maxima, CV is about 0.09 for QM5 and 0.06 for M . Only for the conditional simulations in Table 12, there are a number of cases where the relative differences are $2 \times$ CV or more (Freudenstadt, area II). The observed relative differences between the multi-day snowmelt maxima at Freudendstadt in Table 13 are less than $2 \times$ CV (CV ≈ 0.12 for QM5 and CV ≈ 0.10 for M).

The largest values of α/ξ (up to about 1.5) occur at lowland stations with little snow. For some of these stations the computed N -day maximum snowmelt is zero in about 10% of the years. The fit of an extreme-value distribution to such data

	α/ξ	$\kappa = -0.1$	$\kappa = 0$	$\kappa = 0.1$
QM5	0.25	0.110	0.085	0.066
	0.50	0.154	0.121	0.097
	0.75	0.177	0.142	0.115
	1.00	0.191	0.155	0.126
	1.25	0.201	0.164	0.135
	1.50	0.209	0.170	0.141
M	0.25	0.057	0.055	0.053
	0.50	0.105	0.101	0.098
	0.75	0.146	0.141	0.136
	1.00	0.181	0.175	0.169
	1.25	0.211	0.204	0.198
	1.50	0.238	0.230	0.224

Table B2: Coefficient of variation (CV) of the upper quintile mean QM5 and the median M in samples of size $J = 34$ for a GEV distribution with $\alpha/\xi = 0.25, 0.50, 0.75, 1.00, 1.25$ or 1.50 and $\kappa = -0.1, 0$ or 0.1.

needs special care. One possibility is to treat zeros as left-censored values. Buisland (1986) fitted the Gumbel distribution to censored samples of annual maximum snow cover depths at lowland stations around the Netherlands using the maximum likelihood method. Wang (1996) discusses the use of partial PWMs to fit extreme-value distributions to censored samples. A notable point in Table B2 is that the CV of M exceeds that of QM5 at large α/ξ . The relatively large percentage differences for the medians at Essen and Frankfurt can partly be explained by the large CV of M for these stations.