Estimation of extreme floods of the River Meuse using a stochastic weather generator and a rainfall–runoff model

ROBERT LEANDER[1], ADRI BUISHAND[1], PAUL AALDERS[2*] & MARCEL DE WIT[3]

1 *Royal Netherlands Meteorological Institute (KNMI), PO Box 201, NL-3730 AE De Bilt, The Netherlands*
leander@knmi.nl

2 *Wageningen University, Sub-department Water Resources, Nieuwe Kanaal 11, NL-6709 PA Wageningen, The Netherlands*

3 *Institute for Inland Water Management and Waste Water Treatment (RIZA), PO Box 9072, NL-6800 ED Arnhem, The Netherlands*

Abstract A stochastic weather generator has been developed to simulate long daily sequences of areal rainfall and station temperature for the Belgian and French sub-basins of the River Meuse. The weather generator is based on the principle of nearest-neighbour resampling. In this method rainfall and temperature data are sampled simultaneously from multiple historical records with replacement such that the temporal and spatial correlations are well preserved. Particular emphasis is given to the use of a small number of long station records in the resampling algorithm. The distribution of the 10-day winter maxima of basin-average rainfall is quite well reproduced. The generated sequences were used as input for hydrological simulations with the semi-distributed HBV rainfall–runoff model. Though this model is capable of reproducing the flood peaks of December 1993 and January 1995, it tends to underestimate the less extreme daily peak discharges. This underestimation does not show up in the 10-day average discharges. The hydrological simulations with the generated daily rainfall and temperature data reproduce the distribution of the winter maxima of the 10-day average discharges well. Resampling based on long station records leads to lower rainfall and discharge extremes than resampling from the data over a shorter period for which areal rainfall was available.

Key words design discharge; HBV model; nearest-neighbour resampling

Estimation des crues extrêmes de la Meuse à l'aide d'un générateur stochastique de variables météorologiques et d'un modèle pluie–débit

Résumé Un générateur stochastique de variables météorologiques a été développé pour simuler de longues séries journalières de pluies régionales et de températures stationnelles pour les sous-bassins belges et français de la Meuse. Ce générateur météorologique repose sur le principe de ré-échantillonnage des plus proches voisins. Dans cette méthode, les données de pluie et de température sont échantillonnées simultanément à partir de multiples séries historiques, avec remplacements de sorte que les corrélations temporelles et spatiales soient bien preservées. Une attention particulière est portée sur l'utilisation dans l'algorithme de ré-échantillonnage d'un petit nombre de longues séries stationnelles. La distribution du maximum hivernal sur 10 jours de la pluie moyenne sur un bassin est assez bien reproduite. Les séries générées ont ensuite été utilisées comme entrées du modèle pluie–débit semi-distribué HBV pour des simulations hydrologiques. Bien que ce modèle puisse reproduire les débits de pointe des crues de décembre 1993 et de janvier 1995, il a tendance à sous-estimer les débits de pointe journaliers moins extrêmes. Cette sous-estimation n'apparaît pas pour les débits moyens sur 10 jours. Les simulations hydrologiques issues des données générées de pluie et de température journalières reproduisent bien la distribution du maximum hivernal du débit moyen sur 10 jours. Les extrêmes de pluie et de débit obtenus sont inférieurs suite à un ré-échantillonnage basé sur de longues séries stationnelles par rapport à ceux obtenus suite à un ré-échantillonnage de données de pluie régionale disponibles pour des périodes plus courtes.

Mots clefs débit de projet; modèle HBV; ré-échantillonnage des plus proches voisins

* *Now at:* De Dommel Water Board, PO Box 10.001, NL-5280 DA Boxtel, The Netherlands.

Open for discussion until 1 June 2006

INTRODUCTION

The Meuse is one of the largest rivers in northwest Europe. Its basin covers an area of more than 30 000 km^2, including parts of France, Belgium, Germany and The Netherlands. Protection against flooding is a matter of continuous concern. In The Netherlands the design discharge for flood protection works is based on the extrapolation of the distribution of historical annual discharge maxima at Borgharen, where the river enters the country. Disadvantages of this method are that strong extrapolation is required, and that discharge records are potentially inhomogeneous. Furthermore, considering annual discharge maxima gives no insight into the shape and duration of the flood peaks. As an alternative, the simulation of long-duration discharge sequences by running a hydrological model with generated sequences of rainfall and temperature has been suggested (Parmet *et al.*, 1999).

A weather generator based on nearest-neighbour resampling has been developed to provide long daily sequences of rainfall and temperature for the Meuse basin. This weather generator is similar to that in Buishand & Brandsma (2001) for the contiguous Rhine basin, except that in this study areal rainfall of the sub-basins is simulated rather than station rainfall. A particular problem is to make use of a number of station records of daily rainfall that were longer than those of the available areal rainfall. It is desirable to take advantage of these long records to reduce the influence of recent years.

In this study the simulation of both extreme multi-day rainfall and river discharges are addressed. The weather generator is described first. Then rainfall–runoff modelling and the simulation of extreme discharges using the generated rainfall and temperature data are discussed. The paper ends with a presentation of the conclusions.

WEATHER GENERATOR

Nearest-neighbour resampling

In nearest-neighbour resampling, daily weather variables are sampled with replacement from the historical data to generate long daily sequences. The most important advantage of nearest-neighbour resampling, compared to other stochastic weather generators, is that it automatically preserves the correlation between variables on the same day. Furthermore, due to the conditioning of new daily values on the preceding days, the autocorrelation of those variables can also be well preserved, though there is a random influence in the selection process. The reproduction of the autocorrelation of daily rainfall is in particular crucial for the Meuse basin, because large river flows are generally induced by persistent rainfall over a multi-day period. The method requires no assumptions concerning statistical distributions and relationships and is completely data-driven.

In the resampling process, both in the generated sequence and the historical data, each day is characterized by a feature vector, which summarizes the weather conditions for the region of interest. The extent to which two days t and u differ is quantified by the weighted squared Euclidean distance $\delta^2(\boldsymbol{D}_u,\boldsymbol{D}_t)$ between their feature vectors $\boldsymbol{D}_u$ and $\boldsymbol{D}_t$, i.e.

$$\delta^2(\boldsymbol{D}_u, \boldsymbol{D}_t) = \sum_{j=1}^{p} w_j (D_{tj} - D_{uj})^2 \qquad (1)$$

where the index j refers to the jth of the p elements of the feature vector and w_j is the weight of this element. In each cycle of the algorithm one of the k closest historical days ("nearest neighbours") of the latest simulated day is selected at random. The weather variables of the historical successor to the selected nearest neighbour are then added to the sequence. Following a suggestion of Lall & Sharma (1996), a decreasing kernel is used to select one of the k nearest neighbours. This kernel gives more weight to closer neighbours. The probability of selecting the jth closest neighbour is given by

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

In order to account for the seasonal variation in the data and to prevent that typical winter and summer days are mixed up in the simulation, the search for nearest neighbours is usually restricted to a "window" of W calendar days, centred on the last simulated day.

A more elaborate discussion on the nearest-neighbour resampling of daily weather variables can be found in Rajagopalan & Lall (1999) and Buishand & Brandsma (2001).

The Meuse basin and available data

Figure 1 shows the Meuse basin upstream of Borgharen (20 830 km^2), which is partly situated in France (45%) and partly in Belgium (55%). The Meuse is dominated by a rainfall–evaporation regime that produces low flows during summer and high flows during winter. For Borgharen a daily discharge record was available beginning in 1911. The record was corrected for the influence of the Albert Canal, Zuid-Willemsvaart and Juliana Canal, branching off just upstream of the gauging station.

Fifteen sub-basins were defined for hydrological modelling. One of these, the Sambre sub-basin, was further subdivided into a French (9F) and a Belgian (9B) part. For all sub-basins, daily areal rainfall was available for the period 1961–1998. The areal rainfall for the Belgian part was based on the Thiessen method and was obtained from the Belgian Meteorological Institute. For the French part, the areal rainfall was calculated from the data of 63 stations using squared inverse distance interpolation on a 5 × 5 km grid.

Apart from the daily areal rainfall amounts of the 15 sub-basins, daily point rainfall data from a number of stations in and around the Meuse basin were available for this study. The locations of these stations are shown in Fig. 1. In Table 1 the average totals are listed for the winter half-year (October–March) and the summer half-year (April–September) for all rainfall stations. The differences between the winter and summer averages are small. The wettest stations in France (Neufchâteau and Le Chesne) and southern Belgium (Chiny) have relatively large rainfall amounts in winter. The average rainfall amounts in Table 1 refer to the period 1961–1998 for

Fig. 1 Locations of stations in and around the Meuse basin and the definition of the sub-basins. Stations providing temperature data are indicated by [T].

Table 1 Average winter and summer rainfall totals for the stations in Fig. 1 for the period 1961–1998.

Country	Station name	Altitude (m)	Winter total (mm)	Summer total (mm)
France	St. Quentin *	95	358	344
	Nancy *	212	374	378
	Vouziers *	96	380	363
	Chaumont *	317	474	432
	Langres	467	455	423
	Neufchâteau	286	515	428
	Le Chesne	174	536	425
Belgium	Uccle *	100	417	405
	Chiny *	299	728	529
	Stavelot	298	572	520
	Rochefort	193	406	422
Germany	Aachen *	202	388	423

* Homogeneous records available from 1930 (Leander & Buishand, 2004).

which areal rainfall is also available. Seven stations have homogeneous records over the longer period 1930–1998. The year 1940 is missing in most of these records.

Daily temperature records were available for the stations Langres (1949–1998), Reims (1949–1998), Uccle (1930–1998) and Aachen (1930–1998). Additional daily temperature data for the period 1968–1998 were obtained for six stations in Belgium and one in The Netherlands (not shown in Fig. 1).

Resampling of sub-basin rainfall amounts and station temperatures

In the current study, nearest-neighbour resampling is applied to jointly simulate daily sequences of areal rainfall for each of the 15 sub-basins in Fig. 1 and daily temperature at 11 locations in the basin. Simulations based on historical data for the period 1961–1998 (hereinafter referred to as Sim61) were performed, as well as simulations based on the period 1930–1998 (Sim30), each with a length of 3000 years. Daily data from temperature and rainfall stations with complete records for the entire base period were used to form the feature vector in the resampling algorithm. For day t this vector consists of three elements:

– The average daily temperature $\widetilde{T}_t$ at two stations (Sim30) or four stations (Sim61). The temperature records were standardized with the calendar-day mean and standard deviation before averaging.
– The average daily rainfall $\widetilde{P}_t$ at seven stations. The rainfall records were standardized with the (seasonally varying) mean wet-day amount.
– The average of standardized rainfall for the four preceding days, i.e. $\frac{1}{4}(\widetilde{P}_{t-1} + \widetilde{P}_{t-2} + \widetilde{P}_{t-3} + \widetilde{P}_{t-4})$.

The third element was included to improve the reproduction of the autocorrelation of daily rainfall and the standard deviation of monthly totals (Harrold *et al.*, 2003). For Sim61, $\widetilde{T}_t$ was derived from the daily temperature at Langres, Reims, Uccle and Aachen, and $\widetilde{P}_t$ from the daily rainfall at Le Chesne, Langres, Rochefort, Stavelot, Uccle, Chiny and Neufchâteau. The selected rainfall stations were judged to be the most representative of the basin because of their location and their mean annual rainfall. In Sim30, $\widetilde{T}_t$ was based on the temperature records from Uccle and Aachen, and $\widetilde{P}_t$ on the seven long-term daily rainfall records indicated in Table 1.

In contrast to the earlier study of nearest-neighbour resampling for the Rhine basin (Buishand & Brandsma, 2001), the areal rainfall of the 15 sub-basins of the historical successor to the selected nearest neighbour was resampled, rather than the daily rainfall amounts for the seven stations used for the feature vector. This is straightforward in the case of Sim61 for which the records of sub-basin rainfall cover the base period. For Sim30 an additional step in the algorithm is needed, because the sub-basin data are incomplete over the period 1930–1960. Whenever a day from before 1961 is selected in Sim30, the closest neighbour of that day is sought among the days in or after 1961 to serve as an alternative from which the sub-basin rainfall is resampled. For this new search, two-dimensional feature vectors were used, containing only $\widetilde{P}_t$ and $\widetilde{T}_t$. Though the areal values in the resampled sequence correspond to days within the period 1961–1998, the sampling of these days is based on station data for the entire period 1930–1998. Thus, if 1930–1960 is relatively dry, the algorithm will resample more intensely from the drier days in 1961–1998.

For temperature a similar additional nearest-neighbour search was performed to obtain temperature data for 11 locations, instead of two (Sim30) or four (Sim61), using data from the seven additional stations for the period 1968–1998. Since temperature plays only a minor role in the hydrological simulation of extreme floods, only the results for rainfall are presented in this paper.

Besides the composition of the feature vector, a few additional settings are required. The number of nearest neighbours k was set equal to 10. Buishand & Brandsma (2001) show that loops could occur in the simulation in which certain historical days are repeatedly sampled if k is not sufficiently large. On the other hand, the reproduction of the autocorrelation of the generated sequences worsens if k is too large. Buishand & Brandsma (2001) obtained good results with $k = 5$, but in a later study for the Rhine basin (Beersma, 2002), $k = 10$ was used to be better protected against loops.

For the width of the moving window, $W = 61$ was used, in accordance with Buishand & Brandsma (2001). A rather broad window can be chosen because the feature vector was formed from standardized rainfall and temperature data. The areal rainfall data were standardized in the same way as the point rainfall data.

The weights w_j in the Euclidean distance were determined "globally" for each of the feature vector elements as the inverse of their sample variance over the entire series, resulting in a constant set of weights for all days in the year. The use of "local" (i.e. seasonally varying) weights was also considered. This had little effect because the seasonal variation of the weights turned out to be small, due to the standardization of the individual station records in an earlier stage.

Summary of the nearest-neighbour resampling algorithm

The nearest-neighbour resampling algorithm can briefly be summarized as follows:

1. Randomly select a historical day within the moving window centred on 1 January as the first simulated day.
2. Compose a feature vector of average standardized station rainfall and temperature for the latest simulated day.
3. Find the k nearest neighbours of the latest simulated day within a W-day window centred on this day.
4. Select one of these nearest neighbours at random, using the decreasing kernel, given by equation (2). Denote the date of the selected nearest neighbour by i and that of its historical successor by $i + 1$.
5. For each of the simulated variables (areal rainfall or station temperature), check whether data for this variable exist for day $i + 1$.
   – If so, add the standardized historical data to the generated sequence for this variable.
   – If not, form a feature vector of the standardized station data for day $i + 1$, find the nearest neighbour of day $i + 1$ among the days for which the data for the considered variable *do* exist and add the standardized data of *this* nearest neighbour to the generated sequence. The search is restricted to a W-day window centred on day $i + 1$.
6. Repeat steps 2–5 for each simulated day.
7. Transform the resampled standardized variables back to their original scale.

Analysis of generated areal rainfall sequences

The analysis of the generated areal rainfall sequences was focused on the reproduction of the autocorrelation of daily rainfall and the extreme-value distribution of the 10-day rainfall amounts. Large multi-day rainfall amounts in the winter half-year are known to induce high discharges. Figure 2 shows the basin-average autocorrelation coefficients (i.e. an area-weighted average of the coefficients for each individual sub-basin) in the winter half-year for the historical records as well as two 3000-year simulations based on the data for 1961–1998 (Sim61), one with and one without the use of the 4-day memory element in the feature vector. It is seen that the memory element enhances the third- and higher-order autocorrelation coefficients. This leads to a reduction of the bias in the standard deviation of the monthly totals (in winter) from –7.4% to –1.6%.

Figure 3 compares the 10-day winter maxima of basin-average rainfall from the historical data and two 3000-year simulations (one Sim30 and one Sim61). There is a good agreement between the plot of Sim61 and that of the historical maxima. However, the two largest historical extremes lie clearly above the plot of Sim61. These extremes correspond to the Meuse floods of December 1993 and January 1995 (Parmet & Burgdorfer, 1995; Van Meijgaard, 1995; Van Meijgaard & Jilderda, 1996). To investigate whether the events of 1993 and 1995 deviate significantly from the

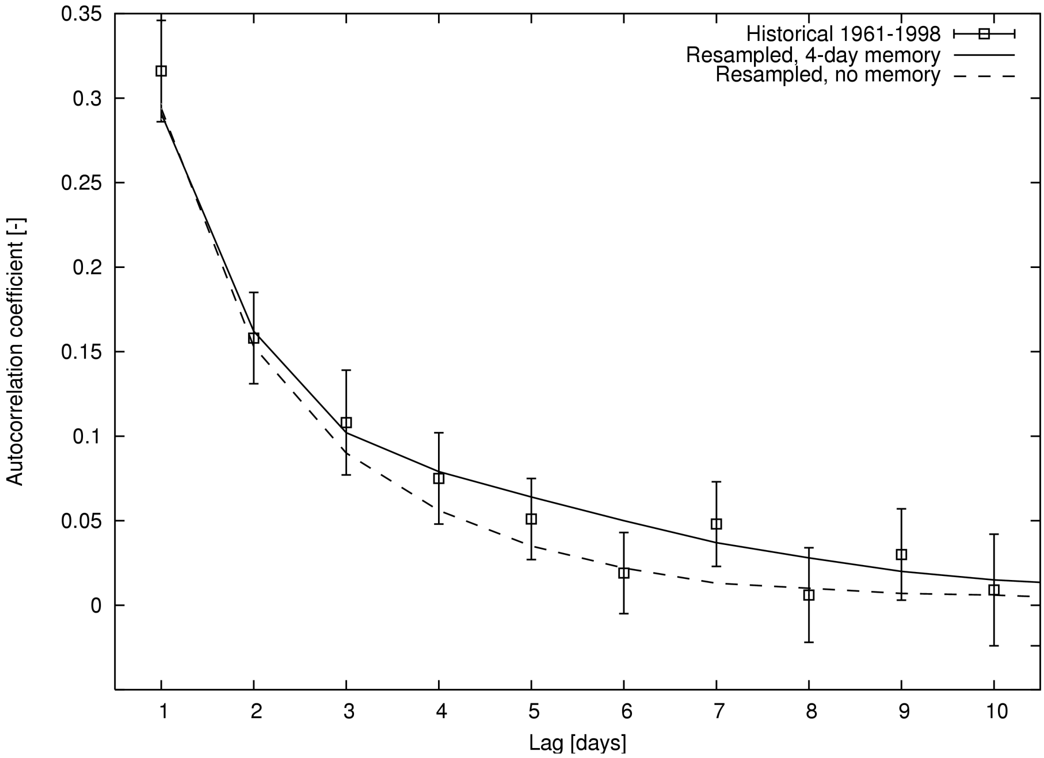

Fig. 2 Basin-average autocorrelation coefficients for the winter half-year of one 3000-year simulation with and one without the 4-day memory element in the feature vector, compared with those of the historical data. Both simulations are based on daily data for the period 1961–1998 (Sim61). The bars correspond to the 2 × *se* intervals. The standard errors *se* were calculated using a jackknife technique (Buishand & Beersma, 1993).

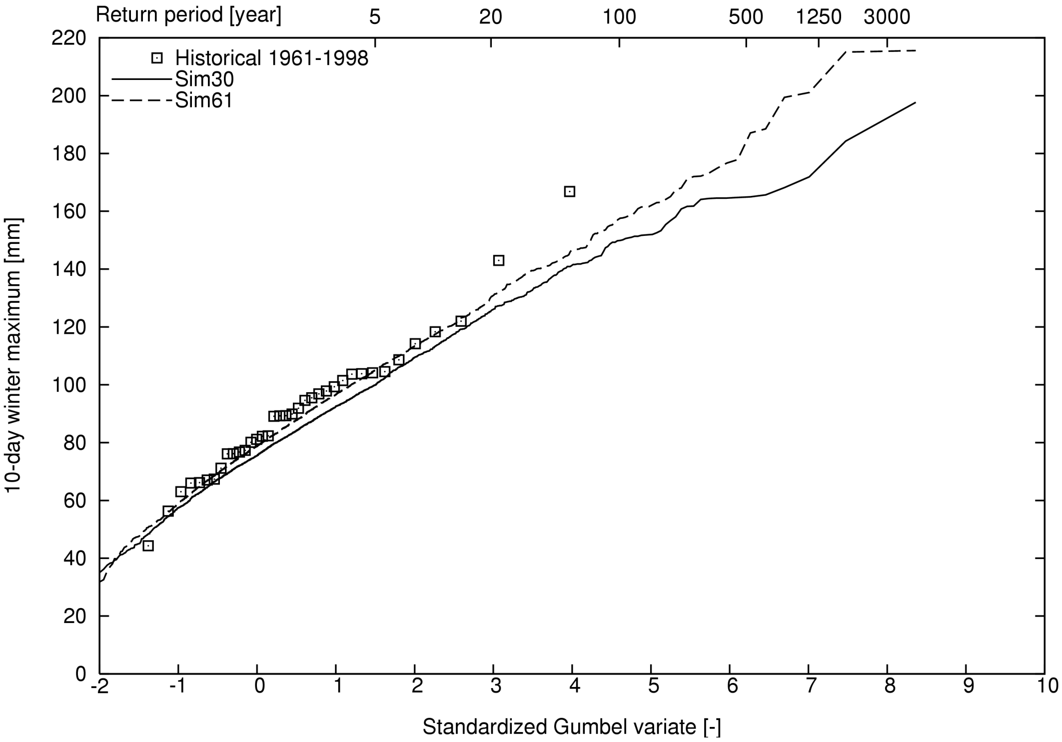

Fig. 3 Winter maxima of 10-day basin-average rainfall from 15 sub-basin records for the period 1961–1998 and from the simulations Sim30 and Sim61.

simulated maxima, Sim61 was split into 78 segments of the same length as the historical record (38 years). It turned out that the maximum of 164 mm in 1995 was exceeded by about 20% of the 38-year segments of Sim61. The observed deviation between the plots of the simulated and historical maxima is therefore not considered to be significant. The plot of Sim30 is somewhat below that of Sim61. This can be ascribed to the fact that the winter half-year is on average drier for the period 1930–1960 than for the years 1961–1998.

The effect of the additional nearest-neighbour step was assessed with a modified version of Sim61, Sim61M. This simulation was also based on the period 1961–1998, but the historical areal rainfall data for the period 1961–1979 were discarded. The additional nearest-neighbour step was used to obtain areal rainfall when a day before 1980 was sampled. Figure 4 compares the Gumbel plot of the 10-day winter maxima of Sim61M with those of:

– four runs of Sim61 with different random number seeds, and
– a simulation based on the period 1980–1998 (Sim80).

It can be seen that the plot of Sim80 does not fall within the spread of those of the four Sim61 runs. On the other hand, Sim61M cannot clearly be distinguished from these four runs, even though the areal rainfall data of this simulation are sampled from the same period as those of Sim80. This indicates that the areal rainfall of Sim61M is representative of the period 1961–1998, just like the areal rainfall of Sim61.

Figure 5 compares the 10-day rainfall extremes for the Semois sub-basin (no. 5). The correspondence between the simulated and historical maxima is similar to that in Fig. 3. From this it can be concluded that the method described here also works for a

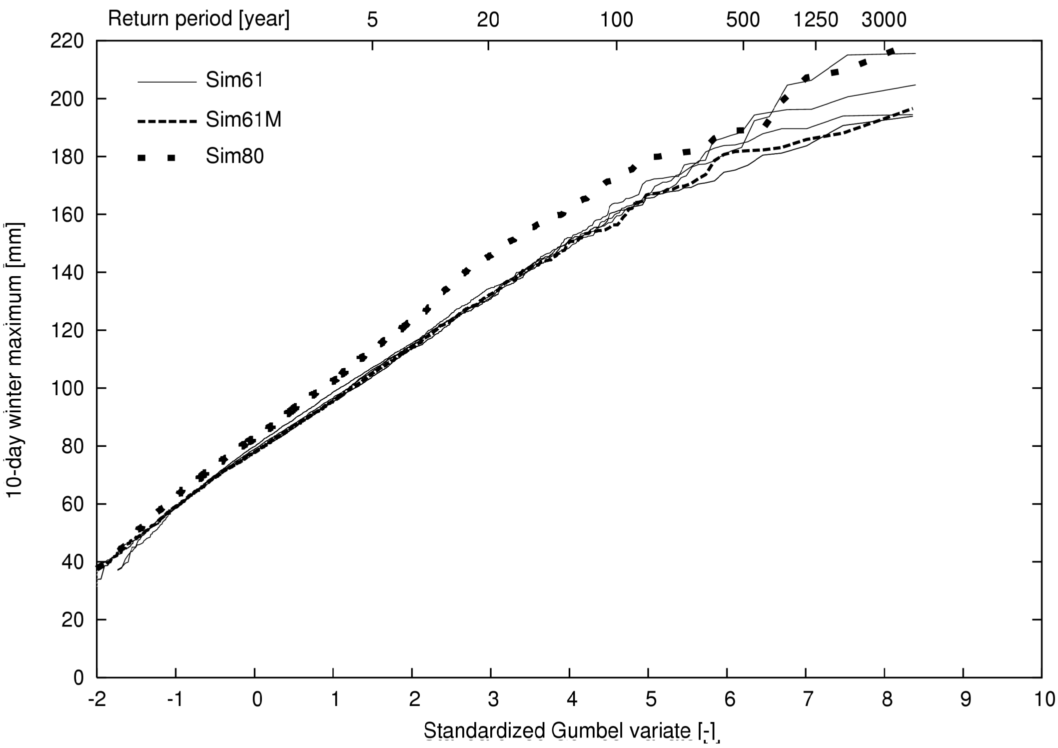

Fig. 4 Winter maxima of 10-day basin-average rainfall for Sim61M (dashed), Sim80 (big dots) and four runs of Sim61 with different random number seeds.

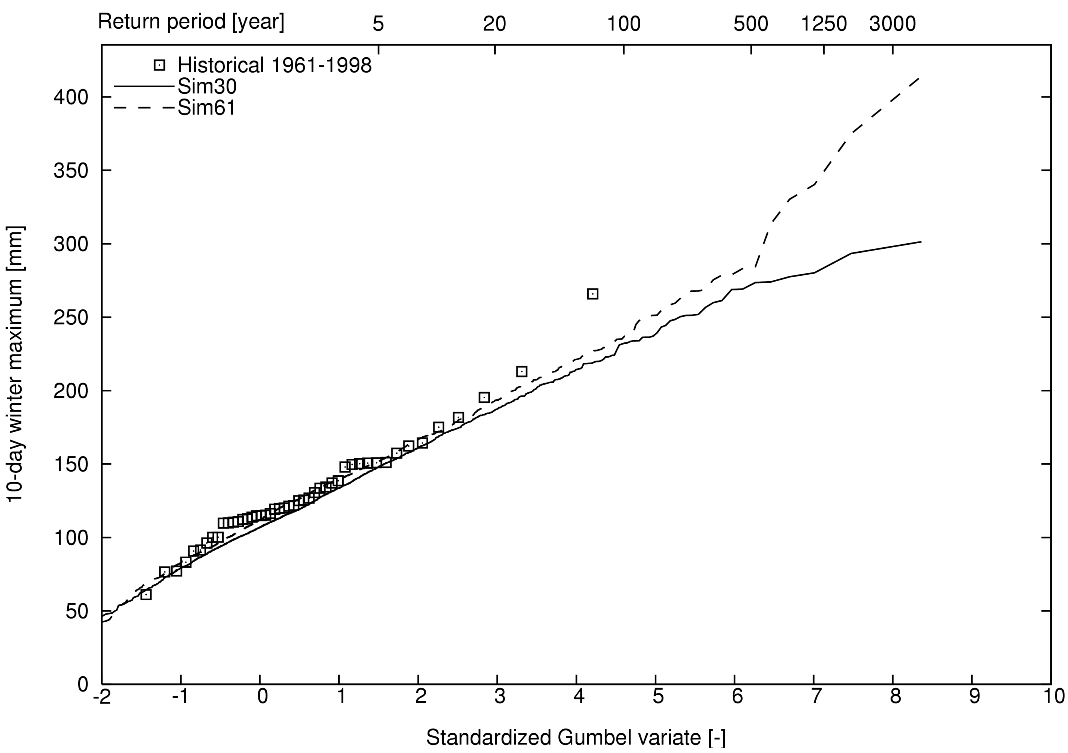

Fig. 5 Winter maxima of 10-day rainfall for the Semois sub-basin from the historical record for the period 1961–1998 and from Sim30 and Sim61.

smaller area, even though only basin-average weather characteristics were used in the resampling algorithm.

RAINFALL–RUNOFF MODEL

In this study the HBV model has been used for rainfall–runoff modelling of each sub-basin. HBV is a conceptual model, developed at the Swedish Meteorological and Hydrological Institute (Lindström *et al.*, 1997). The calculation within HBV is organized into several "routines". The "snow routine" represents snow accumulation and snowmelt; the "soil moisture routine" controls which part of rainfall and melt water forms excess water and how much is evaporated or stored in the soil; and the "runoff generation routine" consists of an upper, nonlinear reservoir representing fast runoff components and a lower, linear reservoir representing base flow. Flood routing processes are simulated with a simplified Muskingum approach.

Rainfall, temperature and potential evapotranspiration (*PET*) are required as input for the model. The temperature for each sub-basin was set equal to the average of the four nearest stations using an altitude correction of –6°C km^{-1}. For the Belgian sub-basins, sequences of daily *PET* have been made available for the period 1967–1998 by the Belgian Meteorological Institute. For the calibration of HBV in the French part of the basin, *PET* was set equal to the average *PET* of the Belgian part. In the 3000-year simulations *PET* was obtained from the simulated daily temperature *T* as:

$$PET = [1 + \alpha(T - \overline{T})]\ \overline{PET} \tag{3}$$

with $\overline{T}$ (°C) and $\overline{PET}$ (mm day^{-1}) being the mean daily temperature and mean monthly *PET* for the period 1967–1998 and α a constant factor. Van der Wal (2002) found for α a value of 0.17°C^{-1}.

Calibration and validation

The original calibration of HBV for the Meuse was carried out by Booij (2005). He observed that the most influential parameters were three parameters in the soil moisture routine and three parameters in the fast flow routine. These parameters were optimized for the Lesse, the Ourthe, the Amblève and the Vesdre (respectively sub-basins 8, 12, 13 and 14 in Fig. 1). For the other sub-basins, for which he had no discharge data, the values of these parameters were derived from river basin characteristics, such as slope, area and soil porosity. Default values were taken for the remaining parameters. In the framework of this study, recalibration took place with more detailed meteorological input and additional discharge records for the Semois (sub-basin 5) and a gauging station along the main branch of the river at the French–Belgian border (Chooz, Fig. 1). For the discharge at Borgharen this resulted in a Nash-Sutcliffe efficiency of 0.91 for the calibration period 1969–1984 and 0.93 for the validation period 1985–1998.

Table 2 shows that the mean, the standard deviation and the maximum of the daily discharges from the HBV simulation resemble those of the observed data. However,

Table 2 Mean, maximum, standard deviation and the mean of the annual maxima of the simulated and observed daily discharges for the Meuse (period 1968–1998).

	Mean ($m^3 s^{-1}$)	Maximum ($m^3 s^{-1}$)	Standard deviation ($m^3 s^{-1}$)	Mean annual max ($m^3 s^{-1}$)
Recorded discharge	267	3080	269	1474
HBV simulation	277	2976	286	1288

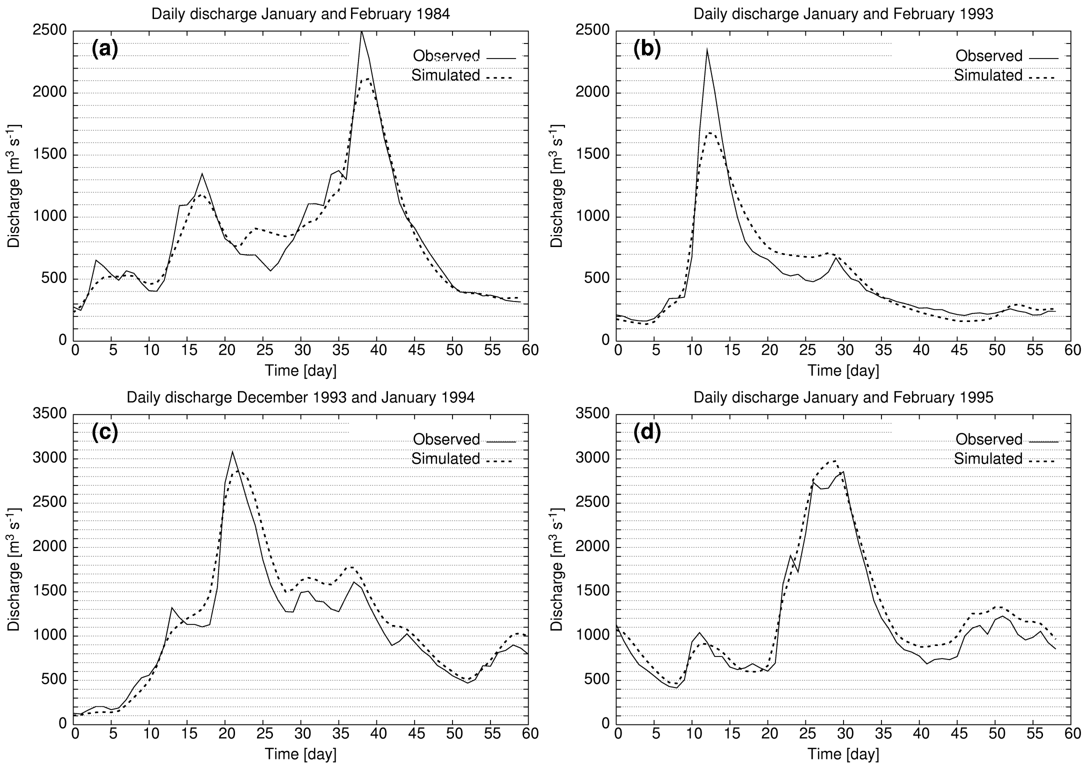

Fig. 6 Observed (solid) and simulated (dashed) daily discharge at Borgharen around the winter maxima of (a) 1984, (b) 1993, (c) 1994 and (d) 1995.

the mean of the annual maximum discharges is underestimated considerably. An underestimation of the mean annual maximum discharge is in line with the results in Eberle *et al*. (2002) for the major sub-basins of the River Rhine. Figure 6 compares the observed and simulated discharges for four historical extreme events. The two highest peaks, in December 1993 (panel (c)) and January 1995 (panel (d)), are well reproduced by the HBV model. The peaks of February 1984 (panel (a)) and January 1993 (panel (b)) are too low in the simulation. The simulated hydrograph is smoother than the observed hydrograph, in particular for January 1993. The total volume of this event, however, is well preserved. The underestimation of the February 1984 peak is partly due to an inappropriate stage–discharge rating curve, which was used during 1984–1987 (Parmet *et al*., 2001). The operation of weirs, sluices and small reservoirs up-stream of Borgharen may be another source of bias in peak flows. The regulation of those reservoirs is not included in the HBV modelling.

Simulated extreme discharges

Two 3000-year generated sequences (one Sim30 and one Sim61) of areal rainfall and station temperature were used to simulate the daily discharge at Borgharen with the HBV model. Figure 7 shows the highest peak discharge for Sim61 and the corresponding basin-average rainfall amounts, illustrating that high discharge is the result of long periods of persistent rainfall, rather than large rainfall amounts on a single day. This is related to the integrating effect of a large river basin.

Figure 8 shows a Gumbel plot of the winter maxima of the observed daily discharge record at Borgharen, the simulated discharge based on historical meteorological data and the simulated discharge based on Sim30 and Sim61. The HBV simulation using the Sim61 data shows a good correspondence with the simulation using historical meteorological data. However, the simulated maxima for December 1993 and January 1995 are far above the plot of the extremes from the Sim61 data. This deviation is more apparent for the discharges than for the 10-day rainfall amounts shown in Fig. 3. It is rather accidental that two such floods occurred in the relatively short period 1968–1998. These floods are among the three largest daily discharges in the 93-year record. There is only one flood of comparable magnitude within the period 1911–1967 (January 1926). The difference between the results for the two 3000-year simulations is small. As expected, the plot of Sim30 is somewhat below that of Sim61. A point of concern is the systematic difference between the plot of the observed discharges and the plots of the simulations at short and moderate return periods due to a systematic underestimation of the flood peaks by the HBV model in this frequency range. This systematic difference is not apparent in the extreme 10-day average discharges, as shown in Fig. 9.

The largest historical 10-day average discharge occurred in 1995. Although the peak discharge of the 1995 event was somewhat below that of the 1993 event, its volume was much larger, due to the relatively long duration of discharges exceeding 2500 $m^3 s^{-1}$. The attenuation of the flood wave was therefore less in 1995, which resulted in higher water levels downstream of Borgharen. It should be noted that, in the 3000-year simulations, much larger 10-day average discharges are found than the historical 1995 maximum.

It is unclear how well HBV can describe flood peaks outside the range for which it was calibrated. The model does not consider the possibility of inundations upstream of Borgharen, which may limit the amount of water that can reach The Netherlands.

CONCLUSION

A stochastic weather generator for the Meuse basin upstream of Borgharen based on nearest-neighbour resampling has been developed. Daily sequences of areal rainfall for 15 sub-basins and station temperatures were simultaneously generated using nearest-neighbour resampling. Several 3000-year simulations were performed, driven by the historical rainfall and temperature for the periods 1961–1998 and 1930–1998. An additional nearest-neighbour step was applied to resample from records which did not completely cover the base period. The generated rainfall and temperature sequences were used to perform 3000-year simulations of the daily discharge at Borgharen with the HBV rainfall–runoff model.

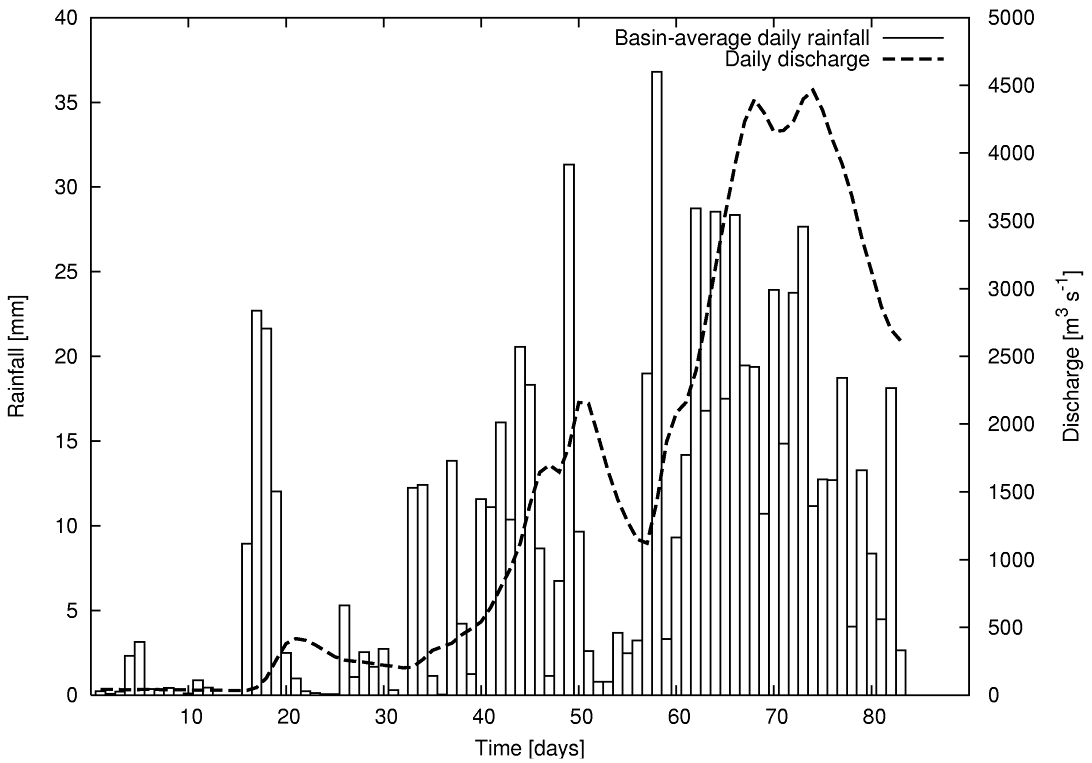

Fig. 7 The daily discharge (dashed curve) around the highest discharge maximum (4464 $m^3 s^{-1}$) in the HBV simulation driven by Sim61 and the corresponding daily rainfall (bars).

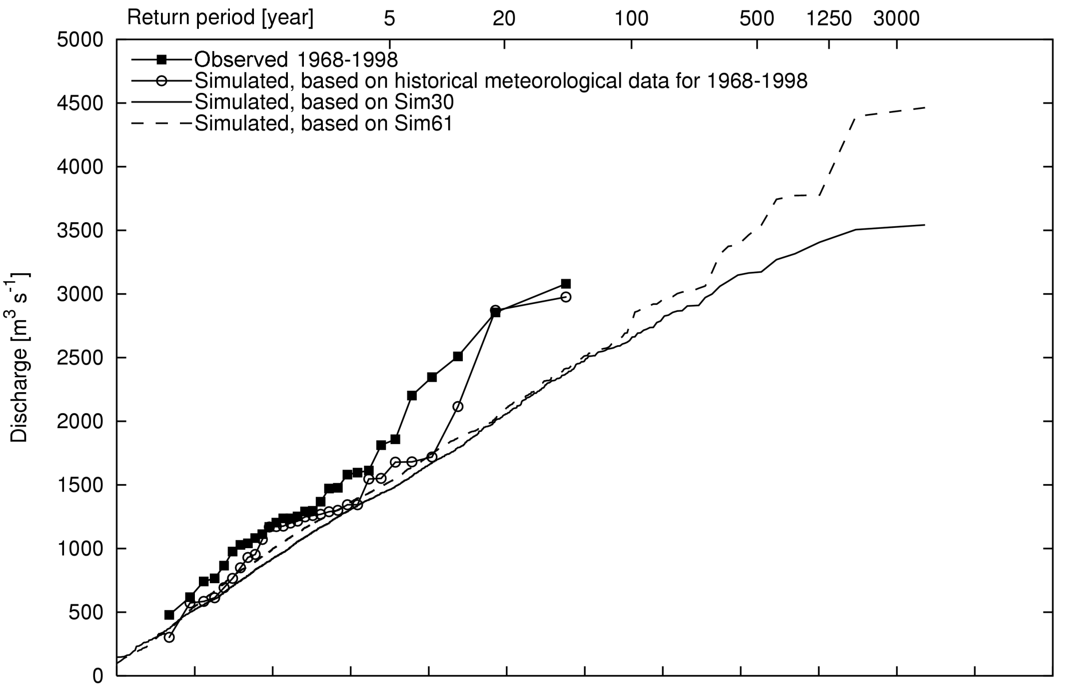

Fig. 8 Winter maxima of the observed daily discharge at Borgharen (1968–1998) and the simulated daily discharge based on the historical meteorological data for 1968–1998, Sim30 and Sim61.

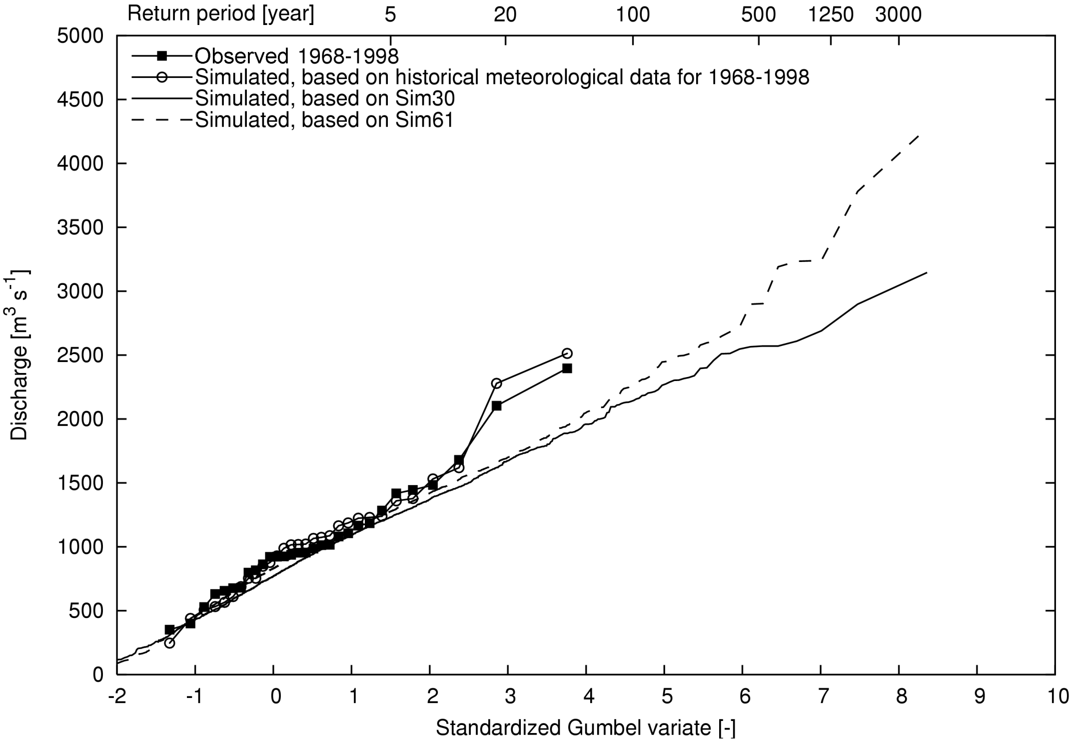

Fig. 9 Winter maxima of the observed 10-day average discharge at Borgharen and those simulated using respectively historical meteorological data or data from Sim30 and Sim61.

The weather generator reproduces the distribution of the extreme 10-day rainfall quite well both for individual sub-basins and the entire basin. The choice of base period 1930–1998 leads to somewhat lower extreme 10-day winter rainfall because of the drier winters that occurred during 1930–1960. The distributions of the discharge winter maxima from the 3000-year simulations resemble those from the HBV simulations with historical meteorological data. The influence of the base period for resampling turns out to be small. For the daily discharges there are significant differences between the extreme-value distributions from the observed and simulated data because of the tendency of HBV to underestimate flood peaks. Improvement requires more detailed modelling (a finer temporal resolution, inclusion of the reservoir regulation and coupling with a hydraulic model for flood routing) for which not all required data are readily available. However, the major floods of December 1993 and January 1995 are already adequately reproduced by the current HBV model. Moreover, the discrepancies between the observed and simulated extreme-value distributions disappear if the 10-day average discharges are considered. This indicates that the weather generator is able to provide reliable estimates of the volumes and durations of extreme floods.

Acknowledgements The station records and sub-basin data for the Belgian part of the Meuse basin were kindly provided by the Royal Meteorological Institute of Belgium (RMIB). The French station data were made available by Météo France. The Belgian discharge data were provided by MET-Sethy. The paper benefited from the critical comments of Dr A. Sharma and an anonymous reviewer on an earlier version. The authors thank Willem van Deursen and Martijn Booij for the calibration of HBV. Furthermore, the authors are grateful to Paul Torfs, Piet Warmerdam, Leonie Bolwidt, Hendrik Buiteveld and Wout van Vuuren for their support.

REFERENCES

Beersma, J. (2002) Rainfall generator for the Rhine basin: description of 1000-year simulations. KNMI Publ. 186-V, De Bilt, The Netherlands.

Booij, M. J. (2005) Impact of climate change on river flooding assessed with different spatial model resolutions. *J. Hydrol.* **303**, 176–198.

Buishand, T. A. & Beersma, J. J. (1993) Jackknife tests for differences in autocorrelation between climate time series. *J. Climate* **6**(12), 2490–2495.

Buishand, T. A. & Brandsma, T. (2001) Multisite simulation of daily precipitation and temperature in the Rhine basin by nearest-neighbor resampling. *Water Resour. Res*. **37**(11), 2761–2776.

Eberle, M., Buiteveld, H., Beersma, J. J., Krahe, P. & Wilke, K. (2002) Estimation of extreme floods in the River Rhine basin by combining precipitation–runoff modelling and a rainfall generator. In: *Proc. Int. Conf. on Flood Estimation* (Berne, Switzerland, 6–8 March 2002) (ed. by M. Spreafico & R. Weingartner), 459–467. CHR Report II-17, CHR Secretariat, Lelystad, The Netherlands.

Harrold, T. I., Sharma, A. & Sheather, S. J. (2003) A nonparametric model for stochastic generation of daily rainfall occurrence. *Water Resour. Res*. **39**(10), 1300, DOI 10.1029/2003WR002182.

Lall, U. & Sharma, A. (1996) A nearest neighbor bootstrap for resampling hydrologic time series. *Water Resour. Res*. **32**(3), 679–693.

Leander, R. & Buishand, T. A. (2004) Rainfall generator for the Meuse basin: inventory and homogeneity analysis of long daily precipitation records. KNMI Publ. 196-II, De Bilt, The Netherlands.

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

Parmet, B. W. A. H. & Burgdorfer, M. (1995) Extreme discharges of the Meuse in The Netherlands: 1993, 1995 and 2100—operational forecasting and long-term expectations. *Phys. Chem. Earth* **20**(5/6), 485–489.

Parmet, B. W. A. H., Buishand, T. A., Brandsma, T. & Mülders, R. (1999) Design discharge of the large rivers in The Netherlands—towards a new methodology. In: *Hydrological Extremes: Understanding, Predicting, Mitigating* (ed. by L. Gottschalk, J.-C. Olivry, D. Reed & D. Rosbjerg), 269–272. IAHS Publ. 255, IAHS Press, Wallingford, UK.

Parmet, B. W. A. H., van de Langemheen, W., Chbab, E. H., Kwadijk, J. C. J., Lorenz, N. N. & Klopstra, D. (2001) Analyse van de maatgevende afvoer van de Maas te Borgharen (Analysis of the design discharge of the River Meuse at Borgharen). RIZA Report 2002.013, Arnhem, The Netherlands.

Rajagopalan, B. & Lall, U. (1999) A *k*-nearest-neighbor simulator for daily precipitation and other weather variables. *Water Resour. Res*. **35**(10), 3089–3101.

Van der Wal, K. (2002) Neerslag-afvoermodellering met gegenereerde neerslagdata in het HBV-model (Rainfall–runoff modelling using generated precipitation data in the HBV model). Civil Engineering & Management Research Report 2002W-002/WMW-005, University of Twente, Enschede, The Netherlands.

Van Meijgaard, E. (1995) Excessive rainfall over the Belgian Ardennes in December 1993: evaluation of model predictions. *Meteorol. Applic.* **2**(1), 39–52.

Van Meijgaard, E. & Jilderda, R. (1996) The Meuse flood in January 1995. *Weather* **51**(2), 39–45.

Received 18 March 2005; accepted 2 September 2005