

Comparison of the WGEN and LARS-WG stochastic weather generators for diverse climates

Mikhail A. Semenov^{1,*}, Roger J. Brooks¹, Elaine M. Barrow², Clarence W. Richardson³

¹IACR-Long Ashton Research Station, Department of Agricultural Sciences, University of Bristol, Bristol BS41 9AF, United Kingdom

²Climatic Research Unit, University of East Anglia, Norwich NR4 7TJ, United Kingdom

³Grassland, Soil and Water Research Laboratory, USDA/ARC, Temple, Texas 76502, USA

ABSTRACT: Stochastic weather generators are used in a wide range of studies, such as hydrological applications, environmental management and agricultural risk assessments. Such studies often require long series of daily weather data for risk assessment and weather generators can produce time series of synthetic daily weather data of any length. Weather generators are also used to interpolate observed data to produce synthetic weather data at new sites, and they have recently been employed in the construction of climate change scenarios. Any generator should be tested to ensure that the data that it produces is satisfactory for the purposes for which it is to be used. The accuracy required will depend on the application of the data, and the performance of the generator may vary considerably for different climates. The aim of this paper is to test and compare 2 commonly-used weather generators, namely WGEN and LARS-WG, at 18 sites in the USA, Europe and Asia, chosen to represent a range of climates. Statistical tests were selected to compare a variety of different weather characteristics of the observed and synthetic weather data such as, for example, the lengths of wet and dry series, the distribution of precipitation and the lengths of frost spells. The LARS-WG generator used more complex distributions for weather variables and tended to match the observed data more closely than WGEN, although there are certain characteristics of the data that neither generator reproduced accurately. The implications for the development and use of stochastic weather generators are discussed.

KEY WORDS: Weather generators · Model validation · Model comparison · Climatic diversity

1. INTRODUCTION

A stochastic weather generator is a numerical model which produces synthetic daily time series of a suite of climate variables, such as precipitation, temperature and solar radiation, with certain statistical properties (Richardson 1981, Richardson & Wright 1984, Racsco et al. 1991). There are several reasons for the development of stochastic weather generators and for the use of synthetic weather data instead of observed. The first one is the provision of weather data time series long enough to be used in an assessment of risk in hydrological or agricultural applications. Observed daily weather is one of the major inputs into mathematical models in hydrology, agriculture and land use, but the length of the time series is often insufficient to allow a

good estimation of the probability of extreme events. Moreover, observed time series represent only one 'realisation' of the climate, whereas a weather generator can simulate many 'realisations', thus providing a wider range of feasible situations. The second reason is to provide the means for extending the simulation of weather to locations where observed weather data is not available. This can be achieved by interpolating the parameters of a weather generator between sites using an interpolation technique such as kriging or thin-plate smoothing splines (Hutchinson 1995). A third area of application has recently arisen from climate change studies. The output from Global Climate Models (GCMs), which are the main tools for predicting the evolution of climate on Earth, cannot be used directly at a site because of their very coarse spatial resolution (even in a high resolution GCM, one grid box represents an area of greater than 50 000 km²). A

*E-mail: mikhail.semenov@bbsrc.ac.uk

weather generator can serve as a computationally inexpensive tool to produce site-specific climate change scenarios at the daily time-step. The changes in both climatic means and climate variability predicted by the GCM experiments can be applied to the parameters derived by the weather generator for the

current climate at the site in question. Daily scenario data can then be obtained by running the weather generator using this revised set of parameters (Wilks 1992, Semenov & Barrow 1997).

Weather generators are now widely used by researchers from many different backgrounds in conjunc-

Table 1. Comparison of the procedures used in WGEN and LARS-WG to generate each weather variable

Weather variable	WGEN	LARS-WG
Precipitation status		
Definition of wet day	Precipitation >0 mm.	Precipitation >0 mm.
Determination of precipitation status for a given day	Transition probabilities of a first-order 2-state Markov chain applied to the previous day's status. Separate probabilities are calculated for each month.	Lengths of alternate wet and dry sequences chosen from a semi-empirical distribution fitted to the observed series. Separate parameters are calculated for each month.
Precipitation		
Daily distribution	Two-parameter gamma distribution.	Semi-empirical distribution.
Parameters	Separate parameters are calculated for each month.	Separate parameters are calculated for each month.
Correlation	None.	None.
Minimum temperature		
Daily distribution	Normal distribution.	Normal distribution.
Parameters	The mean and standard deviation of the normal vary daily. These parameters are obtained by fitting cosine functions to the means and coefficient of variations of the observed data throughout the year (grouped into 28 d periods).	The mean and standard deviation of the normal vary daily. These parameters are obtained by fitting Fourier series to the means and standard deviations of the observed data throughout the year (grouped into months).
Conditioned on precipitation status?	No.	Yes. Separate Fourier series are fitted for wet and dry days.
Correlation	Constant lag 1 auto-correlation. Constant cross-correlation between maximum temperature, minimum temperature and radiation.	Constant lag 1 auto-correlation. Pre-set cross-correlation between maximum and minimum temperature.
Maximum temperature		
	Same procedure as for minimum temperature except that maximum temperature is conditioned on precipitation status with separate cosine functions for dry and wet days.	Same procedure as for minimum temperature.
Radiation		
Daily distribution	Normal distribution.	Semi-empirical distribution.
Parameters	The mean and standard deviation of the normal vary daily. These parameters are obtained by fitting cosine functions to the means and coefficient of variations of the observed data throughout the year (grouped into 28 d periods).	Separate parameters are calculated for each month.
Conditioned on precipitation status?	Yes. Separate cosine functions for wet and dry days.	Yes. Separate parameters are calculated for wet and dry days for each month.
Correlation	Constant lag 1 auto-correlation. Constant cross-correlation between maximum temperature, minimum temperature and radiation.	Constant lag 1 auto-correlation.

tion with their impact models. They are becoming a standard component of decision support systems in agriculture, hydrology and environmental management. There is a danger, however, that generators will be used 'as supplied', i.e. without sufficient validation being carried out for the sites at which they are applied. Most weather generators have been tested intensively, but usually only for one country or one region.

The purpose of the work described here was to test and compare 2 weather generators, WGEN, developed in the USA, and LARS-WG, developed in Europe, for a diverse range of climates. The following section describes the structure of the 2 generators, and the choice of sites at which they were tested is explained in Section 3. The statistical comparison tests and their results are set out and discussed in Section 4 and finally Section 5 considers the implications for the future development and use of stochastic weather generators.

2. DESCRIPTION OF GENERATORS USED IN THE COMPARISON

The 2 weather generators being compared here, WGEN and LARS-WG, are used by many research teams world-wide. Both generators produce daily values of minimum and maximum temperature, precipitation and solar radiation. The generators work in a similar way by analysing certain statistical properties of the input observed daily weather data for the chosen site and then by using these properties, along with a pseudo-random number generator, to produce simulated weather data one day at a time.

Both generators specify daily probability distributions for each weather variable and also certain statistical relationships between the variables. The input observed weather data is used to determine the parameters of the probability distributions and the correlation coefficients between the variables. The parameters for each distribution are varied either on a monthly or a daily basis to express the seasonal component of each weather variable. Both temperature and radiation are related to the amount of cloud cover. For this reason, as is the case with most weather generators, both WGEN and LARS-WG use separate temperature and solar radiation distributions for wet and dry days, and the daily generation procedure in each case, therefore, starts by determining the precipitation status of the day. This governs which parameters are to be used for that day. The synthetic data for the day is generated by choosing a random value from the specified distribution, and by applying the correlation coefficients. Table 1 summarises the way in which the generators model each weather variable and further details of the models are set out below.

2.1. WGEN model

The publicly available version of WGEN as described in Richardson & Wright (1984) was used. This is based on the model presented in Richardson (1981), but contains the following amendments and simplifying assumptions:

(1) Daily precipitation is modelled by a 2-parameter gamma distribution, which tends to match observed data significantly better than the simple exponential distribution used in the original description.

(2) The precipitation parameters for the Markov chain and the gamma distribution are constant for a given month, but vary between months. In Richardson (1981) they were interpolated by a Fourier series and varied continuously during the year on a day-to-day basis.

(3) The means and coefficients of variation for temperature and solar radiation are interpolated by a cosine curve, which could be considered as a Fourier series of order 1. In Richardson (1981) they were interpolated by a Fourier series with 3 harmonics.

(4) The cosine functions fitted to the mean and coefficient of variation of the temperature and solar radiation values require annual mean, amplitude and phase parameters. The phase parameter actually corresponds to the day with the maximum value plus 14 d (an adjustment that is required due to the way the cosine function is calculated). However, separate parameter values are not calculated for all functions. For example, a single function is used for the minimum temperature mean values, rather than distinguishing between dry and wet days, since the means on these days were not considered to be significantly different (Richardson 1981), and all the temperature functions use a pre-set phase value of 200 d.

(5) The auto-correlation and cross-correlation coefficients between minimum and maximum temperature and solar radiation are assumed to be constant. The values for these are pre-set in the model to the average values calculated using a number of American sites (Richardson & Wright 1984).

When running WGEN 3 further changes were made:

(1) It was found that the average phase parameter for temperature could be very different from the pre-set value of 200 d. It was as high as 211 d for Seville, Spain, and as low as 194 d for Tashkent, Uzbekistan, and so, to improve the performance of the model, the calculated values were used for each individual site.

(2) WGEN behaves very poorly for temperature when mean values for some of the 28 d periods are close to or less than zero. Such values cause the coefficient of variation to have a very large magnitude and, for temperature values less than zero, to be negative. This behaviour prevents WGEN from adequately fit-

ting a cosine curve to the coefficient of variation. This problem was overcome by adding 100 to the observed data and then subtracting 100 from the generated temperature. This problem, although not its cause, had been noted in Hanson et al. (1994), but no corrections had been made to the version of WGEN in the public domain.

(3) The program was also amended to ensure that it could cope with missing values in the observed weather files (coded as -99).

Several other weather generators based on the model described in Richardson (1991) exist. Their performance may therefore differ from the public domain version of WGEN used here. Two of these weather generators, USCLIMATE and CLIMGEN, were recently compared for 6 sites in the USA (Johnson et al. 1996).

2.2. LARS-WG model

LARS-WG is based on the series weather generator described in Racsko et al. (1991) and in Semenov & Barrow (1997). Version 2.6 of LARS-WG, which is implemented in C++ with a full Windows interface¹, was used for this comparison. This version utilises semi-empirical distributions for the lengths of wet and dry day series, daily precipitation and daily solar radiation. The semi-empirical distribution $Emp = \{a_0, a_i, h_i, i = 1, \dots, 10\}$ is a histogram with 10 intervals, $[a_{i-1}, a_i)$, where $a_{i-1} < a_i$, and h_i denotes the number of events from the observed data in the i th interval. Random variables from the semi-empirical distributions are chosen by first selecting one of the intervals (using the proportion of events in each interval as the selection probability), and then selecting a value within that interval from the uniform distribution. Such a distribution is flexible and can approximate a wide variety of shapes by adjusting the intervals $[a_{i-1}, a_i)$. The cost of this flexibility, however, is that the distribution requires 21 parameters to be specified compared with, for example, 3 parameters for the mixed-exponential distribution used in an earlier version of the model to define the dry and wet series (Racsko et al. 1991).

The intervals $[a_{i-1}, a_i)$ are chosen based on the expected properties of the weather data. For solar radiation, the intervals $[a_{i-1}, a_i)$ are equally spaced between the minimum and maximum values of the observed data for the month, whereas, for the lengths of dry and

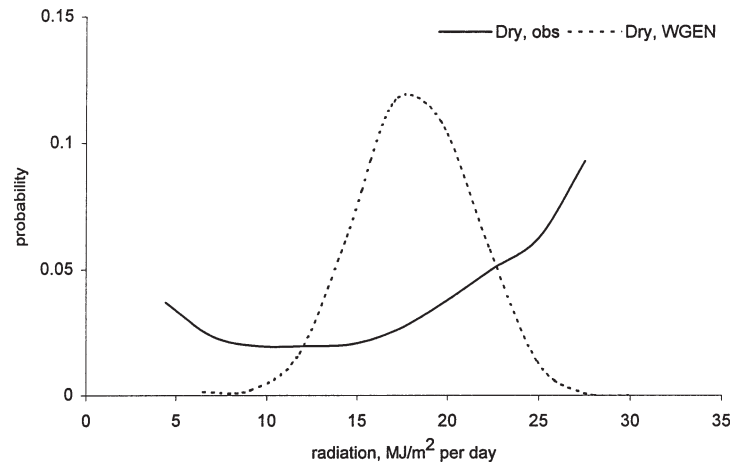


Fig. 1. Probability distribution functions for solar radiation for dry days in June at Rothamsted, for 30 yr of observed weather (solid line) and 300 yr of synthetic data generated by WGEN (dashed line)

wet series and for precipitation, the interval size gradually increases as i increases. In the latter 2 cases, there are typically many small values but also a few very large ones and this choice of interval structure prevents a very coarse resolution being used for the small values.

The simulation of precipitation occurrence is modelled as alternate wet and dry series, where a wet day is defined to be a day with precipitation >0 mm. The length of each series is chosen randomly from the wet or dry semi-empirical distribution for the month in which the series starts. In determining the distributions, observed series are also allocated to the month in which they start. For a wet day, the precipitation value is generated from the semi-empirical precipitation distribution for the particular month independently of the length of the wet series or the amount of precipitation on previous days.

Daily minimum and maximum temperatures are considered as stochastic processes with daily means and daily standard deviations conditioned on the wet or dry status of the day. The technique used to simulate the process is very similar to that presented in Racsko et al. (1991). The seasonal cycles of means and standard deviations are modelled by finite Fourier series of order 3 and the residuals are approximated by a normal distribution. The Fourier series for the mean is fitted to the observed mean values for each month. Before fitting the standard deviation Fourier series, the observed standard deviations for each month are adjusted to give an estimated average daily standard deviation by removing the estimated effect of the changes in the mean within the month. The adjustment is calculated using the fitted Fourier series already obtained for the mean.

¹LARS-WG is available from <http://www.lars.bbsrc.ac.uk/model/larswg.html> or from <ftp://ftp.lars.bbsrc.ac.uk/models/lars-wg/>

The observed residuals, obtained by removing the fitted mean value from the observed data, are used to analyse a time auto-correlation for minimum temperature and a time auto-correlation for maximum temperature. For simplicity both of these are assumed to be constant through the whole year for both dry and wet days with the average value from the observed data being used. Minimum and maximum temperature residuals have a pre-set cross-correlation of 0.6. Occasionally, the simulated minimum temperature is greater than the maximum, in which case the program replaces the minimum temperature by the maximum less 0.1.

The analysis of daily solar radiation over many locations showed that the normal distribution for daily solar radiation, commonly used in other generators, is unsuitable for certain climates (see Fig. 1). The same results have been found in Australia for the distribution of cloud cover (Chia & Hutchinson 1991). The distribution of solar radiation also varies significantly on wet and dry days. Therefore, separate semi-empirical distributions were used to describe solar radiation on wet and dry days. An auto-correlation coefficient was calculated for solar radiation and assumed to be constant throughout the year. Solar radiation is modelled independently of temperature. LARS-WG accepts sunshine hours as well as solar radiation data, in which case, for the days when solar radiation data are missing, sunshine hours are automatically converted to solar radiation using the regression relationships between these 2 variables as described in Rietveld (1978).

3. CHOICE OF SITES USED IN THE COMPARISON

The characteristics of those sites selected for this comparison are given in Table 2. They were chosen using the following 2 criteria.

3.1. Diversity of climates

The sites were chosen to represent as many as possible of the major types of climate in the world. The ecosystem classification presented in Bailey (1989) was used to define different climates. This is a hierarchical classification obtained by identifying aggregates of ecosystems within larger ecobiomes. The classification consists of 4 domains—polar, humid temperate, dry and humid tropical—which are split into 6, 12, 8 and 4 principal divisions respectively based mainly on large ecological climate zones. These are further subdivided into provinces on the basis of macro features of the vegetation. The selected sites are located in 12 of

the 30 climate divisions (1 polar, 3 dry and 8 humid temperate) from subarctic in the north to tropical/sub-tropical in the south, and from marine coastal to temperate desert, although no humid tropical site is included. The quality and availability of long daily weather data limited the choice of sites mainly to the USA and Europe, although 2 sites from Asia were also included. Nevertheless, the sites used should be sufficient to identify the main differences in performance between the 2 generators and their limitations.

3.2. Quality of data

Relatively long records of daily weather variables (minimum and maximum temperature, precipitation and solar radiation) must be available for each selected site in order to make a reasonable comparison of the generators. There is no formal constraint on the number of years of observed data, and, for example, LARS-WG can operate with as little as 1 yr of data. However, fairly long records are required to calculate robust and representative generator parameters for the site. The number of years of available data ranged from 17 at Seville, Spain, to 51 at Caribou, ME, USA (Table 2). Ideally, weather variables should be measured directly at each site, rather than being estimated from the other variables. Unfortunately, measured solar radiation is not available at several sites in the earlier years of the observed data record (Jokioinen, Finland, and Munich, Germany, and most of the sites in the USA and Russia). However, where sunshine hours or cloud cover were available, the missing solar radiation values were estimated from sunshine hours or cloud cover using the regression methods of Rietveld (1978) and Doorenbos & Pruitt (1984), respectively. In the absence of this information, the solar radiation values for those years were treated as missing and the solar radiation parameters calculated from the remaining data. LARS-WG automatically ignores missing values and converts sunshine hours to radiation, and the appropriate amendments were also made to WGEN.

4. COMPARISON TESTS

The aim in designing weather generators is to produce synthetic weather data which is statistically similar to the observed data. The 2 weather generators investigated here were run at the 18 selected sites and a number of statistical tests comparing the synthetic and observed data were carried out.

Initially the observed data for the particular site were run through each generator to produce a site parameter file. The generators can then use their respective

Table 2. Characteristics of 18 sites selected for use in comparison with the ecological climate-zone classifications from Bailey (1989)

Site	Country	Lat. (°N)	Long. (°E)	Alt. (m)	Period of record	Ecoregion division	Ecoregion province
Athens	Greece	37.97	23.72	107	1965–90	Mediterranean Regime Mountains	Shrub-forest-meadow
Bismarck, ND	USA	46.77	-100.75	506	1949–91	Temperate Steppe	Dry steppes of continental climate
Boise, ID	USA	43.57	-116.22	874	1950–95	Temperate Desert	Desert of continental climate
Bologna	Italy	44.53	11.30	49	1952–90	Mediterranean Regime Mountains	Forest-alpine meadows of western oceanic climate
Caribou, ME	USA	46.87	-68.02	190	1941–91	Subarctic	Eastern oceanic tundra
Debrecen	Hungary	47.48	21.63	112	1960–90	Prairie	Broadleaf-wooded and meadow steppes of moderately continental climate
Indianapolis, IN	USA	39.73	-86.27	246	1949–91	Hot Continental	Moderately humid broadleaf forest in moderately continental climate
Jokioinen	Finland	60.82	23.50	103	1961–90	Marine	Western oceanic coniferous and mixed forests
Mobile, AL	USA	30.68	-88.25	67	1949–91	Subtropical	Oceanic mixed constantly humid forests
Montpellier	France	43.60	3.90	81	1961–90	Mediterranean	Western oceanic mixed sclerophyll forests and shrub
Moscow	Russia	55.83	37.62	156	1951–80	Warm Continental	Moderate continental mixed forest
Munich	Germany	48.13	11.70	527	1951–80	Marine Regime Mountains	Forest-alpine meadows
Rothamsted	UK	51.80	-0.35	128	1960–90	Marine	Permanently humid western oceanic broadleaf forests
Seville	Spain	37.42	-5.88	12	1975–91	Mediterranean	Western oceanic mixed sclerophyll forests and shrub
Tashkent	Uzbekistan	41.27	69.27	477	1951–80	Temperate Desert	Deserts of continental climate
Tucson, AZ	USA	32.12	-110.93	779	1949–91	Tropical/Subtropical Steppe	Shrub and semi-shrub semi-deserts of continental climate
Verhojansk	Russia	67.55	133.38	136	1951–80	Subarctic	Continental light deciduous needleleaf open forest
Wageningen	Netherlands	51.97	5.67	7	1954–90	Marine	Permanently humid western oceanic broadleaf forests

parameter files to produce a time series of synthetic data of any length. For each of the 18 sites, 300 yr of daily weather data were generated using WGEN and LARS-WG. Such a long series of data was used so that the statistical properties of the synthetic data would be close to the true distribution of the data produced by the generators (which in theory, may be calculated analytically, although the complex interactions between the parameters makes this impractical). A longer data series makes the statistical tests more powerful since they are then more likely to give a significant result when there is a difference between the observed and synthetic data.

A number of statistical tests were carried out to compare a variety of characteristics of the data. It is important not only for the synthetic data to be similar to the observed data on average, but also for the distribution

of the data to be similar across their whole range. The χ^2 goodness-of-fit test was used to compare the probability distributions for the lengths of wet and dry series for each season (the year was split into quarters starting on 1 December) and for the distribution of precipitation for each month. A weakness of the χ^2 test is that its results may vary considerably depending on the intervals used for the test. Where the test was applied to a variable that LARS-WG models using a semi-empirical distribution, the same intervals as the semi-empirical distribution were used for ease of calculation (intervals were combined if there were less than 5 observed events). This choice of interval will tend to give the best possible fit for LARS-WG and alternative interval choices may give many more significant results. A change in the intervals used would be unlikely to have much effect on the WGEN results.

Weather generators are often used with crop growth simulation models to predict crop yield. Sequences of days with frost or high temperatures are important for such applications. For this reason the lengths of series of frost days (days with minimum temperature less than 0°C) and series of hot days (days with maximum temperature greater than 30°C) were recorded for each season and compared with the observed data using the χ^2 test.

Monthly means for precipitation, minimum temperature, maximum temperature and solar radiation were compared using the t -test. For each month, F -tests were carried out on the variances of all the daily values for the month across all the years ('day var'), and on the variances of the monthly mean values for the different years ('mon var'). The former variance value measures daily variability and the latter measures the inter-annual variability in the monthly means.

These tests are based on the assumption that the observed and synthetic weather data are both random samples from existing distributions and they test the null hypothesis that the 2 distributions are the same. In the case of observed weather data, such a distribution represents the 'true' climate at the site which would, in the absence of any changes in climate, be the distribution of observed data over a very long time period. Each test produces a p-value which measures the probability that both sets of data come from the same distribution (i.e. that there is no difference between the 'true' and synthetic climate for that variable). Hence, a very low p-value means that the synthetic climate is unlikely to be the same as the 'true' climate and so the generator is probably behaving poorly. Such tests cannot prove that the distributions are the same, and indeed the simplified nature of the generators means that there must be at least a small difference. A large p-value indicates that the differences are small enough that there is insufficient evidence to reject the null hypothesis. The required closeness of the synthetic and observed data depends upon the application in which the synthetic data are to be used.

4.1. Results

The number of tests that gave a p-value of less than 0.05 (i.e. a significant result at the 95% confidence level) are shown in Table 3. Even if the generator modelled the 'true' climate perfectly, on average 1 in 20 tests would still give a p-value of less than 0.05. However, a large number of such values indicates poor performance of the generator. Table 4 shows a sample of the output of the statistical tests for monthly maximum temperature means and variances for 2 sites, Athens, Greece, and Bologna, Italy. LARS-WG reproduced

these characteristics well at Athens, but failed to consistently reproduce the monthly variances at Bologna (5 out of 12 p-values for the F -test are less than 0.05).

Table 3 shows that WGEN gave a considerable number of significant results for all variables tested except for mean monthly precipitation. For LARS-WG, on the other hand, only the tests on the inter-annual variability in monthly means and on the distribution of frost and hot spells gave a large number of p-values less than 0.05, although, as mentioned in the previous section, the χ^2 test is biased in favour of LARS-WG. The results for each variable are discussed below.

4.2. Wet and dry series

WGEN models wet and dry series indirectly using a first-order 2-state Markov chain, whereas LARS-WG simulates them directly using semi-empirical distributions. The results show that the ability of the Markov chain to reproduce the full characteristics of the distribution of the dry and wet series is limited. The χ^2 test gives between 2 and 7 significant values out of the 8 tests for each site. For 8 of the sites, 4 or more of the 8 tests for WGEN were significant at the 1% level. The direct empirical approach of LARS-WG means that it is not surprising that it models the distributions well.

The Markov chain model uses an average transition probability and so the distribution of the series is geometric. This distribution is not able to reproduce accurately enough (from the point of view of the χ^2 test) the shapes of the distributions for the wet and dry series (Fig. 2a, b). It can also underestimate (Fig. 2c) or overestimate (Fig. 2d) the maximum series length. For example, the maximum length of a dry series in winter (DJF) at Athens, for the 300 yr of synthetic data produced by WGEN was 36 d, whereas the actual observed maximum over 25 yr for dry series was 65 d (Fig. 2c). A rare event, such as an unusually long dry series in Athens, does not significantly affect the calculation of the transition probabilities for the Markov chain and therefore remains undetected by this model. On the other hand, every single series from the observed data is represented in the semi-empirical distributions of LARS-WG.

4.3. Precipitation

Precipitation was tested in 3 different ways by comparing: (1) the rainfall distributions for each month of the year using a χ^2 test, (2) the monthly means using the t -test, and (3) the monthly variances using the F -test. Accurate simulation of the distribution of daily rainfall, particularly of large precipitation amounts,

Table 3. Results of the statistical tests comparing the observed data for 18 sites with 300 yr of synthetic data generated by (a) WGEN and (b) LARS-WG for the seasonal distributions of wet and dry series (Series), distributions of daily rainfall (Rain), monthly total rainfall (mn) and its variances (Mon var), monthly means, monthly variances and daily variances (Day var) for maximum (MAX) and minimum (MIN) temperature and solar radiation (RAD), and seasonal distributions of spells of temperature below 0°C (Frost) and above 30°C (Hot). Distributions were compared using the χ^2 test, and means and variances were compared using the *t*-test and *F*-test, respectively. The numbers in parentheses show the total number of tests performed for each category and the numbers in the table show how many tests gave significant results at the 5% significance level. A large number of significant results indicates a poor performance of the generator

Site	Series (8)	Rain (12)	mn (12)	Mon var (12)	MIN (12)	Mon var (12)	Day var (12)	MAX (12)	Mon var (12)	Day var (12)	RAD (12)	Mon var (12)	Day var (12)	Frost/ Hot (8)
(a) WGEN														
Athens	4	5	0	3	7	2	6	1	6	6	8	7	10	1
Bismarck	4	12	0	8	6	7	10	5	7	2	8	4	11	3
Boise	2	9	0	6	9	5	10	7	7	7	11	4	9	4
Bologna	5	12	3	3	7	4	12	9	5	10	10	12	12	3
Caribou	3	12	0	2	8	1	12	8	1	10	11	3	11	2
Debrecen	3	9	0	7	5	5	7	8	8	5	9	11	12	1
Indianapolis	5	12	0	2	1	4	7	9	1	6	10	2	9	5
Jokioinen	5	10	0	1	6	7	12	7	4	12	11	10	11	2
Mobile	3	11	0	1	6	4	6	7	5	6	8	2	10	4
Montpellier	2	12	0	5	8	4	10	9	5	4	7	9	10	3
Moscow	5	9	0	7	4	5	11	8	6	7	12	12	11	1
Munich	4	9	0	5	3	2	8	6	4	7	10	2	12	1
Rothamsted	5	12	0	9	10	9	9	8	7	7	9	12	12	1
Seville	4	0	0	4	7	6	6	7	6	8	8	5	10	3
Tashkent	2	2	0	1	3	4	6	4	8	6	11	9	12	3
Tucson	6	5	0	3	10	6	6	6	8	9	9	1	12	3
Verhojansk	3	2	0	4	8	0	8	8	1	8	6	7	11	1
Wageningen	7	11	0	9	6	4	8	9	5	9	11	11	12	1
(b) LARS-WG														
Athens	0	0	0	0	0	1	0	1	0	1	0	7	2	0
Bismarck	0	0	0	2	1	5	7	2	8	2	0	4	0	2
Boise	0	1	0	2	0	4	4	2	2	0	0	4	0	1
Bologna	0	0	0	1	0	2	3	0	5	2	0	12	1	1
Caribou	0	0	0	0	2	3	8	2	2	0	0	6	0	2
Debrecen	0	0	0	3	0	3	2	1	2	3	3	8	4	0
Indianapolis	0	0	0	2	0	2	1	0	1	2	0	1	0	4
Jokioinen	0	0	0	4	0	6	5	0	2	2	0	9	3	3
Mobile	0	0	0	0	1	3	1	2	3	1	0	4	0	1
Montpellier	0	0	0	1	0	3	0	0	5	0	0	4	0	1
Moscow	0	0	0	2	0	2	2	2	4	3	0	10	8	3
Munich	0	0	0	3	0	2	2	0	1	1	0	3	4	0
Rothamsted	0	0	0	4	0	4	1	1	4	0	0	4	1	2
Seville	0	0	0	1	0	7	3	1	0	1	0	7	1	3
Tashkent	0	0	0	0	0	0	2	0	2	2	0	6	6	1
Tucson	1	0	0	1	2	1	1	3	1	0	0	2	1	3
Verhojansk	0	0	0	3	1	1	8	3	0	7	1	3	5	0
Wageningen	0	0	0	5	0	5	0	0	2	1	1	2	1	1

plays an important role in the assessment of flood and soil erosion risks.

LARS-WG uses a semi-empirical distribution to model precipitation, whereas the gamma distribution is used by WGEN. Both models generate synthetic values to 1 decimal place. Fig. 3 shows cumulative probability and Table 5 shows probability density values for Wageningen, The Netherlands, in December. The points in Fig. 3 are plotted at the end of the intervals of the semi-empirical distribution, whereas the intervals

in Table 5 are of an equal size of 1 mm. A comparison of the cumulative distribution functions for a few months at several sites indicated that the gamma distribution underestimates the probability of very small precipitation amounts (less than 1 mm) and probably tends to overestimate the probability of larger values. Precipitation events of less than 1 mm have negligible importance for most applications, since very little moisture remains after evapotranspiration, but their occurrence is so high that the fitted gamma distribution may

Table 4. An example of the output data from the statistical tests, showing the comparison of the observed maximum temperature monthly means and variances with those of 300 yr of synthetic data generated by LARS-WG at Athens and Bologna. Probability levels (p-value) calculated by the *t*-test and *F*-test for the monthly means and variances are shown. A probability of 0.05 or lower indicates a departure from the observations that is significant at the 5% level

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Athens												
Obs. mean	12.7	13.7	16.0	20.2	25.3	29.7	32.5	32.2	29.0	23.1	18.3	14.7
Obs. variance	4.54	11.49	8.76	6.86	2.86	0.77	2.96	2.28	2.96	6.86	4.16	1.69
Sim. mean	13.3	13.7	16.1	20.3	25.1	29.4	32.3	32.1	28.7	23.4	18.2	14.8
Sim. variance	4.16	10.50	4.28	2.53	1.74	1.82	1.37	1.12	1.25	2.69	3.53	2.96
p-value for <i>t</i> -test	0.041	1.000	0.738	0.705	0.401	0.201	0.374	0.641	0.176	0.263	0.722	0.706
p-value for <i>F</i> -test	0.822	0.816	0.178	0.056	0.346	0.208	0.143	0.178	0.108	0.074	0.706	0.411
Bologna												
Obs. mean	4.9	8.2	13.3	17.9	22.6	26.8	29.9	29.3	25.4	19.0	11.2	6.1
Obs. variance	4.58	6.45	5.15	1.72	2.62	1.42	1.99	3.50	4.12	2.13	2.07	2.19
Sim. mean	5.1	8.2	13.1	17.7	22.4	26.6	29.6	29.1	25.3	18.9	11.4	6.1
Sim. variance	2.07	2.82	2.46	1.99	1.72	1.56	1.51	1.44	1.49	1.35	1.51	1.74
p-value for <i>t</i> -test	0.485	1.000	0.52	0.45	0.43	0.395	0.203	0.407	0.687	0.661	0.397	1.000
p-value for <i>F</i> -test	0.001	0.000	0.002	0.648	0.083	0.777	0.264	0.000	0.000	0.064	0.199	0.346

give too large a probability of the other more important events. Although the WGEN values are fairly close to the observed values, the differences are large enough to give significant values for many of the χ^2 tests, which shows the strength of the test. Indeed, for WGEN, 9 or more out of the 12 tests gave p-values less than 0.01 for 10 of the sites.

The LARS-WG semi-empirical distribution models precipitation as a step function and therefore its shape only approximately follows the shape of the observed values. However, the relative number of events within each interval for the 300 yr of generated data is very close to the number of observed events, since the probabilities are obtained from the observed data (Table 5). This allows LARS-WG to model the data accurately throughout its range, preventing the type of systematic errors produced by WGEN. For this reason, and because the same intervals as the semi-empirical distribution were used for the test, the χ^2 test gives only 1 significant value for LARS-WG. An alternative choice of intervals may give many more significant values but the χ^2 test used here does indicate the absence of major systematic errors. LARS-WG also accurately reproduces the relative number of large precipitation events that are contained in the observed

data, but the small number of these events means that there must be considerable uncertainty as to their 'true' probability for the given climate.

Observed monthly mean precipitation was duplicated well by both generators. The tests on the WGEN values were significant only for 3 months at Bologna, and for LARS-WG only for 1 test at Verhojansk. WGEN

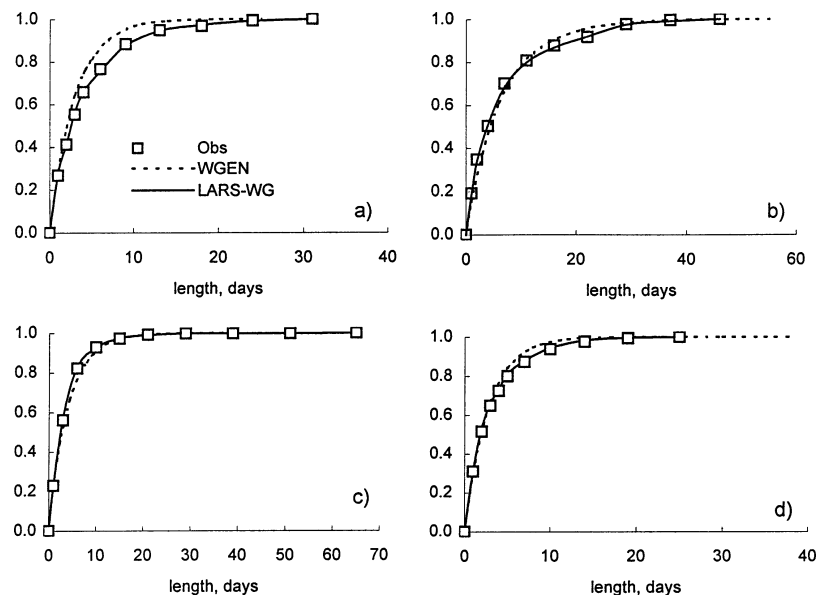


Fig. 2. Cumulative probability functions for the distributions of dry and wet series for observed data and synthetic data generated by WGEN and LARS-WG: (a) wet series in Jokioinen in winter (DJF), (b) dry series in Athens in autumn (SON), (c) dry series in Athens in winter (DJF), and (d) wet series in Moscow in winter (DJF)

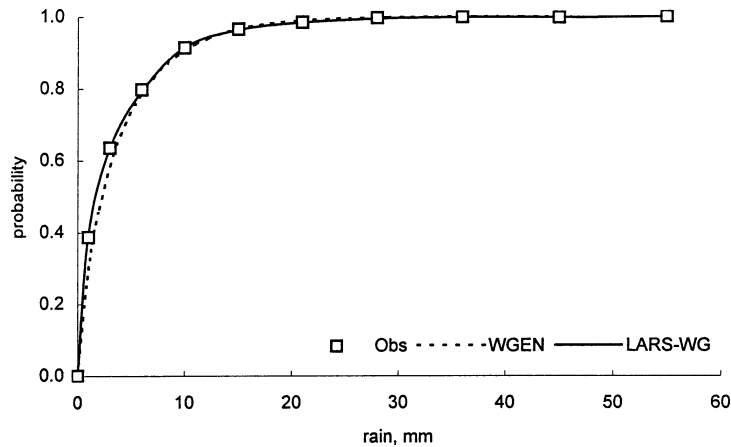


Fig. 3. Cumulative probability function of daily precipitation at Wageningen for December for observed data and synthetic data generated by WGEN and LARS-WG

does not accurately reproduce the distributions of wet days and daily rainfall, but the methods used are able to match the average behaviour and so the performance for mean rainfall is good.

The results for the inter-annual variances of monthly mean precipitation vary considerably from site to site from a total of 1 significant value for the 2 generators at Mobile to a total of 14 at Wageningen. The tendency is for the synthetic data to have a lower inter-annual variance than the observed data, and the main reason for this is probably that both models treat daily precipita-

Table 5. Probability density values for the first 20 intervals of size 1 mm for daily December precipitation at Wageningen for the observed data and for the synthetic data generated by WGEN and LARS-WG

Interval (mm)	Observed	WGEN	LARS-WG
0.1–1.0	0.368	0.298	0.390
1.1–2.0	0.158	0.168	0.127
2.1–3.0	0.107	0.121	0.123
3.1–4.0	0.072	0.085	0.052
4.1–5.0	0.040	0.068	0.056
5.1–6.0	0.046	0.050	0.054
6.1–7.0	0.046	0.040	0.031
7.1–8.0	0.033	0.030	0.029
8.1–9.0	0.019	0.026	0.025
9.1–10.0	0.017	0.023	0.026
10.1–11.0	0.022	0.016	0.011
11.1–12.0	0.007	0.016	0.010
12.1–13.0	0.009	0.010	0.010
13.1–14.0	0.009	0.009	0.010
14.1–15.0	0.007	0.008	0.010
15.1–16.0	0.009	0.006	0.004
16.1–17.0	0.000	0.004	0.003
17.1–18.0	0.002	0.005	0.003
18.1–19.0	0.003	0.002	0.003
19.1–20.0	0.005	0.003	0.004

tion as independent events from the given monthly distribution. In fact, at some sites, a series of several rainy days either tends to consist of a series of days with heavy rain or of a series of days with light rain, depending upon the prevailing meteorological conditions. The generators simulate such a series using independent random values from their monthly distributions so that the total precipitation for the series is less variable. This is a principal limitation of the framework used by both generators. One possible solution would be the incorporation of circulation patterns into the weather generators, with the distribution of precipitation conditioned on the type of circulation occurring (e.g. Bardossy & Plate 1991). Circulation patterns have been categorised subjectively for several regions in the world, such as Lamb (1972) for the British

Isles, but there are no general formalised rules on producing such a categorisation and any such rules would probably require records of daily pressure fields for the region around the site.

4.4. Minimum and maximum temperature

Both models use very similar underlying techniques to simulate maximum and minimum temperature. The means and variances are conditioned on the wet and dry status of the current day and the temperatures are reduced to normalised residuals, with the auto-correlations and cross-correlations between the temperature residuals assumed to be constant through the year. Means and variances of temperature in LARS-WG are approximated by Fourier series through the year, but the means and coefficients of variation are modelled by cosine curves in WGEN. WGEN also contains the further simplifications described in Section 2.1. WGEN demonstrated relatively poor performance in reproducing the monthly means and variances of maximum and minimum temperature (Table 3) not only for European sites, but also for the sites in the USA. The possible source of errors could be the use of a cosine function rather than a Fourier series or the use of pre-set values for correlation coefficients for all sites. Site-specific values of the correlation coefficients were calculated for 3 sites, but this gave little change in WGEN performance. On the other hand, Fig. 4a shows that a cosine curve is not a good approximation for mean minimum temperature at Boise. At Caribou, the amplitudes for maximum temperature on dry and wet days are significantly different, 31 and 25°C respectively, but WGEN uses the dry series amplitude for both wet and dry functions. This results in the July mean maxi-

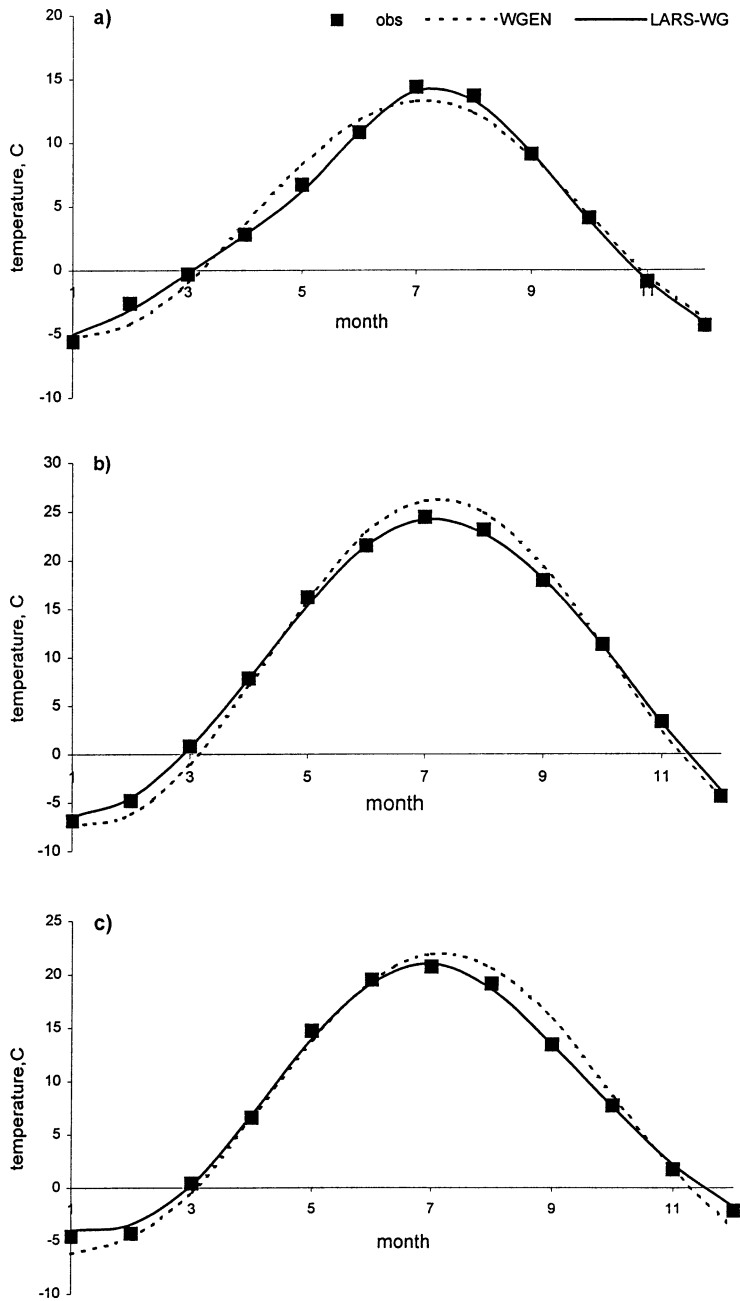


Fig. 4. Comparison of monthly mean temperatures for observed data and synthetic data generated by WGEN and LARS-WG: (a) minimum temperature in Boise; (b) maximum temperature in Caribou; (c) maximum temperature in Jokioinen

imum temperature being overestimated by about 2°C (Fig. 4b). The use of the original procedures for temperature and solar radiation, described in Richardson (1981), would be likely to improve significantly the simulation results.

LARS-WG reproduced the monthly means of maximum and minimum temperature well for all sites but gave mixed results for the daily variances. The average

monthly values for the daily variance, even after the adjustment for the effect of changes in the mean during the month (described in Section 2.2), do not follow a smooth periodic curve and so the Fourier series is unable to fit all the points well, resulting in significant differences for some months. Even where the Fourier series does provide a good fit, using the 12 monthly means gives small errors arising from fitting a daily curve to average monthly values. This also applies to the Fourier series for the means. The average of the values of the daily curve for a particular month can differ significantly from the observed monthly mean value, particularly for those months with the annual maximum and minimum values. This error would be reduced, where there is sufficient data, by calculating the average values over a shorter period. The fitting process could also be amended.

As for precipitation, both generators tend to underestimate the annual variance in monthly means. Although they do include auto-correlation, the structure is very simple and, as for precipitation, the strength of the correlation between successive values will vary considerably according to the prevailing circulation pattern. In particular, there are probably certain periods in which the correlation between temperature on successive days is very strong.

4.5. Solar radiation

The normal distribution used by WGEN is unable to reproduce the solar radiation distribution well, whereas the results for LARS-WG using semi-empirical distributions are much better. As for temperature, the generators tend to underestimate the inter-annual variance. The results are particularly poor for Bologna. However, analysis of the observed data showed that there is a decreasing trend in average annual solar radiation at Bologna over the period of the data (1958 to 1988) of nearly 1% per year, which has also been observed at other sites (e.g. Liepert & Kukla 1990, Russak 1990). In contrast, the Athens data show an increasing trend. The fundamental assumption of weather generators is that the climate is stationary and so, before using the generators in such circumstances, the data should be detrended.

4.6. Extreme temperature events

The seasonal distributions of the length of spells with minimum temperatures below 0°C (frost) and maximum temperatures above 30°C (hot) for observed and synthetic data were compared using the χ^2 test. The performance of WGEN and LARS-WG varies from site to site. For 9 sites for WGEN and 13 sites for LARS-WG the number of failed tests does not exceed 2 (out of 8). For some sites, such as Indianapolis, the number of failed tests was large for both generators. The most likely explanation for this is the same as for the inaccuracy in the simulation of monthly temperature variances described in Section 4.4.

5. CONCLUSIONS

The 2 generators investigated, LARS-WG and WGEN, have a similar structure in that they both use observed data to fit the parameters for the daily distributions of the variables for minimum and maximum temperature, precipitation and solar radiation. Both generators analyse dry and wet days separately and so include a mechanism for selecting the precipitation status of each day in the generated weather data.

The generators differ mainly in the choice of the daily distributions used. WGEN uses simple standard distributions, whereas LARS-WG tends to use the more flexible semi-empirical distributions. One advantage of using a standard distribution is that it will have a smoothing effect on the observed data. If there are good theoretical reasons for adopting the distribution, this will tend to make the synthetic climate closer to the 'true' climate, particularly when the amount of observed data is small. Errors and noise in the observed data will also be smoothed out and the interpolation effect of fitting the distribution allows the weather generator to simulate feasible values outside the range of the observed data. Such distributions require only a few parameters, thus making it easier to spatially interpolate to new sites. However, if the climate cannot be fitted by the theoretical distribution the generator will be unable to model it well, and a simple distribution is unlikely to be able to match the full range of climates across the world. A semi-empirical distribution is flexible enough to closely fit any shape of distribution, although any errors in the data will be modelled directly. For example, a day with a very large precipitation value erroneously included in the data will occasionally be reproduced from a semi-empirical distribution (although LARS-WG does include simple checks for extreme values). The data are smoothed a little, however, by using a uniform distribution within each interval. A semi-empirical distribution does

require more parameters to be recorded, although the size of a parameter file for LARS-WG is still very small. The limitations of using simple distributions are evident in the results of testing the generators over the 18 diverse sites, with LARS-WG able to match the observed data much better than WGEN. The greater number of parameters of LARS-WG may mean that it requires longer series of observed data for a good parameterisation. It is proposed to investigate the effect of the choice of the period of observed data and its length on the robustness of weather generator parameters.

Both generators had difficulty in reproducing the annual variability in monthly means of the climate variables and in reproducing the distribution of frost and hot spells. This is probably because both models have only a simple auto-correlation structure and, indeed, for precipitation, there is no correlation between the amounts on successive days. Both generators tend to underestimate the inter-annual variance in monthly means, which is likely to be due to the observed data containing many periods in which successive values are highly correlated, whereas the synthetic data vary more randomly. It is proposed to investigate more complex methods of incorporating correlation in future versions of LARS-WG.

Neither generator performed uniformly well in simulating the daily variances of the climate variables. Neither a cosine nor a Fourier curve was able to accurately fit the observed monthly values and the direct use of the monthly average variance values (adjusted for the effect of changes in the mean during the month) would improve the performance of the generators.

As with any model, it is always important to validate the weather generator by testing it at each site before the synthetic data are used in an application. Even if the generator has been extensively tested, the climate for the site in question may be different to that of the test sites. In addition, the accuracy required for each variable will vary according to the sensitivity of the application in which the data will be used to that variable. For some applications, for example, extreme weather events may be particularly important. The statistical tests used here are implemented in the LARS-WG model, allowing them to be carried out automatically for any new site.

Acknowledgements. The work of R.J.B. and E.M.B. was supported by the European Commission's Environment Programme under Contract Number ENV4-CT95-0154. The European data used here were supplied as part of the above contract. We are grateful to the other project participants for making the data available, namely: Tim Carter, Finnish Meteorological Institute, Helsinki (Jokioinen); Marco Bindi, Institute for Agrometeorology and Environmental Analysis, Florence (Bologna); Theodore Karacostas, Department of Meteorology and Climatology, University of Thessaloniki

(Athens); Zsolt Harnos, Department of Mathematics and Informatics, University of Horticulture and Food Industry, Budapest (Debrecen); Richard Delécolle, Station de Bioclimatologie, INRA-Avignon (Montpellier); Ana Iglesias, Escuela Técnica Superior de Ingenieros Agrónomos, Madrid (Seville); Joost Wolf, Department of Theoretical Production Ecology, Wageningen Agricultural University (Wageningen). The Russian station data were supplied by the Carbon Dioxide Information and Analysis Centre (CDIAC), Oak Ridge, Tennessee, USA. IACR-Long Ashton receives grant-aided support from the Biotechnology and Biological Sciences Research Council in the United Kingdom. We also thank the anonymous reviewers for their helpful comments.

LITERATURE CITED

- Bailey RG (1989) Ecoregions of the continents. USDA Forest Service, Washington, DC
- Bardossy A, Plate EJ (1991) Modelling daily rainfall using a semi-Markov representation of circulation pattern occurrence. *J Hydrol* 122:33–47
- Chia E, Hutchinson MF (1991) The beta distribution as a probability model for daily cloud duration. *Agric For Meteorol* 56:195–208
- Doorenbos J, Pruitt WO (1984) Guidelines for predicting water requirements. FAO Irrigation and Drainage Paper No. 24. FAO, Rome
- Hanson CL, Cumming KA, Woolhiser DA, Richardson CW (1994) Microcomputer program for daily weather simulation in the contiguous United States. US Department of Agriculture, Agricultural Research Service, ARS-114. USDA, Washington, DC
- Hutchinson M (1995) Stochastic space-time weather models from ground-based data. *Agric For Meteorol* 73:237–265
- Johnson GL, Hanson CL, Hardegree SP, Ballard EB (1996) Stochastic weather simulation—overview and analysis of two commonly used models. *J App Meteorol* 35:1878–1896
- Lamb HH (1972) British Isles weather types and a register of the daily sequence of circulation patterns 1861–1971. *Geophys Mem* 16
- Liepert BG, Kukla GJ (1990) Decline in global solar radiation with increased horizontal visibility in Germany between 1964 and 1990. *J Clim* 10:2391–2401
- Racsko P, Szeidl L, Semenov M (1991) A serial approach to local stochastic weather models. *Ecol Model* 57:27–41
- Richardson CW (1981) Stochastic simulation of daily precipitation, temperature, and solar radiation. *Wat Resour Res* 17:182–190
- Richardson CW, Wright DA (1984) WGEN: A model for generating daily weather variables. US Department of Agriculture, Agricultural Research Service, ARS-8. USDA, Washington, DC
- Rietveld MR (1978) A new method for the estimating the regression coefficients in the formula relating solar radiation to sunshine. *Agric For Meteorol* 19:243–252
- Russak V (1990) Trends of solar radiation, cloudiness and atmospheric transparency during recent decades in Estonia. *Tellus* 42B:206–210
- Semenov MA, Barrow EM (1997) Use of a stochastic weather generator in the development of climate change scenarios. *Clim Change* 35:397–414
- Wilks DS (1992) Adapting stochastic weather generation algorithms for climate changes studies. *Clim Change* 22:67–84

*Editorial responsibility: Mike Hulme,
Norwich, United Kingdom*

*Submitted: November 25, 1997; Accepted: March 5, 1998
Proofs received from author(s): April 24, 1998*