

# A weather generator for obtaining daily precipitation scenarios based on circulation patterns

João Corte-Real\*, Hong Xu, Budong Qian

Institute for Applied Science and Technology (ICAT), Faculty of Sciences, University of Lisbon, Campo Grande, 1700 Lisboa, Portugal

**ABSTRACT:** Based on principal component analysis (PCA) and *k*-means clustering algorithm, 4 daily circulation patterns, which are associated with daily precipitation in southern Portugal, have been identified from observed daily mean sea level pressure (MSLP) fields over the northeastern Atlantic and western Europe. A weather generator based on daily circulation patterns was calibrated with 3 different schemes to simulate daily precipitation occurrence and a 2-parameter gamma distribution was applied to generate daily precipitation amounts on rain days in southern Portugal. Parameters of the weather generator were estimated for every month in the winter half of the year (October to March) subject to each of the 4 daily circulation patterns. The weather generator was validated by hindcast using independent observational data. The synthetic daily precipitation series generated from the weather generator kept most of the important attributes of the observed series, such as the empirical distribution of daily precipitation amounts, the autocorrelation structure of the sequence of wet and dry days, the distributions of durations of wet and dry spells, and even the distributions of precipitation totals on several consecutive days. Furthermore, the daily MSLP fields over the northeastern Atlantic and western Europe simulated by the Hadley Centre's second generation coupled ocean-atmosphere GCM (HADCM2) control run (HADCM2CON) were validated by comparison with observed daily MSLP fields. It is clear that HADCM2 is able to reproduce very well daily MSLP fields and their seasonal variability over the region. Four daily circulation patterns, associated with daily precipitation in Portugal, identified from the observed daily MSLP fields over the area, were also classified well from the simulated daily MSLP fields. The weather generator was then applied to the sequence of daily circulation patterns in simulated MSLP fields of HADCM2CON, and validated. Essentially, the conclusions of previous validation analysis remained the same in the present case. The results imply that the weather generator can reproduce well important features of the 'present climate', including local precipitation in southern Portugal, and can therefore be applied to obtain future precipitation scenarios from daily MSLP fields simulated by the 2 transient experiments of HADCM2, HADCM2GHG (greenhouse gases) and HADCM2SUL (greenhouse gases plus sulphate aerosols). However, it should be kept in mind that the use of the weather generator in obtaining future precipitation scenarios is dependent on the assumption that the present relationship between local precipitation and large-scale atmospheric circulation will remain valid in the future changing climate.

**KEY WORDS:** Weather generator · Downscaling · Precipitation · Daily circulation patterns · Southern Portugal

## 1. INTRODUCTION

It is generally accepted that the increasing concentration of greenhouse gases in the atmosphere can significantly contribute to the change of climate in the near future (Houghton et al. 1990, 1992, 1996). The potential impacts of the induced climate change on various aspects of human environment have been a

source of great concern to scientists and decision-makers (Watson et al. 1996). Many models have been designed to assess the potential impacts of climate change on water resources, vegetation, crop yield (Mearns et al. 1992, Maytín et al. 1995, Semenov & Porter 1995) and so on. However, all these models require as input future climate scenarios with fine resolution on both spatial and temporal scales. One of the most important parameters from those scenarios required as input for such models is precipitation. It is

\*E-mail: jcr@fc.ul.pt

well known that general circulation models (GCMs) are the most powerful tools for obtaining future climate scenarios. However, GCMs are not good enough to simulate 'present climate' and climate change on regional and local scales, although they are able to reproduce well the large-scale characteristics of atmospheric circulation. Therefore, it has been widely agreed that outputs of GCMs cannot be directly used in impact studies and that downscaling is needed; this is especially true with respect to precipitation.

Essentially, downscaling approaches can be divided into 2 categories: statistical and deterministic. The statistical approaches assume that local climatic parameters such as precipitation or temperature are statistically related to large-scale atmospheric circulation and can be designed in many different ways, such as the simple analog method, linear and non-linear regression models and canonical correlation analysis. In addition, some stochastic models, such as a weather generator, can be adopted to generate synthetic time series of the climate element based on the observed relationship between that element and atmospheric circulation (Wilson et al. 1991, Bardossy & Plate 1992, Bogardi et al. 1993, von Storch et al. 1993, Gyalistras et al. 1994, Schubert 1994, Corte-Real et al. 1995, Schubert & Henderson-Sellers 1997). In some cases, another stochastic model is calibrated to simulate transitional regimes of atmospheric circulation (Bardossy & Plate 1991, Hay et al. 1992, Wilson et al. 1992). It is apparent that long-term observational data are needed to train the statistical models and that this technique can be successful only when close relationships are observed between the local climate parameters and large-scale atmospheric circulation. However, the obvious advantageous aspects of the statistical approach are its inexpensiveness and simplicity. This technique is adopted in the present paper. For deterministic approaches to downscaling, dynamic limited area models (LAMs) are nested in GCMs and/or in themselves (multiple nesting) to simulate climate at high resolution scales being driven at each level of nesting by the nearest coarser scale (Giorgi 1990, Giorgi et al. 1992, 1994, Jones et al. 1995). This technique has the advantage that simulations at progressively smaller grid meshes are dynamically consistent with those at larger scales as well as with the large-scale GCM results; also, regional and local forcings (e.g. topography, surface fluxes) are much better accounted for in LAMs than in GCMs. However, this approach is computationally very demanding. A comparison of different downscaling techniques to estimate climate change in southern Europe can be found in Cubasch et al. (1996). It should be pointed out that variable resolution GCMs and time slices with high resolution GCMs are also useful downscaling techniques.

As reported in a previous study (Corte-Real et al. 1998), 4 daily circulation patterns have been identified from daily mean sea level pressure (MSLP) fields over the northeastern Atlantic and western Europe which relate closely to daily precipitation conditions at Évora (southern Portugal). Furthermore, it has been verified that the Hadley Centre's second generation model (HADCM2) (Johns et al. 1997) reproduces very well the abovementioned 4 daily circulation patterns and that their occurrence regimes couple well with corresponding daily precipitation in southern Portugal. Therefore, it is possible to use a weather generator based on these circulation patterns to assess future climate change scenarios on a daily scale, and even some aspects of extreme events such as duration of sustained wet and dry periods. In this paper, a weather generator designed to generate time series of daily precipitation at Évora is calibrated based on historical data in the winter half of the year (October to March). The weather generator is validated not only on an independent observational data set, but on daily MSLP fields obtained from the control run of the HADCM2 model (HADCM2CON) as well. The results from validation imply that the weather generator, coupled with daily circulation patterns simulated by HADCM2CON, can reproduce well the 'present climate' of local precipitation at Évora in the winter half of the year. Therefore, it can be expected that reasonable future precipitation scenarios in southern Portugal can be obtained by applying the weather generator to daily MSLP fields simulated by HADCM2GHG and HADCM2SUL, 2 transient experiments forced respectively by greenhouse gases and greenhouse gases plus sulphate aerosols. Nevertheless, it should be kept in mind that this possibility is only justifiable under the assumption that the relationship between local precipitation and large-scale atmospheric circulation continues to hold in the future changing climate.

## 2. DOWNSCALING PROCEDURE: CALIBRATION

### 2.1. Linkages between local precipitation and large-scale atmospheric circulation

Before calibration of the weather generator, the linkages between precipitation and large-scale atmospheric circulation have to be established. In this study, daily circulation patterns identified from large-scale daily MSLP fields are used to represent large-scale atmospheric circulation. As described in Corte-Real et al. (1998), by applying a *k*-means clustering algorithm coupled with principal component analysis (PCA) to observed daily MSLP fields over the northeastern Atlantic and western Europe, 4 daily circulation patterns have

been classified during the period from January 1, 1946, to August 31, 1992. The optimal number of clusters (circulation patterns) is determined by an information measure associated with daily precipitation at Évora. The quality of daily precipitation data at Évora is quite good and some systematic changes related to changes in large-scale atmospheric circulation have been detected (Zhang et al. 1997, Corte-Real et al. 1998). The details of the clustering technique and the circulation patterns can be found in Corte-Real et al. (1998).

Fig. 1 shows composites of daily MSLP fields corresponding to the 4 daily circulation patterns. The 4 patterns are respectively named as blocking-like, summer dry, winter dry and rainy, according to their seasonal frequencies of occurrences and the way they relate to precipitation. Synoptic structures of the patterns can well explain their relationships with daily precipitation at Évora. Table 1 lists some statistics of daily precipitation at Évora for each of the 4 daily circulation patterns in the winter half of the year. It is obvious that the 4 circulation patterns differentiate well the precipitation conditions at Évora. In fact, precipitation in southern Portugal during the winter half of the year is mostly associated with large-scale atmospheric circulation systems, such as cyclones and the accompanying frontal activities. The synoptic systems usually come from the Atlantic Ocean. A negative trend has been

detected in March precipitation totals, reflecting changes in the frequencies of occurrence of certain circulation patterns; those changes are in turn associated with the observed positive trend in the monthly series of the NAO (North Atlantic Oscillation) index. Moreover, the 4 circulation patterns can also explain a large part of the precipitation variability at other places in continental Portugal (Corte-Real et al. 1998). Therefore, the relationships between the 4 circulation patterns and precipitation in continental Portugal make it possible to calibrate the weather generator not only for Évora but for a variety of locations in the country. In order to describe the calibration and validation details, Évora is used as an example in this paper.

## 2.2. Weather generator

The weather generator has been widely applied in climate change and downscaling studies (Wilks 1992, Semenov et al. 1998). Lettenmaier (1996) conducted a review of stochastic modeling of precipitation with applications to climate model downscaling. In respect of daily precipitation, 2 stochastic models need to be calibrated. One simulates the occurrence of rainfall, and the other simulates rainfall amounts on rain days. First-order Markov chains are usually adopted to sim-

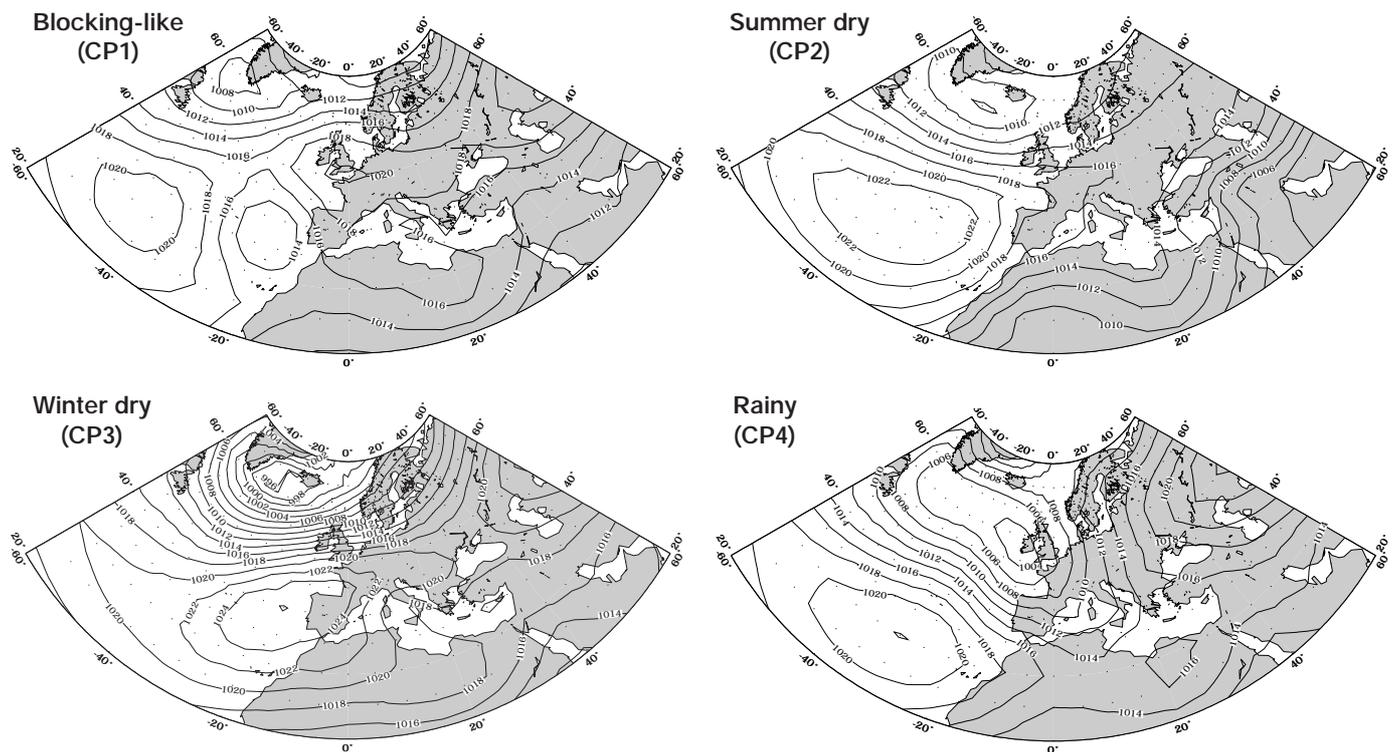


Fig. 1. Composites of daily MSLP fields during 1946 to 1992 corresponding to the 4 daily circulation patterns identified from observations by a *k*-means clustering algorithm coupled with PCA (the interval of the isobars is 2 hPa)

Table 1. Statistics of different circulation patterns and their associated precipitation characteristics at Évora (southern Portugal)

Month	Circulation pattern	Frequency of occurrence (%)	Probability of rain (%)	Mean rainfall amount on a rain day (mm)	Contribution to total rainfall (%)
Oct	CP1	19.8	30.9	7.0	20.7
	CP2	33.8	23.7	3.9	15.1
	CP3	27.9	9.8	2.9	3.8
	CP4	18.5	73.9	9.1	60.3
	Total	100.0	30.5	6.8	100.0
Nov	CP1	22.0	35.0	7.3	21.3
	CP2	8.0	20.0	4.5	2.7
	CP3	41.5	15.7	3.4	8.4
	CP4	28.6	71.6	8.6	67.5
	Total	100.0	36.2	7.2	100.0
Dec	CP1	13.7	51.5	5.2	13.0
	CP2	3.4	26.5	1.9	0.6
	CP3	54.6	22.0	3.1	12.9
	CP4	28.2	79.6	9.3	73.4
	Total	100.0	42.4	6.7	100.0
Jan	CP1	16.5	43.3	7.2	19.1
	CP2	2.3	39.4	8.0	2.6
	CP3	54.7	24.4	2.6	12.8
	CP4	26.6	80.9	8.2	65.5
	Total	100.0	42.8	6.3	100.0
Feb	CP1	19.0	44.0	4.9	14.7
	CP2	6.3	25.0	2.6	1.4
	CP3	42.2	20.5	2.6	8.0
	CP4	32.5	80.0	8.2	75.9
	Total	100.0	44.6	6.3	100.0
Mar	CP1	20.3	32.1	4.7	13.0
	CP2	22.2	27.5	3.3	8.5
	CP3	29.6	18.3	2.9	6.6
	CP4	27.8	70.6	8.6	71.8
	Total	100.0	37.7	6.2	100.0

ulate the occurrence of rain (e.g. Katz 1981, Schubert 1994, Katz & Parlange 1996), although different orders have been chosen in different cases (Coe & Stern 1982, Stern & Coe 1984). Daily precipitation amounts are typically skew distributed. Several probability distributions have been applied to simulate the distribution of daily precipitation, such as the exponential, mixed exponential, gamma and Kappa distributions. In this paper, 3 different schemes are designed to simulate occurrence of wet days, 2 of them being first-order Markov chains, dependent on daily circulation patterns, and the other being only a random process related to circulation patterns. A 2-parameter gamma distribution, also dependent on circulation patterns, is used to generate daily rainfall amounts on wet days.

### 2.2.1. Simulations of rainfall occurrence

Before calibrating the first-order Markov chain, a formal test, similar to the  $\chi^2$  goodness-of-fit test (Wilks 1995), was conducted on the sequence of wet and dry

days for each month in the winter half of the year. The null hypothesis for this test is that the data series is serially independent (i.e. the data are independent binomial variables with  $n = 1$ ); the alternative is that the series was generated by a first-order Markov chain. Results from this test show that the null hypothesis would be rejected at the 0.1% significance level for each month (e.g. the computed  $\chi^2$  value for January is 250.47, while the threshold at 99.9% confidence level is 10.08). In fact, the full autocorrelation function for the Markov chain agrees in each month very closely with the sample autocorrelation function for the first few lags. This match qualitatively supports the first-order Markov chain as an appropriate model for simulating daily rainfall occurrence at Évora.

To calibrate the first-order Markov chain, an indicator series of a discrete random variable  $Y(t)$  is created by classifying the days into dry and wet days. If daily rainfall amount  $r$  is not less than a certain threshold (0.1 mm in this study) the day is wet; otherwise it is dry.

Indicator series  $Y(t) = \{0, r < 0.1 \text{ mm}; 1, r \geq 0.1 \text{ mm}\}$

In the first 2 schemes, the random process  $Y(t)$  is assumed to be a first-order Markov chain but dependent on daily circulation patterns; that is, the probability of rain on Day  $t$  depends not only on the state (wet or dry) of the previous Day  $t-1$  but also on the circulation patterns of Day  $t$  only or both Day  $t$  and Day  $t-1$ . In Scheme 3  $Y(t)$  is considered as a random process related only to daily circulation patterns; therefore, the probability of rain on Day  $t$  depends only on the circulation pattern of that day. The autocorrelation structure of the sequence of wet and dry days is assumed to be physically determined by daily circulation patterns, although reproduction of the observed autocorrelation structure can surely be more satisfactory if the weather state and circulation pattern on Day  $t-1$  are also taken into account. The 3 schemes were designed to make it possible to ascertain whether the probability of rain on Day  $t$  highly depends on the weather state and circulation pattern of Day  $t-1$ . Therefore, in Scheme 1 the probability of rain on Day  $t$  depends on the circulation pattern on that day and both the weather state and the circulation pattern of Day  $t-1$ , while in Scheme 2 it depends on the circulation pattern of Day  $t$  but only on the weather state of Day  $t-1$ ; finally, in Scheme 3, the same probability depends only on the circulation pattern of the considered Day  $t$ . The definitions of the conditional probabilities in the 3 schemes can be found in Appendix 1.

Considering the somewhat different behaviours of circulation patterns in different months, the probabilities of rain and the conditioned transition probabilities are separately calculated for each month from the observational data. In each of the 3 schemes, a day is classified as either wet or dry by comparing the corresponding conditional probabilities with the values of a uniformly distributed random variable  $\omega$  ( $0 < \omega < 1$ ), as described in Appendix 2.

Fig. 2 shows the monthly values of the probabilities defined in Appendix 1, dependent on the occurrence of the rainy pattern on Day  $t$ , calculated from the historical data of 42 yr. It is obvious that Day  $t$  is more likely to be wet if the previous day was wet, no matter what the type of circulation pattern was on Day  $t-1$ . However, the probability of rain on Day  $t$  is higher if the circulation pattern on that day remains the same as the rainy pattern on Day  $t-1$ .

### 2.2.2. Generation of rainfall amounts on wet days

If a day is classified as wet based on any one of the schemes mentioned above, the daily rainfall amounts are fitted by a 2-parameter gamma distribution, whose parameters are also dependent on circulation patterns

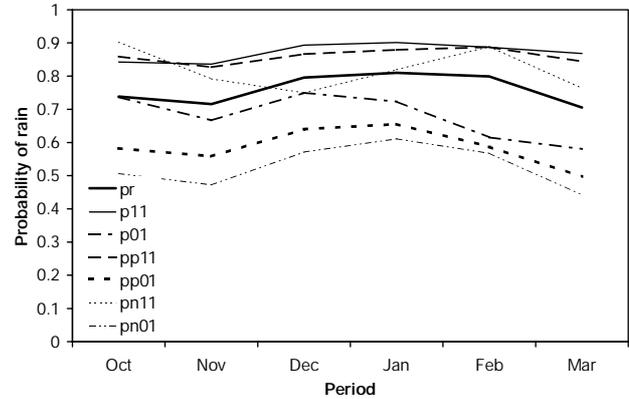


Fig. 2. Different conditional probabilities of rain during the winter half of the year corresponding to the rainy circulation pattern (definitions can be found in Appendix 1)

in different months. The gamma distribution is defined by the probability density function

$$f(x) = \frac{(x/\beta^i)^{\alpha^i-1} \exp(-x/\beta^i)}{\beta^i \Gamma(\alpha^i)} \quad x > 0; \alpha^i, \beta^i > 0$$

$\alpha^i, \beta^i$  being, respectively, the shape and scale parameters of the distribution; superscript  $i$  indicates that the parameters are conditioned on the  $i$ th circulation pattern  $CP^i$ . As discussed by Wilks (1995), the moment estimators for the gamma distribution can give very bad results for small values of  $\alpha^i$ , and a much better approach to parameter fitting for the gamma distribution is to use the method of maximum likelihood. A maximum-likelihood approximation for  $\alpha^i$  and  $\beta^i$ , from Greenwood & Durand (1960), is used in this paper. To match the interval of the gamma distribution, the daily precipitation data on wet days are shifted before fitting the gamma distribution as suggested in Stern & Coe (1984). Furthermore, it should be noted that the estimation of the shape parameter is very sensitive to inaccuracies in the recording of small precipitation amounts. This problem can be avoided by censoring the left tail of the distribution (Coe & Stern 1982, Stern & Coe 1984). In fact, it may also be avoided if a higher threshold is chosen to define the wet days.

The parameters of the gamma distribution, estimated from observed rainfall amounts in each month of the winter half of the year and each circulation pattern, are then used for goodness-of-fit tests. Because these parameters have been fit to the same data used in the test, the original Kolmogorov-Smirnov (K-S) test is not appropriate in this case. However, the modified K-S test, Lilliefors test (Lilliefors 1967), can be used and lower critical values for the test statistic from Crutcher (1975) were applied. In more than 70% of all cases the null hypothesis that the daily rainfall amounts follow a 2-parameter gamma distribution could not be rejected at the 5% level of significance. It implies that, in this

study, the gamma distribution can be accepted as a good candidate to fit the daily precipitation.

### 2.3. Strategy for calibration of the weather generator

The construction of the weather generator and estimation procedure have been described above; details can also be found in Appendices 1 & 2. However, in order to avoid bias in the validation procedure, particular attention should be given to the period of observational data selected for that purpose, especially if non-stationarities, such as trends, are recognized in the data set (Corte-Real et al. 1998). Therefore, the weather generator is calibrated using the data separated into the even years and the odd years during 1946 to 1992 (46 yr in total), and then validated using the corresponding independent data, respectively, in the odd years and the even years. The validation conclusions are essentially the same. Thus, in the validation section of this paper, only the results of simulations for the odd years with the weather generator calibrated from the data in the even years are presented. Moreover, in order to take advantage of all the information provided by observations, the weather generator, calibrated from the data in all years during 1946 to 1992, is applied to the sequences of daily circulation patterns simulated by HADCM2CON.

## 3. DOWNSCALING PROCEDURE: VALIDATION

Cross validation was performed, as mentioned above, on an independent observational data set. In this section, validation results are presented only for the weather generator's simulations in odd years and each one of the 3 schemes described in Section 2. As will be demonstrated, the weather generator with Scheme 1 seems the best one in simulating daily precipitation, especially the distributions of dry and wet spells and precipitation amounts for several consecutive days. For this reason, only the weather generator established with Scheme 1 is applied on the sequence of daily circulation patterns classified from HADCM2CON's MSLP fields.

### 3.1. Validation using observations

Based on the 3 schemes described in detail in Appendix 2, the stochastic models were used to produce 100 realizations (synthetic series) of daily precipitation in the winter half of the year, during the independent validation period (odd years). The statistics presented are averages calculated from the 100 real-

izations. For all schemes, the models reproduce very well the probability of rain, not only for the whole winter half of the year, but also for each specific month. Fig. 3 shows the exceedance probabilities of daily precipitation amounts, simulated by the 3 schemes, compared with the observational ones. The 3 schemes reproduce equally well the observed structure. The results for each month do not alter the conclusions, although, in some cases, the simulations do not match the observations as perfectly as they do in the whole winter half of the year. As an example, the distribution simulated by the weather generator with Scheme 1 in October is displayed in Fig. 3. Fig. 4 displays, for time lags from 1 to 7 d, autocorrelation coefficients (in January as an example) in the synthetic series, generated by the 3 schemes, along with those obtained from observations. It is clear that Schemes 1 and 2 reasonably reproduce the observed autocorrelation structure, for time lags up to 3 d, while Scheme 3 is unable to simulate that structure even for 1 d. Moreover, the autocorrelation in Scheme 1 is the best, although differences in Scheme 2 are slight. The effects of the simulated autocorrelation structure in the sequence of wet and dry days can be found in the distributions of wet and dry spells, as well as in the distributions of precipitation amounts for several consecutive days.

The durations of wet and dry periods are one of the important issues relating to climate change. Fig. 5 shows relative frequencies of durations of wet and dry spells in observations and simulations, during the validation period. As can be concluded, the weather generator with Schemes 1 and 2 can simulate well the durations of wet and dry periods, not only for short but also for long durations, especially the wet spells. As could be expected, the simulations with Scheme 3 are not as good as with the other 2 schemes. As an example the relative frequencies of wet and dry spells in January simulated by the weather generator with Scheme 1 are also shown in Fig. 5. It is clear that the simulations match well the corresponding observational results, although these are not very stable because only a few samples were used to calculate the frequencies on a monthly scale.

Precipitation totals over several consecutive days, such as 3, 5 and 7 d, are important as well, especially for crop growing. Fig. 6 shows the distributions of exceedance probabilities of accumulated precipitation amounts in 2, 3, 5 and 7 consecutive days, simulated by the weather generator with the 3 schemes, compared with the observed ones. It seems that the weather generator can reproduce well the properties of the distributions of accumulated precipitation amounts in several consecutive days, although differences become larger when the number of days increases. In this case, Scheme 1 is the best, while, not surprisingly, Scheme 3

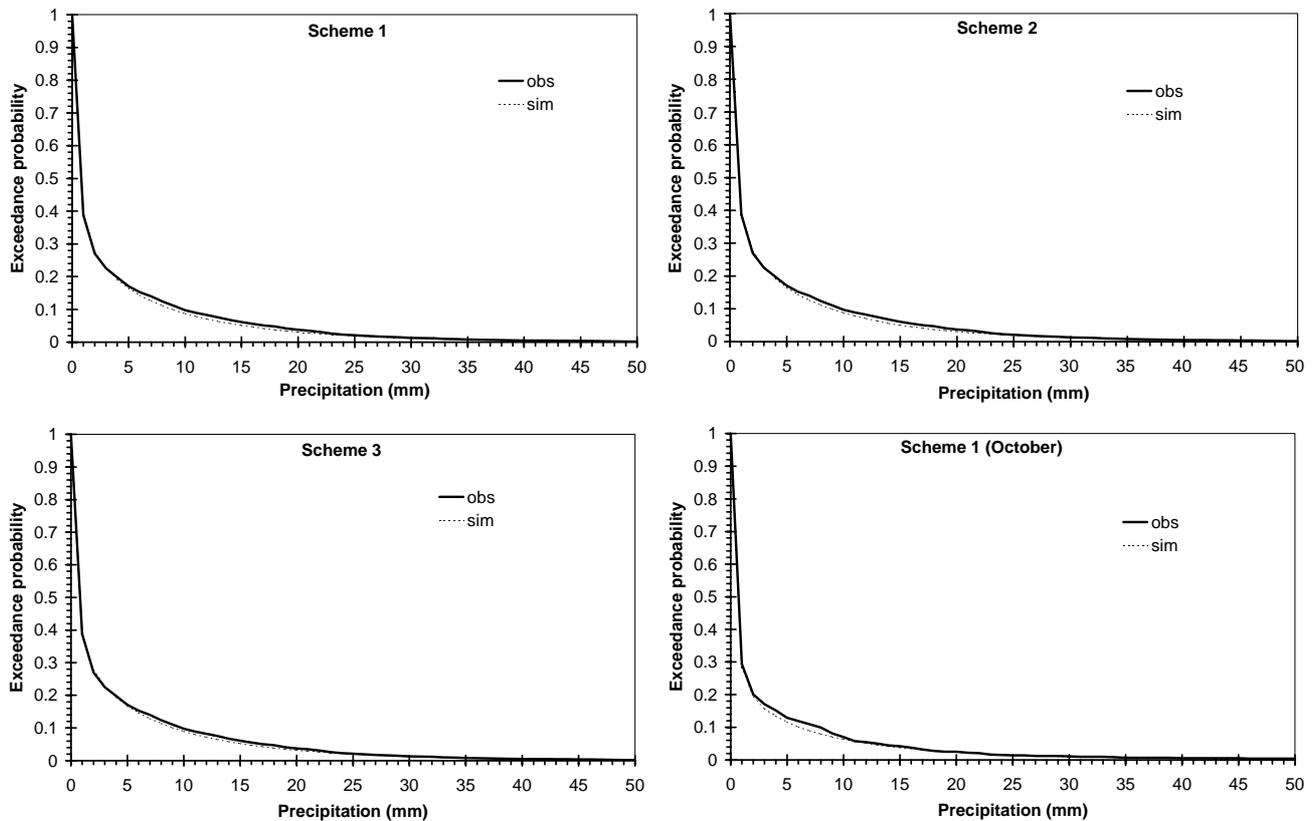


Fig. 3. Observed and simulated distributions of exceedance probabilities of daily precipitation amounts for the winter half of the year and for the month of October

displays the worse performance, particularly regarding wet spells; in fact, it does not reproduce the autocorrelation structures of the sequences of wet and dry days as well as the other schemes. The distributions of exceedance probabilities of accumulated precipitation amounts in 2, 3, 5 and 7 d simulated by the weather generator with Scheme 1 for October are also shown in

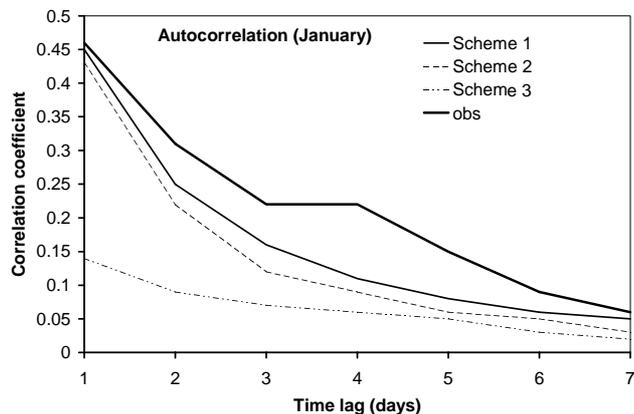


Fig. 4. Observed and simulated autocorrelation structures of the sequences of wet and dry days in January

Fig. 6 as an example of its behaviour on a monthly scale.

In conclusion, the weather generator with Scheme 1 is the best one for satisfactory reproduction of most observed statistical structures of daily precipitation series. This implies that precipitation on Day  $t$  is not only determined by the circulation pattern on the same day but relates closely to both the weather state and circulation pattern of the previous day as well. In these circumstances, only this scheme is chosen for further discussion and application to downscaling.

### 3.2. Validation using HADCM2CON

As already pointed out, the weather generator was designed for downscaling HADCM2 large-scale output for southern Portugal, enabling local future climate scenarios in the region to be obtained. Therefore, it is necessary to verify if the weather generator can reasonably reproduce the 'present climate' in southern Portugal, in respect of daily precipitation structures, by using daily circulation patterns simulated by HADCM2 in its control run. In this regard, it is important to clarify whether the GCM can reproduce well

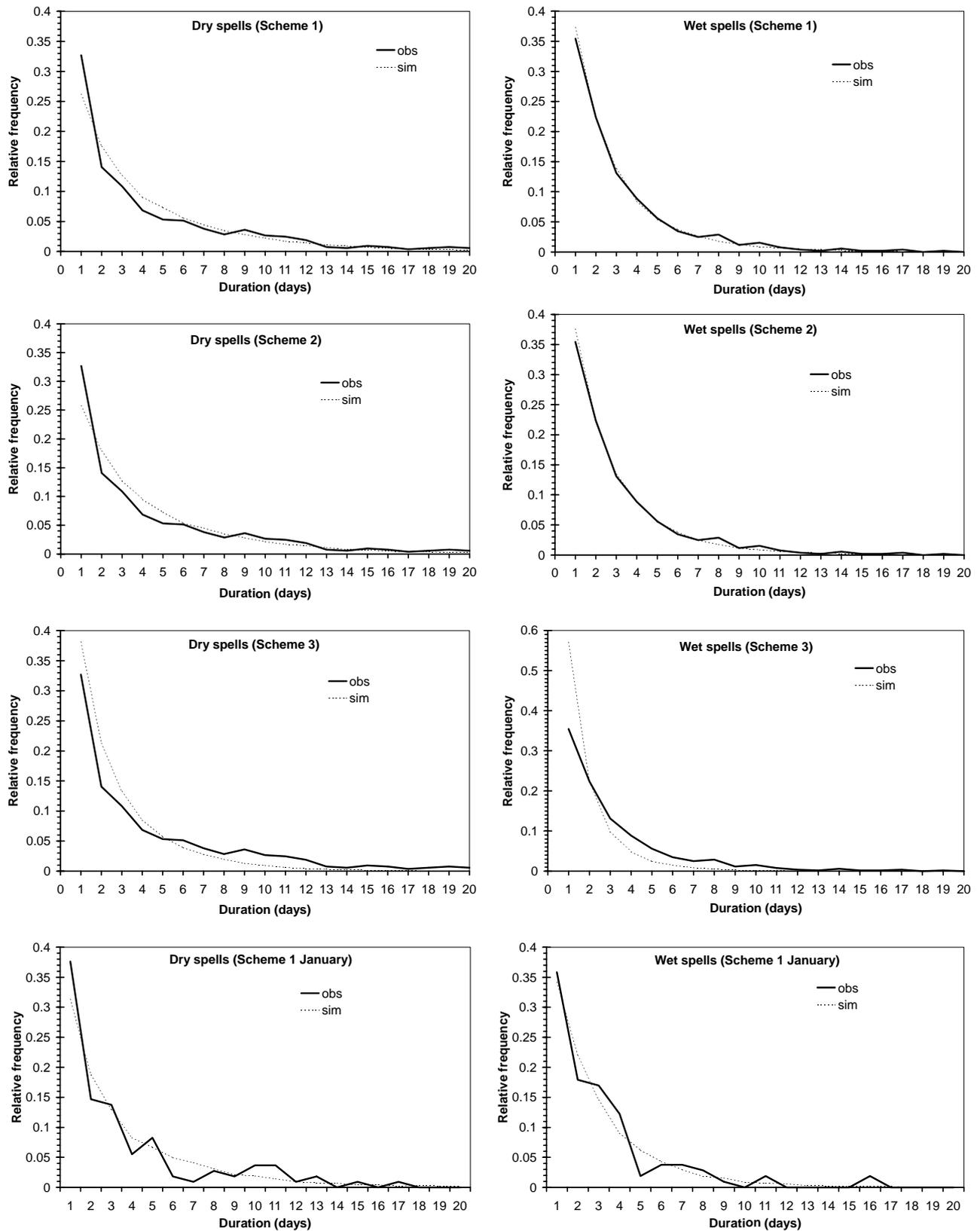


Fig. 5. Observed and simulated relative frequencies of duration of dry (left panels) and wet (right panels) periods in the winter half of the year and in January during the validation period

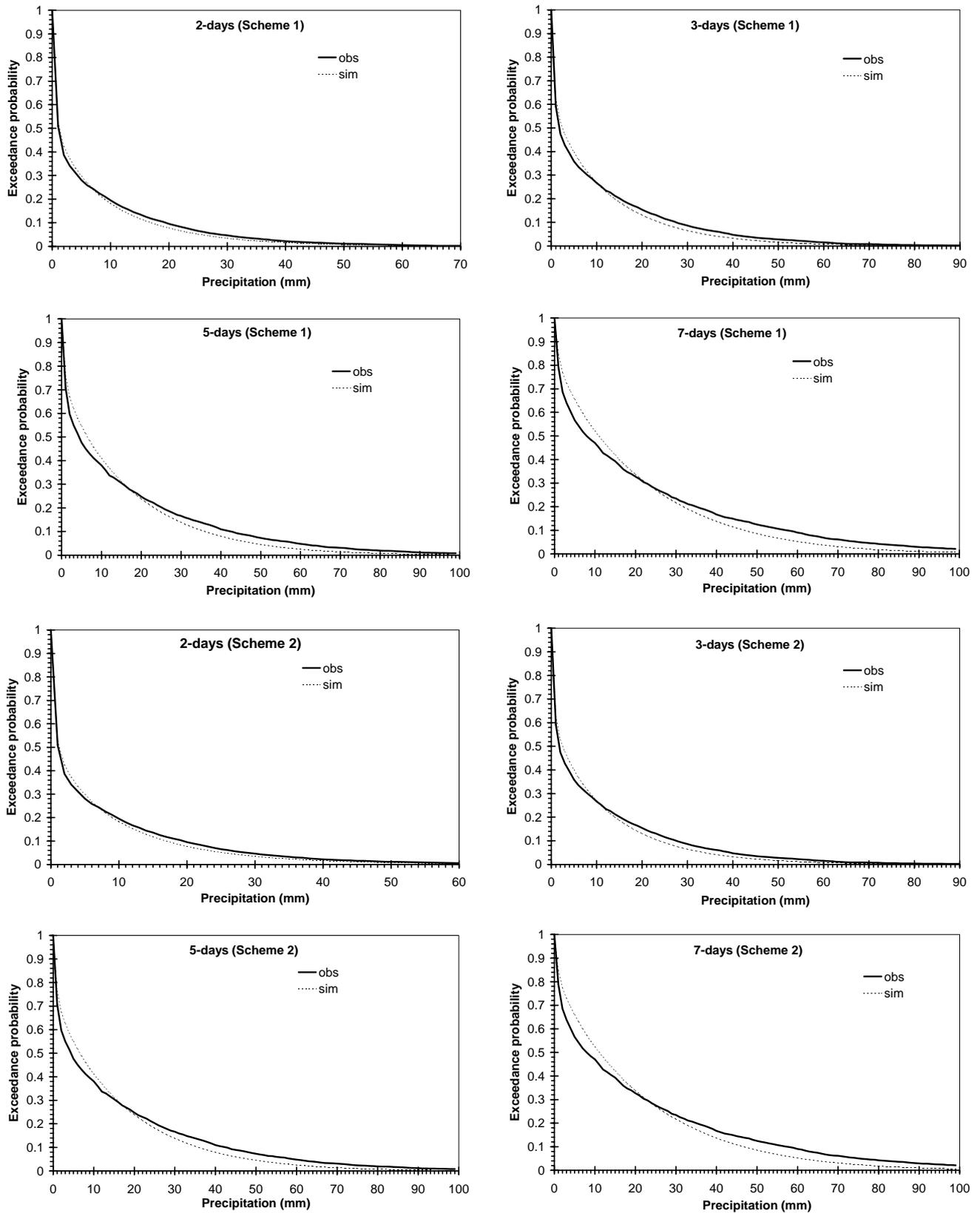


Fig. 6. Observed and simulated distributions of exceedance probabilities of precipitation totals in 2, 3, 5 and 7 consecutive days in the winter half of the year and in the month of October

(Fig. continued overleaf)

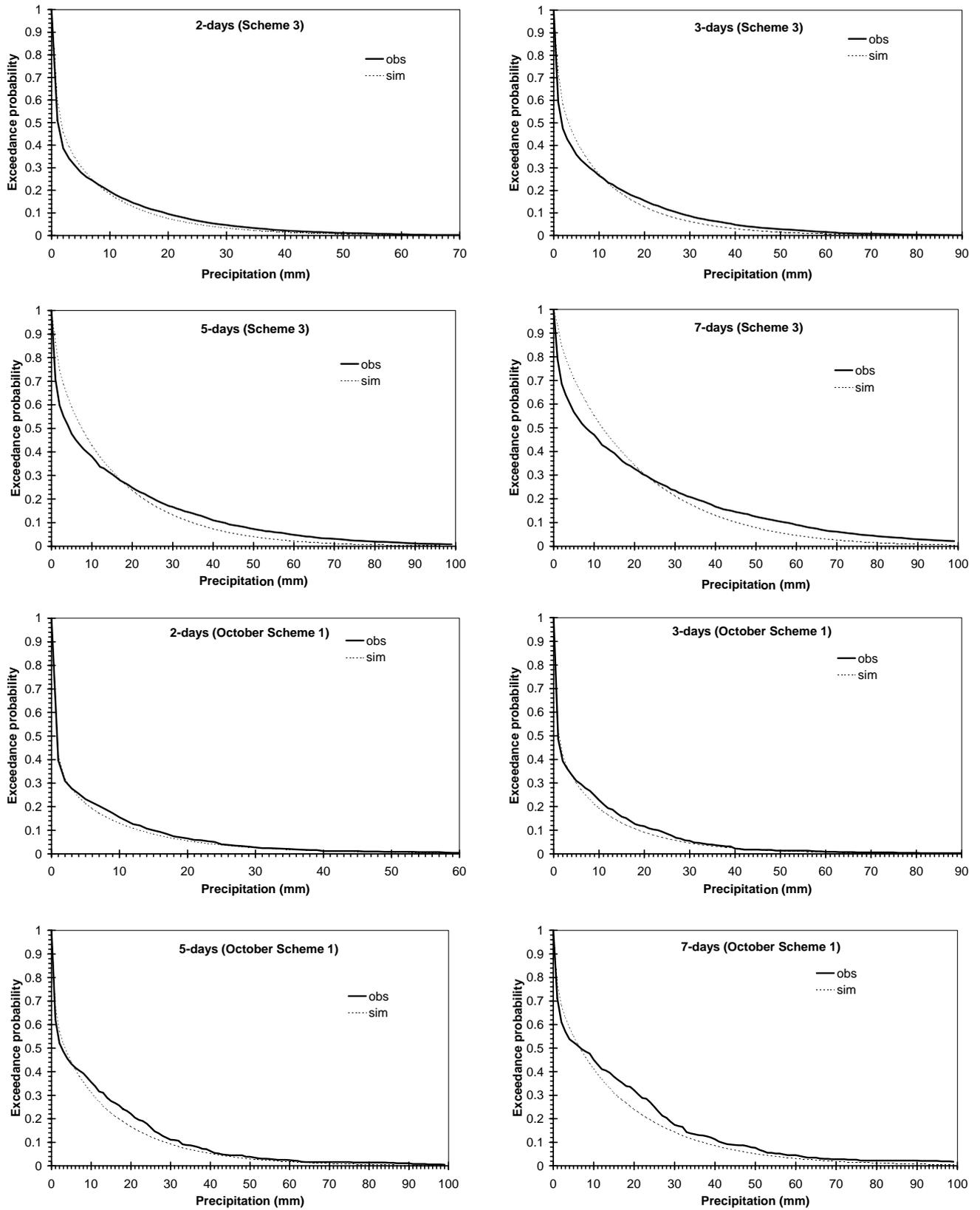


Fig. 6 (continued)

the 4 observed circulation patterns related to precipitation, although a positive confirmation does not guarantee that the GCM is also able to properly simulate future circulation patterns.

### 3.2.1. Validation using the daily MSLP fields simulated by HADCM2CON

To verify if the 4 observed circulation patterns related to precipitation in southern Portugal can be identified in daily MSLP fields simulated by HADCM2CON, the same classification procedure (Corte-Real et al. 1998) was applied to the daily MSLP fields covering the same region and period. The 4 patterns classified from the simulated fields are very similar to the corresponding observed ones, except for the blocking-like pattern. The daily MSLP fields have also been projected into the first 20 EOFs of observed MSLP, and the 20 projections were used to calculate distances, in EOF space, between a sample and the center of each observed cluster. A sample is considered to be a member of the cluster to which it is closest. Fig. 7 shows composites of the simulated daily MSLP fields corresponding to the 4 circulation patterns. It is clear that the 4 circulation patterns are reproduced well, although comparison reveals some slight differ-

ences in detail with Fig. 1. Moreover, HADCM2CON also reproduces well the relationship between precipitation in southern Portugal and daily circulation patterns, since, in the model, the rainy pattern remains the wettest pattern and the summer dry pattern remains the driest pattern. In addition, the frequencies of occurrence of the 4 patterns are very close to the corresponding observational values. All these results indicate that it is reasonable to expect the weather generator to be able to simulate well local daily precipitation structures in southern Portugal from HADCM2CON's daily MSLP fields.

### 3.2.2. Validation using the daily circulation patterns simulated by HADCM2CON

Fig. 8 shows, for the winter half of the year, the distribution of exceedance probabilities of daily rainfall amounts at Évora, from both the observed and the simulated daily precipitation series and their corresponding monthly/seasonal precipitation totals and precipitation intensities in the corresponding period. As already indicated, the parameters of the weather generator were computed from the whole available observational data (1946 to 1992) to retain more information from the historical data. The simulations are the result

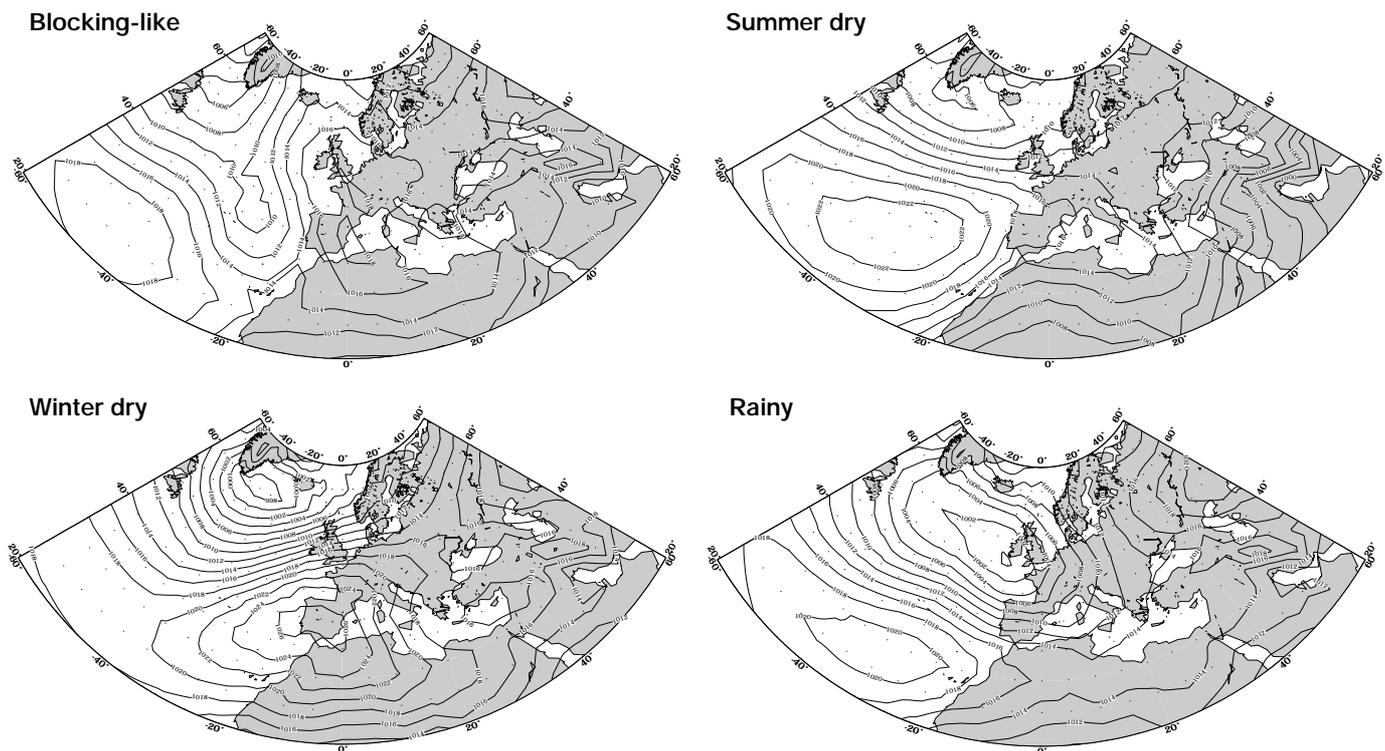


Fig. 7. Composites of daily MSLP fields from the HADCM2CON simulation, corresponding to the 4 circulation patterns identified from observations (Fig. 1) (the interval of the isobars is 2 hPa)

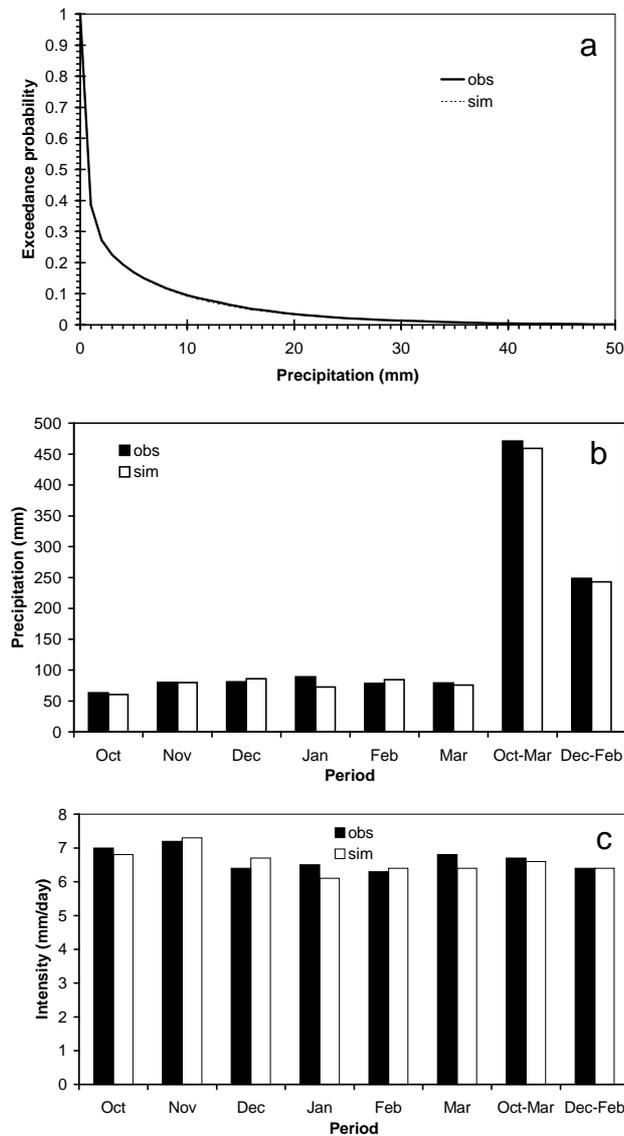


Fig. 8. (a) Observed and simulated distributions of exceedance probabilities of daily precipitation amounts during the winter half of the year, (b) precipitation totals and (c) daily precipitation intensities on rain days for the period 1946 to 1992. Simulations are based on the sequence of daily circulation patterns simulated by HADCM2CON

of coupling the weather generator with daily circulation patterns simulated by HADCM2CON. It can be concluded that simulations are satisfactory. The distributions of relative frequencies of dry and wet spells in this simulation are shown in Fig. 9. The simulations are as good as those obtained from independent observational series of daily circulation patterns. Fig. 10 displays the exceedance probabilities of observed and simulated precipitation amounts in 2, 3, 5 and 7 consecutive days. The simulations are also satisfactory, although for longer durations their quality decreases, a

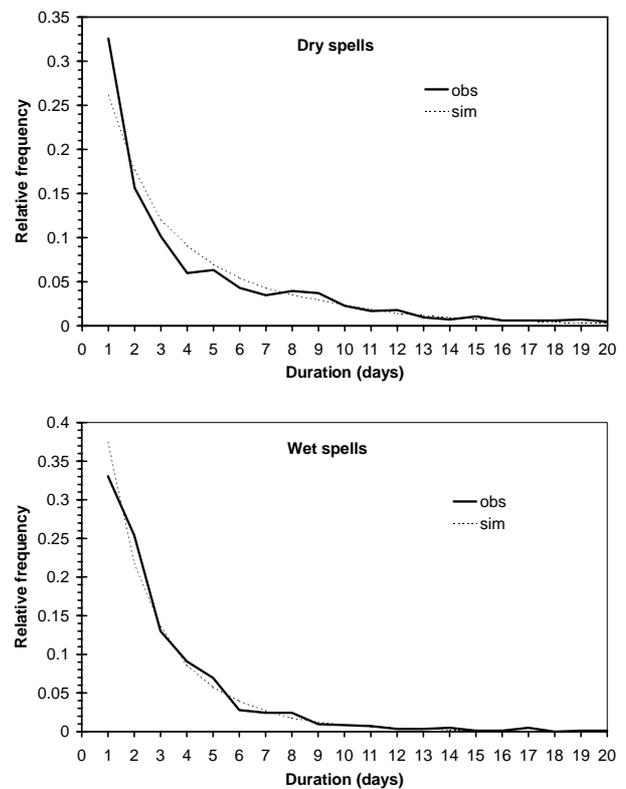


Fig. 9. Observed and simulated relative frequencies of duration of dry and wet spells in the winter half of the year for the period 1946 to 1992. Simulations are based on the sequence of daily circulation patterns simulated by HADCM2CON

characteristic also present when observational data is used.

In summary, it can be concluded that the weather generator coupled with daily circulation patterns classified from daily MSLP fields simulated by HADCM2CON can reproduce well 'present climate' in southern Portugal in terms of local daily precipitation. Therefore, by using the developed weather generator driven by daily circulation patterns obtained from the 2 transient experiments, HADCM2GHG and HADCM2SUL, the prospect of obtaining local precipitation scenarios in southern Portugal for impact studies is good.

#### 4. CONCLUSIONS

In this paper, a weather generator, dependent on 4 daily circulation patterns (weather regimes) related to precipitation in southern Portugal and identified from MSLP fields, was designed, with 3 different schemes for the simulation of rain occurrence, in order to simulate daily precipitation in the same region. It seems that simulations can model the autocorrelation struc-

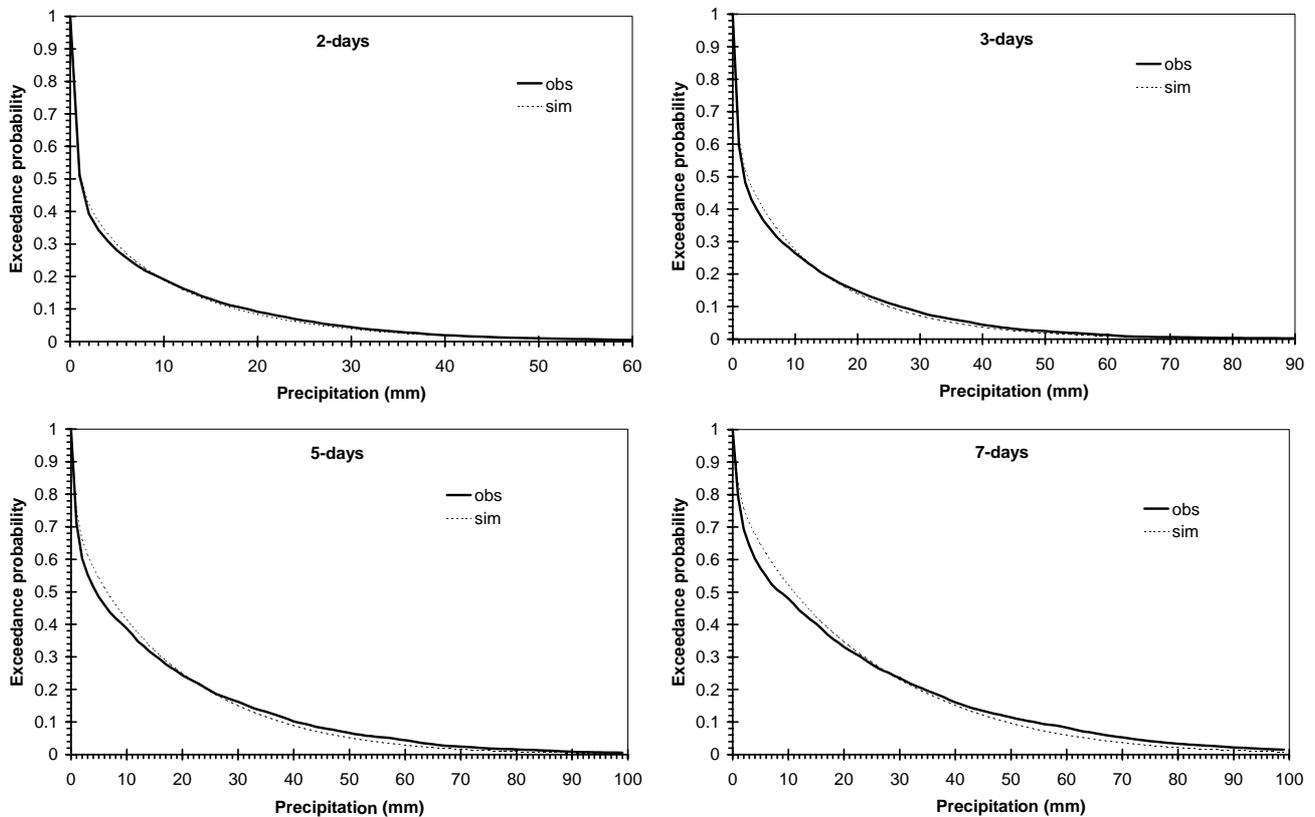


Fig. 10. Observed and simulated distributions of exceedance probabilities of precipitation totals in 2, 3, 5 and 7 consecutive days in the winter half of the year for the period 1946 to 1992. Simulations are based on the sequence of daily circulation patterns simulated by HADCM2CON

ture of the sequence of wet and dry days better when the probability of rain is considered to depend not only on the current circulation pattern but also on the weather state and weather regime of the previous day. In most cases, in the winter half of the year, daily rainfall amounts fit well to the gamma distribution. Validation of the weather generator, in a period independent of the training period, indicates that the weather generator is able to simulate fundamental characteristics of daily precipitation, even features of extreme precipitation, such as the distributions of dry and wet spell durations.

It was demonstrated that HADCM2 in its control run can reproduce well the 4 circulation patterns that affect precipitation in southern Portugal. By using the weather generator driven by daily circulation patterns classified from HADCM2CON's daily MSLP fields, the major structures of daily precipitation at Évora can be reproduced well for the 'present climate'. Therefore, it is hoped that the weather generator, conditioned by the weather regimes simulated by HADCM2GHG and HADCM2SUL, is a reasonably good tool to obtain local precipitation scenarios in southern Portugal for impact studies. It should be stressed that such expectation

relies on the assumption that in a changing climate the observed relationship between local precipitation and weather regimes is preserved, with essentially minor modifications in the weather generator's parameters. Unfortunately, no firm background for such an assumption seems to have been offered.

The procedures of calibration and validation of the weather generator are also suitable for other locations; however, success can only be expected at those sites where local precipitation and large-scale atmospheric circulation are closely related. Furthermore, at those sites, the weather generator can also be used to obtain future precipitation scenarios, under the same assumption mentioned above.

*Acknowledgements.* This research was sponsored by EC project MEDALUS III, under contract ENV4-CT95-0121. The authors are very grateful to Drs Xiaolan Wang and Xuebin Zhang for helpful discussions as well as for stimulating ideas in their PhD theses. B.Q. thanks Professors Peng Gongbing, Fan Zhongxiu and Zhou Enji for their strong encouragement. The authors appreciate very much the critical and constructive comments from 3 anonymous reviewers on the original manuscript.

### Appendix 1

Definitions of different conditional probabilities of rain. In this paper, different conditional probabilities of rain are used in different schemes for simulating the occurrence of precipitation. They are defined as follows:

$$p_{11}^i = P\{Y(t) = 1 \mid Y(t-1) = 1, A(t-1) = A(t) = CP^i\}$$

$$p_{01}^i = P\{Y(t) = 1 \mid Y(t-1) = 0, A(t-1) = A(t) = CP^i\}$$

$$pn_{11}^i = P\{Y(t) = 1 \mid Y(t-1) = 1, A(t-1) \neq A(t) = CP^i\}$$

$$pn_{01}^i = P\{Y(t) = 1 \mid Y(t-1) = 0, A(t-1) \neq A(t) = CP^i\}$$

$$pp_{11}^i = P\{Y(t) = 1 \mid A(t) = CP^i, Y(t-1) = 1\}$$

$$pp_{01}^i = P\{Y(t) = 1 \mid A(t) = CP^i, Y(t-1) = 0\}$$

$$pr^i = P\{Y(t) = 1 \mid A(t) = CP^i\}$$

The superscript  $i$  refers to daily circulation patterns, from 1 to 4.  $CP^1$ ,  $CP^2$ ,  $CP^3$  and  $CP^4$  are the blocking-like, the summer dry, the winter dry and the rainy pattern, respectively

### Appendix 2

Three schemes for simulating the occurrence of rain. The physical meanings of the 3 different schemes are discussed in the paper. Here, only mathematical details are given for each scheme.

Scheme 1:

$$\text{If } Y(t-1) = 1 \text{ and } A(t-1) = A(t) = CP^i, Y(t) = 1 \text{ if } \omega \leq p_{11}^i$$

$$\text{If } Y(t-1) = 0 \text{ and } A(t-1) = A(t) = CP^i, Y(t) = 1 \text{ if } \omega \leq p_{01}^i$$

and

$$\text{If } Y(t-1) = 1 \text{ and } A(t-1) \neq A(t) = CP^i, Y(t) = 1 \text{ if } \omega \leq pn_{11}^i$$

$$\text{If } Y(t-1) = 0 \text{ and } A(t-1) \neq A(t) = CP^i, Y(t) = 1 \text{ if } \omega \leq pn_{01}^i$$

Scheme 2:

$$\text{If } Y(t-1) = 1 \text{ and } A(t) = CP^i, Y(t) = 1 \text{ if } \omega \leq pp_{11}^i$$

$$\text{If } Y(t-1) = 0 \text{ and } A(t) = CP^i, Y(t) = 1 \text{ if } \omega \leq pp_{01}^i$$

Scheme 3:

$$\text{If } A(t) = CP^i, Y(t) = 1, \text{ if } \omega \leq pr^i$$

### LITERATURE CITED

- Bardossy A, Plate EJ (1991) Modelling daily rainfall using a semi-Markov representation of circulation pattern occurrence. *J Hydrol* 122:33–47
- Bardossy A, Plate EJ (1992) Space-time model for daily rainfall using atmospheric circulation patterns. *Water Resour Res* 28:1247–1259
- Bogardi I, Matyasovszky I, Bardossy A, Duckstein L (1993) Application of a space-time stochastic model for daily precipitation using atmospheric circulation patterns. *J Geophys Res* 98(D9):16653–16667
- Coe R, Stern D (1982) Fitting models to daily rainfall data. *J Appl Meteorol* 21:1024–1031
- Corte-Real J, Zhang X, Wang X (1995) Downscaling GCM information to regional scales: a non-parametric multivariate regression approach. *Clim Dyn* 11:413–424
- Corte-Real J, Qian B, Xu H (1998) Regional climate change in Portugal: precipitation variability associated with large-scale atmospheric circulation. *Int J Climatol* 18:619–635
- Crutcher HL (1975) A note on the possible misuse of the Kolmogorov-Smirnov test. *J Appl Meteorol* 14:1600–1603
- Cubasch U, von Storch H, Waszkewitz J, Zorita E (1996) Estimates of climate change in Southern Europe derived from dynamical climate model output. *Clim Res* 7:129–149
- Giorgi F (1990) Simulation of regional climate using a limited-area model nested in a general circulation model. *J Clim* 3:941–963
- Giorgi F, Marinucci MR, Visconti G (1992) A  $2 \times CO_2$  climate change scenario over Europe generated using a limited area model in a general circulation model. II: climate change scenario. *J Geophys Res* 97:10011–10028
- Giorgi F, Shields-Brodeur C, Bates GT (1994) Regional climate change scenarios over the United States produced with a nested regional model. *J Clim* 7:375–399
- Greenwood JA, Durand D (1960) Aids for fitting the gamma distribution by maximum likelihood. *Technometrics* 2:55–65
- Gyalistras D, von Storch H, Fischlin A, Beniston M (1994) Linking GCM-simulated climatic changes to ecosystem models: case studies of statistical downscaling in the Alps. *Clim Res* 4:167–189
- Hay LE, McCabe GJ Jr, Wolock MD, Ayers MA (1992) Use of weather types to disaggregate general circulation model predictions. *J Geophys Res* 97:2781–2790
- Houghton JT, Jenkins GJ, Ephraums JJ (eds) (1990) *Climate change: the IPCC scientific assessment*. Cambridge University Press, Cambridge
- Houghton JT, Callander BA, Varney SK (eds) (1992) *Climate change 1992: the supplementary report to the IPCC scientific assessment*. Cambridge University Press, Cambridge
- Houghton JT, Meira Filho LG, Callander BA, Harris N, Kattenberg A, Maskell K (eds) (1996) *Climate change 1995. The science of climate change. Contribution of Working Group I to the Second Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge
- Johns TC, Carnell RE, Crossley JF, Gregory JM, Mitchell JFB, Senior CA, Tett SFB, Wood RA (1997) The second Hadley Centre coupled ocean-atmosphere GCM: model description, spinup and validation. *Clim Dyn* 13:103–134
- Jones RG, Murphy JM, Noguer M (1995) Simulation of climate change over Europe using a nested regional climate model. I: Assessment of control climate, including sensitivity to location of lateral boundaries. *QJR Meteorol Soc* 121:1413–1449
- Katz RW (1981) On some criteria for estimating the order of a Markov chain. *Technometrics* 23:243–249
- Katz RW, Parlange MB (1996) Mixtures of stochastic processes: application to statistical downscaling. *Clim Res* 7:185–193
- Lettenmaier D (1996) Stochastic modeling of precipitation with applications to climate model downscaling. In: von Storch H, Navarra A (eds) *Analysis of climate variability: applications of statistical techniques*. Springer, Berlin p 197–212
- Lilliefors HW (1967) On the Kolmogorov-Smirnov test for normality with mean and variance unknown. *J Am Stat Assoc* 62:399–402
- Maytín CE, Acevedo MF, Jaimez R, Andressen R, Harwell MA, Robock A, Azócar A (1995) Potential effects of global climatic change on the phenology and yield of maize in Venezuela. *Clim Change* 29:189–211
- Mearns LO, Rosenzweig C, Goldberg R (1992) Effect of

- changes in interannual variability on CERES-wheat yields: sensitivity and  $2 \times \text{CO}_2$  general circulation model studies. *Agric For Meteorol* 62:159–189
- Schubert S (1994) A weather generator based on the European 'Grosswetterlagen'. *Clim Res* 4:191–202
- Schubert S, Henderson-Sellers A (1997) A statistical model to downscale local daily temperature extremes from synoptic-scale atmospheric circulation patterns in the Australian region. *Clim Dyn* 13:223–234
- Semenov MA, Porter JR (1995) Climatic variability and the modelling of crop yields. *Agric For Meteorol* 73:265–283
- Semenov MA, Brooks RJ, Barrow EM, Richardson CW (1998) Comparison of the WGEN and LARS-WG stochastic weather generators for diverse climate. *Clim Res* 10:95–107
- Stern D, Coe R (1984) A model fitting analysis of daily rainfall data (with discussion). *J R Stat Soc Ser A* 147:1–34
- von Storch H, Zorita E, Cubasch U (1993) Downscaling of global climate change estimates to regional scales: an application to Iberian rainfall in wintertime. *J Clim* 6:1161–1171
- Watson RT, Zinyowera MC, Moss RH (eds) (1996) *Climate change 1995: impacts, adaptations and mitigation of climate change: scientific-technical analyses*. Contribution of Working Group II to the Second Assessment Report of Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge
- Wilks DS (1992) Adapting stochastic weather generation algorithms for climate change studies. *Clim Change* 22:67–84
- Wilks DS (1995) *Statistical methods in the atmospheric sciences: an introduction*. Academic Press, San Diego
- Wilson LL, Lettenmaier DP, Wood EF (1991) Simulation of daily precipitation in the Pacific Northwest using a weather classification scheme. *Sur Geophys* 12:127–142
- Wilson LL, Lettenmaier DP, Skillingstad E (1992) A hierarchical stochastic model of large scale atmospheric circulation patterns and multiple station daily precipitation. *J Geophys Res* 97:2791–2809
- Zhang X, Wang XL, Corte-Real J (1997) On the relationships between daily circulation patterns and precipitation in Portugal. *J Geophys Res* 102:13495–13507

*Editorial responsibility: Hans von Storch,  
Geesthacht, Germany*

*Submitted: July 30, 1998; Accepted: May 10, 1999  
Proofs received from author(s): August 5, 1999*