CONDITIONAL STOCHASTIC MODEL TO GENERATE DAILY PRECIPITATION TIME SERIES

Aristita Busuioc* and Hans von Storch**
*National Institute of Meteorology and Hydrology, Bucharest, Romania,
e-mail:busuioc@meteo.inmh.ro
**Institute of Costal Research, GKSS, Geest hacht, Germany, e -mail:storch@gkss.de

Abstract
The purpose of the present paper is the building of a conditional stochastic model to generate daily precipitation time series. The model is a mixture of a two -state first order Markov chain and a statistical downscaling model based on the canonical correlation analysis (CCA). The CCA model links the large-scale circulation, represented by the European SLP field, with the four precipitation distribution parameters: transition probabilities and gamma distributi on parameters. This model has been tested for the Bucharest station at which long observed daily time series were available (1901-1999). The comparison between the skill of the conditional stochastic model and unconditional stochastic model (based only on a Markov chain) is presented using an ensemble of 1000 runs of the two models.
The performance of the conditional stochastic model is analyzed regarding the skill of the CCA model in estimation of the four precipitation distribution parameters and the sto chastic's model performance in reproducing the statistical features of the generated precipitation time series. The CCA model is most skillful for winter and autumn (transition probabilities), slightly skillful for the mean precipitation amount on rainy da ys and unskillful for the shape parameter. There are no significant dissimilarities between the conditional and unconditional models regarding their performance except for the linear trend, which is better captured by the conditional model. Some statistical features are well reproduced by both stochastic models for all seasons, while other statistical features are only partially reproduced by both models or are better reproduced by one of the models. In conclusion, the conditional stochastic model presented in this paper can be successfully used to generate daily precipitation time series for winter and autumn. For the other seasons, the unconditional model can be used to reproduce some statistical features.
Keywords: stochastic model, daily precipitation, M arkov chain, canonical correlation

1. Introduction

Hydrological and crop models usually require daily precipitation time series as input. To evaluate the sensitivity of these models to long -term changes in the precipitation regime an ensemble of input da ta sets are needed. The observed sequences provide only one realization of the weather process. In impact studies that use as input data precipitation time series derived from the simulated climate change scenarios, the number of these sequences are still limited due to high computational cost of these scenarios. To evaluate the range of results that may be obtained with other statistically equivalent series it is desirable to generate synthetic sequences of precipitation data based on the stochastic struct ure of the meteorological process. Richardson (1981) presented such a technique to simulate daily values of precipitation, maximum and minimum temperature, and solar radiation. For the precipitation component, a two state first -order Markov chain has been used to describe the precipitation occurrence and the exponential distribution has been used to approximate the distribution of rainfall amount. This model has also been used by Wilks (1992) with gamma distribution instead of exponential distribution. In this case the model has been adapted for climate change studies.

Such models may be conditioned on large -scale meteorological conditions, which incorporate the cause -effect information about the probability of wet, or dry, conditions. There are various ways to define the large-scale conditions such as: large -scale circulation indexes

(Katz and Parlange ,1993) circulation classification (Zorita et al., 1995; Goodess and Palutikov, 1998) and analogs (Zorita et al., 1995). Lettenmaier (1995) showing the advant ages and disadvantages of various models has presented a synthesis of the stochastic models of precipitation (conditional and unconditional). Non -linear approaches, such as neural networks, have recently been developed (Cavazos, 1999; Zorita and von Storch , 1999). In a recent synthesis of empirical downscaling methods used in synoptic climatology, Yarnal et al. (2001) have discussed the advantages and disadvantages of stochastic models. They found , in the case of climate change, the conditioning of the sto chastic parameters in a physical meaningful way difficult to achieve.

In this paper a stochastic model conditioned upon large - scale climate characteristics to generate daily precipitation amount is presented. The model uses a first order Markov chain combined with a downscaling model. To link the precipitation distribution parameters with the large-scale circulation, represented by the sea level pressure on the European scale (SLP), a regression model based on the canonical correlation analysis (CCA) has b een used (von Storch et al., 1993; Heyen et al., 1996; Busuioc et al., 1999, 2001). In this way an adjustment of stochastic parameters in a some physical meaningful way is proposed. The model has been tested for Bucharest station at which long daily observ ations are available (1901 -1999). More details about the methodology are presented in Section 2. The skill of the downscaling model in estimation the parameters of precipitation distribution, and the skill of the conditional stochastic model in reproducing the most important statistical features of the generated precipitation time series, are shown in Section 3. The comparison between the skill of the conditional and the unconditional stochastic model is also presented. Compared to other similar models the procedure presented in this paper gives the confidence intervals of the precipitation distribution parameters derived from 1000 Monte Carlo runs of both models. The conclusions are presented in Section 4.

2. Methodology

The model presented in this paper i s a combination between a first order Markov chain and a statistical downscaling model. In the following it is referred to as *conditional stochastic model.* Additionally, *the unconditional stochastic model* based only on a first order Markov chain is used in order to assess the performance of the conditional model for climate change proposes. Both models are tested for the daily precipitation amount at the Bucharest station placed in the southern part of Romania. Observational data refers to the interval 1901 -1999 and they are seasonally stratified: Winter (December -February), Spring (March -May), Summer (June-August), Autumn (September -November).

2.1. Unconditional model

Precipitation occurrence is described by a two -state, first order Marcov chain. The precipitation either occurs or it does not (the two states) and the conditional probability of precipitation occurrence depends only on the occurrence on the previous day. There are two parameters describing the precipitation occurrence process: the transition probability p_{01}, the probability of a wet day following a dry day, and p_{11}, the probability of a wet day following a wet day. As a wet day the case of daily precipitation amount > 0.1 mm is considered in this paper. The choice of the optimum precipitation threshold is an important decision. Dobi -Wantuch et al. (2000) have analyzed the threshold influence on the results and found the 0.1 threshold as appropriate one.

The variation of precipitation amount on wet days is described by the gamma distribution which has two parameters: the shape parameter (k) and the scale parameter (β) (Coe and Stern, 1982; Wilks, 1992). In terms of the two distribution parameters, the mean precipitation amount (considering only wet days) is $\mu=k\beta$. In this paper μ and k are consi dered the gamma distribution parameters. μ is estimated as the sample mean from the observed data set and k is derived as solution of the equation,

$$\ln(k) - \psi(k) = \ln(\bar{x}) - \overline{\ln(x)}$$

where $\bar{x} = [\sum_{i,t} x_i(t)]/n$, $\psi(x)$ is the first derivate of the log Gamma function and $x(t)$ represents in our case the daily precipitation amount for wet days.

The p_{01}, p_{11} transition probabilities were estimated from observed data set. Therefore, the stochastic model to generate daily precipitation depends on four parameters (p_{01}, p_{11}, μ and k) and it is referred in the following as *unconditional model*. The four parameters were computed over the complete interval for every month and their seasonal variation was revealed. This was the reason to divide the data set into the four seasons and to build the model for every season. The highest variability for μ and lowest for the transition probabilities was identified.

In conclusion, the daily precipitation for wet days are generated (in the case of unconditional model) using the four parameters computed from observation over a fixed interval following the procedure presented by Wilks (1992). Knowing whether precipitation occurred on the previous simulated day, the appropriate transition probability, p_{01} or p_{11} is compared to a newly generated uniform [0, 1] random number. A wet day is simulated if the random number is less than the transition probability. If this is the case, a random precipitation amount is generated for the current day using the appropriate gamma distribution. Since the four parameters are kept constant stationary precipitation time series are generated. This is a disadvantage of the unconditional model when used in the climate context. However, this drawback can be overcome by using the conditional stochastic model, which links unconditional stochastic model parameters to large-scale circulation parameter as described below.

2.2. Conditional stochastic model

To link the four parameters of the unconditional stochastic precipitation model to the large-scale circulation a regression model has been constructed with the help of a canonical correlation analysis (CCA; von Storch et al., 1993; Heyen et al., 1996; Werner and von Storch, 1993; Busuioc et al., 1999, 2001). European-scale sea level pressure anomalies are chosen as a representing the large-scale circulation. The monthly SLP data with a resolution of 5^{o}x5^{o} have been provided by the National Centre of Atmospheric Research (NCAR, USA) (Trenberth and Paolino, 1980). The SLP area of 5^{o}-50^{o}E and 30^{o}-55^{o}N was selected, so that the skill of the downscaling model linking seasonal precipitation over Romania and SLP is maximum. It has been shown earlier that seasonal precipitation amount in Romania is strongly connected with the European SLP distribution, especially for winter and autumn (Busuioc and von Storch, 1996; Busuioc et al., 1999). Therefore a strong connection between the four parameters characterizing the daily precipitation distribution and the SLP field were assumed to prevail. In the Section 3 it will be shown that this assumption is correct, especially for winter and autumn.

The stochastic parameters (p_{01}, p_{11}, μ and k) are computed for every season from 90 to 92 daily precipitation amounts in every year. In this way, a time series of the parameters is obtained. Prior to the CCA the four parameters have been standardized. The CCA determines pairs of patterns of two-time-dependent variables (the large-scale SLP and the four stochastic parameters) so that their time components are optimally correlated. Prior to the CCA, the original data are standardized, by subtracting the mean from the original value and by dividing with the standard deviation. The SLP data are projected onto their EOFs (Empirical Orthogonal Functions) to eliminate noise (small-scale features) and to reduce the dimension of the data. Since the time coefficients are normalized to unity the canonical correlation patterns represent the typical strength of the signals. A subset of CCA pairs is then used in a regression model to estimate the four stochastic parameters from the large-scale SLP. The precipitation distribution parameters (p_{01}, p_{11}, μ, k) estimated through the CCA model are then used in the stochastic model in order to generate daily precipitation amounts. These time series are achieved for every season in every year. Since the four parameters should satisfy some conditions ($0 \leq p_{01}, p_{11} \leq 1$ and μ, $k > 0$) the CCA model outputs are processed by applying the reversed operation of standardization before to be used in the stochastic model.

2.3. Skill of stochastic model

The full data set 1901-1999 was split in two intervals, 1901-1949 and 1950-1999. Then, both unconditional and conditional model were fitted one interval and validated with the other interval, so that two models were fitted and validated independently. The skill of the CCA downscaling model is expressed through the variance explained by the reconstructed values as a fraction from the total variance of the observed values or, alternatively, by the correlation between observed and reconstructed values. Additionally, the performance of the models was determined in terms of how well the model reproduces the statistical features of the precipitation time series:

- Appearance/nonappearance of precipitation quantified by: mean and expected maximum duration of wet and dry intervals;
- Daily mean and standard deviation of precipitation for rainy days, expected maximum daily precipitation amount and frequency distributions of daily precipitation;
- Changes (linear trend) in the seasonal precipitation amount induced. The significance of the linear trend is estimated by the Mann-Kendall statistic (Sneyers, 1975).

These statistical features have been computed for the two subintervals.

In this paper an ensemble of 1000 simulations was generated and the statistical parameters mentioned above were expressed as ensemble means with their 90% confidence intervals computed with a bootstrapping procedure.

3. Results

3.1 Precipitation distribution parameters

The CCA analysis has identified pairs of patterns in the SLP fields and in combined vector of the transition probabilities (p_{01}, p_{11}) and gamma distribution parameters (μ, k) whose time series share a maximum of correlation. The correlation coefficients R_1 and R_2 associated to the first two CCA pairs for the four seasons are presented in Table 1. The explained variance of the seasonal SLP anomalies and four parameter anomalies are also presented. The strongest link has been found for winter and autumn. Similar results were achieved when the direct relationship between the SLP field and seasonal precipitation in Romania was analyzed (Busuioc and von Storch, 1996; Busuioc et al., 1999).

Table 1. Canonical correlation coefficients (R1, R2) (x 100) and explained variance of the first two CCA pairs of seasonal mean SLP and the four parameters of precipitation distribution.

Season	1901-1949						1950-1999					
	R1	R2	Explained variance (%)				R1	R2	Explained variance (%)			
			SLP		Parameters				SLP		Parameters	
			CCA1	CCA2	CCA1	CCA2			CCA1	CCA2	CCA1	CCA2
Winter	77	49	45	7	29	27	78	42	38	21	33	9
Spring	66	53	31	6	24	25	51	40	22	8	29	22
Summer	67	55	11	10	30	22	65	44	13	7	30	21
Autumn	82	48	16	34	34	18	80	52	10	37	29	32

The first CCA pair generally shows similar physical mechanisms even if it is more difficult to be explained in terms of statistics of precipitation distribution such as transition probabilities and gamma distribution parameters. The CCA analysis has been done for two subintervals: 1901-1949, 1950-1999. The first CCA pair is almost identical for the two subintervals for all seasons except for a slight shift and spatial extension of the pattern nucleus. Figure 1 shows an example for the winter season he first two CCA pairs for all seasons.

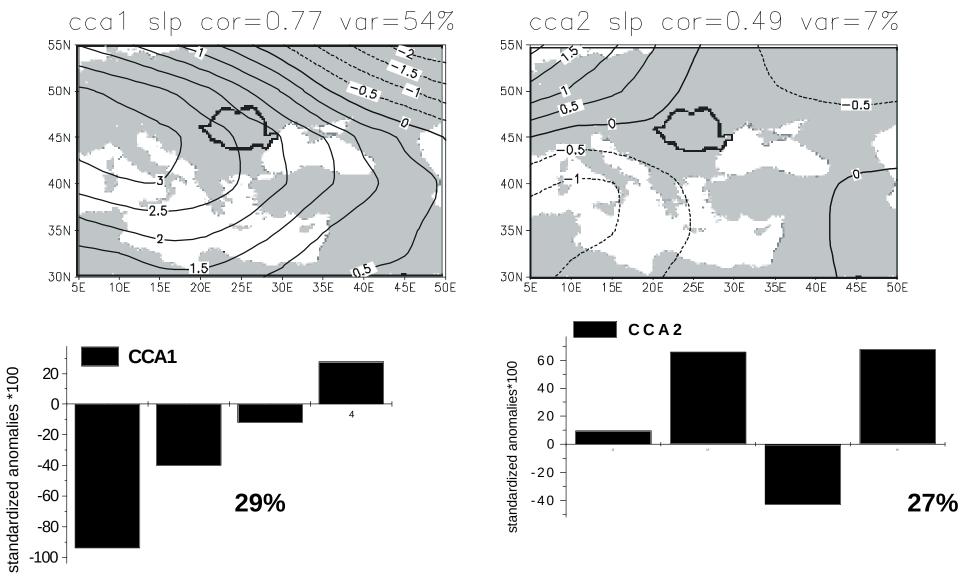


Figure 1. The patterns of the first two CCA pairs of the winter mean SLP and winter parameters of precipitation distribution as derived from observation (1901 -1949).

A southerly /northwesterly circulation over Romania in winter is associated with above/below normal daily mean precipitation within wet days and higher transition probabilities to rainy days at Bucharest. This mechanism seems to be reasonable from a physical point of view; southwesterly circulation brings moisture air mass from Mediterranean basin to Romania (especially in the southern part where Bucharest station is located) and more precipitation is recorded. As a result, the daily precipitation mean on rainy days (μ) and the probability to have a wet day ($p_{01}+p_{11}$) is higher. However, the link to μ is mostly weak in all seasons, except for autumn. This seems to contradict the presence of strong link between SLP and total seasonal precipitation amount (Busuioc and von Storch, 1996; Busuioc et al., 1999). However, these two observations may be reconciled by noting that the monthly total is dominated by the number of wet days and less dependent on the mean amount on wet days. In fact, the correlation between μ and first SLP EOF (very similar to the first CCA pattern) time series is low, while the correlation with the number of rainy days is high. The link between SLP variations and k variability is unclear, with the sign of the link changing, when different fitting periods are used.

The SLP pattern of the second CCA pair is different over the two subintervals for winter (not shown) but for the other seasons they are stable apart of slight shifts of the center of the SLP pattern (not shown). The mechanisms given by the second CCA pair can not be physically interpreted in a manner presented for the first CCA pair.

The skill of the statistical downscaling model built up by using the time coefficients associated to the four CCA pairs is presented in Table 2. The skill is calculated for the subintervals not used to fit the statistical model. The model is most skillful for winter and autumn (transition probabilities), slightly skillful for μ (winter, especially for 1951-1999 interval) and unskillful for the shape parameter k.

Table 2. Skill of the CCA model (expressed as percentage of explained variance/correlation coefficient) for estimation of the four parameters (P_{01}, P_{11}, μ, k) from the SLP field over the two subintervals considered as independent data set.

Season	1901-1949				1950-1999			
	P_{01}	P_{11}	μ	k	P_{01}	P_{11}	μ	k
Winter	45 / 67	16 / 43	4 / 23	-	39 / 63	24 / 49	11 / 33	-
Spring	-	7 / 26	-7 / 26	-7 / 21	9 / 30	-	-3 / 18	-
Summer	-19 / 14	8 / 30	-	-	2 / 39	6 / 34	-	-
Autumn	50 / 71	21 / 47	-14 / 14	-	51 / 72	15 / 38	-12 / 8	-

Figure 2 shows, as an example, the temporal evolution of the observed and estimated standardized anomalies for winter. As it can be seen the two curves vary coherently for the transition probabilities, the amplitude is sometimes different but the year to year evolution is quite good for μ, but for the shape parameter k the dissimilarities are substantial.

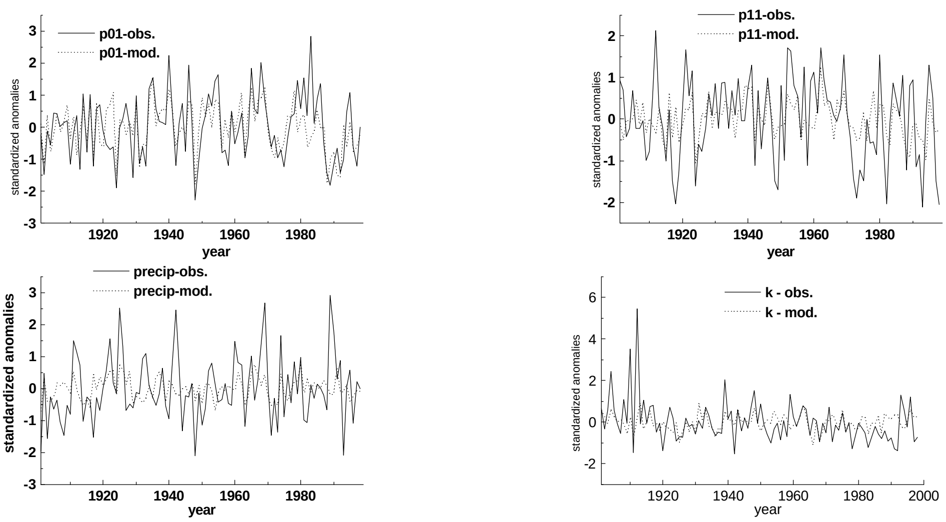


Figure 2. Winter standardized anomalies of the precipitation distribution parameters for the 1901-1999 as derived from observation (solid line) and as derived indirectly from the observed European-scale SLP anomalies by using the downscaling model (dashed line) fitted over the 1901-1949 interval.

3.2. The skill of the stochastic model

The four parameters (p_{01}, p_{11}, μ and k), estimated directly from precipitation data (case of unconditional model) and indirectly from European SLP through the statistical downscaling model presented above (case of conditional model) are then used in the Markov chain model to generate daily time series with appearance and nonappearance of precipitation. The daily precipitation amount is randomly generated using a gamma distribution. The performance of these stochastic models is assessed in terms of how well they reproduce the statistical features of the precipitation time series presented in Section 2. These features are represented by: maximum duration of dry and wet intervals ($\mathbf{d}_{\mathbf{dry}}^{\max}, \mathbf{d}_{\mathbf{wet}}^{\max}$), mean duration of dry and wet intervals ($\mathbf{d}_{\mathbf{dry}}^{\mathbf{mean}}, \mathbf{d}_{\mathbf{wet}}^{\mathbf{mean}}$), daily mean /standard deviation of precipitation within rainy days (pp_{mean}, pp_{sd}), mean number of rainy days (nr), Mann-Kendall statistic (t), expected maximum of daily precipitation amount (pp_{max}) and frequency distributions of daily precipitation within various intervals. After running of unconditional and conditional models 1000 times a distribution of these parameters is achieved. Then, the ensemble mean and their 90% confidence intervals of the respective parameters were computed and these values are considered as expectations for these parameters.

These statistics derived through both stochastic models and directly from observation are presented in Tables 3 and 4. They were computed separately for the two sub intervals. The mean duration and expected maximum duration of wet intervals are very well simulated for all seasons, both subintervals and both models (unconditional and conditional); there are no significant differences between unconditional and conditional models. The expected maximum duration of dry intervals is better estimated by the unconditional model and overestimated by the conditional model even if for both models the observed values are generally covered by the 90% confidence intervals. Mean duration of the dry intervals is well simulated by both models for winter and summer (1901-1949) and autumn (1950-1999). Both models except for spring (both intervals) overestimate the observed values.

The seasonal mean of the number of rainy days is generally well reproduced by both stochastic models, except for spring (1901-1949, both models and 1950-1999, conditional model) and autumn (1950-1999, both models) when the number is underestimated by the conditional model and overestimated by the unconditional model. The daily mean of precipitation for wet days is generally well reproduced by both stochastic models (with small differences for the 1950-1999 interval) while the standard deviation is better estimated for the 1901-1949 interval (both models), except for summer when it is underestimated. For the 1951-1999 interval both models underestimate the standard deviation for all seasons (less for summer). Expected maximum daily precipitation is generally underestimated for all seasons (both models) but the observed values are covered by the 90% confidence intervals, except for summer (1901-1949) and winter and autumn (1950-1999).

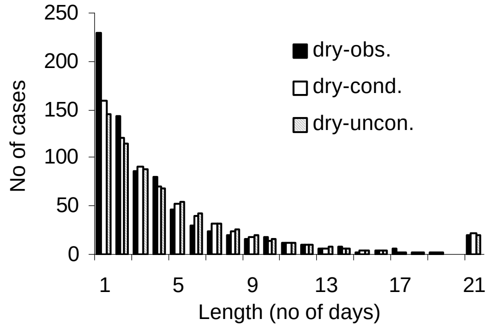
In order to assess if the stochastic model reproduces the linear trend of the observed precipitation time series, the Mann-Kendall statistic (***t***) was computed for the seasonal precipitation amount as derived directly from observation and indirectly from generated daily time series. A ***t***-value greater than 1.96 allows the rejection of the null hypothesis with a risk of 5% and less, provided that the data are not serially correlated. This ***t***-statistic was computed for every experiment of the 1000 run ensemble. Table 3 contains the ensemble mean of this statistic with its 90% confidence interval.

As it was expected the unconditional model does not show any significant change. The conditional model reproduces better the winter trend even if it is underestimated for the first interval and overestimated for the second one. In springtime the conditional model also overestimates the trend. One reason explaining these results could be the failure of the CCA model in estimating of the **k** parameter and the relatively low skill for μ. Another reason could be the urbanization effect in Bucharest. For the 1901-1949 interval the winter European SLP variation, given by the time series associated to the first EOF, reveals a slightly decreasing trend (not significant using the Mann-Kendall test) that leads to simulated precipitation which are close to normal. For the 1950-1999 interval the situation is reversed: the time series associated to the first SLP EOF present a very strong decreasing trend (associated to less frequent southwesterly circulation over Romania) that induces less precipitation, but much less than observations where the urbanization effect is added, which is associated with an increase in precipitation

Figure 3 shows the frequencies of dry and wet intervals with various lengths derived from observation and from unconditional and conditional stochastic models as ensemble means over 1000 runs. Only the results for winter are displayed. The frequencies of the extreme events such as dry intervals longer than 15 days are separately presented. Generally, there are no significant differences between the results achieved with unconditional and conditional models, except for spring (short intervals) when the unconditional model is better. For the other seasons the unconditional model is slightly better than conditional one for shorter dry intervals. The best agreement with observations is obtained for winter and autumn (especially for wet intervals). For all seasons, the dry intervals of one - two day length are less frequent in simulation then in reality. Generally, the shorter dry intervals are underestimated and the longer dry intervals (greater than 9 days) are overestimated. The frequency of the extreme events is very well reproduced for winter and summer.

Table 3 Statistics of precipitation regime (maximum duration of dry and wet intervals – ***d^max_dry, d^max_wet,*** mean duration of dry and wet interva ls-, ***d^mean_dry, d^mean_wet,*** daily mean /standard deviation of precipitation within rainy days -p***p_mean / pp_var***; mean number of rainy days - ***nr***; Mann-Kendall statistic - ***t***, expected maximum of daily precipitation amount – ***pp_max***) at the Bucharest station derived directly from observations and indirectly through the stochastic conditional model (*cond*) and stochastic unconditional model (*uncond.*). These statistics are computed as ensemble means for 1000 runs and they are derived over the two subintervals consi dered as independent data set. The 90% confidence intervals for the parameters derived through conditional and unconditional models are also presented.

Season		d^max_dry	d^max_wet	d^mean_dry	d^mean_wet	Pp_mean	Pp_var	Nr	t	Pp_max
					1901-1949					
Winter	Obs.	24	10	4.3	2.1	3.8	5.4	28	2.4	58.6
	Cond.	35	12	4.3	2.1	4.1	5.1	30	1.0	44.1
		25, 51	9, 16	4.1, 4.5	2.0, 2.2	3.9, 4.4	4.7, 5.5	28, 31	-0.2, 2.2	33.6, 60.4
	Uncond	27	12	4.3	2.1	4.2	5.2	29	0.02	45.1
		21, 36	9, 16	4.1, 4.6	2.0, 2.2	4.0, 4.4	4.9, 5.7	29, 31	-1.6, 1.7	34.5, 60.1
Spring	Obs.	25	13	3.7	2.0	4.6	6.4	31	-0.6	61.6
	Cond.	33	12	5.0	2.1	4.9	6.3	27	1.4	57.8
		25, 45	9, 16	4.7, 5.3	2.0, 2.2	4.6, 5.2	5.8, 6.9	26, 29	0.2, 2.7	42.8, 81.9
	Uncond	24	12	3.9	2.2	4.9	6.3	34	0.01	54.8
		19, 33	10, 16	3.7, 4.1	2.1, 2.3	4.7, 5.2	5.9, 6.7	32, 3.5	-1.6, 1.7	41.6, 72.8
Summ	Obs.	29	9	4.0	1.9	7.0	10.8	28	0.1	136.6
	Cond.	27	10	3.9	1.8	6.7	8.6	29	0.9	75.6
		20, 35	8, 13	3.8, 4.1	1.8, 1.9	6.3, 7.1	7.9, 9.3	28, 30	-0.6, 2.3	56.7, 101.6
	Uncond	25	10	4.0	1.8	6.7	8.7	29	0.5	74.4
		20, 34	8, 13	3.8, 4.2	1.8, 1.9	6.3, 7.1	8.0, 9.3	28, 31	-0.9, 2.0	55.8, 99.7
Autunn	Obs.	30	12	4.7	2.0	5.2	7.7	25	-1.1	59.6
	Cond.	40	10	6.0	1.9	6.0	7.6	23	0.1	66.8
		29, 54	8, 14	5.3, 5.6	1.8, 2.0	5.6, 6.3	6.9, 8.4	22, 25	-1.0, 1.2	49.0, 91.9
	Uncond	36	10	5.6	1.9	5.9	7.6	23	0.1	63.6
		27, 47	8, 13	5.3, 6.0	1.8, 2.0	5.5, 6.3	7.0, 8.2	22, 25	-1.3, 1.5	48, 85.7
					1950-1999					
Winter	Obs.	31	11	4.0	2.1	4.2	6.0	29	-1.0	60.9
	Cond.	36	12	4.5	2.1	3.8	4.4	29	-1.8	39.1
		26, 51	9, 16	4.3, 4.7	2.0, 2.2	3.6, 4.0	4.1, 4.8	28, 30	-3.0, -	29.0, 54.2
	Uncond	29	12	4.6	2.2	3.8	4.6	29	0.02	38.5
		23, 39	9, 16	4.3, 4.8	2.1, 2.3	3.6, 4.0	4.3, 4.9	28, 31	-1.7,	30.0, 51.2
Spring	Obs.	28	12	3.7	2.1	4.9	7.2	33	-0.4	58.8
	Cond.	32	12	5.1	2.0	4.6	5.5	26	-2.0	47.6
		25, 43	9, 16	4.8, 5.3	1.9, 2.1	4.3, 4.8	5.1, 5.9	25, 28	-3.3, -	35.3, 64.8
	Uncond	25	11	3.9	2.0	4.6	5.7	32	0.02	49.0
		19, 33	9, 15	3.7, 4.1	2.0, 2.1	4.4, 4.8	5.3, 6.1	30, 33	-1.7,	37.1, 65.0
Summ	Obs.	28	10	3.7	1.8	6.7	9.7	29	1.1	88.2
	Cond.	29	11	4.2	1.9	7.0	8.7	29	-0.2	75.3
		22, 40	8, 15	4.0, 4.5	1.8, 2.0	6.6, 7.4	8.1, 9.4	27, 30	-1.6,	57.1, 101.9
	Uncond	27	11	4.3	1.9	7.0	8.8	29	-0.02	75.5
		21, 36	8, 14	4.1, 4.5	1.9, 2.9	6.6, 7.4	8.2, 9.5	27, 30	-1.6,	58.2, 100.2
Autunn	Obs.	31	11	5.1	1.9	5.9	8.8	23	1.1	93.0
	Cond.	39	11	5.0	2.0	5.0	6.2	26	-0.3	53.1
		28, 56	8, 14	4.8, 5.3	1.9, 2.1	4.7, 5.3	5.7, 6.7	24, 27	-1.6,	39.5, 71.2
	Uncond	33	12	5.2	2.1	5.2	6.8	27	0.02	58.0
		25, 44	9, 15	4.9, 5.5	2.0, 2.2	4.9, 5.6	6.2, 7.4	25, 28	-1.5,	43.7, 78.4



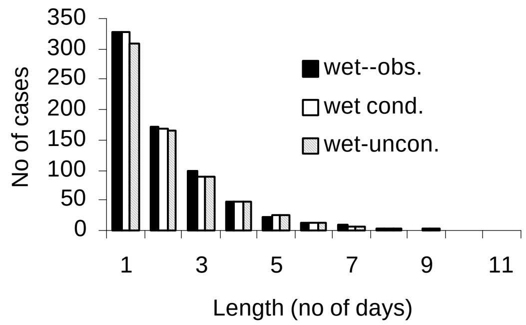


Figure 3. Frequencies of dry (upper) and wet (bottom) intervals with various lengths as derived from observation and from generated time series (with conditional and uncond itional model) at the Bucharest station (winter). The results are obtained for the independent data set over the interval 1950 -1999 with the model fitted over the interval 1901 -1999 as ensemble means over 1000 experiments. The last column for the dry inter vals refers to lengths greater than 15 days.

The performance of the stochastic model in generation of daily precipitation time series was also quantified when the frequencies of days with precipitation amount within or exceeding some thresholds were anal yzed. These results are presented in Table 4. For the 1901 -1949 interval, in spite of some failures of the CCA model for spring and summer, the mean frequencies of days exceeding 20 mm are covered by the 90% confidence intervals of both models for all seas ons. For other thresholds the conditional model is, generally, better except for winter (5 mm - 15 mm) when the values are overestimated by both models and autumn (≤ 5mm) when the values are underestimated. For the interval 1950 -1999 the results are not so good, the observed frequencies are generally underestimated.

Table 4. Seasonal mean frequencies of the daily precipitation amount derived from the observed data set and from the generated time series through conditional and unconditional stochastic model. The values are computed as ensemble means of 1000 runs over the subinterval 1901 -1949 considered as independent data set. The 90% confidence intervals for the conditional and unconditional stochastic models are also presented in parenthesis.

Season		≤ 5 mm	(5mm, 10mm]	(10 mm, 15 mm]	> 10 mm	> 15 mm	> 20 mm
Winter	Obs.	21.4	4.1	1.4	2.8	1.4	0.5
	Cond.	21.5	5.2	1.9	3.2	1.3	0.6
		(20.4, 22.7)	(4.6, 5.8)	(1.6, 2.2)	(2.8, 3.7)	(1.1, 1.6)	(0.4, 0.8)
	Uncond.	21.5	5.1	2.0	3.4	1.5	0.6
		(20.4, 22.8)	(4.5, 5.7)	(1.6, 2.3)	(3.0, 3.9)	(1.2, 1.7)	(0.5, 0.8)
Spring	Obs.	22.1	5.1	1.9	4.0	2.1	1.0
	Cond.	18.6	4.8	2.0	3.9	1.9	0.9
		(17.5, 19.8)	(4.3, 5.4)	(1.7 2.4)	(3.4 4.4)	(1.5 2.2)	(0.7 1.2)
	Uncond.	22.7	5.9	2.6	5.0	2.4	1.2
		(21.5 23.9)	(5.3 6.5)	(2.2 2.9)	(4.4 5.5)	(2.1 2.8)	(0.9 1.4)
Summer	Obs.	17.0	5.1	2.4	6.2	3.8	2.3
	Cond.	17.1	5.5	2.8	6.5	3.7	2.2
		(16.2 18.0)	(5.0 6.1)	(2.4 3.2)	(5.8 7.1)	(3.2 4.1)	(1.8 2.5)
	Uncond.	17.4	5.4	2.8	6.5	3.8	2.2
		(16.4 18.4)	(4.8 6.0)	(2.4 3.2)	(5.9 7.1)	(3.3 4.2)	(1.9 2.6)
Autumn	Obs.	17.3	3.9	1.7	4.0	2.3	1.3
	Cond.	14.3	4.3	2.0	4.4	2.3	1.3
		(13.3 15.2)	(3.9 4.9)	(1.7 2.4)	(3.8 4.9)	(1.9 2.7)	(1.0 1.6)
	Uncond.	14.7	4.3	2.1	4.5	2.4	1.4
		(13.7 15.7)	(3.8 4.8)	(1.7 2.4)	(4.0 5.0)	(2.1 2.8)	(1.1 1.7)

4. Conclusions

The performance of the conditional stochastic model presented in this paper is analyzed in two steps. Firstly, the skill of the CCA model in estimation of the four precipitation distribution parameters is assessed. Se condly, the performance of the conditional stochastic model to reproduce the statistical features of generated precipitation time series is analyzed. In order to see the advantage of the conditional model, it is compared to the unconditional model based only on a Markov chain with the four parameters computed directly from observations. Both stochastic models have been run 1000 times and the precipitation distribution parameters are computed as ensemble means with their associated 90% confidence intervals. The skill of both stochastic models is computed for the two subintervals, which have not been used for fitting the models. The main conclusions, which can be drawn from this analysis, are summarized in the following.

The CCA model is most skillful for wint er and autumn (transition probabilities), slightly skillful for the mean amount for rainy days μ (winter, especially for 1951 -1999 interval) and unskillful for the shape parameter **k**. This result is in agreement with previous studies (Busuioc and von Storch , 1996; Busuioc et al, 1999) when the connection between seasonal precipitation amount at 14 Romanian stations (including Bucharest) and the large -scale circulation was analyzed. The unexpected low skill for daily mean precipitation μ, found in the present paper, suggests that the strong link of the seasonal precipitation (winter and autumn) with the SLP field is given by the strong link between number of rainy days and SLP, a fact supported by the direct correlation between them. The results presented abov e also suggest the fact that the shape parameter of precipitation distribution (**k**) does not depend on the large -scale circulation. One reason of the model failure regarding μ and **k** could be that the two subintervals have different statistics of precipitati on in Bucharest, i.e., two different precipitation regimes. The urbanization could be one reason for this behavior, but it is difficult to separate it from the natural variability, since daily precipitation time series as long as for Bucharest station are not available for other stations in Romania or on One reason of the model failure regarding μ and **k** could be the incoherently temporal variation of the four parameters over the two subintervals suggesting two different precipitation regimes.

The condition al and unconditional stochastic models were found to be similarly skillful in reproducing the statistical properties of precipitation, except for the trend, which by construction only the conditional model can deal with. Thus, for climate change applicatio ns only the conditional model is useful.

Some statistical features were well reproduced by both stochastic models for all seasons, such as: the mean and expected maximum duration of wet intervals; the daily mean of precipitation for rainy days; the expect ed maximum duration of dry intervals; the frequency of days with precipitation amount greater than 20 mm for 1901 -1999 interval.

Other statistical features of the generated precipitation time series are only partially reproduced by both models or are bett er reproduced by one of models such as:

- the mean duration of dry intervals for winter and summer (1901 -1949) and autumn (1950-1999) are well reproduced by both models; in the other cases both models, except for spring when the unconditional model is bet ter, overestimate the observed values;
- the standard deviation is better estimated for the 1901 -1949 interval, except for summer when it is underestimated; for the 1951 -1999 interval the standard deviation is slightly underestimated for all seasons (less for summer);
- the expected maximum daily precipitation is generally underestimated for all seasons (both models) but the observed values are covered by the 90% confidence intervals, except for summer (1901 -1949), winter and autumn (1950 -1999);

• the frequency of the wet intervals (winter and autumn) and frequency of the extreme dry events (winter and autumn) are well reproduced; in the rest of cases, generally, the frequency of shorter dry intervals is underestimated and the longer dry intervals (grater than 9 days) are too frequent.

• the seasonal mean of rainy days is generally well reproduced by both stochastic models, except for spring and autumn when it is underestimated by the conditional model and overestimated by the unconditional model;

• the linear trend of the winter precipitation is well identified but the increase in 1901 -1949 is underestimated and the decrease in 1950 -1999 is overestimated.

In conclusion, the conditional stochastic model presented in this paper can be successfully used to generate daily precipitation time series especially for winter and autumn. This model has the advantage compared to the unconditional model to capture the changes in the local seasonal precipitation induced by changes in the large -scale circulation, represented by the SLP field that makes it useful for climate change scenarios based on GCM outputs (especially for transient version). This model can be improved (especially for spring and summer seasons) adding other large -scale parameters. The moisture variables could be important large -scale predictors but unfortunately they are not available for long time series. The NCEP reanalysis are only available since 1948. For climate change scenarios this problem could be solved developing such kind of conditional stochastic models over shorter interval and to use the cross-validation procedure.

Acknowledgements
This research was partially made during the Dr. Aristita Busuioc's visit at the Institute of Hydrophysics, GKSS Research Center, Geesthacht. The other part has been supported by the Romanian Ministry of Water and Protection of the Environment. The authors thanks to Dr. Reiner Schnur for useful discussions about the stochastic generation of weather time series. Mrs. Mariana Caraman from the National Institute of Meteorology and Hydrology (Bucharest) is acknowledged for preparing the precipitation data. Mrs. Beate Gardeike improved the diagrams.

5. References

Busuioc,A., Von Storch , H., 1996: Changes in the winter precipitation in Romania and its relation to the large-scale circulation, *Tellus* 48 A, 538-552

Busuioc, A., and H. von Storch, 1995: The connection between summer precipitation anomalies in Romania and large-scale atmospheric circulation. *Proc. Atmospheric Physics and Dynamics in the Analysis and Prognosis of Precipitation Fields*, Rome, Italy, 15-16 November 1994, 369-373.

Busuioc, A., H. von Storch and R. Schnur, 1999: Verification of GCM generated regional seasonal precipitation for current climate and of statistical downscaling estimates under changing climate conditions, *Journal of Climate, vol 12, 258-272* .

Cavazos, T, 1999: Large-scale circulation anomalies conductive to extreme precipitation events and derivation of daily rainfall in northeastern Mexico and southeastern Texas *. J of Climate,* vol 12, 1506-1523

Coe and Stern, 1982: Fitting models to daily rainfall data, *Journal of Applied Meteorology*, vol. 21, 1024-1031.

Dobi-Wantuch, I., J. Mika, and L. Szeidl, 2000: Modelling wet and dry spells with mixture distributions, *Meteorology and Atmospheric Physics* 73, 245-256.

Goodness, C and J. Palutikof, 1998: Development of daily rainfall scenarios for southeast Spain using a circulation-type approach to downscaling. *Int. J of Climatol.* 10, 1051-1083.

Heyen, H., E. Zorita, and H. von Storch, 1996: Statistical downscaling of monthly mean North Atlantic air-pressure to sea level anomalies in the Baltic Sea. *Tellus* 48 A, 312-323.

Katz, R. W. and Parlange, M. B., 1993: Effects of an index of atmospheric circulation on stochastic properties of precipitation, *Water Resources Research*, Vol. 29, No. 7, 2335-2344.

Katz, R. W., 1996: Use of conditional stochastic models to generate climate change scenarios, *Climate Change* 32, 237-255.

Lettenmaier, D., 1995: Stochastic modeling of precipitation with application to cl imate model downscaling, *Analysis of Climate Variability: Application of Statistical Techniques*, H. von Storch and A. Navara, Eds., Springer -Verlag, 197-212.

Richardson, C. W., 1981: Stochastic simulation of daily precipitation, temperature, and solar radiation, *Water Resources Research*, Vol. 17, No. 1, 182-190.

Sneyers, R., 1975: Sur l'analyse statistique des series d'observation. *WMO Note Technique*, no. 143, 189 pp.

Trenberth, K.E., and D.A. Paolino, 1980: The Northern Hemisphere sea -level pressure data s et: Trends, errors and discontinuites. *Mon. Wea. Rev*. 108, 855-872.

von Storch, H., E. Zorita, and U. Cubasch, 1993: Downscaling of global climate change estimates to regional scale: An application to Iberian rainfall in wintertime. *J. Climate*, 6, 1161-1171.

Werner, P., and von Storch, H., 1993: Interannual variability of Central European mean temperature in January - February and its relation to large -scale circulation. *Clim. Res*., 3, 195 - 207.

Wilks, D., 1992: Adapting stochastic weather generation algo rithms for climate change studies*,* *Climate Change* 22, 67-84.

Yarnal, B., A. C. Comrie, B. Frakes, and D. Brown, 2001: Development and prospects in synoptic climatology, *International Journal of Climatology* 21, 1923-1950

Zorita, E., J. P. Hughes. D.P. Lette maier and H. von Storch, 1995: Stochastic characterization of regional circulation patterns for climate model diagnosis and estimation of local precipitation, *Journal of Climate*, vol 8, No 5, 1024-1042

Zorita, E. and H. von Storch, 1999: A survey of stati stical downscaling techniques*, J. of Climate* 2, 2474-2489.