Revista Brasileira de Climatologia
ISSN: 2237-8642 (Eletrônica)

A NEW METHODOLOGY FOR STOCHASTIC SIMULATION OF DAILY CLIMATIC DATA PRESERVING THE INTERANNUAL VARIABILITY

KIST, Airton - kist@uepg.br
Universidade Estadual de Ponta Grossa/UEPG

VIRGENS FILHO, Jorim Sousa das - jvirgens@uepg.br
Universidade Estadual de Ponta Grossa/UEPG

LA GUARDIA, Giuliano Gadioli - gguardia@uepg.br
Universidade Estadual de Ponta Grossa/UEPG

ABSTRACT: In this work we propose a new methodology to reproduce, by means of simulations, the interannual variability of climatic variables which included only the minimum air temperature. To evaluate the performance of the proposed method, it was maked a comparison with other two weather generators (i.e., PGECLIMA_R and LARS-WG). Moreover, it was utilized the historical series of thirty years of five meteorological stations of the state of Parana - Brazil to generate ten sets of thirty years for each model, which were confronted with the respective historical series. The performance of the proposed model as well as weather generators was evaluated by applying tests of central tendency, variability and distribution. Furthermore, was utilized the statistical measures RMSE, MBE and Willmott agreement index (d). In the stations investigated, the proposed methodology reduced the total error and eliminated the negative bias of interannual variability. In only four (of 600) generated sequences the interannual variability differs significantly from the observed one. The series generated by PGECLIMA_R and LARS-WG presented rejection rate of 99% in the variability test. In this case, the bias was ten times greater and the RMSE was twice times greater than the proposed methodology. The d index was always greater than 0.98 for the five locations in the proposed methodology and around 0.83 in other models. Based on these results, the new methodology provides a relevant contribution concerning the interannual variability of climatic variables.

Keywords: Simulation of climatic data, interannual variability, minimum air temperature.

UMA NOVA METODOLOGIA PARA SIMULAÇÃO ESTOCÁSTICA DE DADOS CLIMÁTICOS DIÁRIOS PRESERVANDO A VARIABILIDADE INTERANUAL

RESUMO: Este trabalho objetivou propor uma nova metodologia para reproduzir, por meio de simulações, a variabilidade interanual de variáveis climáticas que incluiu apenas a temperatura mínima do ar. Para avaliar o desempenho da metodologia proposta fez-se a comparação com outros dois geradores de clima: PGECLIMA_R e LARS-WG. Foram utilizadas séries históricas de trinta anos de cinco estações meteorológicas do estado do Paraná - Brasil para gerar dez conjuntos de trinta anos para cada modelo, que foram confrontados com a respectiva série histórica. O desempenho do modelo proposto e dos geradores de clima foi avaliado através da aplicação de testes de tendência central, variabilidade e distribuição. Além disso, utilizaram-se as medidas estatísticas RMSE, MBE e índice de concordância de Willmott (d). Nas estações analisadas, a metodologia proposta reduziu o erro total e eliminou o viés negativo da variabilidade interanual. Em apenas quatro (das 600) sequências geradas a variabilidade interanual difere significativamente da observada. As sequências geradas pelo PGECLIMA_R e LARS-WG apresentaram taxa de rejeição de 99% no teste de variabilidade. O viés nesses modelos foi dez vezes maior e o RMSE foi duas vezes maior do que na metodologia proposta. O índice d foi sempre maior do que 0,98 nas cinco estações na metodologia proposta e em torno de 0,83 nos outros modelos. Com base nestes testes , a nova metodologia fornece uma contribuição relevante relativa à variabilidade interanual de variáveis climáticas.

Palavras-chaves: Simulação de dados climáticos, variabilidade interanual, temperatura mínima do ar.

1. INTRODUCTION

The analysis, understanding and simulation of climatic data, are of great importance in the assessment of the decision making in the human activities, such as agriculture and environmental monitoring. Several researchers have extensively investigated this area of interest and proposed several computational models to simulate daily synthetic series of precipitation, air temperature, relative humidity of air, among other climatic variables (VIRGENS FILHO et al., 2014). However, even being suitable tools in this research area, it is important to point out that simulated climatic data are preferred to be applied only in the case where the historical data are not available or also in the case to make prognostic analysis (HOOGENBOOM, 2000). Moreover, climatic data simulated are utilized to fill incomplete historical series.

Weather generators are computationally inexpensive tools to produce multiple-year climate change scenarios at the daily time scale, which are used to assess the impact of future climate change. However, several of them tend to under predict interannual variability of generated sequences of precipitation and other variables (CHEN et al., 2010).

For many years, several weather generators were developed, such as WGEN (Richardson, 1981; Richardson and Wright, 1984), USCLIMATE (Hanson et al., 1994), CLIGEN (Nicks et al., 1995), LARS-WG (Semenov and Barrow, 1997), ClimGen (Stockle et al., 1999), PGECLIMA_R (Virgens Filho et al., 2013), WeaGETS (Chen et al., 2012). It is expected that the data generated by these models provide statistical properties similar to the observed data such as: mean, variability and distribution.

The generators LARS-WG, CLIGEN, WeaGETS and PGECLIMA_R, when utilized in the simulation of agricultural models tend to underestimate the interannual variability of climatic variables in the corresponding generated series (SOLTANI and HOOGENBOOM, 2003). Great part of weather generators provides good estimates to the mean of climatic variables. However, according to Hansen and Mavromatis (2001), these weather generators are not good to reproduce oscillations that occur from one year to another, producing monthly means with few variability, generating therefore bias around the mean. One problem with daily weather generators is that they underestimate monthly and interannual variances because they do not take into account the low-frequency component of climate variability (CHEN et al., 2010).

Several models dealing with climatic variables preserving the variability have been proposed in the literature (Hansen and Mavromatis, 2001; Chen et al., 2010 and Khazaei et al., 2013). In other works such as Khazaei et al. (2013) and Virgens Filho et al. (2013), the authors have compared generators of climatic data. Minuzzi et al. (2010) have investigated trends of monthly climatic variables, such as minimum temperature in some stations in the state of Minas Gerais - Brazil. The authors observed a great variability of monthly means in some stations. They also have verified changes in the non gradual temporal variability of the maximum and minimum temperature in some months.

ISSN: 2237-8642 (Eletrônica)

Hansen and Mavromatis (2001) presents a stochastic weather generator to improve interannual variability characteristics by perturbing monthly parameters using a low-frequency stochastic model, and evaluates the effectiveness of the low-frequency component on interannual variability of generated monthly climate. For monthly precipitation and maximum and minimum temperatures, the low-frequency correction reduced total error and eliminated negative bias of interannual variability of monthly climatic means. It was deduced the number of station-months with significant differences between observed and generated interannual variability. The authors concluded that further refinement is needed to better represent interannual variability of both precipitation occurrence and intensity processes, and to rectify over-correction of interannual temperature variability.

An approach to correct the low frequency variability of climate variables in a climate generator was proposed in Chen et al. (2010) that evaluated its ability to reproduce key statistical parameters at the daily, monthly and yearly scales. Low-frequency variability was modeled based on observed power spectra of monthly and annual time series. While weather generators are good at preserving the precipitation quantity, they however underestimate low-frequency variations.

Khazaei et al. (2013) proposed a daily weather generator to preserve extremes and low-frequency variability. Addresses also the unanticipated effects of changes to precipitation occurrence under climate change scenarios on secondary variables. The weather generator is composed of three major components, including a stochastic rainfall model, a multivariate daily temperature model conditioned to the rainfall occurrence, and a suitable multivariate monthly generator to fit the low-frequency variability of daily maximum and minimum temperature series. The performance was tested by comparing statistical characteristics of the simulated and observed weather data, and by comparing statistical characteristics of the simulated runoff outputs by a daily rainfall-runoff model fed by the generated and observed weather data. Furthermore, compared the simulated data with those of the LARS-WG weather generator. It concluded that the performance of the weather generator is acceptable, better than LARS-WG in the majority of tests.

Many stochastic weather generator use all the variability observed a month (a period) and they do not consider that one year to another the climate variable may show great variability. Therefore, these generators tend to simulate monthly series with few variability around the mean, and this fact has not been observed in the historical series. According to Minuzzi et al. (2010) the climatic changes that occur from one year to another can be caused by climatic phenomena such as El Niño or La Niña or even by phase shifts of interdecadal oscillations, among other phenomena. This view point of the problem suggests that a potential approach to improve the characteristics of interannual variability could disturb the monthly parameters by applying stochastic models which take into account such variation. Among the climatic variables, the air temperature is of great relevance in the human activities, mainly in agriculture where air temperature exerts a great influence on plant phenology.

Most of the stochastic simulation models of daily minimum and maximum temperatures utilizes Gaussian distributions conditioned to occurrence of precipitation. Thus, the models cannot reproduce interannual great variations

of air temperature contained in the historical series. In fact, there exist atypical months with minimum/maximum temperature which are much above or below the mean of the historical period. In this context, in order to reproduce historical series with the same statistical properties, was proposed a new methodology to generate synthetic series of daily climatic variables. This methodological proposal was just applied to the minimum air temperature.

2. MATERIALS AND METHODS

This research was developed at the Laboratory of Computational and Applied Statistics, at Department of Mathematics and Statistics - State University of Ponta Grossa (UEPG). In this work it was utilized the climatic data of minimum air temperature of thirty years (1980-2009) with respect to five weather stations (which belong to the Agronomic Institute of Parana - IAPAR) located in the state of Parana - Brazil.

2.1 REGION OF STUDY

The size of the time series, the integrity of the series, data consistency and spatial coverage were important factors in the choice of locations. The meteorological stations are located between latitudes -26°29' and -23°05', longitudes -53°55' and -48°32', and with elevation 40 and 1100m according Table 1. According to the classification of Köppen, in the state of Parana the predominant climatic types are Cfa (subtropical climate) and Cfb (temperate climate). The locality of Guarapuava, with Cfb climate, presents total annual precipitation of 1956mm and minimum and maximum temperatures of 12.86 and 23.54°C, respectively. In Guaraqueçaba, with Cfa climate, the total annual rainfall is 2423mm and minimum and maximum temperatures are 17.40 and 26.28°C, respectively. The locality of Palmas, with Cfb climate, presents total annual precipitation of 2098mm and minimum and maximum temperatures of 12.01 and 22.59°C, respectively. In Palotina, with Cfa climate, the total annual rainfall is 1665mm and minimum and maximum temperatures are 16.09 and 28.71°C, respectively. The locality of Paranavai, with Cfa climate, presents total annual precipitation of 1478mm and minimum and maximum temperatures of 17.83 and 28.46°C, respectively.

Table 1 – Weather stations used in the study.

Station	Latitude (ºS)	Longitude (ºW)	Elevation (m)
Guarapuava	-25° 21'	-51° 30'	1058
Guaraqueçaba	-25° 16'	-48° 32'	40
Palmas	-26° 29'	-51° 59'	1100
Palotina	-24° 18'	-53° 55'	310
Paranavai	-23° 05'	-52° 26'	480

2.2 STATISCAL ANALYSIS

After this initial analysis, the data was separated by month, and then it was applied a hypothesis test for equality of means to verify if the monthly mean minimum temperature observed in each year equals to the monthly

historical mean corresponding to thirty years. More precisely, the following hypothesis was tested:

H_0: *the mean monthly minimum air temperature of each year is equal to its monthly mean of thirty years*;

H_1: *the mean monthly minimum air temperature of each year is different from its monthly mean of thirty years*.

Symbolically, for an observed sequence of a years, let $\bar{y}_{ij},\ j = 1, 2, 3, \ldots, a,$, denote the mean monthly minimum air temperature in a given month $i,\ i = 1, 2, \ldots, 12,$, and $\bar{y}_{i\cdot}$ denote the mean minimum air temperatures of all a years of the corresponding month i. It was tested the following hypotheses:

$$H_0: \bar{y}_{ij} = \bar{y}_{i\cdot} \quad \text{versus} \quad H_1: \bar{y}_{ij} \neq \bar{y}_{i}. \ .$$

In weather generators, including USCLIMATE (Hanson et al., 1994), LARS-WG (Semenov and Barrow, 1997), ClimGen (Stockle et al., 1999), and CLIGEN (Nicks and Lane, 1989), it is common to assume that daily maximum and minimum air temperature are normally distributed, since these models utilize Gaussian distribution (equation 1) for simulating them, parameterized separately for dry and wet days.

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right],\ x \in (-\infty, \infty) \qquad (1)$$

where μ (mean) and σ (standard deviation) are the parameters of the Gaussian distribution.

In order to verify if the distribution of the minimum air temperature for each month on wet and dry days has the same functional form, each sequence of values was standardized and the two sample Kolmogorov-Smirnov (KS) test was applied to historic data of the five stations analyzed. The KS test was performed on a total of 180 months (3 tests for each of the 12 months and five stations) and there were 10 rejections (Paranavai: 2, Guarapuava: 6 and Guaraqueçaba: 2) at 5% level of significance. This shows that, for local analysis, the minimum temperature distribution has the same functional form in the three cases investigated. These last two facts justify the use of Student´s t test in the proposed method shown in this work.

Based on all these facts, assuming that under null hypothesis and also by considering that the minimum temperature of the month i of each year has a Gaussian distribution with the same mean of the monthly historical series and with unknown variance. The two-tailed Student t test with significance level α = 0.05 is applied in the monthly series under analysis of each station. Several monthly series may not be significant at level α, thus rejecting the hypothesis of equality of means.

As the two-tailed t-test was applied, these rejections occurred because some months the means were smaller (or larger) when compared to the historical monthly mean. Assuming that the daily temperature for the dry/wet days of each month has a Gaussian distribution (because it is used by several

weather generators), it was proposed a new procedure to adjust the probability distribution of such variable. The difference of our procedure lies in the form to adjust the distribution.

In sequence the data was separated per month. After this, it was applied the t test in order to separate the monthly data into three groups: a group in which the null hypothesis was not rejected and two groups (lower/higher means) in which the null hypothesis was rejected. In the sequence, it was separated the data of each group into dry and wet days, fitting dry/wet days to a Gaussian distribution with mean and variance estimated from the observed data. In this manner, instead of working with two distributions, one for wet days and another for dry days, it was considered six distributions, three for dry days and three for wet days.

More generally, for historical data of a years, it has been for each month a_1,a_2 rejections due to the fact that the mean is lower or higher than the mean of a years and the acceptance has occurred for $a_3 = a - (a_1 + a_2)$ years. Note that $a_i, i = 1,2,3$, can differ from month to month. For each of them, (i.e. a_1, a_2, a_3), it was fitted a Gaussian distribution to dry days and another to wet days. Therefore, it has been six distributions to generate the minimum daily temperatures for a given month of a given location.

In order to simulate a series of daily minimum temperatures in a given month, it was chosen (for each year) a pair of distributions adjusted for this month. Let us now consider $\alpha_i = \frac{a_i}{a}, i = 1,2,3$. Then the daily temperature for a given month is generated with probability ɑ_i,i=1,2,3, given by the mixture of two Gaussian distributions parameterized separately for dry (index 0) and wet (index 1) days. This resulting distribution for the air temperature is given by Hansen and Mavromatis (2001):

$$f_{\alpha_i}(x) = (1 - \pi) f_0(x|\bar{X}_0, S_0) + \pi f_1(x|\bar{X}_1, S_1),\ i = 1,2,3 \quad (2)$$

where π is the probability of occurrence of wet days and $\bar{X}$ and S denote, respectively, the mean and standard deviation computed from historical data. Therefore, the daily monthly temperature simulated has the following distribution:

$$f(x) = \alpha_1 f_{\alpha_1}(x) + \alpha_2 f_{\alpha_2}(x) + \alpha_3 f_{\alpha_3}(x) \quad (3)$$

The interannual variability of the simulated monthly mean values is related to the choice of distribution that generates temperatures of the month, one of the $f_{\alpha_i}(x)$

2.3 SIMULATION OF MINIMUM AIR TEMPERATURE

In this study, as in Richardson (1981), Nicks et al. (1995), Semenov and Barrow (1997), Chen et al. (2010), Virgens Filho et al. (2013) the minimum air

temperature it was conditionated at the precipitation occurrence using a two-state first order Markov chain model to generate the number of precipitation events.

The first order Markov chain model requires the calculation of the probability of a wet day following a dry day and the probability of a wet day following a wet day. The two-state Markov chain for the combination of conditional probabilities can be defined in terms of two transition probabilities p_01 and p_11:

$$p_{01} = P(\text{precipitation on day } t \mid \text{no precipitation on day } t-1) \quad (4)$$

$$p_{11} = P(\text{precipitation on day } t \mid \text{precipitation on day } t-1) \quad (5)$$

Since precipitation either occurs or does not on a given day, the two complementary transition probabilities are $p_{00} = 1 - p_{01}$ and $p_{10} = 1 - p_{11}$. Therefore the transition probabilities are fully defined given p_{00}, $p_{11}$and the wet or dry state on day $t-1$. These probabilities are computed for each month of the year from the analysis of the historic long-term precipitation data.

The simulated monthly daily minimum temperature is conditioned on wet and dry states of the first-order Markov model, where a wet day is defined to be a day with precipitation > 0.0mm. Since there exist three pairs of fitted distributions, the choice of distribution is random but proportional to the number of times that these $\alpha_i{}'$s were observed in the historical series utilized for the adjustment. Thus, once chosen the pair of distributions, one for dry days and another one for wet days, temperatures of one month with this pair are generated. After this, repeat this procedure every month for the entire simulation period.

To better evaluate the proposed model (PM), the proportion of times that each pair of adjusted distributions was used to simulate the temperature was exactly the same as observed.

In order to evaluate and validate the simulation of the proposed model, it was compared it with two other weather generators available in the literature: the PGECLIMA_R (PG) and LARS-WG (LW).

For each station in Table 1, each of weather generator models was parameterized using all valid observations. Ten stochastic sequences of daily climatic data with size equal to the observed period (thirty years) were generated. The simulation procedure described above and the weather generators were evaluated by comparing historical monthly means sequences and generated sequences.

To test the hypothesis of equality interannual variability between observed and generated series, it was applied the F-test to monthly means. An independent t-test was utilized to test the hypothesis of equality of the central tendency conditioned on the equality of variances test. Finally, the non-parametric Kolmogorov–Smirnov two-sample test was employed to test the hypothesis of identical distributions. Additionally, it was also performed an analysis of the generated series daily.

For cumulative distribution functions F and G on the real line, define the Kolmogorov-Smirnov distance between F and G to be $d_K(F,G) = \sup_x |F(x) - G(x)|$, where sup is the supremum function and F and G are the empirical distribution functions. The Kolmogorov-Smirnov test rejects the null hypothesis if $d_K(F,G) > s_{n,1-\alpha}$,, where $s_{n,1-\alpha}$ is the $1-\alpha$ quantile of the null distribution.

The performance of the weather generators as well as the proposed model were also characterized by means of three descriptive statistics that indicate the degree of consistency among results based on the observed and generated climatic data for the mean and standard deviation, given by:

root-mean-squared-error (RMSE)

$$RMSE = \left(\frac{1}{a}\sum_{i=1}^{a}\left(\bar{y}_{O,i} - \bar{y}_{G,i}\right)^2\right)^{1/2}; \quad (6)$$

mean bias error (MBE)

$$MBE = \frac{1}{a}\sum_{i=1}^{a}\left(\bar{y}_{O,i} - \bar{y}_{G,i}\right); \quad (7)$$

and Willmott agreement index (d)

$$d = 1 - \frac{\sum_{i=1}^{a}\left(\bar{y}_{O,i} - \bar{y}_{G,i}\right)^2}{\sum_{i=1}^{a}\left(|\bar{y}_{G,i} - \bar{y}_{O,\cdot}| - |\bar{y}_{O,i} - \bar{y}_{O,\cdot}|\right)^2} \quad (8)$$

where, $\bar{y}_{O,i}$ and $\bar{y}_{G,i}$ denote, respectively, the ordered values observed and generated monthly in a years, and $\bar{y}_{O,\cdot}$ denotes the mean of temperatures observed in a given month in the period of a years.

Similar procedures for evaluation and comparison of models were also used by Willmott et al. (1985), Hansen and Mavromatis (2001), Mavromatis and Hansen (2001) and Coutinho et al. (2016).

3. RESULTS AND DISCUSSION

Initially, an analysis of the minimum temperature observed on thirty years (1980-2009) in the five stations (Guarapuava, Guaraqueçaba, Palmas, Palotina and Paranavai) was done. It was observed that the monthly mean minimum air temperature in each month presents a large variation. To illustrate that the monthly mean for January, April, July and October for Paranavai station is shown in Figure 1. It was also observed that the other stations (and months) present similar characteristic (not shown). Although, from one year to another, the mean can fluctuate significantly, there exist periods of two to five years in which the mean has low variation (Figure 1).

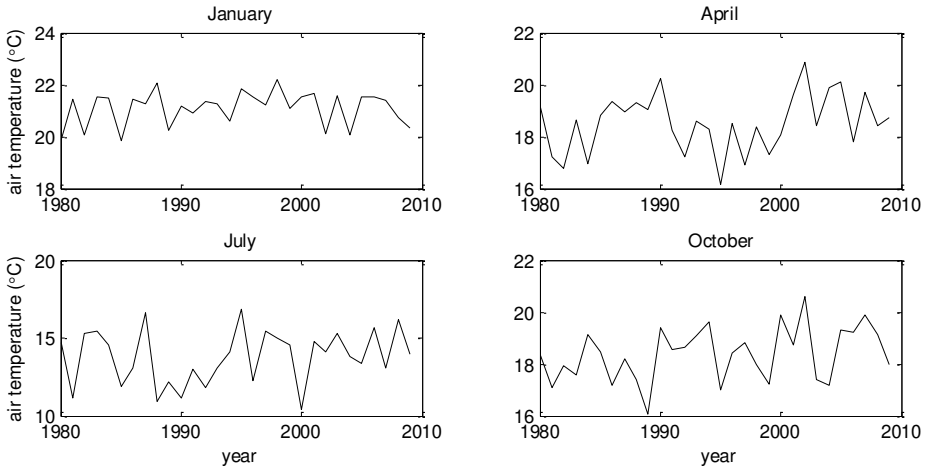

Figure 1 - Mean minimum air temperatures observed during the 1980 to 2009 period, at the station of Paranavai in January, April, July and October.

The two-tailed Student t test with significance level α (α = 0.05 in this work) is applied in the monthly series under analysis of each station. Several monthly series may not be significant at the significance level α, thus rejecting the hypothesis of equality of means.

It was applied the Student's t test, previously mentioned, with significance level α = 0.05 in all the stations. Several monthly series were not significant for the level of significance considered. In Table 2, it can be seen the number of times that the referred test were not significant for the level of significance 5% for these five stations.

The number of times that the test rejected the null hypothesis was great. More specifically, the test rejected (on mean) 12.08, 13.42, 10.42, 13.67 and 12.58 times for the stations Guarapuava, Guaraqueçaba, Palmas, Palotina, Paranavai, respectively. In the case of the stations Guaraqueçaba, Palotina and Paranavai, in some months, more than a half of times there was rejection of the equality of the means. The high proportion of rejections can be one of the reasons for which the weather generators do not reproduce the interannual variability of the historical mean.

Table 2 - Number of times the test t were not significant at the 5% significance level, in each month, for the minimum air temperature for each station for the 1980 to 2009 period.

Station	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Guarapuava	17	15	10	16	13	10	12	8	10	14	11	9
Guaraqueçaba	16	14	11	19	15	13	15	12	11	11	10	14
Palmas	13	14	8	13	11	10	13	6	11	14	6	6
Palotina	23	20	11	15	14	7	13	11	10	14	13	13
Paranavai	16	15	15	14	11	9	16	8	13	15	8	11

From a total of 600 sequences generated by PM, in the t test all of them were significant at the 5% significance level (Table 3). This means that PM reproduced accurately the monthly means observed of minimum temperature.

In only four generated sequences the interannual variability differs significantly from observed by F-test and there was no rejection on the distribution test (K-S test). Analyzing the descriptive statistics (Table 4), that indicate the degree of consistency between results observed and generated, it was noted the occurrence of small bias, the RMSE is acceptable and the index of Willmott (d) is greater than 0.98.

On the other hand, the generated sequences in PG and LW had less than 1% of acceptance by the variability test (Table 3). Additionally, 22% in the distribution test and 31.2% of the sequences generated in PG and LW, respectively, differed significantly from those observed, although in the t-test both had good performance. The bias in these two generators were approximately ten times greater than PM (Table 4). The RMSE was two times greater than PM, and the d index was around 0.83.

Concerning the standard deviation, the performance of PM was also better than the other two weather generators. Thus, the temperatures generated monthly were more accurate and had lower variability than those observed.

The results of central tendency, variability and distribution tests of the daily series are shown in Table 5. The proposed model had less than 2% of rejection on the central tendency test, while PG had 10% of rejection and LW almost 100% of rejection. In variability test, the PM had 2 rejections in 600 tests, while PG had 57 and LW 178 rejections, respectively. The distribution of data generated by PM were different from those observed in 22% of cases, while in the PG and LW were found 66% and 76% of significative differences, respectively.

The results of the tests and descriptive measures showed the capacity of PM to fit to observed data and to simulate daily data with interannual variations of the mean of the minimum air temperature.

Hansen and Mavromatis (2001) proposed a stochastic weather generator which attempts to improve interannual variability characteristics by perturbing the monthly parameters (Model 1 and 2). In 300 months simulated, the Models 1 and 2 had 6 and 2 sequences rejected, respectively, in central tendency test. With respect to variability test, the Models 1 and 2 had 78 and 22, respectively. In the distribution test, both had 7 rejections. The bias for the minimum temperature was -0.044 for Model 1 and -0.023 for Model 2, while the value RMSE was 0.220 and 0.251, respectively. Concerning standard deviation, the bias was -0.0653 and RMSE 0.372 in Model 1 and 0.0746 and 0.234 in Model 2, respectively.

Comparing the results obtained by PM with the results of Hansen and Mavromatis (2001), it was observed that PM have superior performance in the tests of central tendency, variability and distribution, once the bias was lower in PM when compared with the models of Hansen and Mavromatis (2001). The results of the RSME for the mean and standard deviation in the models of Hansen and Mavromatis (2001) were similar to those proposed in this work.

ISSN: 2237-8642 (Eletrônica)

Table 3 - Number of calendar months in which central tendency, interannual variability and distribution of generated climate variables differed significantly from the observed $n = 120$

	Central tendency			Variability			Distribution		
Station	PM	PG	LW	PM	PG	LW	PM	PG	LW
Guarapuava	0	0	7	0	117	117	0	22	32
Guaraqueçaba	0	1	3	0	120	120	0	36	50
Palmas	0	0	5	1	120	118	0	22	33
Palotina	0	0	1	3	120	120	0	32	34
Paranavai	0	0	12	0	120	119	0	20	38
Total	0	1	28	4	597	594	0	132	187

Table 4 - Mean bias error (MBE), root-mean-squared-error (RMSE) and and Willmott agreement index (d) of mean and interannual standard deviations of observed and generated monthly minimum temperatures in each station $n = 120$.

		Mean			Standard deviation		
Station		PM	PG	LW	PM	PG	LW
Guarapuava	MBE	0.0110	0.0996	0.1395	0.0468	0.1321	-0.3529
	RMSE	0.2876	0.6654	0.6699	0.2397	0.3472	0.4626
	d	0.9828	0.8338	0.8435	0.9529	0.8799	0.8553
Guaraqueçaba	MBE	0.0004	-0.1643	0.1262	0.0461	0.1689	-0.2969
	RMSE	0.2277	0.6338	0.6240	0.2322	0.3635	0.3967
	d	0.9861	0.8039	0.8179	0.9439	0.8492	0.8503
Palmas	MBE	0.0121	-0.1343	0.1189	0.0469	0.1149	-0.3758
	RMSE	0.2911	0.7065	0.6988	0.2372	0.3413	0.4764
	d	0.9825	0.8332	0.8462	0.9545	0.8895	0.8551
Palotina	MBE	0.0315	-0.0894	0.1565	0.0429	0.2134	-0.3512
	RMSE	0.3593	0.8566	0.8543	0.2533	0.4180	0.4980
	d	0.9805	0.8112	0.8244	0.9598	0.8714	0.8679
Paranavai	MBE	0.0007	-0.0853	0.0738	0.0469	0.1579	-0.3775
	RMSE	0.2550	0.6121	0.6397	0.2436	0.3568	0.4706
	d	0.9844	0.8337	0.8309	0.9431	0.8697	0.8383

ISSN: 2237-8642 (Eletrônica)

Table 5 - Number of calendar months in which central tendency, interannual variability and distribution of generated daily climate variables differed significantly from the observed $n = 120$.

Station	Central tendency			Variability			Distribution		
	PM	PG	LW	PM	PG	LW	PM	PG	LW
Guarapuava	1	12	117	0	6	30	24	69	72
Guaraqueçaba	0	10	120	0	28	45	17	95	109
Palmas	2	24	118	0	11	30	22	74	91
Palotina	5	8	118	1	5	33	38	87	84
Paranavai	2	7	120	1	7	40	32	69	104
Total	10	61	593	2	57	178	133	394	460

After doing the analysis of the performance of weather generators in tests and in the statistics, it was also performed graphical analyses. The Figures 2, 3, 4 and 5 show the distribution of monthly means observed and generated by season. In these figures, the box plot Ob represents the distribution of monthly mean minimum temperatures observed (30 values). The other box plots represent, respectively, the distribution of monthly mean minimum temperatures of ten simulations of 30 years (300 values) for models PM, PG and LW in each month and station. Analyzing the figures it can be observed that the means of models PG and LW presented little variability when compared to the observed data and the generated by the PM model. For example, the interquartile range of these models is very small, as does the amplitude of the mean, simulated by these models. In the months of May, June and September, the weather generator LW does not reproduce the variability and also presents a bias when compared to the observed data. This shows that the distribution of the observed means and the generated means by PG and LW are different. Thus it is concluded that the mean minimum temperatures generated by PG and LW does not have the same variability of the observed means. On the other hand, the mean minimum temperatures generated by PM can reproduce the interquartile range as well as the variability of the observed means.

Another feature of the generators PG and LW is that both present symmetry of the box plots, while the observed means does not have symmetry. Note that our method does not present symmetry and this fact corroborates with observed means.

ISSN: 2237-8642 (Eletrônica)

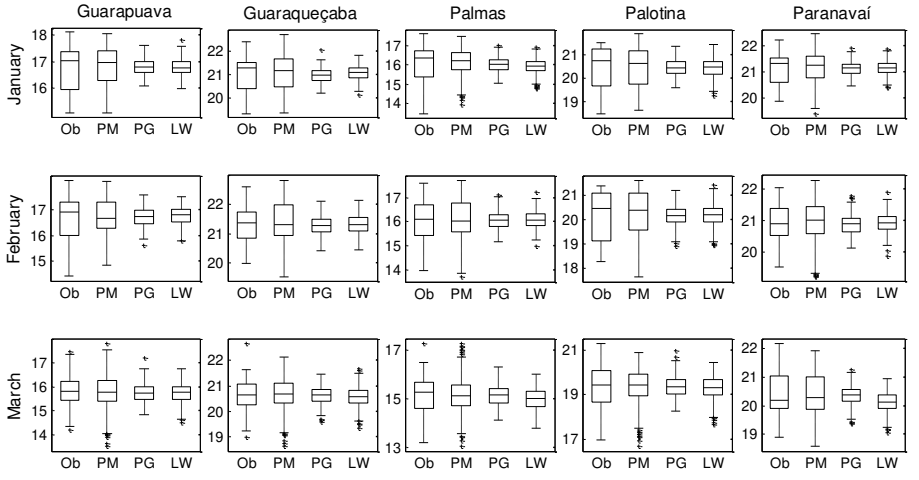

Figure 2 - Box plots of thirty minimum means observed (Ob) and (10 × 30 = 300) minimum mean simulated of air temperatures (ºC) for the summer months in each station.

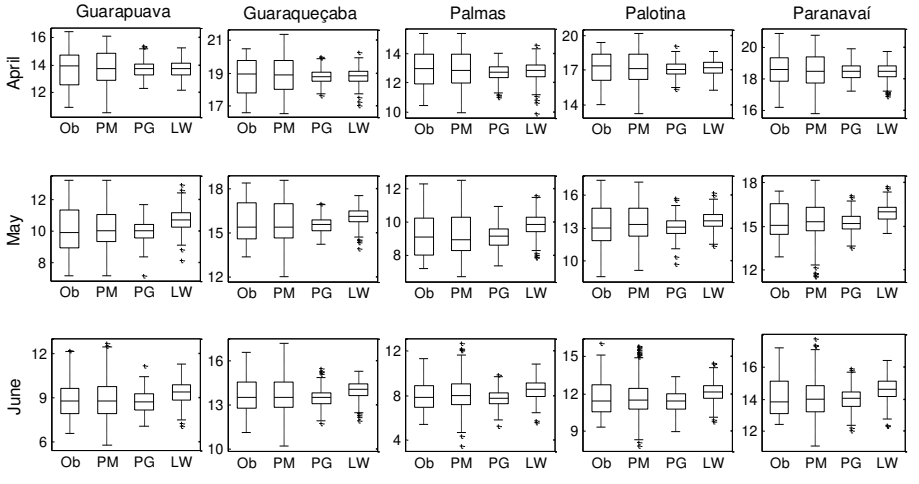

Figure 3 - Box plots of thirty minimum means observed (Ob) and (10 × 30 = 300) minimum mean simulated of air temperatures (ºC) for the autumn months in each station.

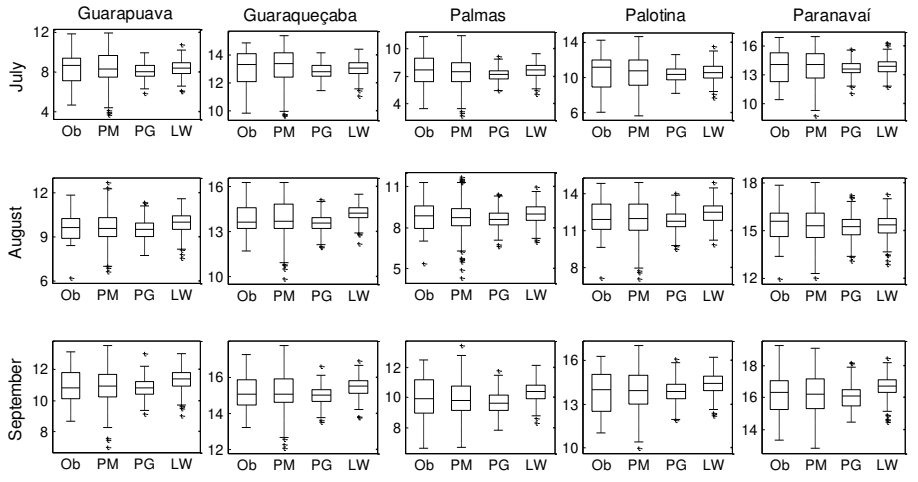

Figure 4 - Box plots of thirty minimum means observed (Ob) and (10 × 30 = 300) minimum mean simulated of air temperatures (ºC) for the winter months in each station.

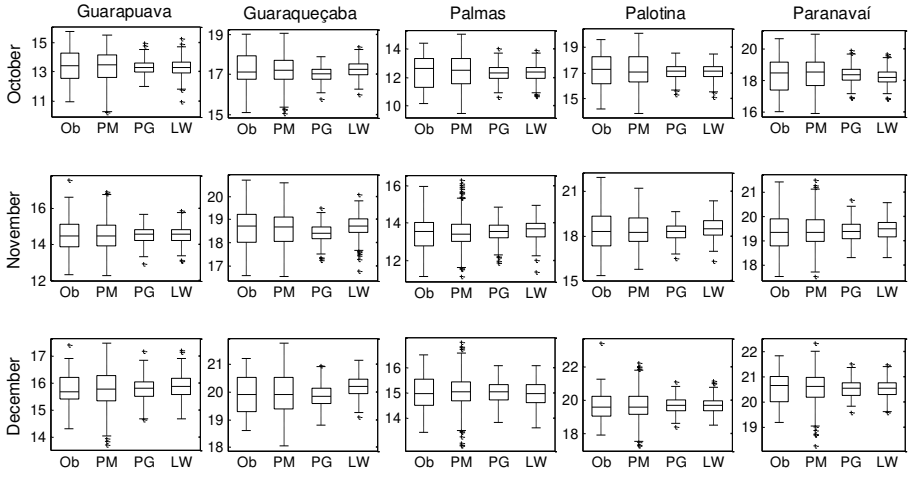

Figure 5 - Box plots of thirty minimum means observed (Ob) and (10 × 30 = 300) minimum mean simulated of air temperatures (ºC) for the months of spring in each station.

4. CONCLUSIONS

The results indicated that the proposed methodology in this research generated climatic series with interannual variability similar to the historical series. Mean minimum air temperatures simulated did not differed significantly from the historical mean minimum air temperature. The descriptive statistics, MBE, RMSE and Willmott agreement index, indicated a good degree of consistency between minimum air temperatures observed and generated by the proposed methodology in all station analyzed.

In all aspects analyzed (central tendency, variability and distribution form), the new methodology performed better than, when comparated with the

methodologies implemented in the wheather generators LARS-WG and PGECLIMA_R. The statistical properties, mean and standard deviations of monthly series were better reproduced in the series generated with the proposed methodology, and moreover, the model reproduced with good accuracity the observed interannual variability, reflecting the stochastic structure in the historical series.

The proposed methodology has the potential to be applied in other climatic variables, since it does not depend on the probability distribution used. Furthermore, is possible to generate easily scenarios by increasing or decreasing the proportion of months where the monthly average was statistically higher or lower than the historical average.

5.ACKNOWLEDGEMENTS

The authors thank to the CNPq and Fundação Araucária by the financial support of the research.

6. REFERENCES

CHEN, J.; BRISSETTE, F. P.; LECONTE, R. A daily stochastic weather generator for preserving low-frequency of climate variability. Journal of hydrology, v. 388, n. 3, p. 480-490, 2010.

CHEN, J.; BRISSETTE, F. P.; LECONTE, R. WeaGETS-a Matlab-based daily scale weather generator for generating precipitation and temperature. Procedia Environmental Sciences, v. 13, p. 2222-2235, 2012.

COUTINHO, M. D. L.; LIMA, K. C.; SANTOS AND SILVA, C. M. Improvements in precipitation simulation over South America for past and future climates via multi-model combination. Climate Dynamics, p. 1-19, 2016.

HANSEN, J. W.; MAVROMATIS, T. Correcting low-frequency variability bias in stochastic weather generators. Agricultural and Forest Meteorology, v. 109, n. 4, p. 297-310, 2001.

HANSON C. L.; CUMMING K. A.; WOLLHISER D. A.; RICHARDSON C. W. Microcomputer program for daily weather simulation in the contiguous United States. U.S. Dept. of Agriculture, Agricultural Research Service, Publication No. ARS-114, 1994.

HOOGENBOOM, G. Contribution of agrometeorology to the simulation of crop production and its applications. Agricultural and Forest Meteorology, Amsterdam, n.103, p. 137–57, 2000.

KHAZAEI, M. R., AHMADI, S., SAGHAFIAN, B., ZAHABIYOUN, B. A new daily weather generator to preserve extremes and low-frequency variability. Climatic Change, p. 1-15, 2013.

MAVROMATIS, T.; HANSEN, J. W. Interannual variability characteristics and simulated crop response of four stochastic weather generators. Agricultural and forest meteorology, v. 109, n. 4, p. 283-296, 2001.

MINUZZI, R. B.; VIANELLO, R. L.; SEDIYAMA, G. C. Oscilações climáticas em Minas Gerais. Revista Brasileira de Meteorologia, v. 25, n. 2, p. 227-236, 2010.

NICKS A. D.; LANE L. J. Weather Generator. In: Profile model documentation: Lane LJ, Nearing MA (1989). WEPP, 2.1-2.19. NSERL Report No.2, USDAARS, W. Lafayette, IN, 1989.

NICKS, A. D.; LANE, L. J.; GANDER, G. A. Weather generator: USDA-water erosion prediction project (WEPP). West Lafayette: USDA-ARS, chapter 2, 1995. 22p.

RICHARDSON, C. W. Stochastic simulation of daily precipitation, temperature, and solar radiation. Water Resources Research, v. 17, n. 1, p. 182-190, 1981.

RICHARDSON, C. W.; WRIGHT, D. A. WGEN: A model for generating daily weather variables. US Dept. Agric., Agricultural Research Service. Publ. ARS-8, 1984.

SEMENOV, M. A.; BARROW, E. M. Use of a stochastic weather generator in the development of climate changes scenarios. Climatic Changes, Amsterdam, v.35, p. 397-414, 1997.

SOLTANI, A.; HOOGENBOOM, G. A statistical comparison of the stochastic weather generators WGEN and SIMMETEO. Climate Research, Norwich, v.24, p.215-30, 2003.

STOCKLE, C. O., CAMPBELL, G. S., NELSON, R. ClimGen Manual. Biological Systems Engineering Department, Washington State University, Pullman, WA, 1999.

VIRGENS FILHO J. S.; OLIVEIRA R. B.; LEITE M. L.; TSUKAHARA R. Y. Desempenho dos modelos CLIGEN, LARS-WG e PGECLIMA_R na simulação de séries diárias de temperatura máxima do ar para localidades do estado do Paraná. Engenharia Agrícola, Jaboticabal, v.33, n.3, p. 538-547, 2013.

VIRGENS FILHO, J. S.; LEITE, M. L.; DAL GOBBO, B. L.; POBB, K.; FRUTEIRA, R. S. Analysis of the accuracy of daily series of global solar radiation simulated by the weather generator PGECLIMA_R, in the State of Parana, Brazil. Revista Brasileira de Geografia Física, v. 7, n. 1, p. 180-192, 2014.

WILLMOTT, C. J.; ACKLESON, S. G.; DAVIS, R. E., FEDDEMA; J. J., KLINK, K. M.; LEGATES, D. R.; ROWE, C. M. Statistics for the evaluation and comparison of models. Journal of Geophysical Research. v. 90(C5), p. 8995-9005, 1985.