

# Application of weather generation to high frequency and high resolution gridded datasets in Sao Paulo

## Abstract

The adjusted stochastic weather generator to disaggregate the monthly climate averages into daily timescale is investigated due to the limitations of Climate Models in representing the daily weather fluctuations. We use weather generator for daily precipitation and maximum and minimum temperature are investigated for Sao Paulo State. We use the adjusted stochastic weather generator to disaggregate the monthly climate averages into daily timescale due to the limitations of Climate Models in representing the daily weather fluctuations. The weather generator WGEN was adjusted considering the statistical distribution of temperature and precipitation from gauges to the present climate. This derived dataset was first applied to investigate the ability of a Regional Climate Model version 4 (RegCM4) in reproduce the spatial variability of those distributions. Then, we developed to generate realizations of daily temperature and precipitation for gridding datasets over Sao Paulo while preserving the spatial and temporal distribution as well as produce gridded dataset of the proprieties that define the precipitation and temperature intensity and frequency. For temperature, we consider topographies to take the interpolation using distance square inverse. Comparison and evaluating generated time series with observations and regional model shows that performance of the weather generator tested in term of the frequency and extreme events reproduced the observed statistical distribution.

**Keywords:** daily precipitation and temperature, weather generator

Volume I Issue 2 - 2014

Michel Nobre Muza,<sup>1</sup> Santiago Vianna Cuadra,<sup>2</sup> Rosmerir Porfirio da Rocha,<sup>3</sup> Marta Llopart,<sup>3</sup> Shigetoshi Sugahara<sup>4</sup>

<sup>1</sup>Federal Institute of Santa Catarina, Brazil

<sup>2</sup>Embrapa Temperate Climate, Brazil

<sup>3</sup>University of Sao Paulo, Brazil

<sup>4</sup>Program and Institute of Meteorological Research and Graduate Studies, University of Sao Paulo, Brazil

**Correspondence:** Michel Nobre Muza, Federal Institute of Santa Catarina, Florianopolis, State of Santa Catarina, Brazil, Tel +55 48 32210568, Fax +55 3224-0727, Email michel.muza@ifsc.edu.br

**Received:** May 16, 2014 | **Published:** June 13, 2014

**Abbreviations:** WGEN, weather generation; RegCM4, regional climate model version 4; CRU, climatic research unit; TRMM, tropical rainfall measurement mission

## Introduction

High frequency variability requires daily weather data for many applications like high resolution regional climate change impact assessment studies. High resolution datasets are major features to assess climate change impact.<sup>1</sup> In view of this, weather generation based in parameters of gauge stations has been used in order to maintain representativeness of distribution of frequency in daily data created in each weather event. To development of weather generator, a good estimation of the statistical parameters of probability distribution of each station or grid point is essential to produce synthetic data like precipitation, temperature and solar radiation. One useful application of daily precipitation and temperature is the capacity to provide an indication of the extremes, which is essential for risk assessment and for generating climate change scenarios, and evaluation of productivity impact to generate scenarios under extreme temperature.<sup>2</sup> Over several decades, weather generator has been shown to be extremely versatile for the useful analysis for many versions and adaptations of remote sites including area of poor gauges station and specific scenarios including agriculture area.<sup>3-5</sup> These stochastic weather generators require a number of local parameters as input to be able to generate a weather series for a specific site and they have been incorporated in a suite of crop models and computer programs integrated into a single software package in order to facilitate the application of crop-simulation models in research and decision-making.<sup>6</sup>

The weather generators can be used to disaggregate the monthly averages generated by Global or Regional Climate Models and

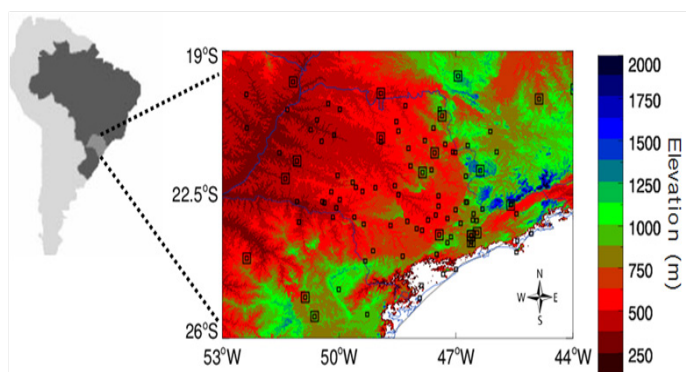
produce the overlap of the future climate anomalies on the time series of the present climate to investigate the impacts on agriculture, e.g.<sup>7</sup> Moreover, the changes in frequency distribution of extreme events have come under intense interest for their potential to impact agricultural productivity more than changes in mean climate.<sup>8</sup> These previous researches have demonstrated that for the determination of weather generation, different models have specific accurate for each region and purposes. It produces more robust daily dataset when more stations there is on site for parameters are properly adjusted. However, uncertainties can be more affected by process of temporal disaggregation than by the assessment or the variability of dataset.

High resolution surfaces are major features to assess climate change impact. In this way, downscaling regional methods has taken on particular evidence. However, the development downscaling methods also might substantially increase uncertainties then reducing the accuracy of the future climate scenarios. In view of all this, weather generation based in parameters of gauge stations have been used in order to maintain representativeness of distribution of frequency in daily data created in each daily precipitation event.<sup>1</sup> The weather generators methods is based in fitted of probability distribution function which include statistical parameters of according to each station or grid to produce synthetic precipitation.<sup>9</sup>

The objective of this study was to adjust the stochastic weather generator (daily precipitation and maximum and minimum temperature) to reproduce the weather stations observations as well as produce gridded dataset of the proprieties that define the precipitation and temperature intensity and variability. This derived dataset were first applied to investigate the ability of a Regional Climate Model in reproduce the spatial variability of those distributions.

## Materials and methods

The precipitation and maximum and minimum temperature data were used in this analysis came from a combined set of National Institute of Meteorology (Instituto Nacional de Meteorologia-INMET) and Agronomic Institute of Campinas in Sao Paulo (Instituto Agronomico de Campinas-IAC) meteorological station. The time periods considered were 1961 to 2011 for INMET and 1991 to 2011 for IAC stations. These time periods were selected based in the observations length record (time span), stations density (number of stations) and observation quality (missing data). Additional information from these data can also be obtained from<sup>10</sup> that used the same weather stations dataset. The station location and the topography available<sup>11</sup> over the Sao Paulo State are illustrated in (Figure 1). The study area is located over Southeastern Brazil and grid points over the Atlantic Ocean are masked out. Figure 1 shows that the topography over the region is dominated by the Serra do Mar, system of mountain ranges along the Southeastern coast, with ranging from 700 to 1200m. The highest peaks (nearly 2.5km) are located in southern region of the Minas Gerais State located to northern from Sao Paulo. The elevation decreases toward Tiete River in west of Sao Paulo, over Parana River Basin ranging 500 to 200m. The state of Sao Paulo with a very structured topography ranging from flat to the high mountains has an area of 248.000km<sup>2</sup> containing 108 stations onto the study area used here. The data were checked extensively in which missing data were not considered in the analysis. The original data were interpolated onto two regular grid 0.1x0.1 and 0.5x0.5 latitude-longitude and we selected a box that covered the Sao Paulo State sector in which the domain of analysis is limited to 26°S to 19°S, 53°W to 44°W.



**Figure 1** Relief map of 8km Wordclim (digital elevation in meters) over the State of Sao Paulo study area southeastern Brazil. Square denote meteorological station locations (108 totals), in which 15 are INMET station (double-square).

Daily simulations from RegCM4 and monthly data from CRU also are used in this study for comparison in which describe in details can be obtained from,<sup>12,13</sup> respectively. CRU includes monthly maximum and minimum air temperature, precipitation and it was used to compare the analysis against of weather generation dataset (<http://badc.nerc.ac.uk>). Tropical Rainfall Measurement Mission-TRMM<sup>14</sup> was used to compare with specific days during extreme events.

This study makes use of publicly available version weather generation model (WGEN), one of the broadest these used models. It was selected for following reasons: full method; data compression techniques;<sup>9</sup> the model can handle local and spatial parameter, most widely used and has shown successful result in future scenarios analysis. Others weather generation models have been developed and

their accuracy and skill differences are dependent for the region where it is applied.

The WGEN model determines the occurrence precipitation event based on two-state and first-order simplest from Markov chain. It considers the two transitional probabilities from a wet day to a wet day and from a dry day to a wet day. That is, the Markov chain determines the occurrence of rain on any given day. The model generates intensity of the synthetic precipitation time series based basically on the two parameter gamma distributions,  $\alpha$  (shape) and  $\beta$  (scale). The simulation of the occurrence of daily precipitation can be accomplished using the computer-generated random time series with these estimated parameters. The four precipitation parameters are constant for a given month but are varied from month to month. Since it is known that the precipitation pattern depends on the seasonality.

The precipitation scale factor for a given month is calculated as the mean monthly precipitation from actual data divided by the mean monthly precipitation generated with the Markov chain and gamma distribution algorithms. The generated daily precipitation amounts are multiplied by the precipitation scale factor for the appropriate month to obtain a precipitation amount.

The model uses the precipitation parameters for the Markov chain determines the occurrence of precipitation on any given day. Minimum temperature is sampled from a normal distribution dependently of precipitation occurrence. Maximum temperature is then generated depending on whether the generated day is wet or dry. Scale factor of temperature is based on the mean monthly temperature and it is calculated as the differences between the mean monthly temperature for the location and the mean monthly temperature theoretically generated using the parameters for the location. In this study, the solar radiation variable was supplied of the comparative analysis, because observed data can be hardly accomplished in this area.

The values of each of the four parameters were determined for the 108 stations in the around Sao Paulo State. The parameters were defined using two distinct datasets: 51years (1961-2011) and 21years (1991-2011) of daily precipitation and maximum and minimum temperature data for each station. These parameters that are required for each dataset being gridded data, RegCM4 and CRU-WGEN during every month, but here analyzed on period the Austral Summer (December to February) and Winter (June to August). The main feature of the wet season in southeastern Brazil is the enhanced convective activity and heavy precipitation in association with South American monsoon system, which typically starts in October–November, is fully developed during December–February and retreats in late April or early May.<sup>15</sup>

The analysis in thresholds of distribution revealed the extreme daily precipitation and minimum and maximum temperature calculated consistent with each fitted distribution are listed in (Table 1). This is not an aspect relative to sample restriction of the data that are discrete, because it is also observed for 50years of data on 20 stations from 1961 to 2011. The comparison between last 51 and 21years dataset indicated a small reduction in the precipitation for the smaller percentiles include median. On the other hand, it showed increase in the 90<sup>th</sup> percentiles of precipitation between last 51 and 21years. For temperature, the distribution showed an increase in the overall thresholds percentiles, except 85<sup>th</sup> percentiles for maximum and 99<sup>th</sup> percentiles for minimum temperature.

The Cressman interpolation method was used to derive the

precipitation fields. The Cressman interpolation method combines multiple passes made through the grid and a weighting factor.<sup>16</sup> Discussed that in an area where the first pass was poor, very discontinuities tended to develop there. In this case, we used a climatological field based in CRU monthly data.

The maximum and minimum temperature was spatially interpolated using an inverse-square-distance weighting function and then mapped to the elevation surface of a gridded elevation data. The inverse-square-distance interpolation was chosen because it is easy to implement the effect of elevation. In the case where some neighbors are very far from the interpolation point, the  $1/d^2$  weighting function ensures that distant stations receive proportionally little weight. Missing values were ignored during the interpolation when was encountered that station. As temperature is dependent on atmosphere pressure, which in turn decreases with altitude, the effect of elevation on temperature was account by consider the difference of elevation between the grid point and the station elevation. We consider the typical value used for the global mean environmental lapse rate is  $-6.5^\circ\text{C}/\text{km}$ .<sup>17</sup> The technique for incorporating elevation effects considers the temperature values normalized to sea-level equivalents, using the station elevation and a constant linear lapse rate adjustment. The temperature observations were then interpolated using inverse square distance smooth surface fitting approach. Finally, the interpolated sea-level temperature was adjusted back to actual temperature using the same lapse rate function, but temperature is converted to actual elevation of high resolution data.

The error analysis was performed to assess the interpolated precipitation by Cressman, Kriging and ISD techniques (Table 2). The sum of the differences divided by the total sample (BIAS) shows that the Cressmann presented an overestimation of 2.3mm/day compared to Kriging and 0.3 to ISD. Values less than 1mm/day mean that the error is smaller than the standard deviation, based on the normalized

Root Mean Squared. Although the bias of precipitation for three methods has different value, the RMS errors showed less than the standard deviation of the precipitation. In general, error is smaller than  $1.7^\circ\text{C}$  for daily maximum temperature and mean air temperature for three techniques.

Monte Carlo test was applied for evaluate the gridded data of the daily precipitation and maximum and minimum temperature on study area. We replace the grid point of the dataset with the stochastic series and repeating the same procedure as for the gridded data. For performed this analysis the dataset (from the 1991 to 2011) were divided into five sets, a random number generator was used to get the same period uniformly distributed random numbers. According to the each period, the dataset were reshuffled. Then, by using the thresholds of 95% confidence level, it was calculated correlation between gridded and simulations data for each grid point and the correlation of the area was compared with number significant of grid point. If the number of grid point at which the correlations are significant exceeds critical values for correlation of Pearson coefficient, the correlation can be considered reliable; otherwise it might occur by chance. The test was designed for the analysis of area in which wet day following a wet day (W/W) for precipitation were selected from the each 4 years data, a random number generator was used to get uniformly distributed random dataset. According to the test, correlations significant at 95% confidence level are greater than Monte Carlo test, except for first experiment during 1991 to 1994 (Table 3). This test was performed for a total of 5 experiments and the W/W event in the period for each test was generally different.

The daily precipitation and minimum and maximum temperature were interpolated onto a  $0.1^\circ$  latitude-longitude grid and it was selected an area that covered the State of Sao Paulo within  $26^\circ\text{S}$  to  $19^\circ\text{S}$ ,  $53^\circ\text{W}$  to  $44^\circ\text{W}$ . We included a grid of spatial resolution  $0.5 \times 0.5$  latitude-longitude for compare with CRU analysis and RegCM4 model.

**Table 1** Differences in the percentiles of the distributions of precipitation (fitted by Gamma) and minimum-maximum temperature (fitted by Gaussian) associated with the daily extremes in State of Sao Paulo) from Gauges from 1961 to 2010 (50y) and 1991 to 2010 (20y)

	Precipitation (Gamma)				Minimum   Maximum		Temperature (Gaussian)	
	50y	20y	20y-DJF	20y-JJA	50y	20y	20y DJF	20y JJA
5	4.5	4.2	2.2	0.4	9.7   18.6	10.3   22.2	17.2   26.1	8.1   20.4
10	4.5	4.2	2.2	0.4	11.3   22.9	11.7   23.9	17.8   26.9	9.6   21.7
MED	4.5	4.2	2.2	0.4	16.5   27.2	16.9   28.4	19.5   29.8	12.7   26.3
85	4.5	4.2	2.2	0.4	19.3   32.2	19.6   31.2	20.3   31.8	14.6   28.8
90	4.5	20.4	20.4	0.4	19.7   30.9	19.9   31.8	20.5   32.2	15.0   29.4
95	25.	20.4	20.4	0.5	20.2   31.8	20.3   32.5	20.8   32.7	15.6   30.5
99	23.7	23.7	23.7	2.2	21.0   33.3	20.9   33.9	21.2   33.6	16.6   32.2
AVE	4.1	4.	7.5	1.2	15.9   27.	16.3   28.1	19.3   29.7	12.4   25.9
STD	6.3	6.	7.2	3.2	3.3   3.2	3.3   3.1	1.1   2.1	2.3   3.1

**Table 2** Average results of bias and normalized Root Mean Squared by Cressmann and Kriging and ISD inverse-square-distance methods of Prec (mm), Tmax ( $^\circ\text{C}$ ) and Tmin ( $^\circ\text{C}$ ) at study area, during 1991-2011.

	Cressmann and Kriging		Cressmann and ISD	
	Bias	RMS.S	Bias	RMS.S
Prec (mm/day)	2.3	0.92	0.3	0.92
Tmax ( $^\circ\text{C}$ )	1.1	0.88	1.4	0.84
Tmin ( $^\circ\text{C}$ )	1.7	0.85	1.2	0.86

**Table 3** Correlation grid point number with t statistic reaching confidence test standards and thresholds for Monte Carlo tests

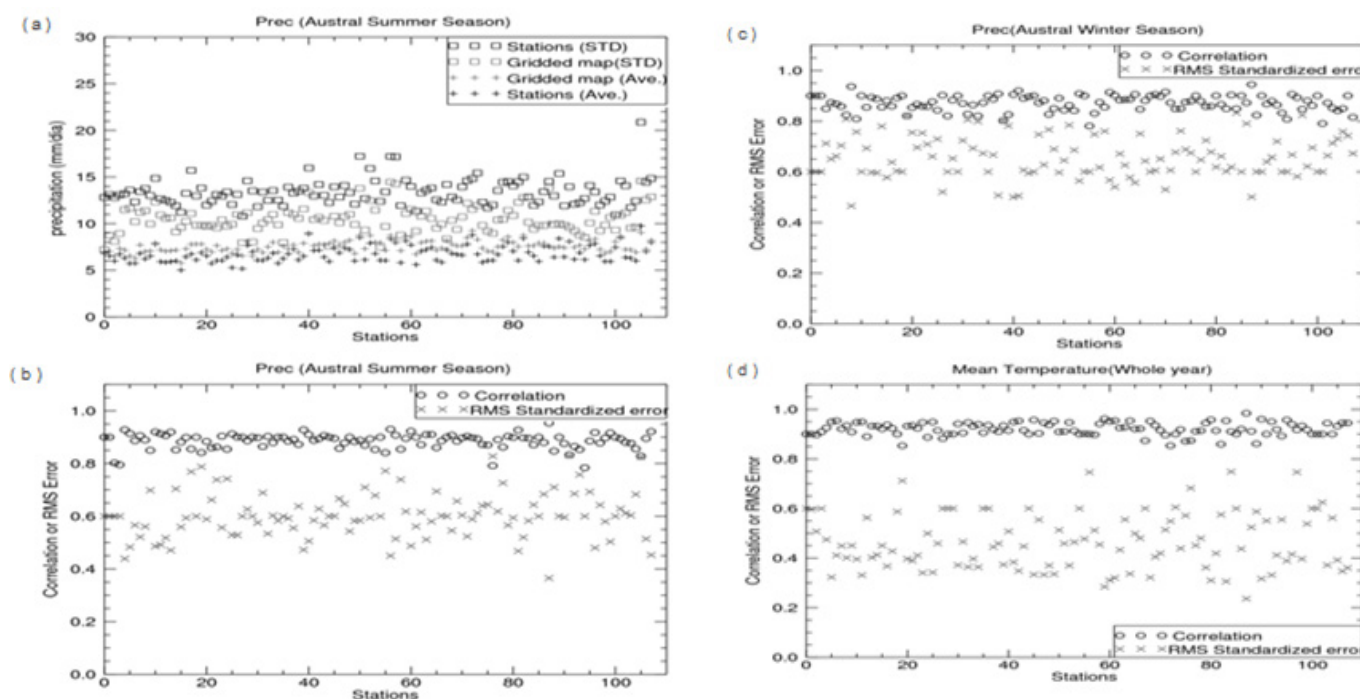
Variable	1991-94	1995-98	1999-2002	2003-06	2007-10
Correlation values	0.80	0.79	0.79	0.83	0.76
Monte Carlo test	0.81	0.75	0.66	0.83	0.70

## Results and discussion

The comparison of interpolated and observed precipitation showed an overestimate of 1.5mm for average daily precipitation on Summer Season (December to February), that is, gridded lightly overestimate the data compared station data (Figure 2A). While average gridded precipitation overestimate station data, the standard deviation ( $\sigma$ ) daily gridded precipitation underestimate variability of station. (Figures 2B) illustrate the linear correlation and normalized Root Mean Squared (RMS) error between interpolated and observed precipitation in each station during wet season. The correlation is high in all stations, with a typical value 0.9 in both wet and dry season (Figure 2A–2C), respectively. On the other hand, we observed that range of normalized RMS error was around 0.6 to 0.7 in wet and dry season, respectively. We noted that the stations close boundary exhibit high normalized RMS error between interpolated and observations precipitation data, sometimes with more than 0.6 values. This RMS error is less than  $\sigma$  and the same is observed on dry season. Different of the precipitation, comparative statistics for minimum and maximum temperature were stable across seasons and over years, for this only mean is shown

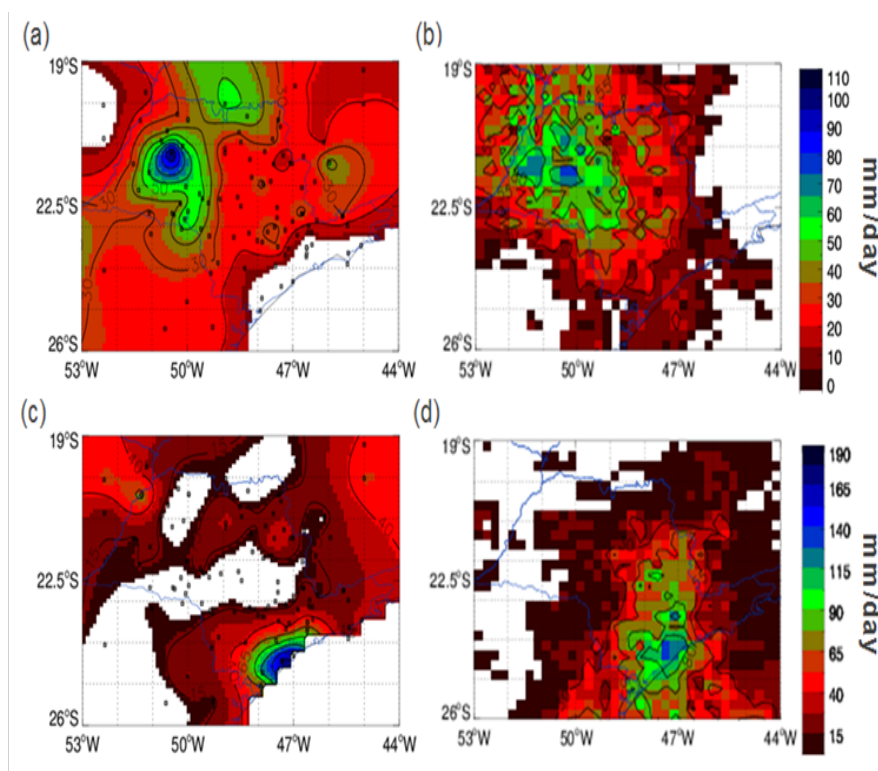
(Figure 2D). The normalized RMS error values for minimum and maximum temperature averaged around 0.5°C and the correlation were generally higher than 0.9.

Figure 3A shows interpolated precipitation for a period leading to one daily extreme precipitation on 95<sup>th</sup> percentile in the record, based on the station dataset. There were highest precipitation on station was 104.7mm. The comparison demonstrates the interpolated dataset showed as daily heavy precipitation occurred, especially the magnitude. It is demonstrated that the interpolated dataset to capture during daily extreme. During one of the wet daily extremes in the record, registered in 6<sup>th</sup> of January of 2006, gridded precipitation shows the most extreme values over western of the region, while relative less precipitation were recorded in the coast. On the other hand, during one of the wet daily extremes in the coast (Figure 3C), gridded precipitation shows the most extreme values coherently over coast region (registered in 13<sup>th</sup> of January of 2008). The precipitation in TRMM during both periods is shown in (Figure 3B & Figure 3D), such as comparison of daily values and poorer fine resolution of 0.25° latitude×0.25° longitude. The maximum is such as shown for gridded precipitation over western Sao Paulo and near coast, respectively.

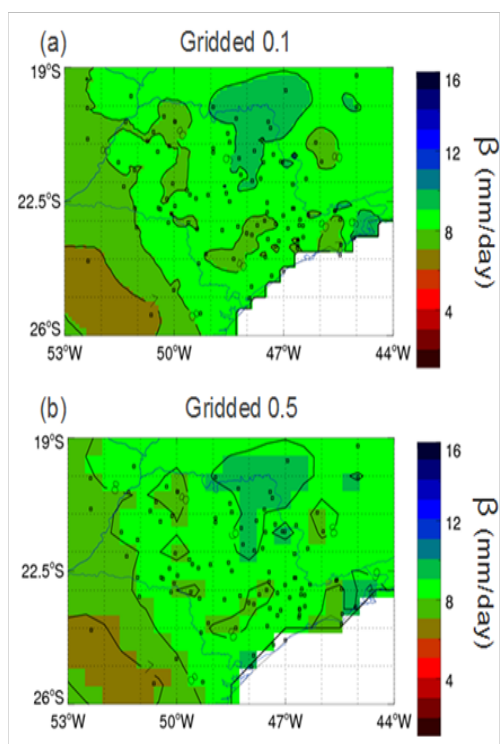


**Figure 2** Comparison between the gridded precipitation and observations from station: (A) average and standard deviation (B) correlation and RMS error precipitation during the summer, (C) winter, (D) whole year of the mean temperature.



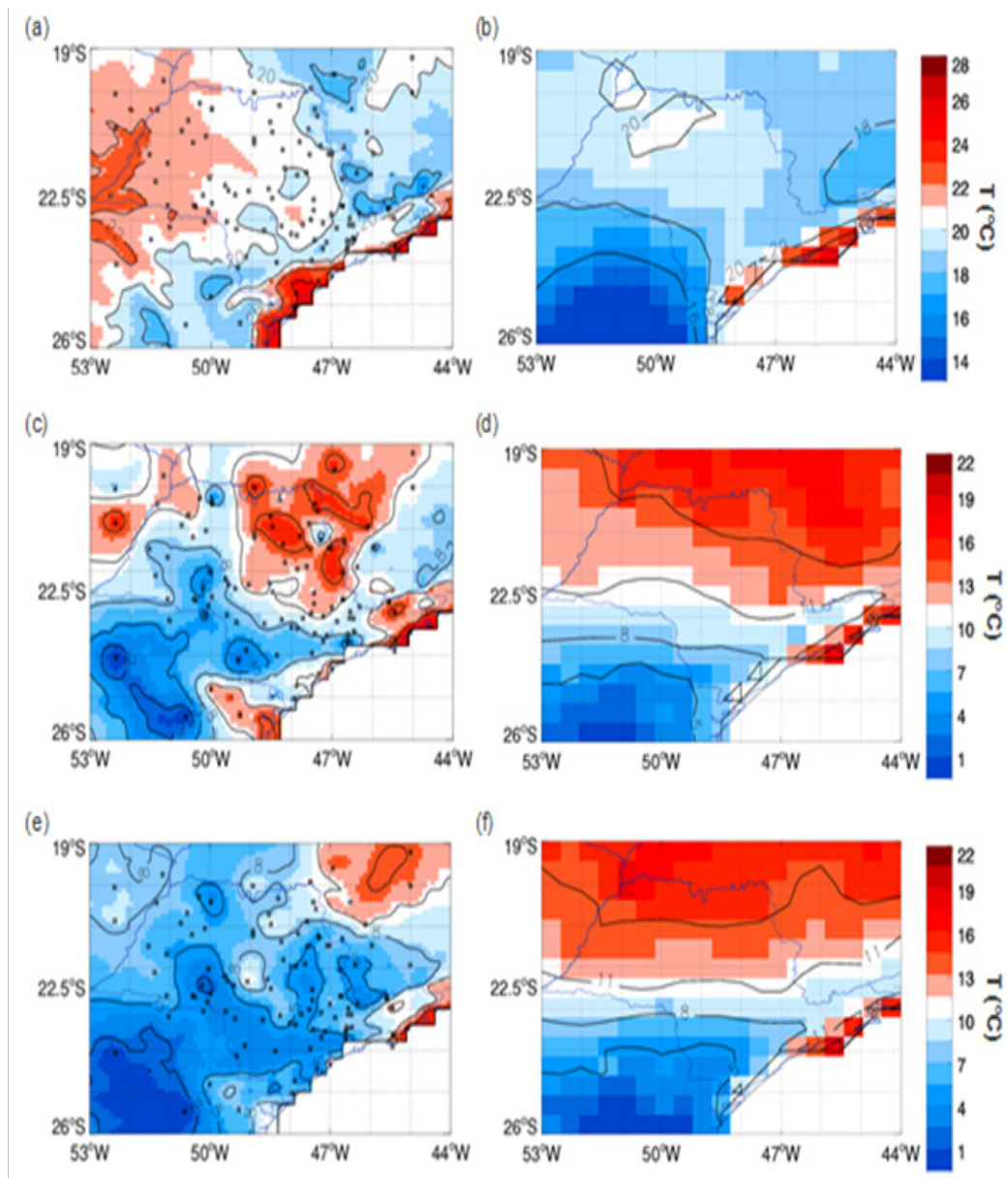


**Figure 3** Interpolated precipitation (mm/day) in 6 of January of 2007 (A) gridded and (B) TRMM dataset and in (C,D) 13 of January of 2008. Square denote meteorological station locations in gridded precipitation. Color scale is different for each one. Data grid spacing is  $0.1^\circ$  ( $0.25^\circ$ ) lat/lon in gridded precipitation (TRMM).



**Figure 4** Shape  $\beta$  parameters (units are mm/day) of gamma probability distribution functions fitted to (A) gridded precipitation spacing  $0.1^\circ$  lat/lon (B) Gridded data  $0.5^\circ$  lat/lon over State of Sao Paulo for DJF. Contour interval is 2mm/day. Square denote meteorological station location

Figure 5 compares the spatially the daily minimum temperature from gridded data and RegCM4 model during hot and cold extreme in the record (95<sup>th</sup> and 5<sup>th</sup> percentile, respectively). It is notable how the gridded dataset is able to reproduce relative minimum values in regions of higher topography, such as the Serra do Mar, in eastern part of the domain (Figure 5A). The RegCM4 model shows be more colder than gridded minimum temperature in around  $4^\circ\text{C}$  in southern and western from Sao Paulo regions (Figure 5A & Figure 5B), according with.<sup>12</sup> However, gridded minimum temperature and RegCM4 model reproduce interested aspects over the coast and northwestern regions. For example, differences between both datasets are less than  $2^\circ\text{C}$ , but considerable differences exist, however, on northwestern from Sao Paulo. During lag one day of the cold daily extremes, in 16th of June of 2008, gridded minimum temperature shows the lowest extreme values over plain of southwestern of the region, while relative high minimum temperature were recorded in northward (Figure 5C–5E). This situation change abruptly during next day and a relative less minimum temperature is observed. In general, comparison between datasets shows relevant features such as optimal resolution for validation studies, which are the baseline for future climate scenarios studies. The results above indicate that, although several aspects of the general characteristics of a specific case are well represented, some important inconsistencies are also found. In fact,<sup>12</sup> showed that RegCM4 is systematic colder slightly greater in the DJF than in the JJA. The gridded dataset is of great value contributing in the validation of climate models. Climate model simulations give information as area mean and comparison between observed and simulated data should be carry out carefully, since weather stations represent variables at specific locations.



**Figure 5** Interpolated minimum temperature (°C) in (A) 6 of January of 2007 and (B) 16 of June of 2008 and (C) 17 of June of 2008. Note that color scale is different between January and June. Square denote meteorological station locations. Data grid spacing is  $0.1^\circ$  ( $0.5^\circ$ ) lat/lon in interpolated minimum temperature (RegCM4).

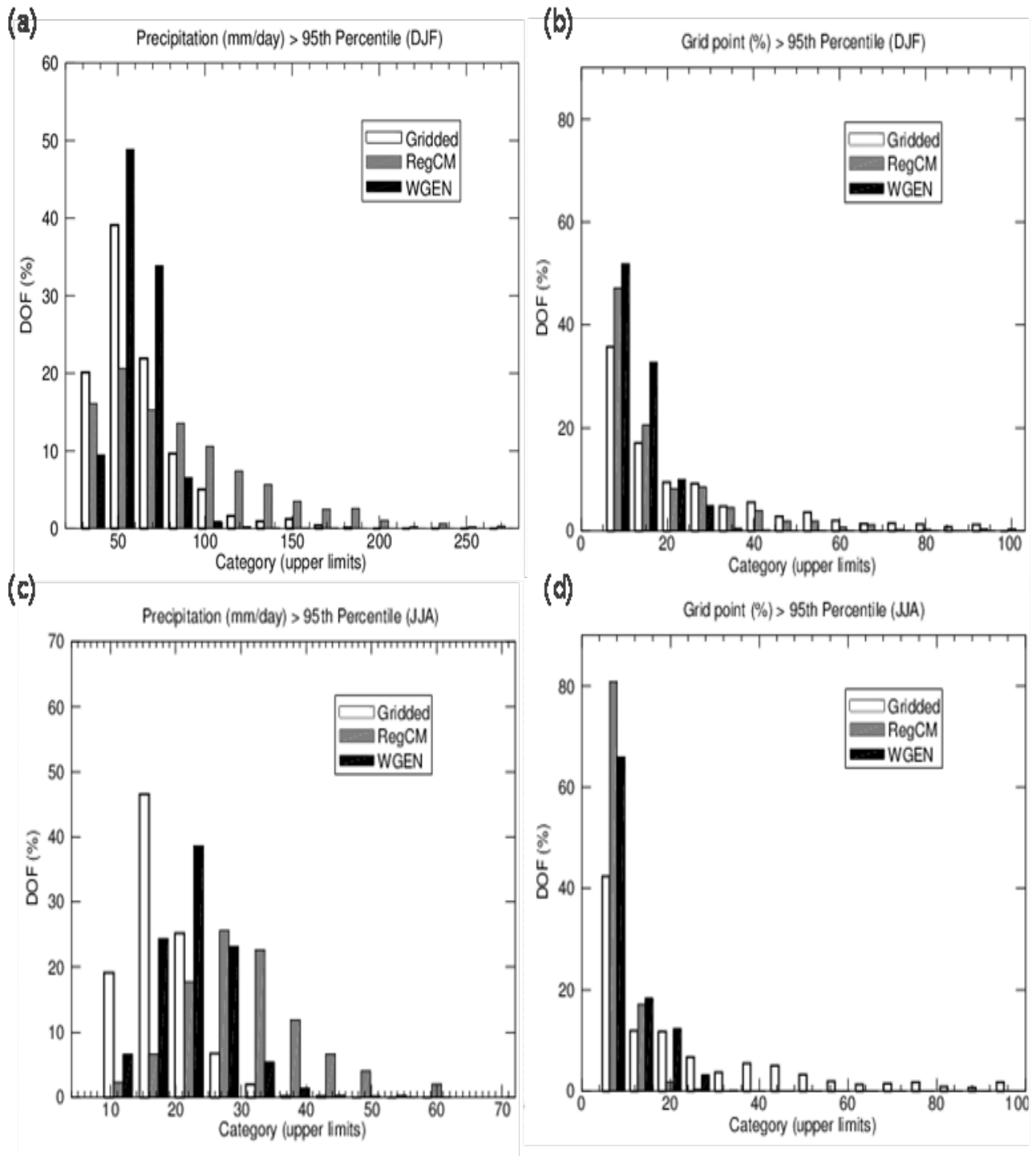
Figure 6 shows the Gridded precipitation and RegCM4 and WGEN (fitted for CRU dataset) that exceeds the 95<sup>th</sup> percentile distribution of precipitation during summer (Figure 6A) and winter (Figure 6C). That is, daily extremes obtained by averaging over grid point that exceeds this threshold into the boxes. There appears to be a slight positive skew (a tail to large values), except for RegCM4 which clearly departs from a Gaussian distribution. In this context, gridded precipitation and WGEN distribution were more symmetrical with lower frequencies at the lower and upper ends of the distribution. The results the precipitation exceeds the 95<sup>th</sup> percentile distribution of precipitation during winter appears to be a small positive skew for Gridded precipitation and WGEN-CRU, comparable to that of RegCM4 (Figure 6C). In fact, distribution of RegCM4 extremes during winter season shows a relatively large positive skew, which tend to a

Gaussian distribution. (Figure 6B) (DJF) and (Figure 6D) (JJA) show the frequency distribution of Gridded precipitation and RegCM4 and WGEN categories separated according to grid point percentage. The total over Sao Paulo area of extremes (exceeding 95<sup>th</sup> percentile) in Gridded precipitation is approximately 10% smaller than RegCM4 and WGEN extremes with similar frequency distributions (Figure 6B). About 40% of Gridded precipitation occurred in less than 10% Sao Paulo area (Figure 6D) while are 63% and 81% for RegCM4 and WGEN, respectively.

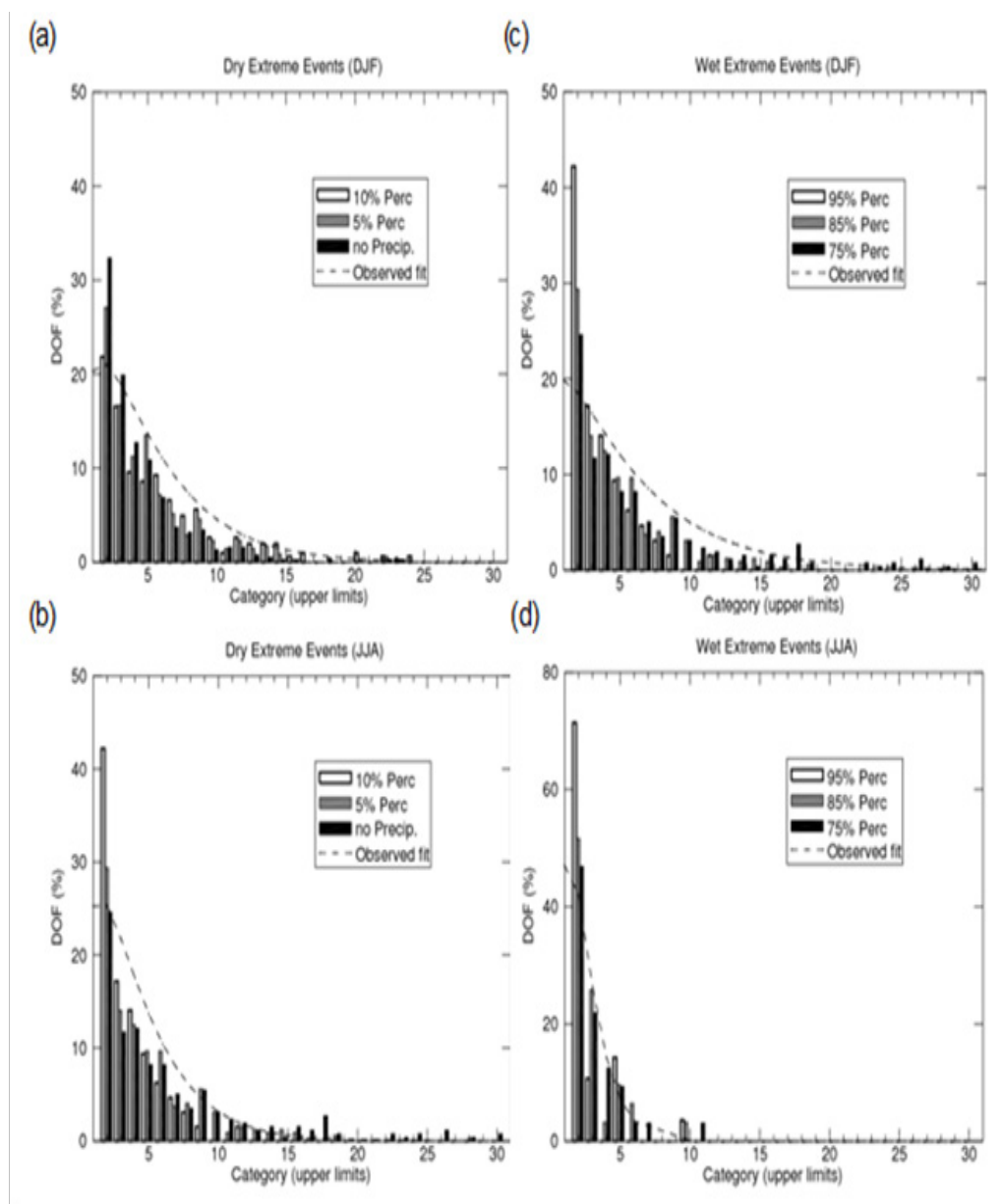
An important issue related to the WGEN extremes is its persistence here investigated on 2 to 30 days. The persistence was reported for five different thresholds: less than the 5<sup>th</sup> and 10<sup>th</sup> percentile and wetter than the 85<sup>th</sup>, 90<sup>th</sup> and 95<sup>th</sup> percentile. Moreover, we include the persistence of events in which there was not precipitation. The

persistence is analyzed by (Figure 7A-7D) with dry and wet extreme events in both season (JJA and DJF). It is interesting to note that the frequency distribution in upper threshold extremes during summer season are comparatively equals with skew and upper tails similar (Figure 7A & Figure 7C). That is, it can be recognized which persistence of events receiving substantially less or slightly less as well as slightly more, or substantially more precipitation are

comparable considering climatological distribution. The probability of dry and wet extreme events of persistence more than 10days is less than four percent. The frequency distribution during the winter produced two main categories below and above of 5mm. (Figure 7C) shows wet extreme events that it is substantially lower on winter as compared to summer or with dry extreme events.



**Figure 6** Frequency distribution of 95<sup>th</sup> percentile of extreme daily (A,C) precipitation amount and (B,D) Grid percentage partitioned for gridded and RegCM4 and WGEN-CRU dataset. Time period is (A,B) DJF and (C,D) JJA.



**Figure 7** Histogram illustrating the distributions of persistence 2 to 30 days of dry (A,B) and wet (C,D) extreme events on the threshold of 5 and 10% for dry events (included no precipitation) and the threshold of 75, 85 e 95% for wet extreme events.

## Conclusion

The stochastic method for generating daily precipitation and minimum and maximum temperature data preserving frequency distribution reproduced the main characteristic of the observed data in individual weather station as well as the gridded data and regional model. The distributions for observed gridded precipitation and synthetic data generated WGEN by three dataset, Gridded data and RegCM4 and CRU, are able to reproduce accurately enough the shapes of the distributions. As expected, the model did not show frequency and variance in the seasonal to high frequency timescale like observed data. Disaggregation of monthly data as CRU appears to be useful since to be possible to obtain good parameters of probability in this case wet to wet days and wet to dry days and fitted distribution Gamma for precipitation and Gaussian for temperature.

An interpolation procedure for the gridded precipitation and temperature maximum and minimum for use weather generator spatially has been described. The interpolation data of stations yielded grids of precipitation and minimum and maximum temperature for each day analyzed here in two compared sample from 1961 to 2010 and from 1991 to 2010. We used spatial interpolation method with Cressman and constructed gridded data in 1991 to 2010 at two spatial resolutions (0.1 and 0.5lat-lon). Analysis shows similar results in both resolutions. For temperature, the interpolation procedure, based on a combination of an inverse-square-distance method was used incorporating particularly landscapes with elevation. In gridded minimum and maximum temperature, adjust for elevation was made using typically the lapse rates. We compared the average and standard deviation based on the daily data gridded and stations. We observed that the variance of the interpolated values was of the less order than



the variance of the observations. The station located nearly boundary exhibit major differences between gridded and station. In general, all stations are very well reproduced by interpolation method.

Analysis was performed to assess comparatively the newly estimated precipitation by CRU-WGEN with daily extreme event and RegCM4 simulations. Overall, both datasets reproduce the extreme daily data, except in broad around the extreme precipitation. Result shows that the frequency distribution agrees with observed precipitation and RegCM4 as well as for dry and wet extreme. Moreover, the frequency distribution during summer and winter shows similar aspects except for different intensity. RegCM4 precipitation tends to produce overestimate in relation gridded precipitation. Based on percentage area with daily extreme, WGEN shows smaller (larger) percentage than RegCM4 and Gridded precipitation during summer (winter). Daily precipitation and temperature data are the most interesting climatic variable, due to the combination of high frequency associated with extreme events. The WGEN model reproduced the distribution of frequency in wet and dry events of precipitation, persistence in days, magnitude.

## Acknowledgements

MN. Muza was supported through a CNPq fellowship (160888/2012-3).

## Conflict of interest

The author declares no conflict of interest.

## References

1. Baigorria GA, Jones JW. GiST: A stochastic model for generating spatially and temporally correlated daily rainfall data. *J Climate*. 2010;23(22):5990–6008.
2. Hayhoe HN, Stewart DW. Evaluation of CLIGEN and WXGEN weather data generators under Canadian conditions. *Canadian Water Resources Journal*. 1996;21(1):53–67.
3. Geng S, de Vries PFWT, Supit I. A simple method for generating daily rainfall data. *Agr Forest Meteorol*. 1986;36(4):363–376.
4. Semenov MA, Brooks RJ. Spatial interpolation of the LARS-WG stochastic weather generator in Great Britain. *Clim Res*. 1999;11(2):137–148.
5. Chen J, Brissette FP, Leconte R. A daily stochastic weather generator for preserving low-frequency of climate variability. *J hydrol*. 2010;388(3-4):480–490.
6. Soltani A, Hoogenboom G. A statistical comparison of the stochastic weather generators WGEN and SIMMETEO. *Clim Res*. 2003;24(3):215–230.
7. Haim D, Shechter M, Berliner P. Assessing the impact of climate change on representative field crops in Israeli agriculture: a case study of wheat and cotton. *Clim Change*. 2008;86(3-4):425–440.
8. Lobell DB, Burke MB, Tebaldi C, et al. Prioritizing climate change adaptation needs for food security in 2030. *Science*. 2008;319(5863):607–610.
9. Richardson CW. Stochastic simulation of daily precipitation, temperature, and solar radiation. *Water Resour Res*. 1981;17(1):182–190.
10. Liebmann B, Allured D. Daily precipitation grids for South America. *Bull Amer Meteor Soc*. 2005;86(11):1567–1570.
11. Hijmans RJ, Cameron SE, Parra JL, et al. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*. 2005;25(15):1965–1978.
12. Giorgi F, Coppola E, Solmon F, et al. RegCM4: model description and preliminary tests over multiple CORDEX domains. *Inter-Research*. 2012;52:7–29.
13. New M, Lister D, Hulme M, et al. A high-resolution data set of surface climate over global land areas. *Clim Res*. 2002;21(1):1–25.
14. Huffman GJ, Adler RF, Bolvin DT, et al. The TRMM multi-satellite precipitation analysis: quasi-global, multi-YEAR, combined-sensor precipitation estimates at fine scale. *J Hydrometeorol*. 2007;8(1):38–55.
15. Muza MN, Carvalho LMV, Jones C, et al. Intraseasonal and interannual variability of extreme dry and wet events over southeastern South America and subtropical Atlantic during the austral summer. *J Climate*. 2009;22(7):1682–1699.
16. Cressman GP. An operational objective analysis system. *Mon Wea Rev*. 1959;87(10):367–374.
17. Dodson R, Marks D. Daily air temperature interpolated at high spatial resolution over a large mountainous region. *Clim Res*. 1997;8(1):1–20.