Certainties and Uncertainties in Observational and Reanalysis Data in the Study of Extreme Precipitation Events over La Plata Basin

José Leandro P.S. Campos , Tercio Ambrizzi
Institute of Astronomy Geophysics and Atmospheric Sciences – University of São Paulo

1 – Introduction

Currently many scientific papers have been using reanalysis data produced by weather and climate prediction centers and regular gridded interpolated data derived from surface stations. Due to this wide use and lack of reliable and easy access time series, is advisable to check the capacity of this data in represent climate extremes of precipitation. The aim of this study is to analyze the existing deviations in these data sets, when considering extreme precipitation events.

2 - Methodology

Six data sets of daily precipitation were used in this work, the NCEP/NCAR and ECMWF/ERA40 reanalysis; three interpolated gridded data, the CU/CIRES with resolution of 1.0° and 2.5°, the CPC with resolution of 0.5° and data from 108 meteorological stations along the La Plata basin from CLARIS-LPB project, in the period of 1979-2002, where all the data sets were compared with the CLARIS-LPB data (observational data). Some specific indices, such as, the PRCPTOT, Rx1day, SDII and CWD were calculated and statistical analysis like correlations, scatter plots and empirical cumulative density functions (CDF).

To analyze the departure between the CDF's the gridded data and the station data, was computed here the Kolmogorov-Smirnov test, where is computed the difference between two CDF's.

$$D_S = \max_X |F_n(x_1) - F_m(x_2)| \ , \qquad (eq.1)$$

Where $F_n(x_{(i)}) = i/n$ and $F_m(x_{(i)}) = i/m$. The K-S test looks the largest difference between the empirical distribution function of n_1 observations and n_2 observation of x_2 (Wilks, 2011). Where the hypothesis test of the two samples was draw by the same distribution is rejected at $\alpha \cdot 100\%$ level if:

$$D_S > \left[-\frac{1}{2}\left(\frac{1}{n_1}+\frac{1}{n_2}\right) ln\left(\frac{\alpha}{2}\right)\right]^{1/2}, \qquad \text{(eq.2)}$$

Then here is suitable that the D_s be rejected, in other words the two CDF's will be draw with the same distribution if D_s be rejected at level $\alpha \cdot 100\%$.

Were chosen nine points along La Plata basin to do the index time series analysis and CDF analysis, showed in the fig.1.

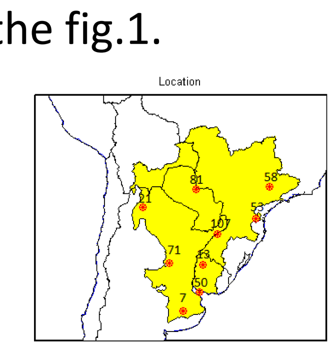

Fig.1 – Nine points chosen to do the analysis of extremes and CDF.

Tab.1 – Extreme Precipitation Indices for climate change

ID	Name	Definition	Unit
Rx1day	Max 1-day amount precipitation	Let RR_{ij} be the daily precipitation amount on day i in the period j. The maximum 1 day amount in the period j $$Rx1day_j = \max(RR_{ij}) \qquad (eq.3)$$	mm
Prcptot	Annual total wet day precipitation	Let RR_{ij} be the daily precipitation amount on day i in period j. If I represent the number of days in j then: $$prcptot_i = \sum_{i=1}^{I} RR_{ij} \qquad (eq.4)$$	mm
SDII	Simple intensity index	Let RR_{wj} be the daily precipitation amount on wet days, w (RR>1mm) in period j. If W represent the number of wet days in j, then: $$SDII_j = \frac{\sum_{w=1}^{W} RR_{wj}}{W} \qquad (eq.5)$$	mm/day
Cwd	Consecutive wet days	Let RR_{ij} be the daily precipitation amount on day i in the period j. Count the largest number of consecutive days where $RR_{ij} \geq 1$mm.	Days

3 – Cumulative Density Function and K-S test

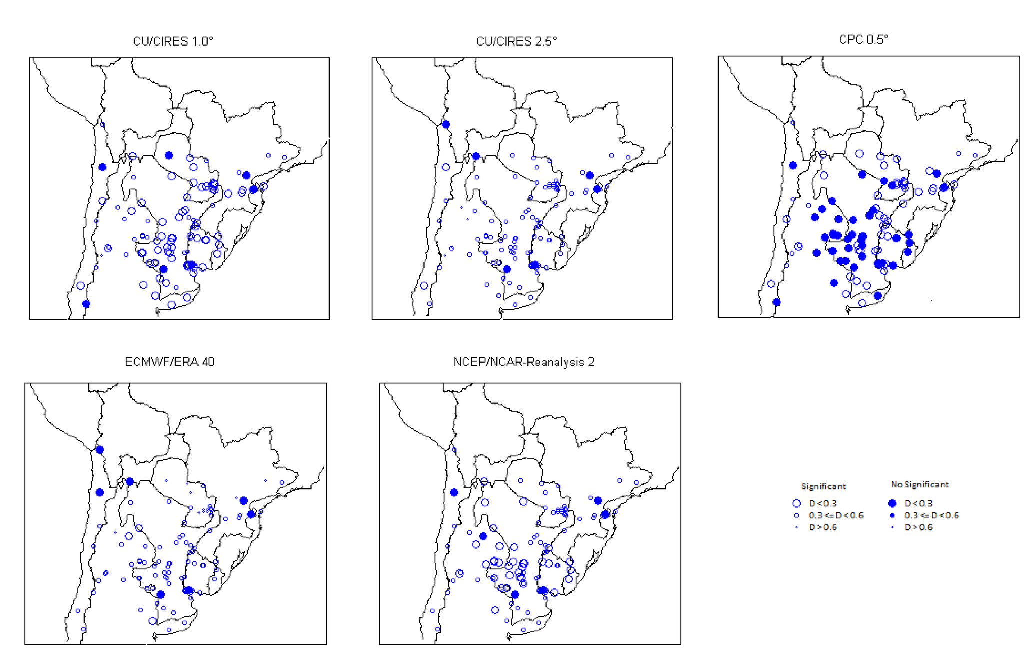

Fig.2 - Kolmogorov-Smirnov test between station data and gridded data (specified in the title of the maps) in the period of 1979-2002, values of D_s no significant at 5% level are filled and significant are not, no significant values means that the CDF of station data and gridded data can be represented by the same CDF at 5% level.

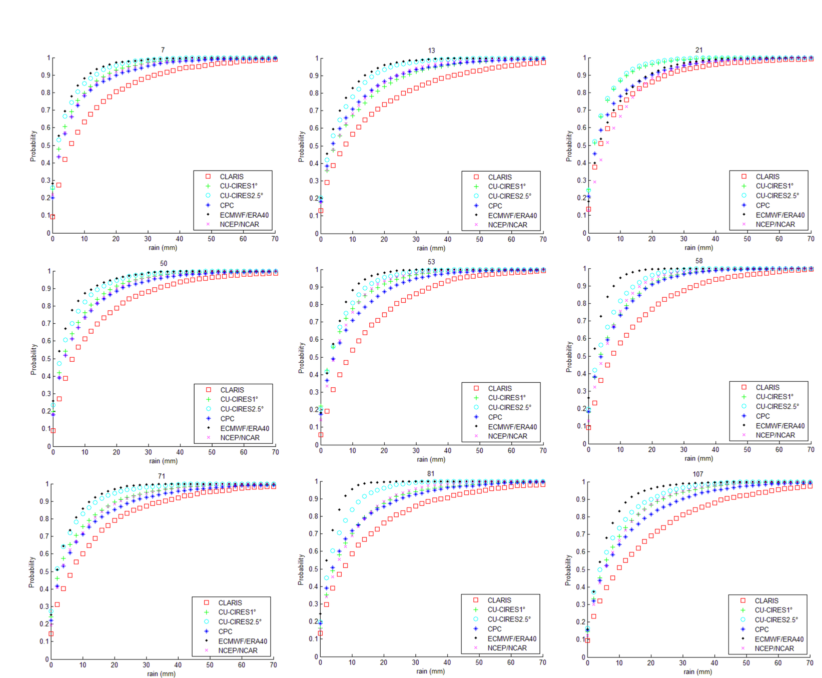

Fig.3 – Cumulative Density Function for daily precipitation in the period of 1979-2002, for the nine points chosen

Fig.2 shows that the CPC data has more no significant data than others gridded data, it can also be seen that the most no significant data in CPC data set are concentrated in central Argentina and Uruguay.

In the fig.3 it is observed that the curves corresponding to the gridded data grows faster than the station data and the precipitation is underestimated by the gridded data.

5 – Extreme Precipitation Indices

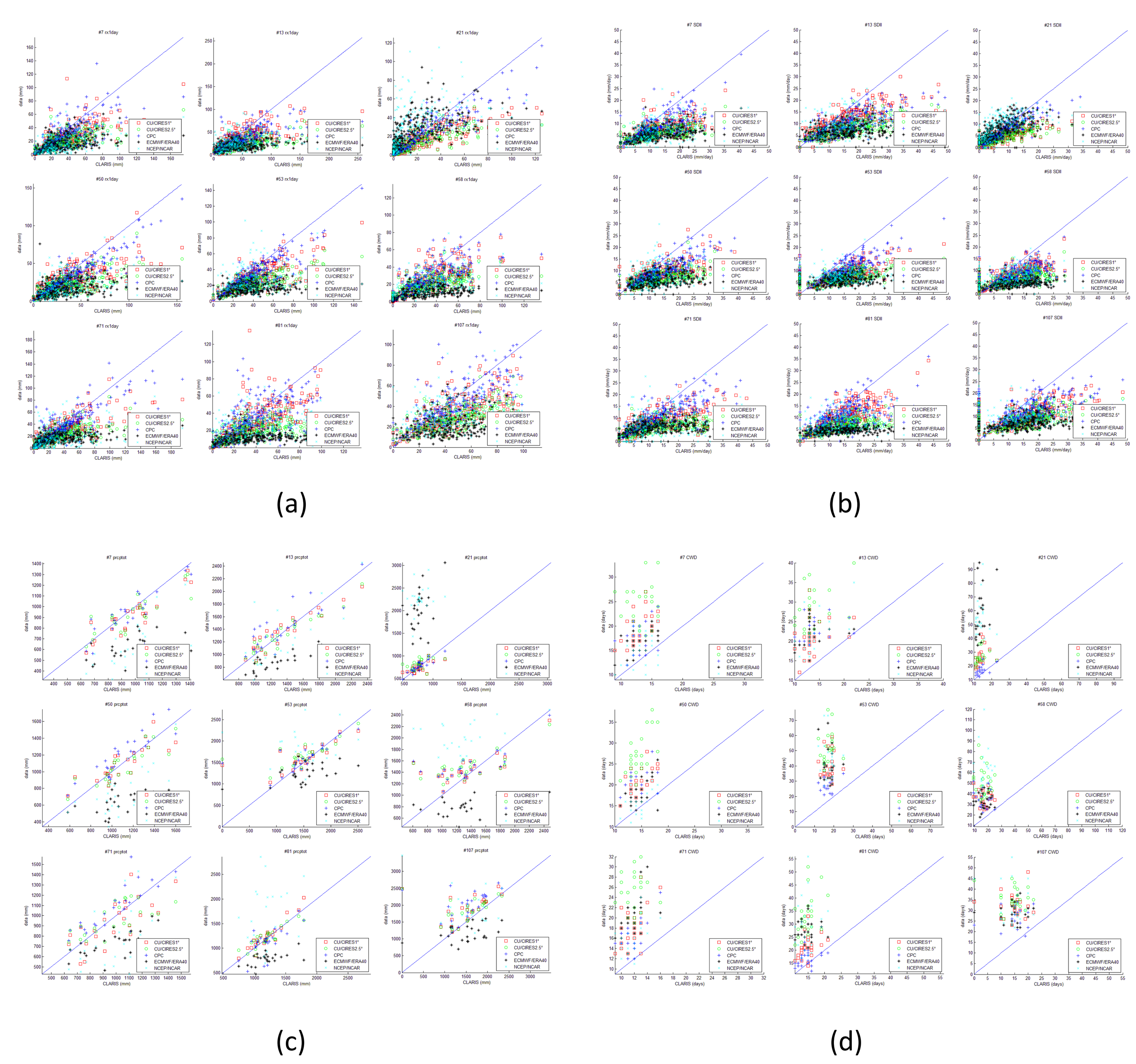

Fig.4 - Scatterplots for the precipitation indices in the period 1979-2002, between the station data (x axis, labeled here as CLARIS) and gridded data (y axis, labeled as data), (a) Rx1day, (b) SDII, (c) PRCPTOT and (d) CWD.

Analyzing the scatter plot figures, one can see that the monthly precipitation indices underestimate the daily precipitation and the yearly indices shows overestimation. In all plots is observed that the nearest points of the line 1:1 is the CPC analysis, and the PRCPTOT index is the one with the best agreement between the station and gridded data.

6 – Conclusions

- The results presented suggest that the gridded data underestimate the extreme precipitation events been a result well demonstrated by the CDF's;
- In the gridded data there is more days with light rain than in station data. This result is obtained by analyzing the "growth" of CDF's;
- The monthly precipitation indices are underestimated by the gridded data in relation to the station data;
- The yearly indices show overestimation by the gridded data;
- Other analyzes are currently being made.

7 – References

- Wilks, D.S., 2011,*Statistical Methods in the Atmospheric Sciences* 3rd ed, Elsevier, 705pp
- Karl, T.R., N. Nicholls, and A. Ghazi, 1999: CLIVAE/GCOS/WMO workshop on indices and indicators for climate extremes: Workshop summary. Climatic Change, 42, 3-7.
- Liebmann, B., and D. Allured, 2005: Daily precipitation grids for South America. Bull. Amer. Meteor. Soc., 86, 1567-1570.

Corresponding Authors

Name: José Leandro P.S. Campos and/or Tércio Ambrizzi
Organization: Department of Atmospheric Sciences - University of São Paulo
Address: Rua do Matão, 1226
São Paulo, 05508-090
Brazil
Email Adress: campos@model.iag.usp.br
ambrizzi@model.iag.usp.br