



EPiC Series in Engineering

Volume 3, 2018, Pages 626–634

HIC 2018. 13th International
Conference on Hydroinformatics



Chaotic Statistical Downscaling (CSD): Application and Comparison in the Bogotá River Basin

Freddy Duarte¹, Gerald Corzo², Germán Santos¹ and Oscar Hernández³

1. Escuela Colombiana de Ingeniería Julio Garavito, Bogotá, Colombia

2. UNESCO-IHE Institute for Water Education, Delft, The Netherlands.

3. The University of Iowa, Iowa City, United States.

freddy.duarte@escuelaing.edu.co, g.corzo@unesco-ihe.org,
german.santos@escuelaing.edu.co, ohernandezmurcia@uiowa.edu

Abstract

This study presents a new statistical downscaling method called Chaotic Statistical Downscaling (CSD). The method is based on three main steps: Phase space reconstruction for different time steps, identification of deterministic chaos and a general synchronization predictive model. The Bogotá river basin was used to test the methodology. Two sources of climatic information are downscaled: the first corresponds to 47 rainfall gauges stations (1970-2016, daily) and the second is derived from the information of the global climate model MPI-ESM-MR with a resolution of $1,875^\circ \times 1,875^\circ$ daily resolution. These time series were used to reconstruct the phase space using the Method of Time-Delay. The Time-Delay method allows us to find the appropriate values of the time delay and the embedding dimension to capture the dynamics of the attractor. This information was used to calculate the exponents of Lyapunov, which shows the existence of deterministic chaos. Subsequently, a predictive model is created based on the general synchronization of two dynamical systems. Finally, the results obtained are compared with other statistical downscaling models for the Bogota River basin using different measures of error which show that the proposed method is able to reproduce reliable rainfall values (RMSE=73.37).

Introduction

Global models for the management and planning of water resources in different parts of the world are being developed and used to simulate climate future scenarios. Global Climate Models (GCM) of different organizations in the world provide meteorological information for multiple future scenarios at large scale. In this sense, new challenges have been presented to be able to use this information at

the basin level. GCM generate climate information of fundamental variables at a cell scale that may need to be transform to a smaller basin scale. Different techniques exist in this area, from physical models to stochastic models, as observed in (Corzo, Jonoski, Yimer, Xuan, & Solomatine, 2009), among the most used techniques are highlighted: ANN (Artificial Neural Network), SDSM (Statistical Downscaling Model), ADC (Advanced Delta Change Method) and WRF (The Weather Research and Forecast model).

However, considering that the climate systems and its associated dynamical processes are essentially non-linear, and possibly chaotic, the effectiveness of these techniques may be limited and the deterministic estimation of the precipitation obtained from GCMs for hydrological modeling is difficult, as described in (Sivakumar & Ronny, 2010). In view of this, it is necessary the creation of a downscaling model that considers explicitly the properties of the chaotical systems that have the potential to improve downscaling of rainfall.

The analysis of deterministic chaos in time series has been developed in different study areas, being widely used in forecasts of time series and estimation of missing data as presented in (Gallego, 2010) and (Siek, 2011), also allowing the creation of predictive models, such as those presented in (Hernández, 2009). The application of the concepts of nonlinear dynamics and chaos theory in hydrological studies in recent decades has been considerable, especially in the study of the presence of deterministic chaos in precipitation time series. One of the pioneers was Hense, who in 1987 analyzed a series of 1008 monthly rainfall records in Nauru, observing deterministic chaos in a low dimension of the phase plane. Rodríguez-Iturbe (1989); Shafiri (1990); Tsonis (1993); Jayawardena and Lai (1994); Waelbroeck (1998); Berntsson (1994); Georkakos (1995); Sivakumar (2000) and Gaume (2002) have also reported the presence of deterministic chaos in precipitation time series.

Material and methods

The Bogotá river originates in the central part of Colombia in South America, specifically in the department of Cundinamarca. It's basin covers 32% of Cundimarca's territory with an area of 5891 km^2 . Figure 1 shows the general location of Bogotá River Basin. The hydrological regime of the river in its main channel is bimodal with an average flow of 50 m^3/s , with wet seasons between March-May and September-November, and dry seasons between December-February and June-August.

The MPI-ESM-MR model of the Max Planck Institute was used as the Global Climate Model in its first historical realization (1850-2005) with the RCP8.5 experiment (2006-2100) in its first realization for a total of 91676 daily precipitation data, based on the report of the Magdalena River Basin (Angarita, 2014), which concludes that this GCM reflects adequately most of the local climate state patterns for the study region, in total, two cells of the model were used (as it is presented in Figure 2), with an average value of 5.13 mm, an average standard deviation of 6.12mm and an average maximum value of 45.82mm.

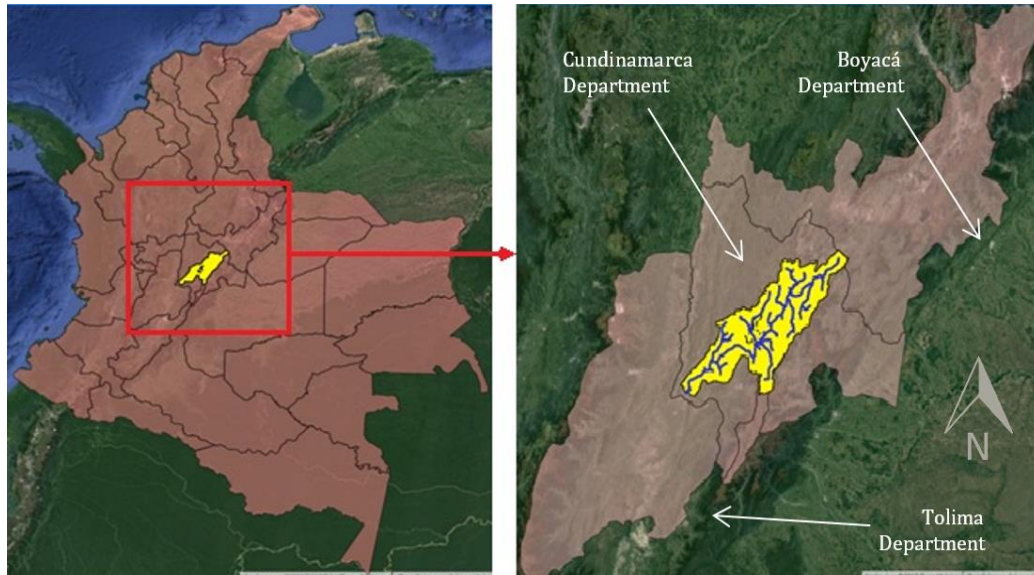


Figure 1: Location of the Bogotá River Basin in the country (left) and in the department (right)

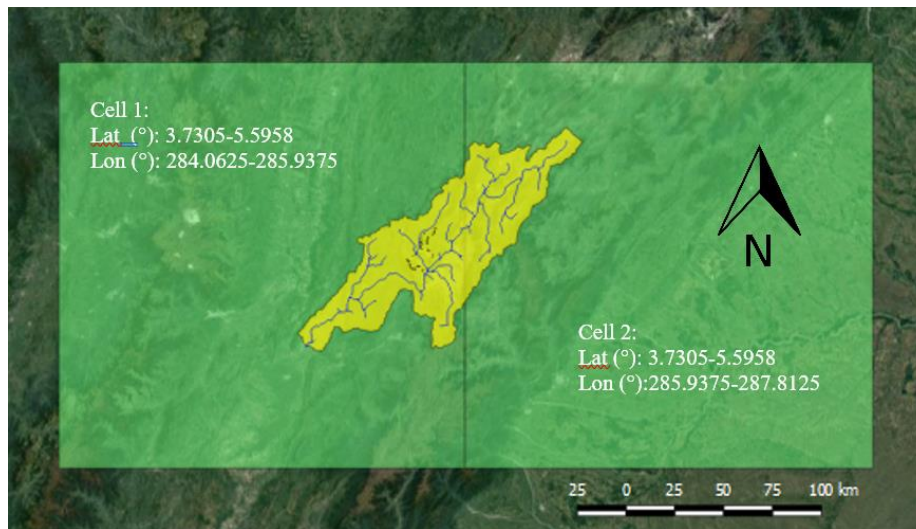


Figure 2: Cells of the Global Climate Model: MPI-ESM-MR, in the Bogotá river basin.
Coordinate reference system: WGS-84

For the local data, 47 ground rainfall gauges stations selected from IDEAM (Institute of Hydrology, Meteorology and Environmental Studies, Colombia) were used for the period 1958-2016, as shown in the Figure 3, observing that there is an average of 12239 daily precipitation data per station, for use as predictors, with an average value of 2.76 mm, an average standard deviation of 6.69 mm and an average maximum value of 105.54 mm.

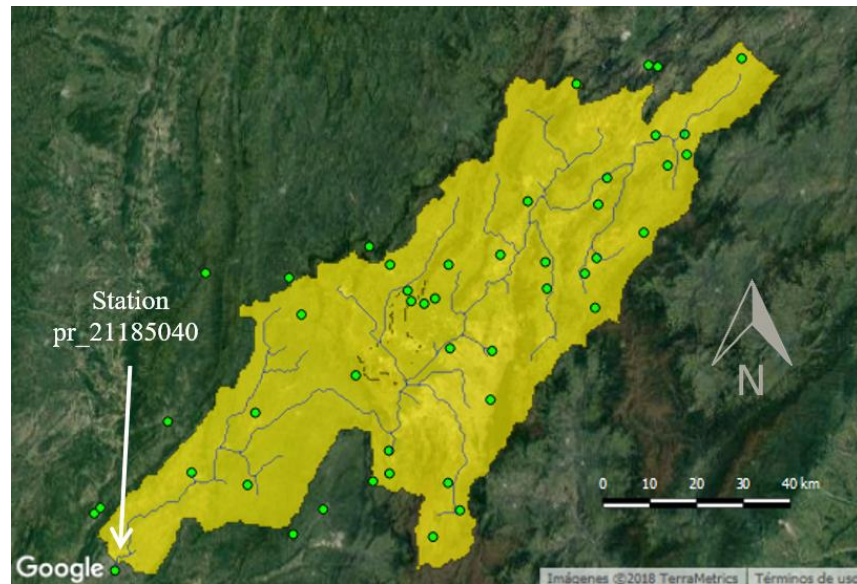


Figure 3: Ground rainfall gauge stations selected for the downscaling in the Bogotá River Basin

The Chaotic Statistical Downscaling model (CSD) evaluate the presence of deterministic chaos for different rainfall accumulation intervals for both of the time series (rainfall gauge stations and the global climate model). Then a chaotic predictive model is constructed with the results of the time delay, the embedding dimension and the Lyapunov exponents found for the "optimal" precipitation accumulation interval of each dynamic system. The predictive model is based on the synchronization between two dynamical systems, given by the parameter μ of the mutual false nearest neighbor's method (MFNN).

According to (Sivakumar & Ronny, 2010): "The phase space is essentially a graph, whose coordinates represent the variables necessary to fully describe the state of the system at any time, the trajectory of the phase space diagram describes the evolution of the system for an initial state. The "region of attraction" of these trajectories in the phase space provides important qualitative information when determining the degree of complexity of the system ", as shown in Figure 4 where the phase space is observed in three dimensions for different accumulation intervals of precipitation. The limiting set that brings together asymptotic trajectories close to equilibrium is known as 'attractor'. The attractors of deterministic chaotic systems can exhibit an unusual type of self-similarity and present structures at all their scales and it is therefore necessary to find an appropriate dimension of the phase plane, such that the structure of the attractor remains invariant.

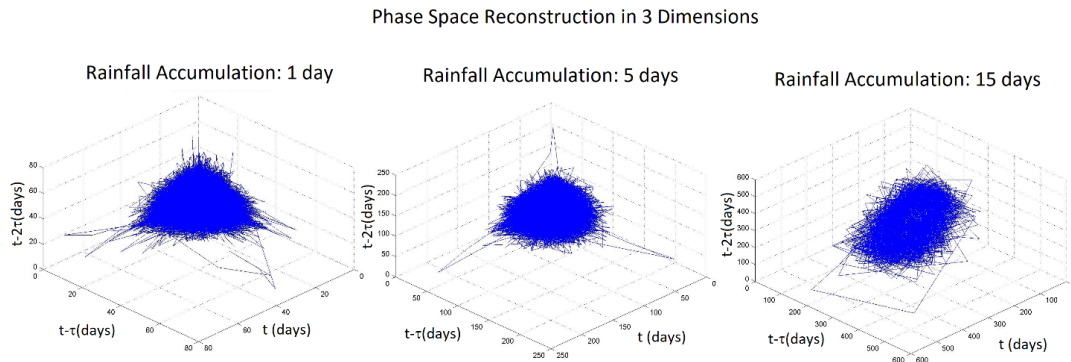


Figure 4: Comparison of the phase space of the precipitation time series for Cell 1 of the Bogotá river basin for rainfall accumulations of 1 day (left), 5 days (center) and 30 days (right).

For the evaluation of the presence of deterministic chaos in the dynamical systems Lyapunov exponents were used, for which it was necessary to reconstruct the phase space with the Method of Time-Delay, which finds the appropriate values of the time delay (τ) and embedding dimension (m) to capture the attractor dynamics. The autocorrelation function and the mutual information were used for the selection of the time delay, and the embedding dimension was selected using the correlation dimension method, the False Nearest Neighbors method (FNN) and Cao's method, which were successfully used in (Gallego, 2010), (Hernández, 2009) and (Siek, 2011).

Within the different phenomena of dynamical non-linear systems, the synchronization of dynamical systems has had a great importance in the last decade, being used in a wide variety of practical and experimental applications, such as electrical and electronic circuits, lasers, telecommunications systems and chemical reactions. The synchronization of two or more dynamical systems is a fundamental phenomenon that is obtained when at least one of the systems changes its trajectory due to a coupling (unidirectional or bidirectional) with another system, allowing a coherent behavior in coupled systems. In chaotic systems, given the dependence of the initial conditions and their evolution in different attractors and dimensions, the synchronization process allows the systems to follow a common and defined trajectory.

The mutual false nearest neighbor's method created by (Rulkov, Sushchik, Tsmiring, & Abarbanel, 1995), is a statistical technique based on the calculation of the parameter μ , which evaluates the local neighborhoods between two time series, so that μ takes values of the order of 1 if exist complete general synchronization, otherwise, μ has to be a number whose magnitude is comparable with the product of the attractor size divided by the product of the distance between the nearest neighbors in the time series x and y . The formula for the calculation of μ is given in Equation 1 and the relation between points of the attractors is shown in Figure 5,

$$\mu = \frac{1}{N} \sum_{n=1}^N \frac{\|y_n - y_{n_{NND}}\| \|x_n - x_{n_{NNR}}\|}{\|x_n - x_{n_{NND}}\| \|y_n - y_{n_{NNR}}\|}$$

Equation 1: Mutual false Nearest Neighbors

Where,

x_n and y_n , are points of the drive and response system at an instant n

$x_{n_{NND}}$ and $y_{n_{NNR}}$ are the closest neighbors of x_n and y_n in their respective systems.

$x_{n_{NNR}}$ and $y_{n_{NND}}$ are points of the opposite system at time n for ynnR and xnnD.

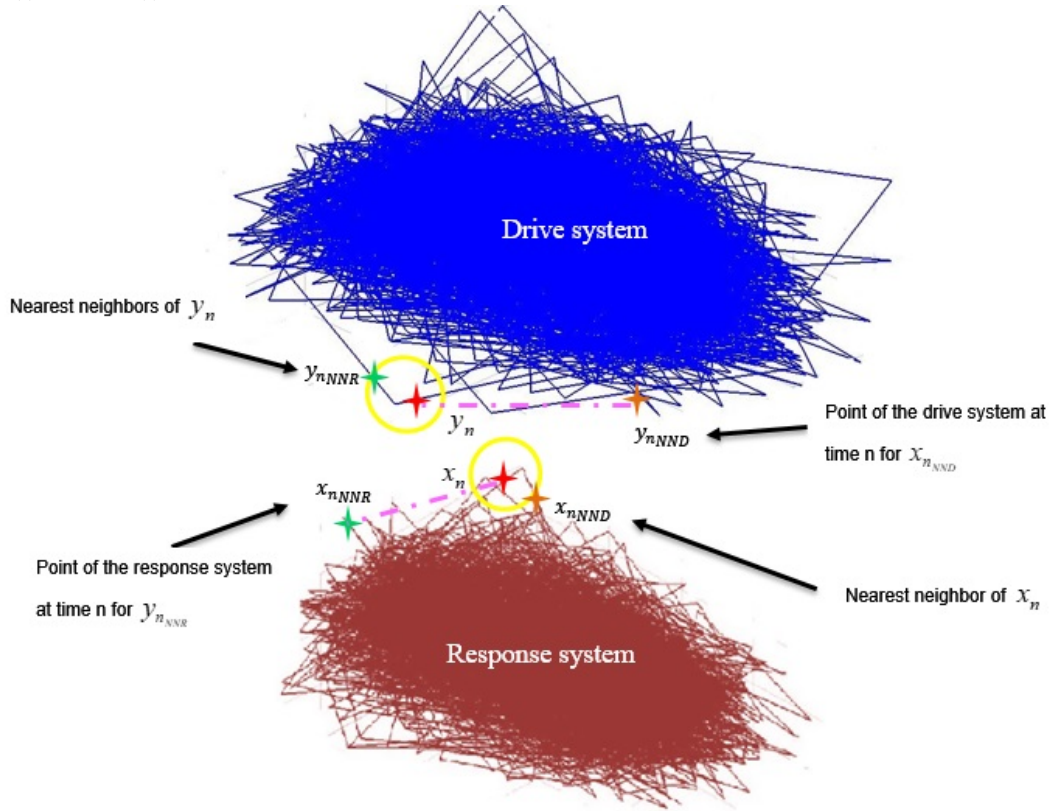


Figure 5: Relation between x_n , y_n , $x_{n_{NND}}$, $y_{n_{NNR}}$, $x_{n_{NNR}}$ and $y_{n_{NND}}$ for drive and response systems

Results and Discussion

With the purpose of obtain a wider range of analysis, the evaluation of the presence of deterministic chaos was performed in intervals of 1,3,5, 7, 10,15 and 30 days. Analyzing the results, it

is observed that in the cells of the GCM and in the rainfall gauge stations time series, starting from a rainfall accumulation of 5 days the type of movement of the dynamic system is no longer random (noise) and becomes mainly deterministic chaos, ensuring the short-term predictability for climate projections. It was also observed that in both time series the type of movement of the dynamic system changes from deterministic chaotic to stable for a rainfall accumulation interval of 15 to 30 days, ensuring the existence of long-term predictability for monthly climatic projections.

The phase space reconstruction for local gauge stations and GCM cells shows that the embedding dimensions are 6 and 8, respectively. This indicates that the GCM is more complex than the dynamic system shown by the local gauge stations. The complexity is also reflected when calculating the average attractor size of 3.5 for the local stations and 4.86 for the GCM. These results are of great importance, since they allow to identify the degree of difficulty to synchronize the two chaotic systems.

The values obtained for the parameter μ of the mutual false nearest neighbor's method are greater than the theoretical value of $\mu = 1$ ideal for general synchronization. However, it is possible to observe that using the CSD model it was possible to decrease the difference with the ideal theoretical value by finding the minimum possible distances for $x_n - x_{n_{NNR}}$ and $y_n - y_{n_{NND}}$, which represents a slight increase in the synchronization of the two systems, as presented in Figure 6 for the station pr_21185040

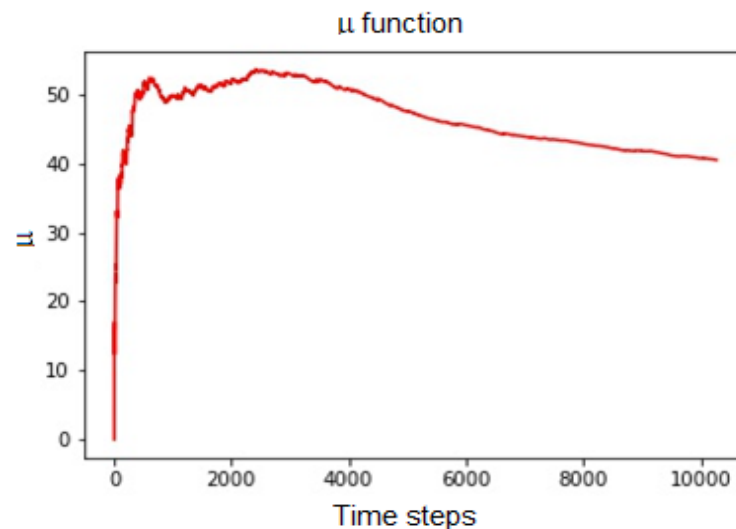


Figure 6: Results of the parameter μ of mutual false nearest neighbor's method for station pr_21185040

In the process of compare the downscaling model, a wide variety of statistical downscaling methods were used in the Bogotá river basin, for a total of 16 techniques, including: CSD, k-NN Bootstrapping, Delta methods, Analog methods, Quantile Mapping method, Weather Type methods and Generalized Linear Methods, most of the methods were include in the Meteolab toolbox (Cofiño, et al., 2013). These methods were evaluated under three different measures of error (Mean value difference, Max value difference and Root-Mean Square Error (RMSE)), in the Figure 7, the RMSE comparison is shown.

Finally, it was possible to compare the monthly multiannual precipitation, the annual precipitation, and the return period for a wide variety of downscaling techniques with the reference period, and use these results to estimate the future river flows using a hydrological model. In the Figure 8 is shown the spatial variability for the percentage difference between the historical reference and the projected time series using the CSD model.

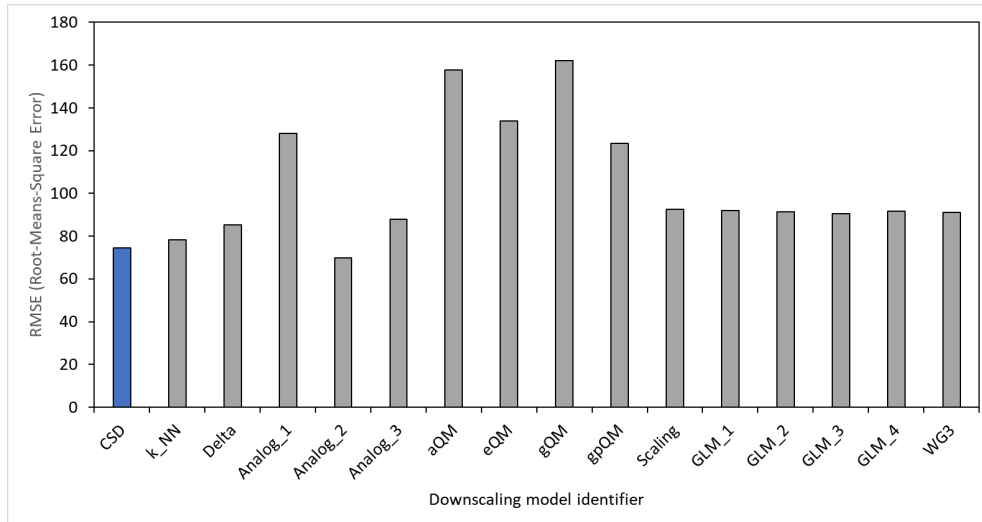


Figure 7: Comparison of the RMSE error obtained for different downscaling techniques

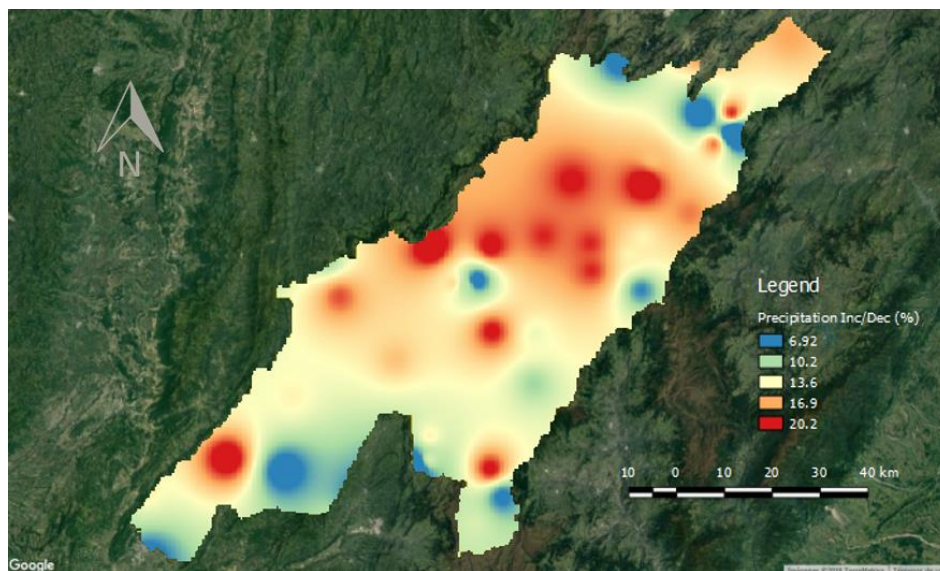


Figure 8: Spatial variability of the differences in average precipitation between the historical reference records 1970-2000 and the series projection using the MPI-ESM-MR RCP8.5 realization for 2070-2100 using the Chaotic Statistical Downscaling Model

Conclusions

It was found the presence of deterministic chaos in all the local stations time series for 5-day accumulation intervals (46 stations) and 7 days (1 station) for the Bogotá river basin. Likewise, deterministic chaos was found in the cells of the GCM for accumulation intervals greater than 5 days. This result is very important since it shows that it is possible to formulate models to improve the transformation of precipitation data between different scales.

In the validation process of the CSD model, it was found that it shows an optimum performance in the simulation of the mean values of the time series with the RMSE. However, it is necessary to improve the modeling of extreme events (droughts and floods) with a more efficient calibration of the model parameters, that consider better the bias of the mean and extreme values.

In the constructing process of the predictive model, it was not found complete general synchronization between the two dynamic systems, despite having found a slight degree of synchrony between them. However, given the results obtained by the scale reduction technique, the degree of synchrony found is sufficient, and it is recommended to continue exploring to find a better synchronization between the dynamic systems that allow better results.

References

- Angarita, H. (2014). *Metodología para incluir variabilidad climática y escenarios de cambio climático en el modelo WEAP de la macro Cuenca del Rio Magdalena y resultados de las simulaciones*. Bogotá.
- Cofiño, A., Ancell, R., San-Martín, D., Herrera, S., Gutiérrez, J., & Manzananas, R. (2013). *MeteoLab: An open-source Matlab toolbox for Meteorology & Climate*. Cantabria .
- Corzo, G., Jonoski, A., Yimer, G., Xuan, Y., & Solomatine, D. (2009). *Downscaling Global Climate Models Using Modular Models and Fuzzy Committees*.
- Gallego, J. (2010). *Aplicación de la teoría de caos para el análisis y pronóstico de series de tiempo financieras en Colombia*. Bogotá.
- Hernández, O. (2009). *Analysis and Optimization of Chaotic Models for Storm Surge Prediction*. Delft.
- Rulkov, N., Sushchik, M., Tsmiring, L., & Abarbanel, H. (1995). Generalized Synchronization of chaos in directionally coupled chaotic systems. *Physical Review E Volume 51 Number 2*, 980-994.
- Siek, M. (2011). *Predicting Storm Surges: Chaos, Computational Intelligence, Data Assimilation, Ensembles*. Delft.
- Sivakumar, B., & Ronny, B. (2010). *Advances in Data-Based Approaches for Hydrologic Modeling and Forecasting*. Davis, USA: World Scientific Publishing.