



Study of the Rainy Season Modification in the South of Madagascar: Case of Weather Drought

Fideline Béatrice HAOVA¹, Jean Eugène RANDRIANANTENAINA², Jacques Chrysologue RATSIMAVO³, Adolphe Andriamanga RATIARISON⁴

^{1,2,3}Laboratory of Atmosphere, Climate and Oceans Dynamics (DyACO), University of Toliara

⁴Laboratory of Atmosphere, Climate and Oceans Dynamics (DyACO), University of Antananarivo

ABSTRACT: In this work, we analyzed the rainfall data from 1979 to 2017, taken from the website of the European Center Medium Weather Forecast (ECMWF). The interest of this work is to observe the state of drought in the southern part of Madagascar, more precisely in Amboasary Sud. To achieve our goal, we used mathematical tools as a methodology. The time series study informed us that the annual precipitation ranges from 370 to 727 mm with the overall average 529.6189 mm. Then, the PETTITT test showed us that the break date of this series is the year 1999. In addition, the Anomalous Accumulation has informed us that on average, the start of precipitation is on the 296th Days (23 October) and the end is on the 133rd Days (May 13). The SPI allowed us to report that there was a moderate drought which was more prevalent during our study period. Finally, artificial neural network (ANR) modeling warns us that it will have a near-normal drought that will persist in our study area for the next 10 years (2018-2027).

Keywords: Precipitation, Standardized Precipitation Index (SPI), Artificial Neuron Network (RNA), South of Madagascar

I. Introduction

Human activities are one of the factors that cause climate change or climate disruption. This change is the biggest environmental problem of our time. Madagascar was classified in the analyzes carried out at the international level, among the countries vulnerable to climate change [1] Madagascar is a large island located in the southwest of the Indian Ocean, to the east of the Mozambique Channel which separates it from Africa. Madagascar's climate and its terrain are very varied and subject to major seasonal hazards, namely cyclones, torrential rains and prolonged drought [2].

The objective of this work is to analyze the variations in precipitation, Standardized precipitation index and to predict the SPI value by the Amboasary Sud neural network in the southern part of Madagascar.

II. Theoretical methods

II.1. Data presentation

We have studied the precipitation data from January 1, 1979 to December 31, 2017. In this period, there are 14,245 days or 468 months or 39 years. This data comes from the website of the European Center Medium Weather Forecast (ECMWF). It has a spatial-temporal resolution of 0.5 ° X0.5 ° and a three-dimensional matrix (latitudes, longitudes, days). We used MATLAB and EXCEL software to process this data.

II.2 Time or Chronological Series

II.2.1 Central tendency indices

It is conventionally the average and the median which are used. In the context of our study, it is the mean that will be systematically used [3], and it is defined by:

$$\bar{X} = \frac{1}{T} \sum_{t=1}^T X_t \quad (1)$$

II.2.2 Dispersion indices

Very often, it is the empirical variance that is highlighted, it indicates the dispersion of observations around their trend indices, by definition the variance is given by the following relation: $V(X) = \frac{1}{T} \sum_{t=1}^T (X_t - \bar{X}_t)^2$ (2)

It tells us how X_t takes its values, i.e., if the measured quantities are tightly clustered around the mean or if they are scattered, sometimes it turns out that the standard deviation is the square root of the variance, symbol σ , which is used, in this case we have

$$\text{so: } \sigma = \sqrt{\frac{1}{T} \sum_{t=1}^T (X_t - \bar{X}_t)^2} = \sqrt{V(X)} \quad (3)$$

If σ is low the values of $\{X_t\}$ are grouped around the mean; if it is important, on the other hand, they are widely dispersed.

II.2.3 Properties of the autocorrelation function

The sequence $\rho(h)$, of the autocorrelations, which is the quotient of the auto-covariances by the variance is called the autocorrelation function [4].

The functions $\rho(h)$ and $\gamma(h)$ have the following elementary properties:

These sequences are symmetrical: for $h > 0$; $\rho(h) = \rho(-h)$ same thing for $\gamma(h)$.

The autocorrelation function $\rho(h)$ takes its values in $[-1; 1]$ with, obviously $\rho(0) = 1$.

The sign and the value of $\rho(h)$ have direct consequences on the dependence of the series.

Indeed:

- If $\rho(h)$ tends towards zeros (0), then X_t and X_{t+h} are weakly dependent on each other
- If $\rho(h)$ tends to (1), then X_t and X_{t+h} are strongly dependent on each other.

Weak or strong dependence, the positive sign expresses a linear link causing a variation of the two variables in the same direction, as long as the negative sign generates a variation of the variables in opposite directions from each other.

II.2.4 Moving average

The moving average, or sliding average, is a type of statistical average used to analyze ordered series of data, most often time series, removing transient fluctuation in order to highlight longer-term trends [3]. A formula for calculating a simple moving average is:

$$\bar{x}_n = \frac{1}{N} \sum_{k=0}^{N-1} x_{n-k} \text{ où } \bar{x}_n = \bar{x}_{n-1} + \frac{x_n - x_{n-N}}{N} \quad (4)$$

Where $N < n$ and N the number of values in the consecutive subgroup.

II.2 Liebmann method "Anomalous Accumulation"

Goal: We applied the "Anomalous Accumulation" method to determine the start and end of rain each season.

This method is given by the following relation [5]:

$$AA(t) = \sum_{n=1}^t x(n) - \bar{x} * t \quad (5)$$

With AA (t): Anomalous Accumulation at day $t, x(n)$: Daily precipitation and \bar{x} : The average daily precipitation (Liebmann as the relationship between the annual cumulative precipitation and the number of days in the year).

II.3 Standardized Precipitation Index (SPI)

Purpose: The SPI was made to determine the precipitation deficit for a given region during a given period of time.

The calculation of the SPI requires fitting long series of rainfall data to the Gamma distribution which, according to Thom Young, is the distribution that best represents the evolution of the rainfall series [6] [7]. This Gamma distribution is defined by its probability density represented by:

$$g(x) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}} \quad (6)$$

Where: α and β are the shape and scale parameters of the distribution. They are obtained from the least squares method described by Edwards and McKee [8].

The SPI is estimated by fitting the frequency distribution of precipitation at each station to the Gamma density probability distributions. The values of the parameters α and β of the Gamma function are estimated for each station and for each time scale (1; 3; 6; 12; 48 months) and for each month of the year. The classification of drought according to SPI is shown in the following table.

Table 1: Classification of the Standardized Precipitation Index (SPI) [9]

| SPI value | Intensity of Drought |
|----------------|----------------------|
| 0 to -0.99 | Close to normal |
| -1,00 to -1,49 | Moderate |
| -1,50 to -1,99 | Severe |
| ≤ -2 | Extremely dry |

II.4 Hurst exponent test

To test the robustness of our results, we also consider one of the most famous measures of persistence in a time series, the Hurst exponent [10]. This is the R/S statistic, introduced by Hurst in 1951. It is defined as the extent of the partial sums of deviations of a time series from its mean divided by its standard deviation. This R/S procedure is particularly interesting insofar as it gives rise to a coefficient called the Hurst exponent, noted H (or dependence index). This exponent makes it possible to classify the time series according to the nature of their memory, in other words their dependency structure. It is a measure of absolute persistence [11].

Consider a time series of returns $X_t, t = 1, \dots, T$, with mean \bar{X}_t , the R/S statistic, denoted Q_t , is written:

$$Q_T = \frac{R}{s_T} \quad (7)$$

Or :

This statistic is asymptotically proportional to T^H , where the constant H , between 0 and 1, is called the Hurst exponent. The Hurst exponent is thus given by:

$$H = \frac{\log Q_T}{\log T} \quad (9)$$

This exponent then makes it possible to determine the dependency structure of the series as a function of the values of H.

- If $H = 1/2$, the process therefore has no long-term dependence.
- If $1/2 < H < 1$, the process exhibits a persistent form of long memory.
- If $0 < H < 1/2$, the process is anti-persistent; phases of rising tend to be followed by phases of decline.

II.5 Modeling by neuron network

We made the drought forecast by this modeling in the southern part of Madagascar. The neural network shows a great capacity to solve prediction problems compared to classical statistical methods thanks to its power to capture the nonlinear relationships existing in the data series [12].

Mathematical representation

As a general rule, the calculation of the value of this function can be broken down into two steps [13]:

- A linear combination of inputs: $a_k = \sum_{j=1}^n w_{kj}x_{kj} + b_k$ (10)

- The output of the neuron is: $S_k = f(a_k) = f(\sum_{j=1}^n w_{kj}x_{kj} + b_k)$ (11)

Where f : The sigmoid activation function, there are others but the sigmoid is the most used.

w_{kj} : The synaptic weight of the connection between neuron j and k

x_{kj} : The entry of neuron k from neuron j.

b_k : The bias

n : The number of neurons in the layer preceding the layer containing neuron k.

The activation function (or transfer function) is used to convert the result of the weighted sum of the inputs of a neuron into an output value, this conversion is carried out by a calculation of the state of the neuron by introducing a non-linearity in the functioning of the neuron [14].

II.6 Mean Absolute Percentage Error (MAPE)

It is necessary to check whether the forecast is good or bad, whether the errors are large or not. MAPE is often used to compare forecasts along different time series [15]. This is a model that suggests a better forecast.

It is defined by the formula [16]: $MAPE_n = \frac{\sum_{t=1}^n \frac{|\epsilon_t|}{X_t} * 100}{n}$ (12)

With ϵ_t the error of the forecast defined by the following formula: $\epsilon_t = X'_t - X_t$

And the quality of the forecast is defined according to the value of the following MAPE (Table 1):
 2: Classification of research quality [17]

| MAPE value | Forecast quality |
|----------------------|------------------|
| $MAPE < 10\%$ | Excellent |
| $10\% < MAPE < 20\%$ | Good |
| $20\% < MAPE < 30\%$ | Average |
| $30\% < MAPE$ | Bad |

III. Results and discussion

III.1 Presentation of study location

Our place of study is the municipality of Amboasary Sud, Anosy region, in the extreme south part of Madagascar; at the intersection of the geographic coordinates of latitude 24.71 South and longitude 43.47 East. Thus figure 1 shows this place which is in red color.



III. 2 Study of the annual precipitation of Amboasary Sud

We have shown in the following figure 2 the evolution of annual precipitation of Amboasary Sud. The bar in green shows the variations in annual rainfall, the curve in red color determines the moving average and the line in blue color is the global average. We note that in our place of study (Amboasary Sud), from 1979 to 2017, the years with the most rainfall were 1984 with 727mm of precipitation and 1994 with 681mm. In terms of precipitation, the years with the highest deficits are 1980, 2006 and 2016. The last few years have been remarkable for meteorological drought. The moving average decreases below the global average with the value 529.6189mm from the year 2001 until the end of the year 2017. In this period, the rainfall rates have decreased.

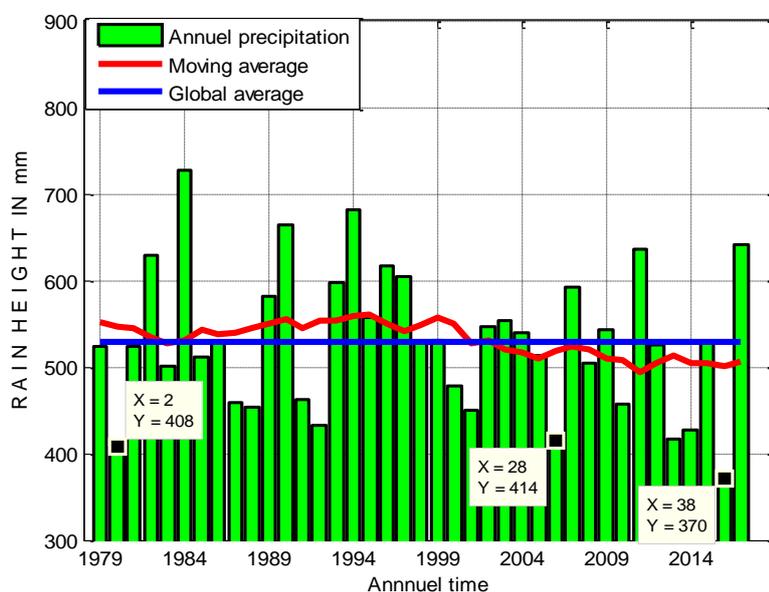


Figure 2: Annual precipitation of Amboasary Sud from 1979 to 2017

III.3 Detection of the rupture date by statistics of Pettitt of Amboasary Sud

Pettitt's test allows us to detect the date of rupture in a given time series. We see in figure 3 below that this breaking date is in 1999. This year, we can see the change in the average value; it is 549.359mm before and 508.839mm after.

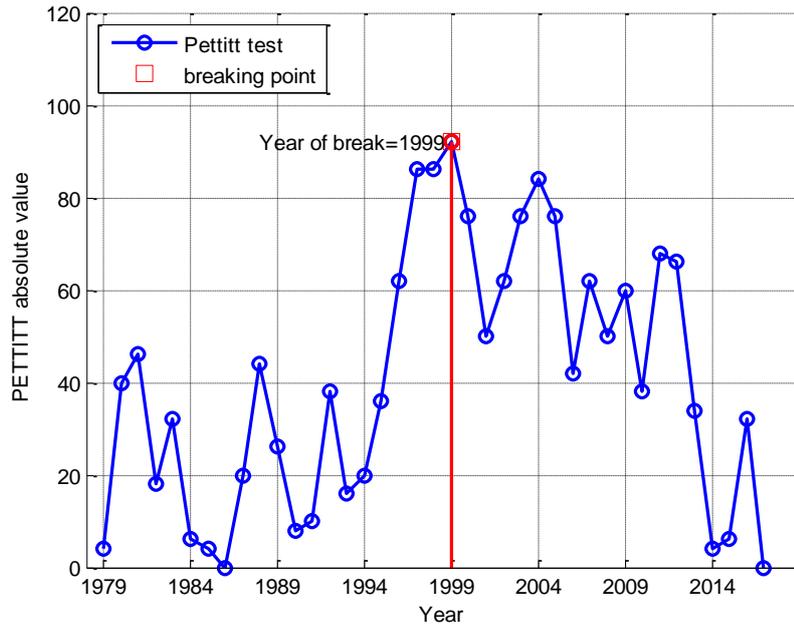


Figure 3: Pettitt test for Amboasary Sud

III.4 Evaluation of the end and start of the rainy season in Amboasary Sud

We will present in Figure 4 below the end and start of rains each season from 1979 to 2017. And talk about two different curves: the pink color curve is the beginning of rain during the season in the subsequent series. We find that the onset of rain varies between October and February but frequently in December. In addition, the overall average substantially or equal to the 296th day (23 October). The red curve is the end of rain for 39 years. We can see that the end of the Amboasary Sud rainy season varies between February and July. Thus, the overall average substantially or equal to the 133rd day (May 13). Finally, the rainy season from 1994 to 1995 (260 days) was very long among the seasons marked on the series and the shortest from 2012 to 2013 (53 days).

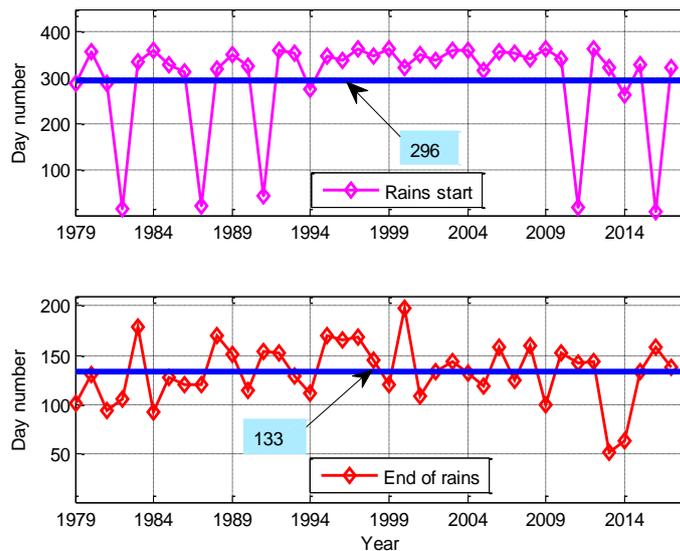


Figure 4 : Beginning and end of the rains from 1979 to 2017 in Amboasary Sud

III.5 SPI-12 value analysis of Amboasary Sud

Figure 5 presents the variation of the SPI-12 values of Amboasary Sud from 1979 to 2017. We illustrate that the moving average curve (in red) circulates below the global average (in green) for the dry periods and it is at the above the overall average for wet periods. The periods of drought in our study location are as follows: 1979-1989, 2005-2006 and 2008-2016. In these periods, drought prevails strongly in relation to humidity. We notice that for the last fifteen (15) years, drought almost appeared in this place.

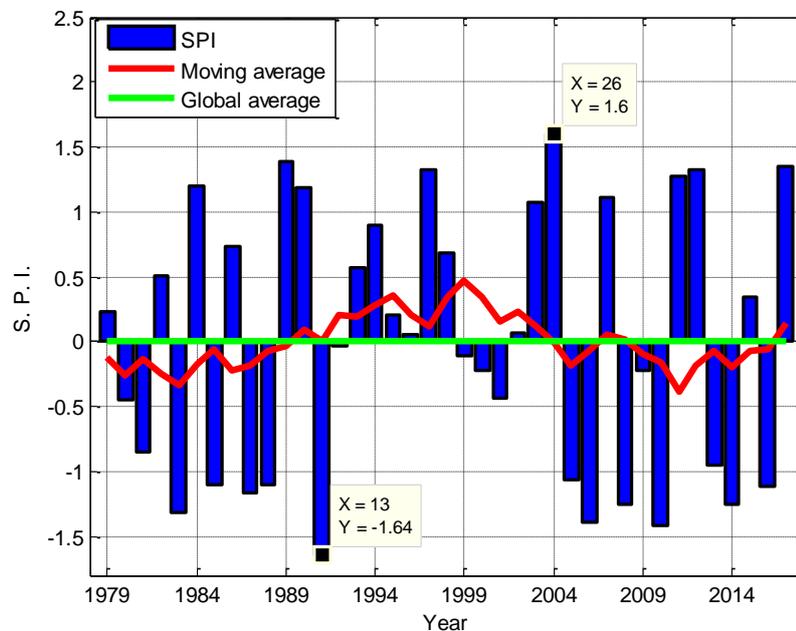


Figure 5 : Amboasary Sud SPI-12 from 1979 to 2017

We present in table 3 below, the frequency in percentage of humidity and drought of this figure 5 After reading this table, we find that the humidity close to normal (at 25.64%) and the moderate drought (at 28.20%) are very dominant during the series.

Table 3: Probability of onset of drought over time

| Classes | Number of years | Percentage in% |
|----------------------|-----------------|----------------|
| Extreme humidity | 0 | 0% |
| High humidity | 1 | 2,56% |
| Moderate humidity | 9 | 23,07% |
| Near normal humidity | 10 | 25,64% |
| Near normal drought | 7 | 17,95% |
| Moderate drought | 11 | 28,20% |
| Severe drought | 1 | 2,56% |
| Extreme drought | 0 | 0% |

III.6 Modeling of SPI-12 of Amboasary Sud by Artificial Neuron Network (RNA)

First, we studied the Hurst exponent to see the memory effect, it is $H = 0.7245$, this means that the process exhibits a persistent form of long memory. So, our forecast is long term. The annual SPI values of Amboasary Sud from 1979 to 2028 are given in figure 6 below. They were obtained by the artificial neuron network. This figure presents three different curves in terms of color: the solid color in red shows the observed SPI, the solid blue is the forecast of the value observed as the modeling and at the end the dotted curve in red the value of our forecast. We found $MAPE = 7.82\%$, which tells us that our forecast is excellent.

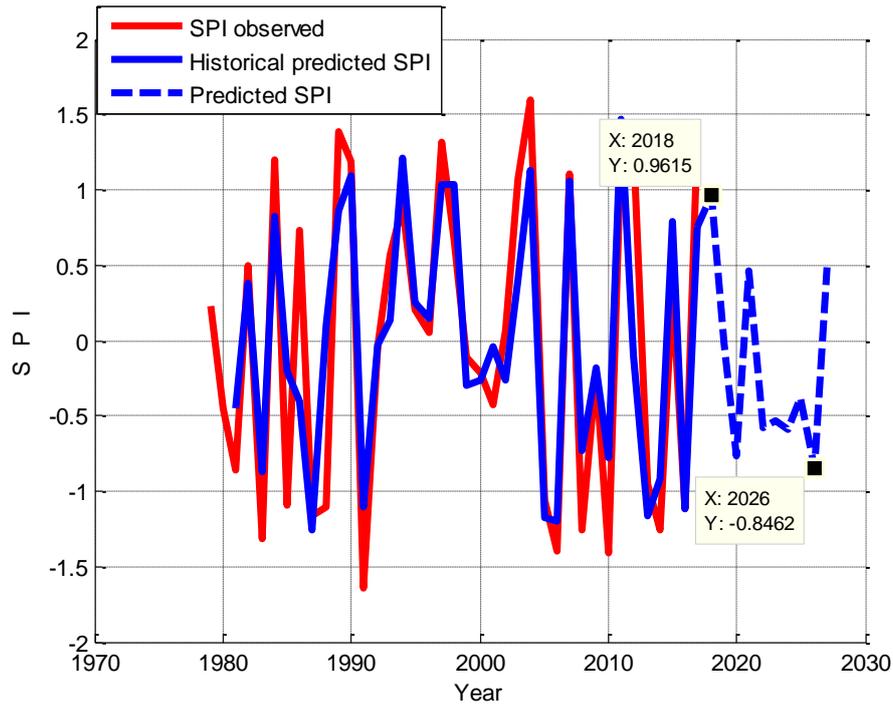


Figure 6 : Forecast of SPI-12 by RNA from Amboasary Sud

We show in table 4 below the forecast values of our data SPI_12 from Amboasary Sud. We find that the predicted value ranges from -0.8462 to 0.9615 then the near normal category as the theory has said but there is near normal humidity and drought. So, this tells us that during the next ten (10) years, drought close to normal will dominate in our place of study (Amboasary-Sud).

Table 4: predicted value with its category from 1979 to 20217

| Year | Value | Category |
|------|---------|----------------------|
| 2018 | 0,9615 | Near normal humidity |
| 2019 | -0,0184 | Near normal drought |
| 2020 | -0,7707 | Near normal drought |
| 2021 | 0,4633 | Near normal humidity |
| 2022 | -0,5736 | Near normal drought |
| 2023 | -0,5286 | Near normal drought |
| 2024 | -0,5848 | Near normal drought |
| 2025 | -0,3778 | Near normal drought |
| 2026 | -0,8462 | Near normal drought |
| 2027 | 0,4879 | Near normal drought |

III.7 Checking the diagnosis

The ACF of the residuals of the RNA model is shown in the following figure 7. All of the peaks are within the confidence interval. This indicates that the residuals are not self-correlated.

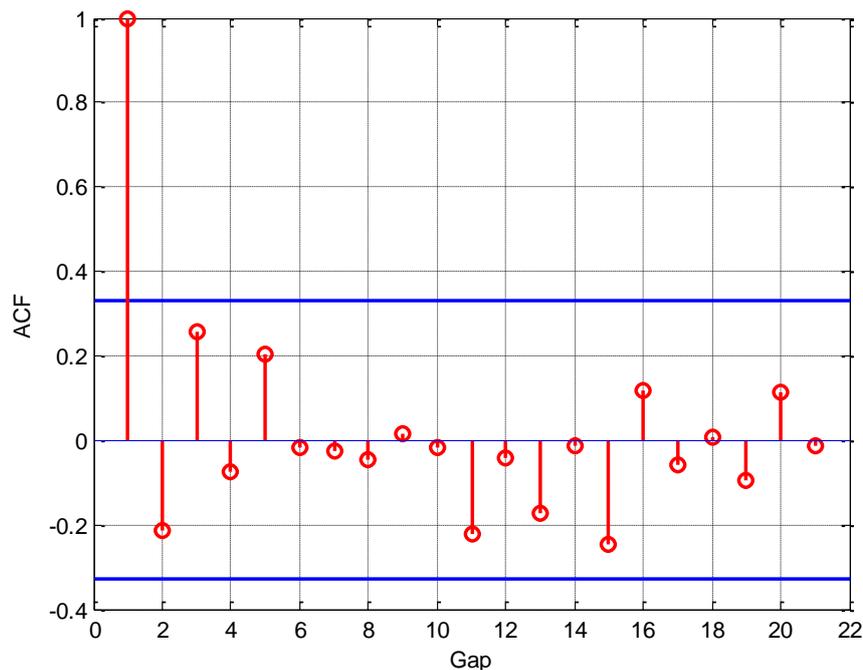


Figure 7 : Autocorrelation function of residuals for SPI-12 from Amboasary Sud

IV. Conclusion

To conclude, the analysis of the time series gave us the possible appearance of the observed values. The Anomalous Accumulation showed us the beginning and end of rain. We used the Standardized Precipitation Index (SPI) to examine the hazard of drought. We found during our work that there were several types of drought for Amboasary Sud. The RNA model was produced for drought forecasting. And we found a good long-term drought forecast by applying Hurst's exponent methods. Thus, for validation, the autocorrelation function (ACF) of the residuals tells us that they are not autocorrelated and the MAPE confirms that our forecast is excellent.

V. References

- [1] ACCLIMATE, Mars 2011. « Etude de la vulnérabilité au changement climatique » Evaluation qualitative – Madagascar.
- [2] DELILLE Hélène, 2011, « Perception et stratégies d'adaptation paysannes face aux changements climatique à Madagascar », pp16.
- [3] GILBERT SAPOTA, 2006. « Probabilités, analyse des données et statistique », Technip.
- [4] ALAIN BACCINI, 2010. « Statistique descriptive multidimensionnelle », Institut de mathématique de Toulouse, Université de Paul Sabatier, pp 15-22.
- [5] ARVOIR D., DUBREUIL V., MEIRELLES M. S. P., 2008. « Apport des données TRMM 3B42 à l'étude des précipitations au Mato Grosso (BRESIL) », Climatologie, vol 5, pp 49.
- [6] THOM H., 1966. Geneva : World Meteorological Organisation
- [7] YOUNG K., 1992. « A three-way model for interolating for monthly precipitation values ». Mon Weather Rev (120) : 2561-9.
- [8] EDWARDS D. & Mc KEE T., 1997. « Characteristics of 20th century drought in the United States at multiple time scales». Climatology report No. 97-2. Paper No. 634. Fort Collins (Colorado) : Colorado State University, Department of Atmospheric Sciences.
- [9] McKEE T., DOESKEN N. & KLEIST J., 1995. « Drought monitoring with multiple time scales ». American Meteorological Society, 9th Conference on applied climatology.

- [10] MIGNON V., 1998. « Methods for Estimating the Hurst Exponent: Application to Stock Market Returns ». *Économie et Prévision*, vol. 132, issue 1, pp 193-214.
- [11] NEWBY W., WEST K., 1987. « A Simple Positive-Definite: Heteroskedasticity and Autocorrelation Consistent Covariance Matrix ». *Econometrica*, Vol. 55, pp 703–708.
- [12] BEIDA MOHAMMED FERHAT TALEB AMAR, 2004. « Systèmes d'Information (SI) », pp6.
- [13] Marc Parizeau, 2004. « Le perceptron multicouche et son algorithme de rétropropagation des erreurs ». Département de génie électrique et de génie informatique Université Laval. 10 septembre.
- [14] CYBENKO, G, 1989. "Approximation by superpositions of a sigmoidal function: *Math. Control, Signals and Systems* », 2, pp303-314.
- [15] THEFER R., 2008. « A proposed technique for the problem of selected the best forecasting Model in time serie ». *A case study. Iraqi Journal of statistical science*, vol.14.
- [16] TANG L. & QIAN Y., 2013. « Exposé oral - model ARMA ».
- [17] RAMUDHIN A., 2009. « Prévision de la Demande ».