American Journal of Sciences and Engineering Research | wwww.iarjournals.com

American Journal of Sciences and Engineering Research

E-ISSN -2348 – 703X, Volume 5, Issue 5, 2022

Modelling of Rainfall Forecasting by Arima Method in the Boeny Region of Madagascar

R. RANDRIANANTENAINA[1]*, R. TELESPHORE[3], F. RAJESITERA[1], R. MANASINA[3] and A. RATIARISON[2]

[1] *Faculty of Technology and Environmental Sciences, Mahajanga University, Madagascar.*
[2] *Laboratory of the Dynamics of the Atmosphere, Climate and Oceans (DyACO), Faculty of Sciences, University of Antananarivo, Madagascar.*
[3] *Higher Institute of Science and Technology of Mahajanga (ISSTM), University of Mahajanga, Building Ex-LOLO Mahajanga-be, Madagascar.*

Abstract: *The study area is the BOENY region, delimited in latitude between 15° South and 18° South and in longitude between 44° East and 48° East.*
We modeled the rainfall from 1979 to 2018 using the ARIMA method. The study leads us that the data follow Gauss's law. According to this method, on average the values of annual rainfall are between 24.77 mm and 43.8 mm. The year 2006 is the year with the least rainfall, the average annual rainfall value is 24.77mm. The wettest year is 2004 with an average value of 43.8 mm.

Keywords: *Rain, forecast, ARIMA, Boeny region.*

I. Introduction

The threat of climate change is the focus of international concern. In order to understand this international problem, knowledge of the meaning of the term « climate change » is necessary. The meaning of the term « climate change » is fairly straightforward and no longer controversial. However, its causes, magnitude and impacts on human well-being and the environment are much debated.

In Madagascar, the impact of climate change, in particular, rainfall, remains a major concern for the Big Island. It is likely to hit the entire country hard in the coming years.

And it is for this reason that this study leads us to model the average annual value of rainfall by the ARIMA method. In this perspective, the proposal of a medium-term forecasting model in a study area is an essential step in conducting this study.

II. Material and methods

2.1 Presentation of the study area

The study area (see Figure 1) is between latitude 15°South and 18°South and longitude 44° East and 48° East.

Received-03-10-2022, Accepted- 12-10-2022

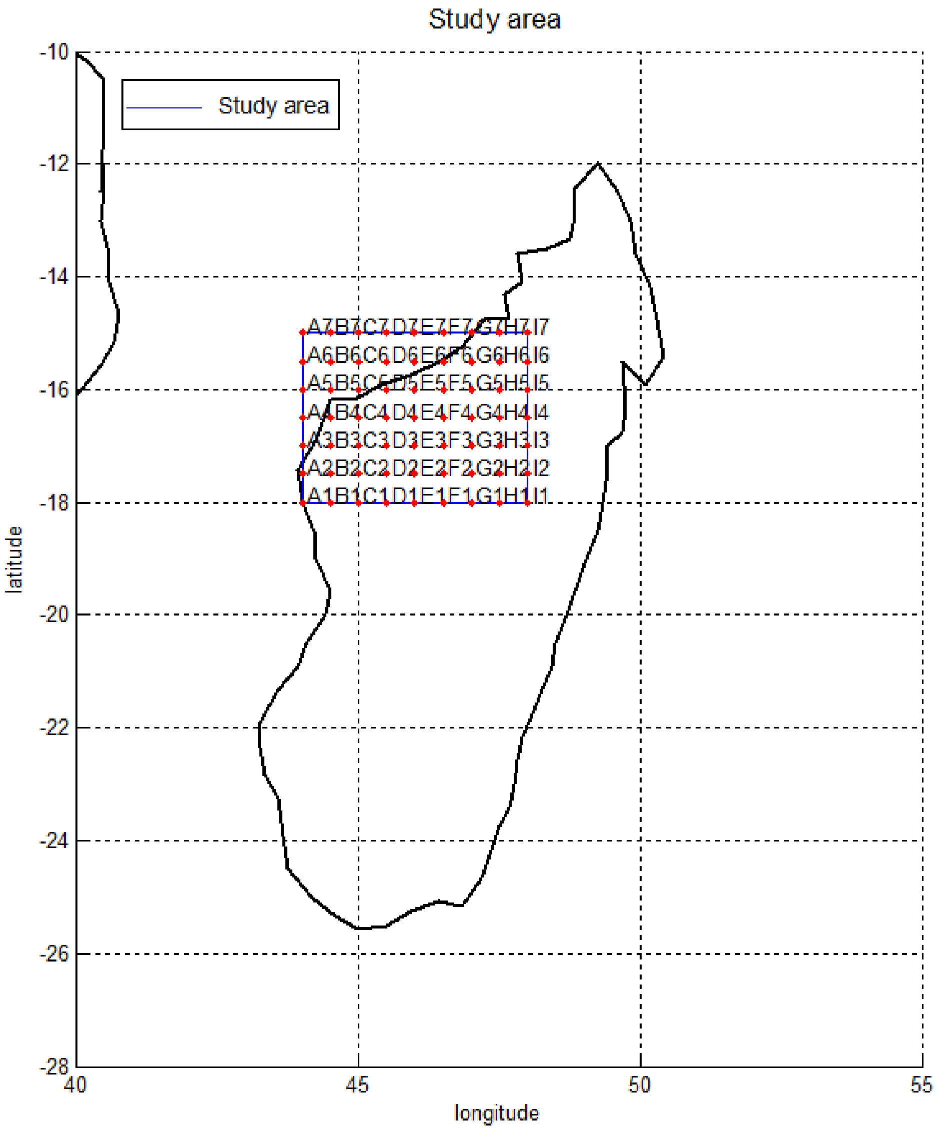

Figure 1: Study area 44° ≤ longitude ≤ 48° and -18° ≤ latitude ≤ -15°

2.2 Databases

The meteorological data we used are from the daily reanalysis data of the European Centre for Medium range Weather Forecasts (ECMWF) experiment (ERA5) at synoptic scale with a 0.5° x 0.5° grid of rainfall over a time depth covering the period 1979-2018.

2.3 ARIMA modelling [1] [2] [3] [4] [5] [6] [7] [8]

Definition: We will say that $X = (X_t)$ is an $ARIMA\ (p,d,q)$ process, if there exists a natural number d such that the process $Y_t = (1 - B)^{d} Xt$ is an $ARMA(p,q)$. In other words, there exist two polynomials $A(z)$ and $C(z)$ of degree p and q respectively and a white noise (ε_t) defined on the same (Ω, Q, P) such that:

$(1-B)^d\ A(B)X_t = C(B)\varepsilon_t$

With the code:

- $e\ <-rnorm(200),$
- $Marchealeatoire\ <-cumsum(e),$
- $ts.plot(Marchealeatoire, type = "l", col = "blue"),$
- $abline(0,0, col = "red"),$
- $acf(Marchealeatoire, 20, "correlation").$

We obtain the trajectory of $X_t = X_{t-1} + \varepsilon_t$ and its autocorrelation function.

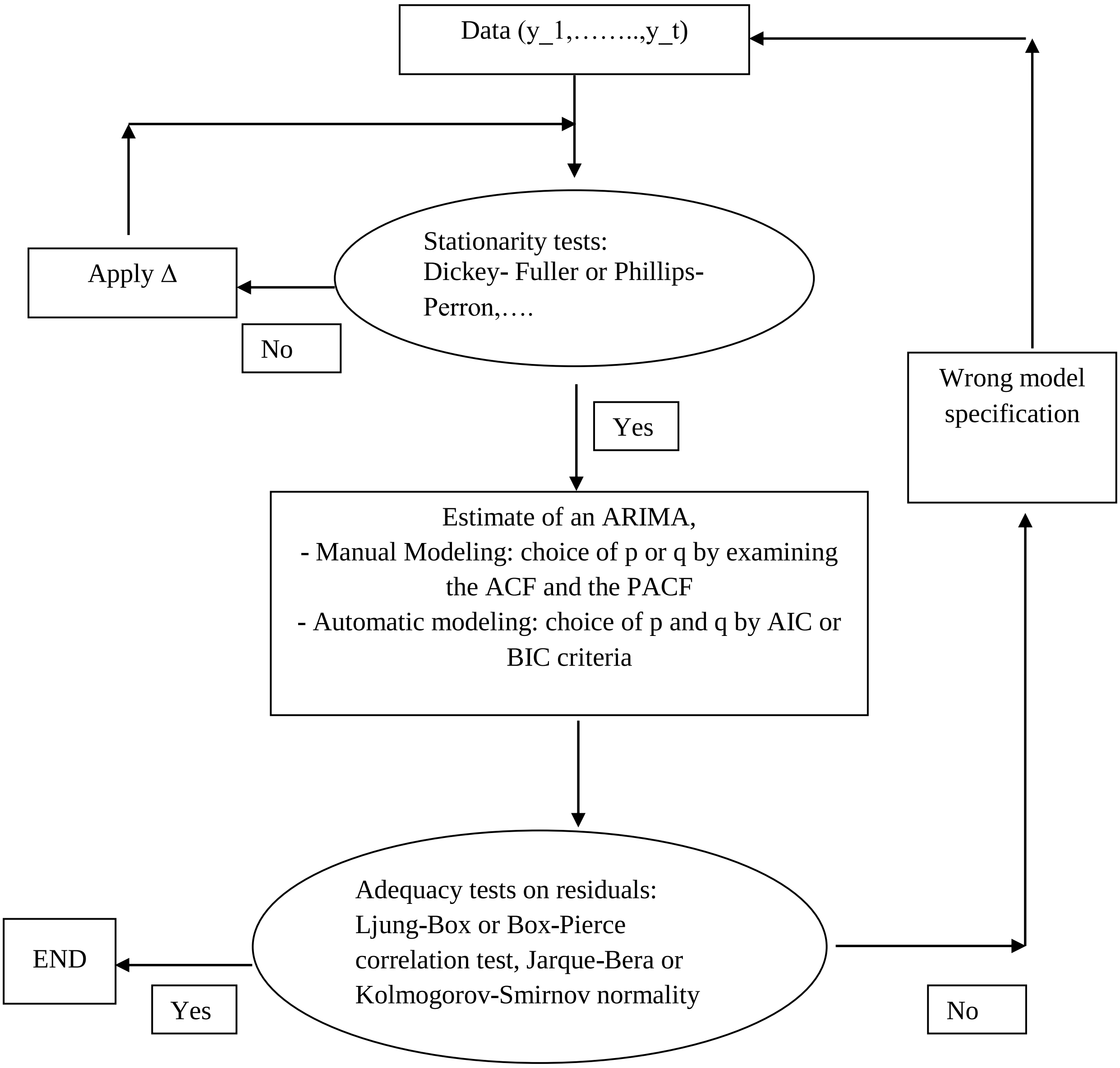

Figure 2: General diagram of the modelling of a time series by an ARIMA model

2.3.1 Modelling steps

Stationarisation by differentiation

The differentiation operator Δ = 1 − B, (or Δ^d) removes trends, d is estimated by performing stationarity tests on the raw series and then on the residual series. An estimator of d is the total number of times stationarity is rejected.

The residual series assumed to be stationary will be modelled by an ARMA (p, q).

Estimation

- R offers the possibility of using two methods:
- the maximum likelihood method "ML",
- and the conditional least squares method "CSS".

-

Validation

- Is the selected model compatible?

We carry out suitability tests on the residuals: absence of autocorrelation, normality, homoscedasticity,...

2.3.2 Non-stationarity test

Augmented Dickey and Fuller test:

Robust to autocorrelation compared to the Dickey-Fuller test

$$\begin{cases} H_0: y_t \sim I(1) \\ H_1: y_t\ n'estpas\ I(1) \end{cases}$$

We run the regression

$\Delta y_t = \beta' D_t + \pi y_{t-1} + \sum_{j=1}^{p} \psi_j \Delta y_{t_j} + u_t$

$ADF = T\hat{\pi} \frac{T\hat{\pi}}{(1-\hat{\psi_1}-\cdots-\hat{\psi_p}}$

Phillips-Perron test

Robust to heteroscedasticity

We run the regression:

$\Delta y_t = \beta' D_t + \pi y_{t-1} + u_t$

$PP = T\hat{\pi} - \frac{1}{2} T^2 \frac{SE(\hat{\pi})}{\hat{\sigma}^2} (\hat{\lambda}^2 - \hat{\sigma}^2)$

Where $\hat{\sigma} = \frac{1}{T-k} \sum_{i=}^{T} \hat{u}_t^2 \hat{u}_t$: residuals of the regression;

$SE(\hat{\pi}) = écart - type\ de\ \hat{\pi}$;

$\hat{\lambda}^2 = \hat{\sigma}^2 + 2 \sum_{j=1}^{q} \left(1 - \frac{j}{q+1}\right) \hat{\gamma}(j), (Newey - West),$

$\hat{\gamma}(j) = \frac{1}{T} \sum_{i=j+1}^{T} \hat{u}_t \hat{u}_{t-j}.$

2.3.3 Estimation :

With the code: $fit7\ < - arima0(ari, order\ =\ c(1,1,0), method = "ML"), fit7$ on we get

$Call:\ arima0(x\ =\ ari, order\ =\ c(1,1,0), method\ =\ "ML")$

Coefficients :

$ar1$

0.6714

$s.e. 0.0368$

$sigma\hat{}2\ estimated\ as\ 0.9678:\ log\ likelihood\ =\ -561.32,\ aic\ =\ 1126.64$

Note: The estimation of the model appears to be of good quality, as the estimator of a_1 is very close to the true value.

2.3.4 Suitability tests:

With the code: $Box.test(fit7\$residuals, lag\ =\ 10, type\ =\ "Box - Pierce")$ we obtain :

Box-Pierce test

data : fit7$residuals

X-squared = 8.728, df = 10, p - value = 0.5581

The p-value is large so the residuals are not correlated.

With the code: $library(nortest), lillie.test(fit7\$residuals)$ we get

$Lilliefors\ (Kolmogorov - Smirnov)\ normality\ test$

$data:\ fit7\$residuals\quad D\ =\ 0.0407, p - value\ =\ 0.1114$

So the data follow a Gaussian distribution.

III. Results and discussion

3.2.5 Graphical representation of the model

Figures 3 and 4 show the forecasts of the average annual rainfall observed during the study period. The observation data curves in purple and the curves in red are the forecast models. The forecast values of the annual average rainfall for the years 2019 to 2028 are shown in Table 1.

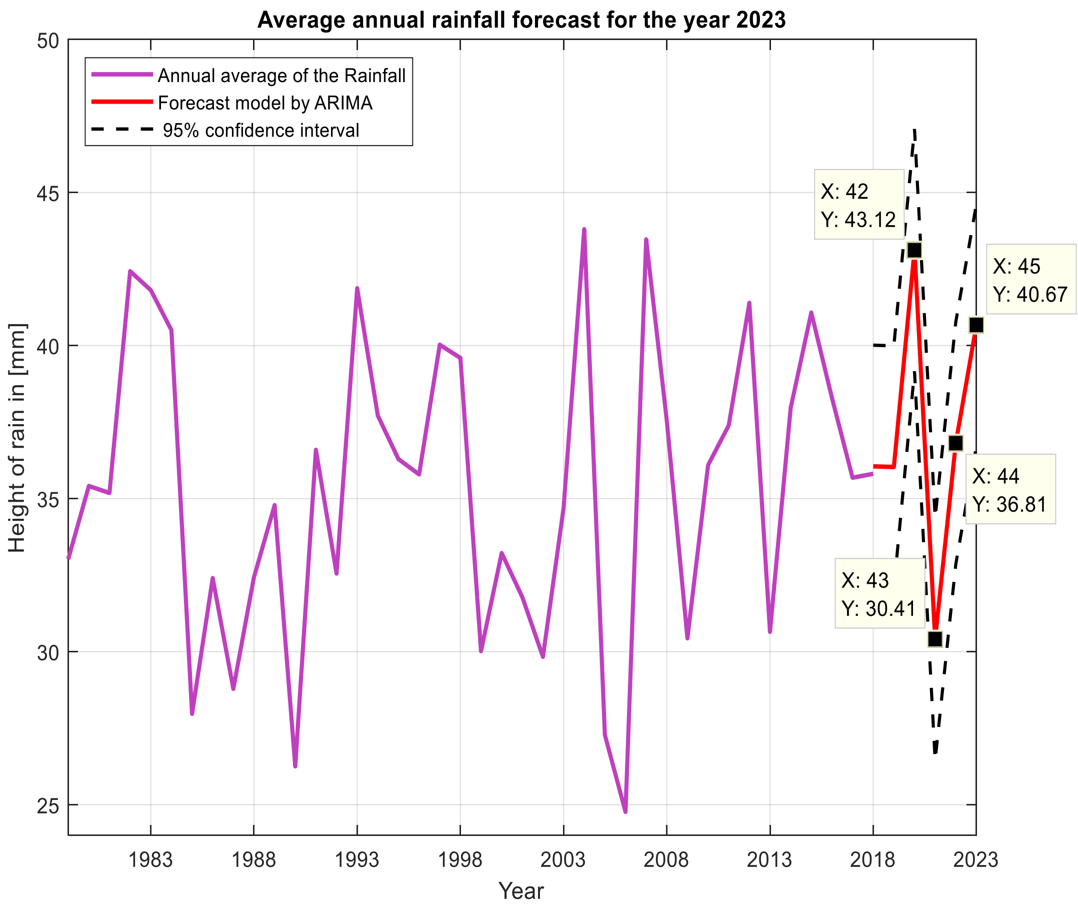

Figure 3: Forecast curve for rainfall in 2023

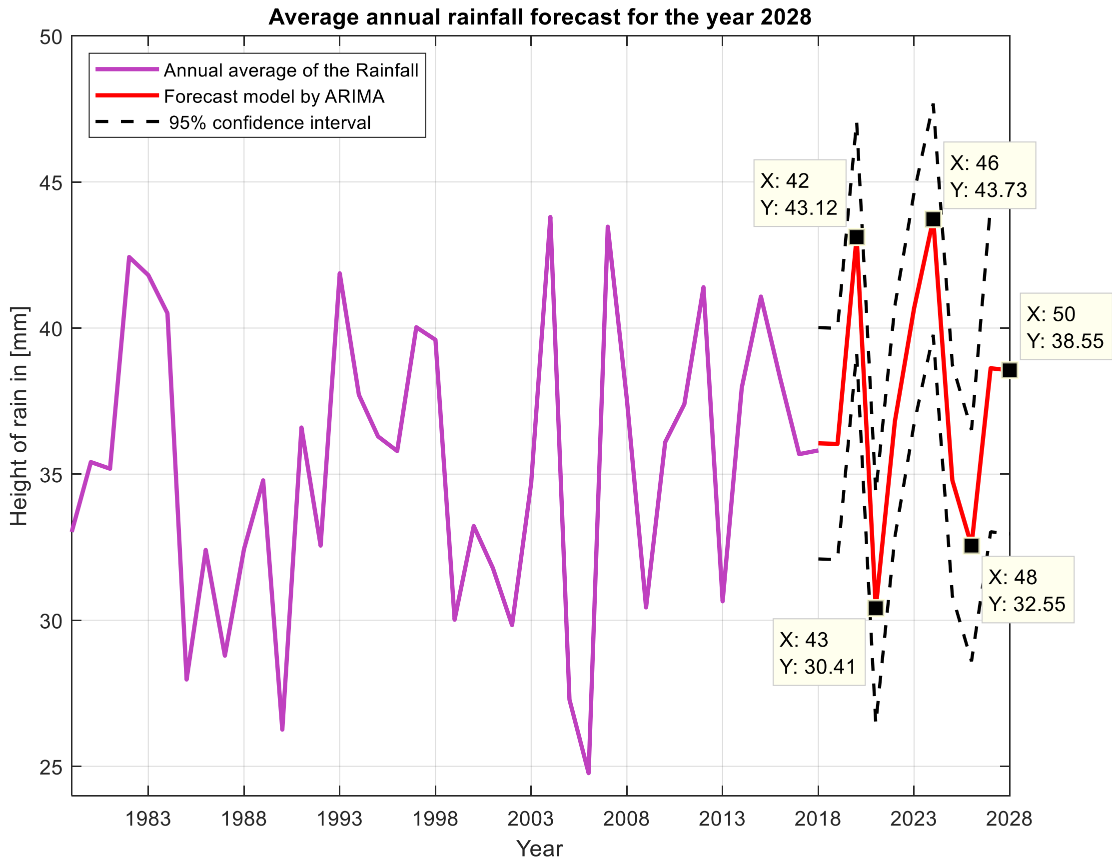

Figure 4: Forecast curve for rainfall in 2028

Table 1: Forecast of annual average rainfall by the ARIMA method

Years of the forecast	Values of the annual average rainfall in [mm]
2019	36.03
2020	43.12
2021	30.41
2022	36.81
2023	40.67
2024	43.73
2025	34.79
2026	32.55
2027	38.62
2028	38.55

Analysing the curves, the average values of the rainfall range from 24.77 mm to 43.8 mm. The minimum value is 24.77mm (in 2006) and the maximum value is 43.8mm (in 2004)

IV. Conclusion

In this article, we are interested in the quantitative analysis of the daily rainfall from 1979 to 2018 in the Boeny region of Madagascar. This part is located between longitude 44° East and 48° East, latitude 18° South and 15° South. To study the predictability of these parameters, it is necessary to make a quantitative study of some climatological parameters. In our case, we proceeded to the use of statistical methods, the ARIMA method.

According to the ARIMA method, the models retained for the average values of rainfall are between 24.77 mm to 43.8 mm. The year 2006 is the year with the least rainfall, the average value of rainfall is 24.77mm. The wettest year is 2004 with an average rainfall value of 43.8mm.

V. References

[1] Méthodes de prévision de série chronologique ARIMA, https://docs.oracle.com/cloud/help/fr/pbcs_common/PFUSU/prhist_arima_intro.htm#PFUSU-pp_user_book_217.

[2] Farid Kadri, Sondès Chaabane, Fouzi Harrou, Christian Tahon, Modélisation et prévision des flux quotidiens des patients aux urgences hospitalières en utilisant l'analyse de séries chronologiques.

[3] Jonas Kibala Kuma, Prévision par l'approche méthodologique de Box et Jenkins : Cas d'une Série Non Saisonnière et Non Stationnaire du type TS (Pratique sur EViews et Stata) [4] L. Chaumont- Statistique descriptive et prévision, Année 2010/2011.

[5] Didier Delignières , Séries temporelles – Modèles ARIMA.

[6] Julien JACQUES, INTRODUCTION AUX SÉRIES TEMPORELLES https://www.coursehero.com/file/43704139/ST-Courspdf/.

[7] JEAN ABADIE, FRANÇOIS MESLIER, Étude de l'utilisation des modèles A.R.I.M.A. pour la prévision à très court terme de l'énergie journalière produite par électricité de France http://www.numdam.org/item?id=RO_1979__13_1_37_0.

[8] Séries temporelles : régression, et modélisation ARIMA(p,d,q), https://avram.perso.univ-pau.fr/sertemp/ser.pdf.