MODELING ANNUAL PRECIPITATION AS A FUNCTION OF CLIMATIC VARIABLES BY FUZZY LOGIC

RANDRIANARIVELO Eddy Flocaudel[1], HARY Jean[2], RANDRIANANDRASANARIVO Raphaëlson Jacques[3], MAXWELL Djaffard[4]

[1] *PhD student, Doctoral School of Life Engineering and Modeling, University of Mahajanga, Madagascar*
[2] *Lecturer, Higher Institute of Science and Technology of Mahajanga, University of Mahajanga, Madagascar*
[3] *PhD student, Doctoral School of Life Engineering and Modeling, University of Mahajanga, Madagascar*
[4] *Professor, Higher Institute of Science and Technology of Mahajanga, University of Mahajanga, Madagascar*

SUMMARY

Currently, the use of artificial intelligence method for forecasting purposes is still lacking in precipitation forecasting research. In this paper, a fuzzy logic-based model is proposed for annual precipitation forecasting in the northwestern part of Madagascar. Several climatic variables were used to build the models. The method is based on the analysis of the correlation coefficient to evaluate the existing relationship between precipitation, temperature, sea surface temperature (SST), atmospheric pressure, specific humidities at 700 hPa and 850 hPa. The models are developed using the multiple linear regression method. Then, the fuzzy inference system is used for series prediction. The study area is subdivided into three rainfall homogeneous sub-zones. Thus, we obtained three different models. The performance of each of these models is evaluated by the value of the mean absolute deviation. All models showed excellent predictive performance. Finally, fuzzy models show an increase in precipitation in all three study sub-areas.

Keywords : *precipitation, multiple linear regression, artificial intelligence, forecasting*

---

1. INTRODUCTION

Precipitation is one of the main indicators of the climate state in the tropical zone. It is the source of all the fresh water available in the biosphere. Food security and water resources are greatly affected by variations in the amount of rainfall. Relevant studies have shown that global warming has exacerbated the instability of the climate system [1]. In Madagascar, fluctuations in rainfall significantly affect agricultural production, food security and water resources. Agricultural production accounts for a significant share of the country's gross domestic product (GDP) and is a major source of employment [2]. Anticipating precipitation is a major challenge for scientists today. Several modeling methods for forecasting purposes have been proposed, namely autoregressive methods, neural methods, hybrid methods and Artificial Intelligence [3][4][5]. However, artificial intelligence methods for forecasting purposes are still lacking in precipitation forecasting research. The objective of this article is therefore to present an analysis to identify the climatic variables closely associated with precipitation in the northwestern part of Madagascar. Then, express the precipitation as a function of the determining climatic variables, followed by a predictive study of the annual precipitation for the future year by the fuzzy inference system.

2. METHODOLOGIES

2.1 Presentation of the study area

Our study area is delimited by longitude between 43°E to 48.5°E and latitude 12°S to 18°S. It includes the regions of Boeny, Sofia and part of the Diana region, Melaky Betsiboka and Alaotra Mangoro.

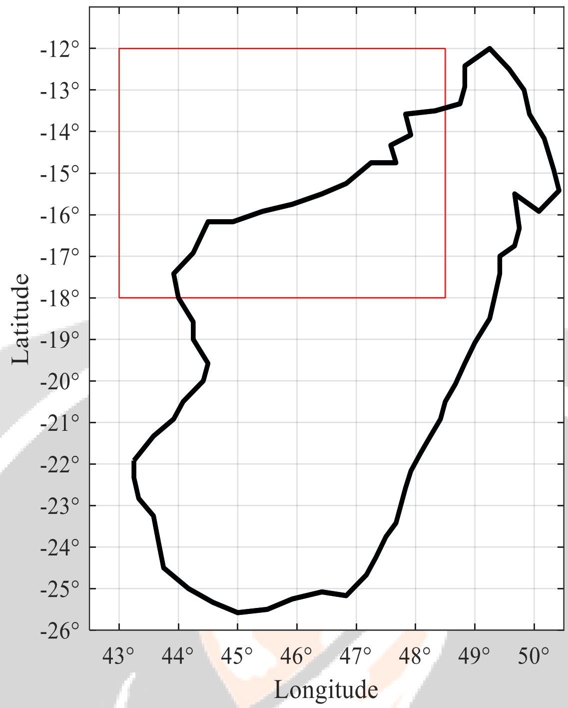

Fig. 1: Representation of the study area

According to the work of Randrianarivelo EF et al [6], in terms of rainfall characteristics, the northwestern part of Madagascar can be subdivided into three rainfall-homogeneous sub-zones: (a) sub-zone 1 is the wettest part, (b) sub-zone 2, corresponding to a zone of heavy rainfall during the dry season, (c) sub-zone 3 relates to the zone of very low precipitation during the dry season and abundant precipitation during the wet season.
In the rest of our work, we will analyze and model the rainfall data from these three different sub-zones.

2.2 Data used

The climate variables used are from the ERA5 hourly data. ERA5 is the fifth generation ECMWF reanalysis. All data were clustered according to a regular 0.25 degree latitude-longitude grid (2D pieces from a 0.25◦x 0.25◦ grid) available at: https://cds.climate.copernicus.eu

V ariables are of different physical nature, characterized by different units. Therefore, it is sometimes desirable to use normalized data in order to avoid the system being parameterized over a particular range of values, thus ignoring extreme values. We normalized the data using the following formula:

$$X_{ij} = \frac{x_{ij} - \overline{x_j}}{\sigma_j} \quad (1)$$

x_{ij} is the value of variable j taken by individual i, $\overline{x}_j$: the mean of variable j

2.3 Linear correlation

Linear correlation allows to detect teleconnections between two climatic parameters because it is based on the common variance between the analyzed variables. It is defined by the following equation [7, 8]:

$$\rho_{XY} = \frac{\frac{1}{N}\sum_{i=1}^{N}\left(X_i - \overline{X}\right)\left(Y_i - \overline{Y}\right)}{\sigma_X \sigma_Y} \quad (2)$$

Où $\overline{X}$ et $\overline{Y}$ sont respectivement les moyennes arithmétiques de variables X et Y, σ_X et σ_Y leurs écart-types.

2.4 Multiple linear regression

The desired linear equation is of the following form:

$$Y = \gamma_0 + \gamma_1 X_1 + \gamma_2 X_2 + \ldots + \gamma_k X_k + \varepsilon \quad (3)$$

Where k corresponds to the number of individuals or observations; Y is the dependent variable; $X_1, X_2, \ldots, X_k$ are the independent variables; ε is the error term; the parameters γ are called partial regression coefficients. They measure the influence of each of the variables on the quantity studied.

2.5 Fuzzy logic

Fuzzy logic is defined as a set of mathematical principles for modeling a complex system by approximate reasoning based on linguistic variables and fuzzy subsets. Fuzzy logic modeling consists of three main blocks (Fig. 2) [9] [10] [11]:

- Fuzzification: This step transforms a real input into a fuzzy subset.
- Fuzzy inference mechanism: application of fuzzy rules to fuzzy values resulting from the fuzzification of the input variable.
- Defuzzification: Extracting a real output value from the output fuzzy subset membership function established by the inference mechanism.

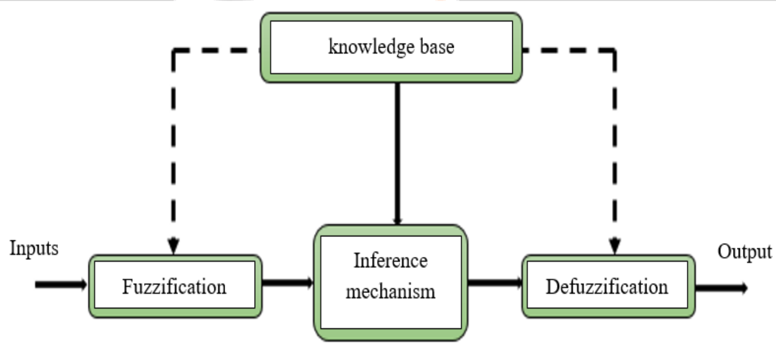

Fig. 2: General structure of a fuzzy model of a dynamic system

2.6 Model validation

- **Mean absolute deviation in percentage**

$$MAPE = 100 \times \frac{1}{N}\sum_{i=1}^{N}\left|\frac{\varepsilon_i}{x_i}\right| \quad (4)$$

where $\varepsilon_i = x_i - \hat{x}_i$ is the residual of the chosen model, x_i is the initial observation, N is the number of observations. The lower the MAPE value, the closer the estimated model is to the observational data. The forecast quality will be judged according to the MAPE value.

- **Accuracy percentage**

The percentage of accuracy is given by equation (5):

$$P = 100 - \sqrt{\frac{\sum_{j}^{N}\left(X_j - \hat{X}_j\right)^2}{N}} \quad (5)$$

Où X_j : observation initiale; $\hat{X}_j$: prévision

3. RESULTS AND DISCUSSIONS

3.1 Correlation between precipitation and climatic variables in the three study sub-areas

Table 1 shows the correlation coefficients between precipitation and climatic variables in the three homogeneous sub-zones. We were able to detect that between rainfall and the three climatic variables (specific humidities, temperature and SST), the correlations are positive while a negative correlation was highlighted between precipitation and atmospheric pressure.

Table 1: Seasonal parameters depending on the ENSO phenomenon

Climate variables	Precipitation in subzone 1	Precipitation in subzone 2	Precipitation in subzone 3
Temperature	0.74	0.76	0.71
Sea surface temperature	0.74	0.81	0.76
Atmospheric pressure	-0.88	-0.90	-0.92
Specific humidity at 700 hPa	0.92	0.94	0.95
Specific humidity at 850 hPa	0.92	0.93	0.95
Wind at 10 m from the ground	0.05	-0.14	-0.67
Wind at 700 hPa	-0.14	-0.26	-0.22
Wind at 850 hPa	-0.07	-0.48	-0.54

3.2 Correlation between precipitation and climatic variables in the three study sub-areas

The equation of multiple linear regression in subarea 1 is given by equation (6):

$$\Delta P_{zone1} = \Omega_1 \text{ Temperature} + \Omega_2 \text{ SST} + \Omega_3 \text{ Pressure} + \Omega_4 \text{ HS700} + \Omega_5 \text{ HS850} + \Omega_C \quad (6)$$

We note that: $\Omega_1 = -141.92$; $\Omega_2 = -275.38$; $\Omega_3 = 4.46$; $\Omega_4 = 311266.60$; $\Omega_5 = 640093.28$ the coefficients of the parameters respectively of the temperature at 2 m from the ground, the sea surface temperature, the atmospheric pressure, the specific humidity at 700 hPa and the specific humidity at 850 hPa. The constant coefficient is $\Omega_C = 105.05$.

ΔP_{zone1} is the annual precipitation in subzone 1 calculated from these five climatic parameters.

The equation of the multiple linear regression obtained in sub-area 2 is given by equation (7):

$$\Delta P_{zone2} = \mu_1 \text{ Temperature} + \mu_2 \text{ SST} + \mu_3 \text{ Pressure} + \mu_4 \text{ HS700} + \mu_5 \text{ HS850} + \mu_C \quad (7)$$

Where the coefficient of temperature at 2 m from the ground is $\mu_1 = -0.745$; the coefficient of sea surface temperature is $\mu_2 = 0.231$; the coefficient of atmospheric pressure is $\mu_3 = 0.060$; the coefficient of specific humidity at 700 hPa is $\mu_4 = 0.352$ and the coefficient of specific humidity at 850 hPa is $\mu_5 = 0.365$. The constant coefficient is $\mu_C = 1{,}3.10^{-15}$.

ΔP_{zone2} is the annual precipitation in subzone 2 calculated from these five climatic parameters.

The multiple linear regression equation in subarea 3 is given by equation (8):

(8)

$\Delta P_{zone3} = \xi_1$ Temperature + ξ_2 SST+ ξ_3 Pressure+ ξ_4 HS700+ ξ_5 HS850 + ξ_C

Where the coefficient of temperature at 2 m from the ground is $\xi_1 = -0.416$; the coefficient of sea surface temperature is $\xi_2 = 0.244$; the coefficient of atmospheric pressure is $\xi_3 = -0.181$; the coefficient of specific humidity at 700 hPa is $\xi_4 = 0.325$ and the coefficient of specific humidity at 850 hPa is $\xi_5 = 0.157$. The constant coefficient is $\xi_C = -8{,}7.10^{-14.}$

ΔP_{zone3} is the annual precipitation in subzone 3 calculated from these five climatic parameters.

In the remainder of this paper, the three constants $\Omega_{C,}$ μ_C and ξ_C are not included in the equations of the multiple regressions. The five variables (temperature at 2 m from the ground, sea surface temperature, atmospheric pressure, specific humidity at 700 hPa and specific humidity at 850 hPa) are sufficient for modeling in our study.

3.3 Universe of discourse and number of partitions of each climatic variable

Table 2 illustrates the discourse universes of each variable in the three study sub-areas. Then the number of partitions of the variables are recorded in Table 3.

Table 2: Universe of discourse of each climatic variable

Climate variables	Subzone 1	Subzone 2	Subzone 3
Temperature	[-3527.85 -3291.04]	[-1,691 1,730]	[-0.933 0.966]
Sea surface temperature	[-7805.25 -7459.62]	[-0.555 0.493]	[-0.588 0.522]
Atmospheric pressure	[4528.92 4535.98]	[-0.165 0.1362]	[-0.388 0.557]
Specific humidity at 700 hPa	[1606.03 1964.12]	[-0.603 1.159]	[-0.550 0.989]
Specific humidity at 850 hPa	[6540.38 7341.33]	[-0.675 1.113]	[-0.383 0.450]
Precipitation	[1733.63 2517.69]	[-1,109 1,195]	[-0.957 1.638]

Table 3: Number of partitions of each climate variable

Climate variables	Subzone 1	Subzone 2	Subzone 3
Temperature	28 scores	22 scores	28 scores
Sea surface temperature	30 scores	27 scores	30 scores
Atmospheric pressure	33 scores	33 scores	33 scores
Specific humidity at 700 hPa	20 scores	28 scores	20 scores
Specific humidity at 850 hPa	29 scores	33 scores	29 scores
Precipitation	36 scores	22 scores	36 scores

3.3 Universe of discourse and number of partitions of each climatic variable

Fig. 3 represents the mapping of the “Mamdani” type fuzzy inference system. The five input variables are on the left and on the right, the output variable (precipitation). The values in the parentheses represent the numbers of partitions to each of the variables. The numbers of partitions of the variables in sub-area 1 and 3 are identical therefore we present only two different FIS system models .

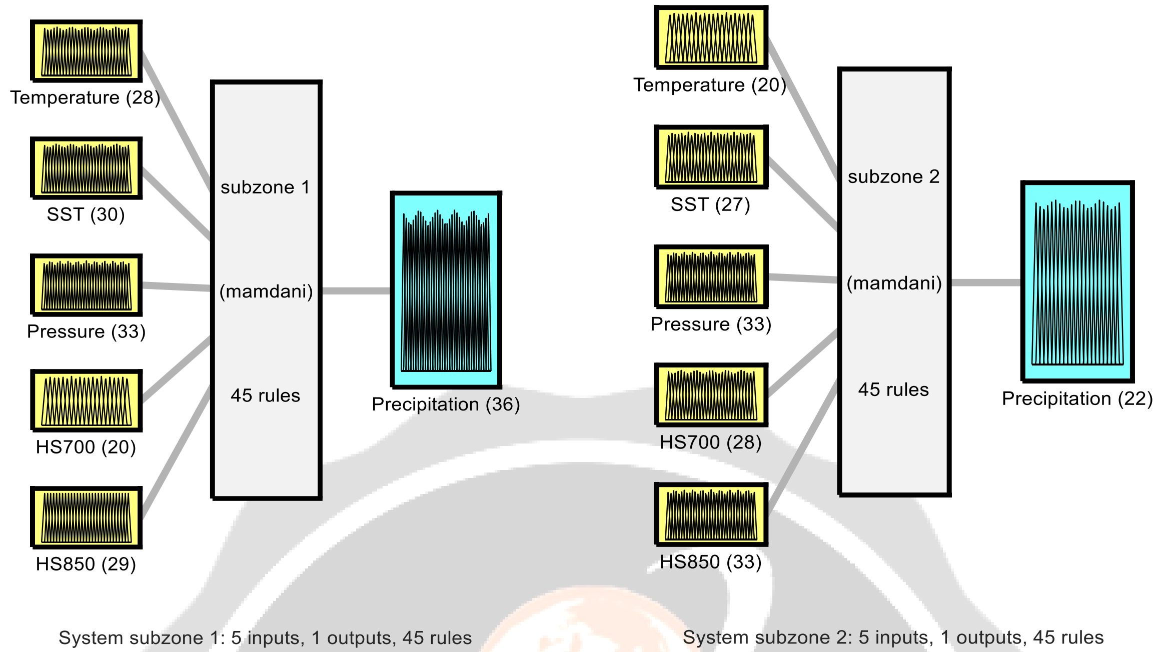

Fig. 3 : Proposed FIS inputs and outputs in subzone 1 and 3

3.4 Presentation of some rules

The rules appear in the order of the chosen model. For a model of order 1, we obtain 45 rules. Below are some examples of these rules applied in sub-area 1, and sub-area 3.

Subzone 1:

1. Si (Température est T18) et (SST est SST17) et (Pression est PR17) et (HS700 est HS9) et (HS850 est hs17) alors (Précipitation est P26) (1)
2. Si (Température est T18) et (SST est SST16) et (Pression est PR16) et (HS700 est HS13) et (HS850 est hs16) alors (Précipitation est P28) (1)
3. Si (Température est T19) et (SST est SST19) et (Pression est PR14) et (HS700 est HS11) et (HS850 est hs17) alors (Précipitation est P29) (1)
4. Si (Température est T17) et (SST est SST9) et (Pression est PR11) et (HS700 est HS17) et (HS850 est hs20) alors (Précipitation est P33) (1)

………..........

40. Si (Température est T10) et (SST est SST10) et (Pression est PR16) et (HS700 est HS8) et (HS850 est hs18) alors (Précipitation est P20) (1)
41. Si (Température est T1) et (SST est SST1) et (Pression est PR13) et (HS700 est HS20) et (HS850 est hs29) alors (Précipitation est P36) (1)
42. Si (Température est T4) et (SST est SST6) et (Pression est PR10) et (HS700 est HS10) et (HS850 est hs21) alors (Précipitation est P21) (1)
43. Si (Température est T14) et (SST est SST11) et (Pression est PR20) et (HS700 est HS8) et (HS850 est hs16) alors (Précipitation est P19) (1)
44. Si (Température est T19) et (SST est SST14) et (Pression est PR16) et (HS700 est HS2) et (HS850 est hs10) alors (Précipitation est P10) (1)
45. Si (Température est T5) et (SST est SST1) et (Pression est PR22) et (HS700 est HS10) et (HS850 est hs20) alors (Précipitation est P16) (1)

Subzone 2:

1. Si (Température est T12) et (SST est SST12) et (Pression est PR18) et (HS700 est HS11) et (HS850 est hs18) alors (Précipitation est P16) (1)
2. Si (Température est T13) et (SST est SST13) et (Pression est PR17) et (HS700 est HS15) et (HS850 est hs22) alors (Précipitation est P22) (1)
3. Si (Température est T14) et (SST est SST11) et (Pression est PR16) et (HS700 est HS10) et (HS850 est hs17) alors (Précipitation est P17) (1)
4. Si (Température est T9) et (SST est SST20) et (Pression est PR11) et (HS700 est HS21) et (HS850 est hs22) alors (Précipitation est P21) (1)
5. Si (Température est T8) et (SST est SST21) et (Pression est PR19) et (HS700 est HS14) et (HS850 est hs23) alors (Précipitation est P17) (1)

……………………………………………………………………………...

43. Si (Température est T8) et (SST est SST18) et (Pression est PR22) et (HS700 est HS9) et (HS850 est hs12) alors (Précipitation est P7) (1)
44. Si (Température est T11) et (SST est SST15) et (Pression est PR19) et (HS700 est HS1) et (HS850 est hs7) alors (Précipitation est P4) (1)
45. Si (Température est T1) et (SST est SST27) et (Pression est PR22) et (HS700 est HS8) et (HS850 est hs16) alors (Précipitation est P1) (1)

Subzone 3:

1. If (Température is T18) and (SST is SST14) and (Pression is PR15) and (HS700 is HS6) and (HS850 is hs18) then (Précipitation is P17) (1)
2. If (Température is T19) and (SST is SST15) and (Pression is PR15) and (HS700 is HS10) and (HS850 is hs19) then (Précipitation is P23) (1)
3. If (Température is T20) and (SST is SST12) and (Pression is PR18) and (HS700 is HS8) and (HS850 is hs18) then (Précipitation is P21) (1)
4. If (Température is T18) and (SST is SST22) and (Pression is PR20) and (HS700 is HS18) and (HS850 is hs24) then (Précipitation is P36) (1)
5. If (Température is T16) and (SST is SST23) and (Pression is PR12) and (HS700 is HS11) and (HS850 is hs22) then (Précipitation is P24) (1)

……..

40. If (Température is T8) and (SST is SST21) and (Pression is PR15) and (HS700 is HS5) and (HS850 is hs13) then (Précipitation is P7) (1)
41. If (Température is T1) and (SST is SST30) and (Pression is PR18) and (HS700 is HS20) and (HS850 is hs29) then (Précipitation is P29) (1)
42. If (Température is T5) and (SST is SST25) and (Pression is PR18) and (HS700 is HS13) and (HS850 is hs18) then (Précipitation is P18) (1)
43. If (Température is T10) and (SST is SST20) and (Pression is PR11) and (HS700 is HS10) and (HS850 is hs14) then (Précipitation is P12) (1)
44. If (Température is T16) and (SST is SST17) and (Pression is PR13) and (HS700 is HS3) and (HS850 is hs9) then (Précipitation is P9) (1)
45. If (Température is T1) and (SST is SST30) and (Pression is PR11) and (HS700 is HS10) and (HS850 is hs2) then (Précipitation is P12) (1)

3.4 Graphical representation of the model and prediction

Once the rules are established, we proceed to the defuzzification of the values of the model output noted ΔSPzone. We can thus calculate the series of the model corresponding to the annual precipitation in the three study sub-zones. From this model we can predict the evolution of the future precipitation.
The MAPE calculations and accuracy percentages in the three sub-areas are given in Table 4. We found that the MAPE values are very low and the accuracy percentages are very high. These results confirm the reliability of the models.

Table 4: Validation of FIS models

Validation	Subzone 1	Subzone 2	Subzone 3
MAPE (%)	2,1338	0.2028	0.0353
P (%)	97.8672	99.9835	99.9682

Fig. 4 shows the precipitation forecast for the year 2024. The blue curve represents the observation data, in red, it is the modeling of the data and then, the green curve is the forecast of the amount of rainfall. The two curves in blue and red are almost contiguous. The rainfall in the northwestern part of Madagascar tends to increase for the following year.

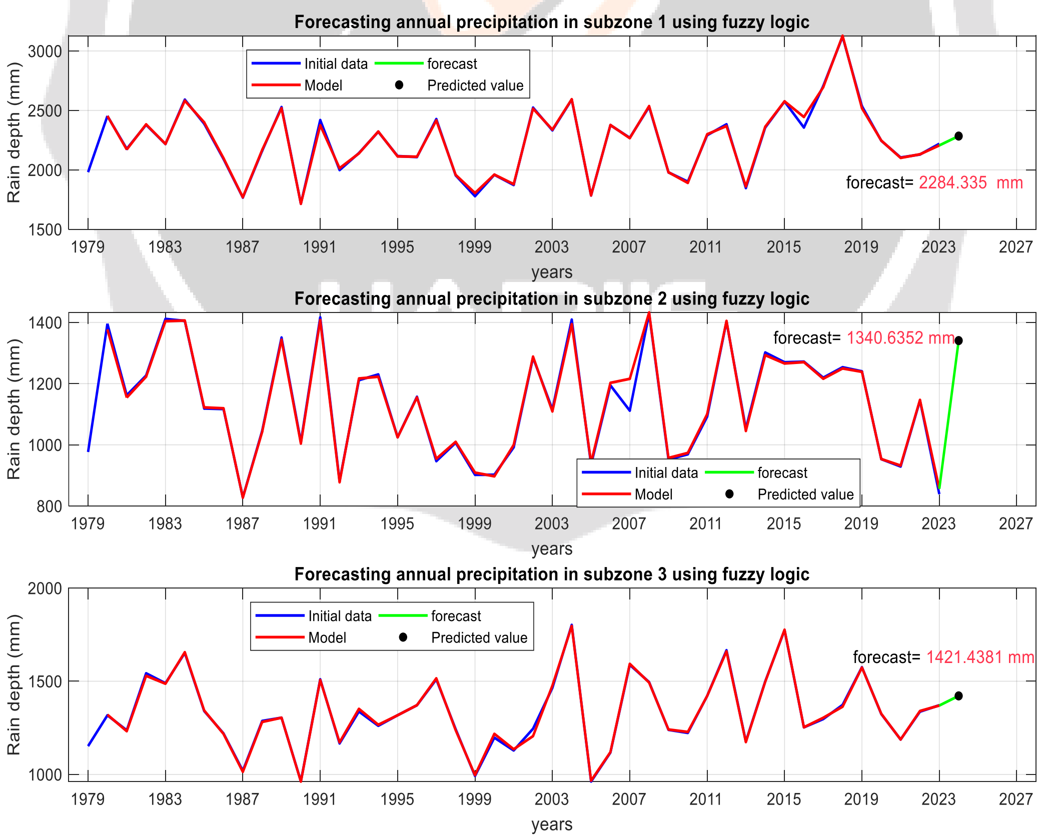

Fig. 4 : Precipitation forecast in the three homogeneous sub-zones

4. CONCLUSIONS

In conclusion, this work showed the relationship between precipitation and climate variables by multiple linear regression. Using the correlation method, we found that five climate variables evolve with precipitation, namely temperature, sea surface temperature, atmospheric pressure and specific humidities at 700 hPa and 850 hPa. The models chosen are the five-input, one-output and 45-rule models, thus showing excellent performance. All the models obtained showed average errors less than 10%. From these models, we could see that the forecast of annual precipitation in our three study sub-areas tends to increase. For next year, we expects a more generous year in terms of precipitation.

5. REFERENCES

[1]. Zhang, W. & Zhou, T. 2019 Significant increases in extreme precipitation and the associations with global warming over the global land monsoon regions. Journal of Climate 32 (24), 1477–1493.
[2]. FAO. Madagascar: Country Indicators. 2024. Available at: https://www.fao.org/world-agriculture-watch/our-program/mdg/en/ (accessed: 20/07/2024).
[3]. Rabeniaina ADU, Rakotovao N A., Ambeondahy, Hary J., Modeling Precipitation as a Function of Several Variables in Northern Madagascar Using the ANFIS Network. AIS. 2024;6(1):135-45.
[4]. Salim Ahmed Ali, "Study of climate variability and the impact of ENSO and IOD modes in northern Madagascar and the Comoros", Doctoral thesis, Physics of the Earth, Energy and the Environment, University of Antananarivo, (2018).
[5]. Rakotoarimanana R. S., Ramahefy T. R., Randrianja S., Prediction De Precipitation Mensuelle Par Reseau De Neurone Recurrent, International Journal of Progressive Sciences and Technologies (IJPSAT), Vol. (42) No. December 2023, pp. 265-276.
[6]. Randrianarivelo EF, Hanta TO, Jakely V., Randrianandrasanarivo R. J. et al. Influence of the ENSO phenomenon on precipitation in the north-western part of Madagascar. Int J Adv Res Innov Ideas Educ. 2024;10(2):132-43.
[7]. Ronald Aylmar Fisher, "The Correlation between Relatives on the Supposition of Mendelian Inheritance", Philosophical Transactions of the Royal Society of Edinburgh, vol. 52, 1918, p. 399–433.
[8]. OPENSHAW, Stan and Peter TAYLOR (1979b). "A million or so correlation coefficients". Statistical methods in the spatial sciences, p. 127–144.
[9]. Gossé JB, Guillaumin C. L'apport de la représentation VAR de Christopher A. Sims à la science économique. Actual Économique. 2013;89(4):305-19.
[10]. Zengui A. E. et Kasmi A. E., Application des systèmes neuro-flous pour l'étude des isolateurs de haute tension, Mémoire de Master, Université Kasdi Merbah Ouargla.
[11]. Brahim W., Benkouider M., Classification neuro-floue par l'identification des défaillances, mémoire de Master, Université BLIDA I, 2015.