Scientific Research Publishing

Open Journal of Statistics, 2024, 14, 394-411
https://www.scirp.org/journal/ojs
ISSN Online: 2161-7198
ISSN Print: 2161-718X

Cameroon Climate Predictions Using the SARIMA-LSTM Machine Learning Model: Adjustment of a Climate Model for the Sudano-Sahelian Zone of Cameroon

Joseph Armathé Amougou[1,2], Isidore Séraphin Ngongo[3,4,5], Patrick Forghab Mbomba[2], Romain Armand Soleil Batha[1,2], Paul Ghislain Poum Bimbar[2,4,6]

[1]Département de Géographie, Université de Yaoundé 1, Yaoundé, Cameroon
[2]Observatoire National sur les Changements Climatiques (ONACC), Yaoundé, Cameroon
[3]Laboratoire d'Analyse et Modélisation Multidisciplinaire Statistique (SAMM), Université de Paris 1 Panthéon Sorbonne, Paris, France
[4]Ecole Nationale Supérieure Polytechnique de Yaoundé 1, Yaoundé, Cameroon
[5]Département de Mathématiques, Ecole Normale Supérieure de Yaoundé, Yaoundé, Cameroon
[6]Département d'Informatique, Université de Yaoundé 1, Yaoundé, Cameroon
Email: bathsol33@yahoo.fr

How to cite this paper: Amougou, J.A., Ngongo, I.S., Forghab Mbomba, P., Batha, R.A.S. and Poum Bimbar, P.G. (2024) Cameroon Climate Predictions Using the SARIMA-LSTM Machine Learning Model: Adjustment of a Climate Model for the Sudano-Sahelian Zone of Cameroon. *Open Journal of Statistics*, **14**, 394-411.
https://doi.org/10.4236/ojs.2024.143016

Received: April 19, 2024
Accepted: June 27, 2024
Published: June 30, 2024

Abstract

It is acknowledged today within the scientific community that two types of actions must be considered to limit global warming: mitigation actions by reducing GHG emissions, to contain the rate of global warming, and adaptation actions to adapt societies to Climate Change, to limit losses and damages [1] [2]. As far as adaptation actions are concerned, numerical simulation, due to its results, its costs which require less investment than tests carried out on complex mechanical structures, and its implementation facilities, appears to be a major step in the design and prediction of complex mechanical systems. However, despite the quality of the results obtained, biases and inaccuracies related to the structure of the models do exist. Therefore, there is a need to validate the results of this SARIMA-LSTM-digital learning model adjusted by a matching approach, "calculating-test", in order to assess the quality of the results and the performance of the model. The methodology consists of exploiting two climatic databases (temperature and precipitation), one of which is in-situ and the other spatial, all derived from grid points. Data from the dot grids are processed and stored in specific formats and, through machine learning approaches, complex mathematical equations are worked out and interconnections within the climate system established. Through this mathematical approach, it is possible to predict the future climate of the Su-

dano-Sahelian zone of Cameroon and to propose adaptation strategies.

Keywords

Adjustment, Calibration, Climate, Sudano-Sahelian Zone, Numerical Model

1. Introduction

The diversity of data sources, the ability to parameterize physical processes, modern programming techniques and the use of supercomputers are all tools that enable the scientific community to better analyze complex physical phenomena. Thanks to mathematical and computer models, it is now possible to predict the future climate of a region and to propose adaptation strategies. These climate predictions are the result of improved numerical methods, the ability of supercomputers to solve complex mathematical and physical equations in grids of completely different points. The dimensioning of these point grids by machine learning (ML) approaches has also facilitated the understanding of the complex interconnections within a single climate system. The present study is based on the use of the SARIMA-LSTM machine learning model to predict the future climate of the Sudano-Sahelian zone of Cameroon. Given the quality of the results obtained, the adjusted SARIMA-LSTM digital learning model is the most appropriate model for the ecological reality of the Sudano-Sahelian zone of Cameroon.

2. Materials and Methods

2.1. Presentation of the Study Area

With a surface area of about 100,353 km^2, or 21% of the national territory, the Sudano-Sahelian zone of Cameroon covers the administrative regions of the North (66,090 km^2, bordering the Adamawa plateau) and the Far North (34,263 km^2, located on the southern borders of Lake Chad). Rainfall conditions in this area are rapidly deteriorating, from 1200 mm of annual rainfall in the Benue basin to less than 500 mm on the borders of Lake Chad (see **Figure 1**).

2.2. Climate Conditions in the Sudano-Sahelian Zone

Generally speaking, Cameroon is under the influence of two main air masses coming from either side of the equator. One comes from the Azores High in the North Atlantic and the second from the St. Helena High in the South Atlantic. This second air mass brings fresh and humid air to the continent. These two air masses merge along a convergence zone, also known as the intertropical front (ITF), which changes latitude according to the season.

Research results show that the ITF rotates during the year, thus contributing to its seasonality. Four positions of the ITF are commonly observed during the year [3]:

1) *between November and February, it lies south of the* 4th *parallel: in this position, the whole of North Cameroon is in the dry season,*

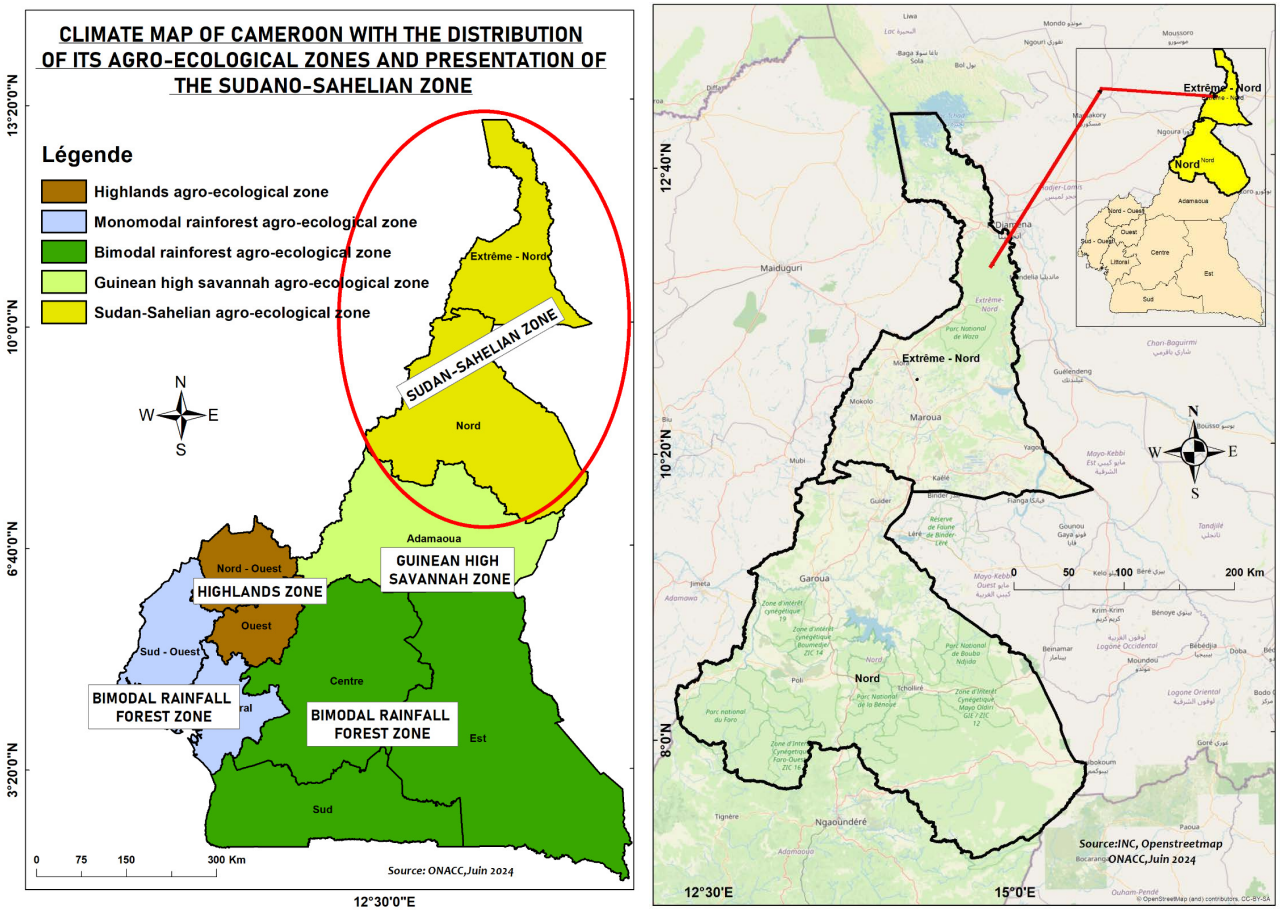

Figure 1. Climate map of Cameroon with the distribution of its agro-ecological zones and presentation of the Sudano-Sahelian zone.

2) *between March and April, it rises above the* 7^{th} *parallel: in this position, it is the extension of the dry season from the Far North to the southern part of Garoua,*

3) *between May and October, it is above the* 13^{th} *parallel: at this level, the entire country is in the rainy season.*

According to Yann L'Hôte (1999), the entire Sudano-Sahelian zone of Cameroon is subject to a tropical climate of the Sudano-Sahelian type, characterized by the following [4]:

• A single rainy season centred on a maximum in August, with average annual totals varying from 400 mm in the far part to 1100 mm in the southern part;

• A dry season which is rigorous and long (at least eight months) as we move northwards and away from the Mandara Mountains;

• Solar radiation and temperatures, often very high as we move closer to the banks of Lake Chad (average annual temperatures around 31˚C, with significant average temperature differences between 8˚C and 20˚C);

• Significant average temperature differences (between 8˚C and 20˚C).

As part of this study, we are beginning a methodical and comprehensive process, starting with the rigorous collection of relevant data. Once this data is gathered, we put it through a series of in-depth analyses, including separating it

into training and testing sets to ensure the objectivity of our approach. Next, we proceed to develop predictive and analytical models, using methods and techniques tailored to our specific research objective. Once the models are developed, we subject them to careful evaluation to determine their performance and accuracy [5]. This evaluation often involves comparative tests with real data and established benchmarks, in order to validate the effectiveness of the proposed models. Finally, we carry out an in-depth interpretation of the results obtained, placing them in the context of our research problem and examining the practical implications of our findings (see **Figure 2**).

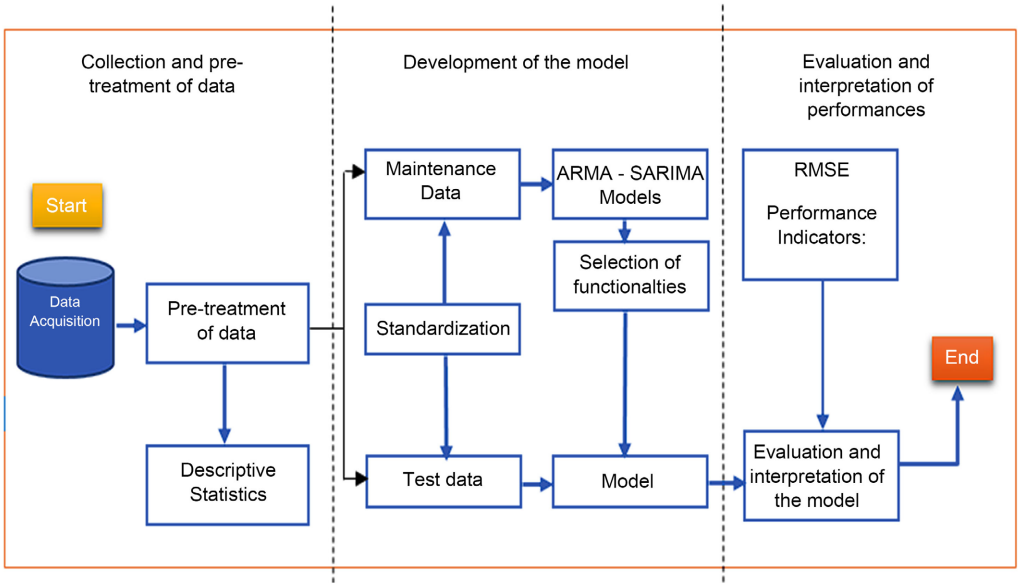

Figure 2. Study framework, from data collection, data separation and model development to model assessment and interpretation.

2.3. Data Collection and Model Construction

The data used are temperature and precipitation time series. The statistical characteristics of these series are shown in **Table 1** and **Table 2**. They were provided by the National Meteorological Centre (NMC).

2.4. LSTM (Long-Short Term Memory) Model

LSTM is one of the Key Neural Network (KNN) architectures used in different domains for the processing of time series data. KNN is mainly used for temporally correlated data [6]. It considers the correlation between previous and current data and predicts future data through past data while having a structure in which signals are cycled to predict future data.

However, there is a problem that past data cannot be stored for a long time. LSTM is an architecture that emerged to overcome these problems. There is a total of six parameters, and through the four-gate structure, not only short-term memory but also long-term memory can be solved, and the structure of LSTM can be confirmed in (**Figure 3**).

Table 1. Climate variables and hazards identified by AEZ [7] [8].

Variables/climatic Hazard	Soudano-Sahelian	Guinea High Savana	Highlands	Bimodal rainfall	Monomodal Rainfall
Temperatures	↑	↑	↑	↑	↑
Heat waves	+++++	++	++	+++	++
Dust storms	++++	NC	NC	NC	NC
Precipitations	↓	↓	↓		↑
Rainfall quantity	+	++	++	++	++++
Rainfall variability	++	++	++	++	+++
Violent winds	++	++	++	+	+++
Extreme Events	↑	↑	↑	↑	↑
Droughts	++++	++	+	+	NC
Floods	+++	++	++	+++	+++
Landslides	++	+++	+++	++	+
Land and Coastal Erosion	+++	++	++	++	+++
Sea Level Rise	NC	NC	NC	NC	↑

Legend: ↑ increase; ↓ decrease; → stable; NC not concern.

Table 2. Stations, coordinates and parameters of localities in the Sudano-Sahelian zone.

Zones	Station names	Geographical information		Data used	Chronicles
		Longitudes	**Latitudes**		
Sudano-Sahelian	Maroua	**14.32**	**10.59**	Temperature	1900-2021
				Precipitation	1953-2021
	Kaele	**14.44**	**10.10**	Temperature	1900-2021
				Precipitation	1953-2021
	Guider	**13.94**	**9.93**	Temperature	1900-2021
				Precipitation	1953-2021
	Kousseri	**15.02**	**12.07**	Temperature	1900-2021
				Precipitation	1953-2021
	Garoua	**9.3**	**13.39**	Temperature	1981-2022
				Precipitation	1981-2022

The LSTM network has the same chain structure as the KNN, but the repetitive modules of the KNN are structured to exchange information with each other through four layers, not just one layer tanh. The state in the LSTM cell is divided into two vectors, where h_t means a short-term state and C_t means a long-term state. Data can be added to or removed from the cell state via sigmoid gates, each gate having an individual weight, similar to a layer or series of matrix operations. It is designed to solve the problem of long-term dependencies, as portals can also retain information from long-term historical data.

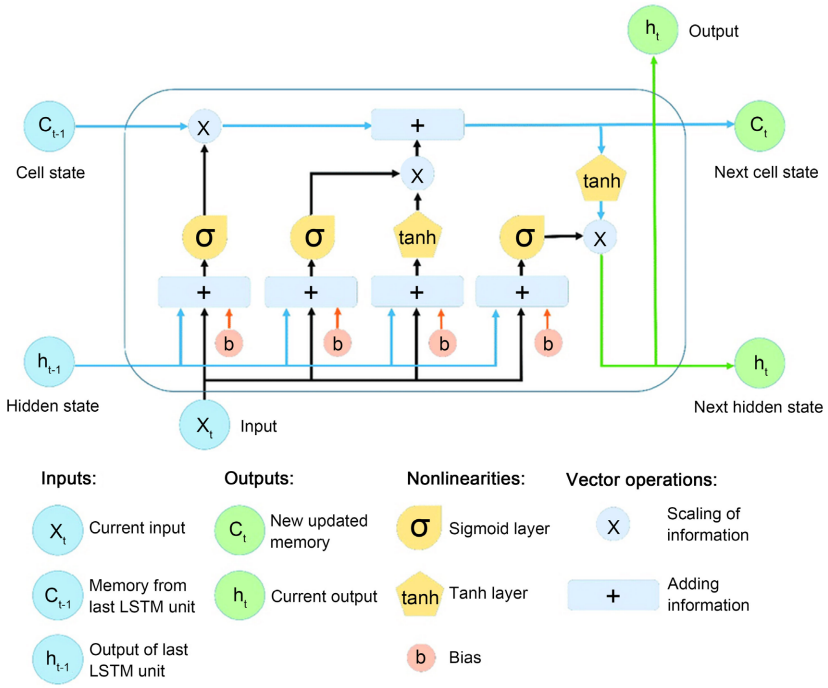

Figure 3. Structure of the Long Short-Term Memory (LSTM).

The first step in an LSTM network is to identify and decide which unnecessary information to omit from the cell. The corresponding cell is called the omit gate, and the process determines the output of the last LSTM cell (h_{t-1}) to $t-1$ and the current input (x_t) at the current time t by a sigmoid function. In this case, the value that comes out of the sigmoid function has a value between (0~1). The bigger the value, the more information about the previous state stored is intact. The smaller the value, the more information about the previous state is forgotten and the part to be omitted from the previous output is decided.

After passing through the forgetting gate, it goes through the process of selecting the information to be stored. Through the omit gate, the memory cell (c_{t-1}) of the previous moment is forgotten, new information to be stored is added, and the value of each item is determined as newly added information. In this case, appropriate choices are made rather than unconditionally accepting new information. A gate that fulfils a corresponding role is called an entrance gate. Hence, the sigmoid function enters through the last LSTM cell (h_{t-1}) and the current value (x_t), and the function tanh, which is an activation function, is added. After passing through the sigmoid layer, the value is between (0 - 1) and represents the degree to which the new information is updated, and the value after going through the tanh has a value between (−1 - 1) and has a weight representative of the importance given to it. Finally, the entrance gate performs the Hadamard product operation on both values, and the corresponding new memory is added to the previous cell state (C_{t-1}) to become C_t.

After assessing the value of the new information via the entrance gate and selecting the information to be stored, the next process is to select the exit infor-

mation. The corresponding gate is called the exit gate, and a sigmoid function is obtained through the current time value (x_t) and the value of the last LSTM cell (h_{t-1}). The result value via the sigmoid function performs the operation of the current cell state (C_t) and the Hadamard product, and the filtering effect of the value occurs and it becomes a hidden state.

$$f_t = \sigma\left(x_t W_f + h_{t-1} W_f + b_f\right), \tag{1}$$

$$g_t = \tanh x_t W_g + h_{t-1} W_g + b_g, \tag{2}$$

$$i_t = \sigma\left(x_t W_i + h_{t-1} W_i + b_i\right), \tag{3}$$

$$c_t = f_t \odot c_{t-1} + g_t \odot i_t, \tag{4}$$

$$o_t = \sigma\left(x_t W_o + h_{t-1} W_o + b_o\right), \tag{5}$$

$$h_t = o \odot \tanh\left(c_t\right), \tag{6}$$

Here, σ denotes a sigmoid function, W denotes a weighting matrix and b denotes a bias. c_t designates the current status of the cell, c_{t-1} refers to the state of the cell at the previous time and stands for the Hadamard product operation. Equation (1) refers to the process of passing through the forget gate, and in the process of passing through the input data, the cell state is updated via Equations (2) to (4). Then it passes through the exit gate indicated by Equation (5) and the LSTM operates in a structure in which the final state of the hidden layer is updated via Equation (6)

2.5. Non-Seasonal ARIMA

Divide the term ARIMA into three terms,

AR, *I*, *MA*:

AR(p) represents the autoregressive model; the parameter p is an integral number that confirms the number of lagged series to be used to predict periods, e.g. yesterday's average temperature correlates with today's temperature [9], so use the parameter *AR*(1) to predict future temperatures.

The model formula *AR*(p) is:

$$y^t = \mu\theta_1 Y_{t-1} \cdots \theta_p Y_{t-p} \tag{7}$$

where μ is the constant term, p is the period to be used in the regression, and θ is the parameter adjusted to data.

I(d) The differentiation part and the parameter d indicate the number of orders of differentiation that will be used. It tries to make the series stationary, for example:

If $d=1$: $y_t = Y_t - Y_{t-1}$ where y_t is the differentiated series, and $Y_t - Y_{t-1}$ is the original series

If $d=2$: $y_t = \left(Y_t - Y_{t-1}\right) - \left(Y_{t-1} - Y_{t-2}\right) = Y_t - 2Y_{t-1}Y_{t-2}$

Note that the second difference is a change in change, which measures local "**acceleration**" rather than trend.

MA(q) Means moving average model, the q is the number of lagged forecast

error terms in the prediction equation, e.g. it is strange, but this term *MA* takes a percentage of the errors between the predicted value and the actual value [9] [10] [11]. It assumes that past errors will be similar in future events.

The model's formula $MA(p)$ is:

$$y_t = \mu + \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} - \cdots - \theta_q\varepsilon_{t-q}$$

where μ is the constant term, q is the period to be used on the term e, and θ is the error-adjusted parameter

$$ET = Y_{t-1} - y^t - e_t = Y_{t-1} - y^t - 1 \tag{8}$$

2.6. Seasonal ARIMA

The parameters p,d,q differ from non-seasonal parameters. $SAR(P)$ Is the seasonal self-regression of the series?

The model formula $SAR(P)$ is $y^t = \mu + \theta_1 Y_{t-s}$ where P is the number of auto-regression terms to be added, usually no more than one term, s is the number of periods to be used as a base, and θ is the parameter adjusted to data [12].

Usually, when the topic is weather forecasting, 12 months ago, you have some information to bring to the current period.

The setting $P=1$ (*i.e.*, $SAR(1)$) add a multiple of Y_{t-s} to the forecast for the moment **Figure 4** represents the temperature variation in Maroua $I(D)$ the seasonal difference MUST be used when you have a stable model.

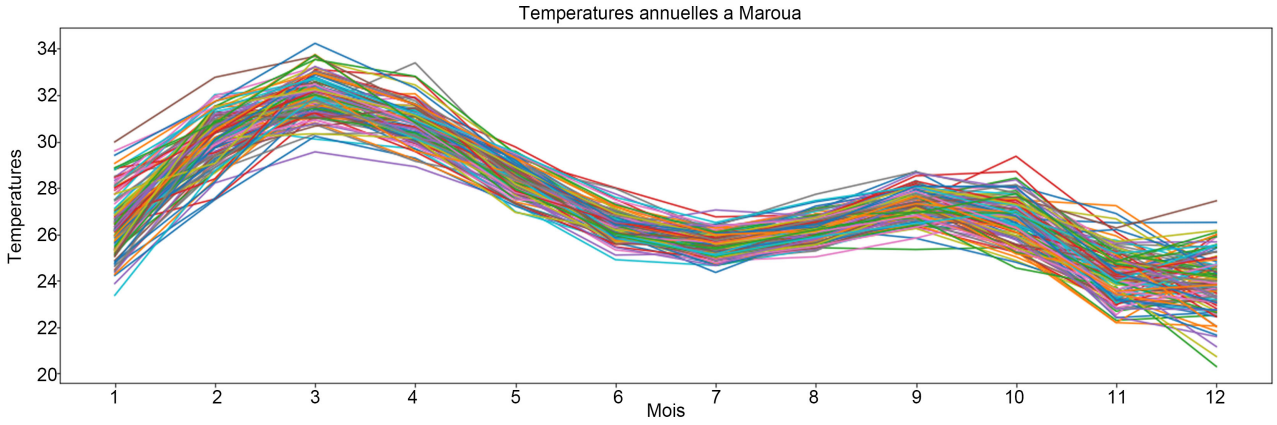

Figure 4. Temperature variation in Maroua.

If $d=0$ and $D=1$: $y_t = Y_t - Y_{t-s}$, y_t is the differentiated series, and $Y_t - Y_{t-s}$ is the original seasonal range.

If $d=1$ and $D=1$:

$$y_t = \left(Y_t - Y_{t-1}\right) - \left(Y_{t-s} - Y_{t-s-1}\right) = Y_t - Y_{t-1} - Y_{t-s} - Y_{t-s-1}$$

2.7. Performance of Indicators

We used three types of indicators to compare the performance of the water level prediction model *MSE* (average squared error), *NSE* (Nash-Sutcliffe efficiency

ratio) and *MAE* (average absolute error). *MSE* Is a widely used method for evaluating the performance of a regression model. It is the square of the difference between the actual observed value and the predicted value. The indicator is sensitive to outliers because the difference between the observed and predicted values is squared. In the case of the hydrological model, the MSE index was selected because it can cause human casualties if an outlier occurs in the prediction [9] [10] [11]. In case of the *NSE*, it is a widely used indicator for assessing the performance of hydrological models. *NSE* Has a value of (−∞ - 1) as an index frequently used to assess the performance of hydrological models. A value closer to means better model performance. In the case of *MAE*, This means the average of all absolute errors of the observed and predicted values and has the advantage of intuitively checking the performance of the model. The equations for the performance comparison metrics are given as follows:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(OBS_i - SIM_i\right)^2 \tag{9}$$

$$NSE = 1 - \frac{\sum_{i=1}^{N}\left(OBS_i - SIM_i\right)^2}{\sum_{i=1}^{N}\left(OBS_i - OBS\right)^2}\ NSE = 1,\ 0 < NSE < 1,\ NSE < 0 \tag{10}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|OBS_i - SIM_i\right|, MAR > 0 \tag{11}$$

3. Results

Annual Temperatures in Maroua;

The highest temperatures occur in February and April, while the lowest temperatures occur in July and December. Taking the monthly average levels of each of these lines and creating a single line, (**Figure 4**) represents the monthly variation in Maroua.

Here are some statistics from this series to see a trend over the years. The average temperature increased by about 5.25% over more than a century, from 26.5 to 28.5 degrees Fahrenheit. First, divide the data into training, validation and test sets to see how it all works together. (**Figure 5**) shows the annual change in Maroua: a confirmation forward over 48 months, then extrapolate the future over another year to compare with the test set.

The RMSE temperature of the baseline is 2.17˚C. An increasing trend is shown in the data, and seasonality is apparent, with higher temperatures at the beginning and end of each season and lower temperatures in the middle. To create a time series forecast (constant mean, variance and autocorrelation). The function occupied can determine whether the series is static. We can safely build our model if the series has less than 5% of P value.

(**Figure 6**) shows the autocorrelation, partial correlation and distributed graph. The partial autocorrelation function, or PACF, is used to describe this. Without considering the impact of past lags, it displays the correlation between the current temperature and the type covered. For example, in the case of tem-

perature, it shows only the impact of lag three without considering the effect of lags 1 and 2.

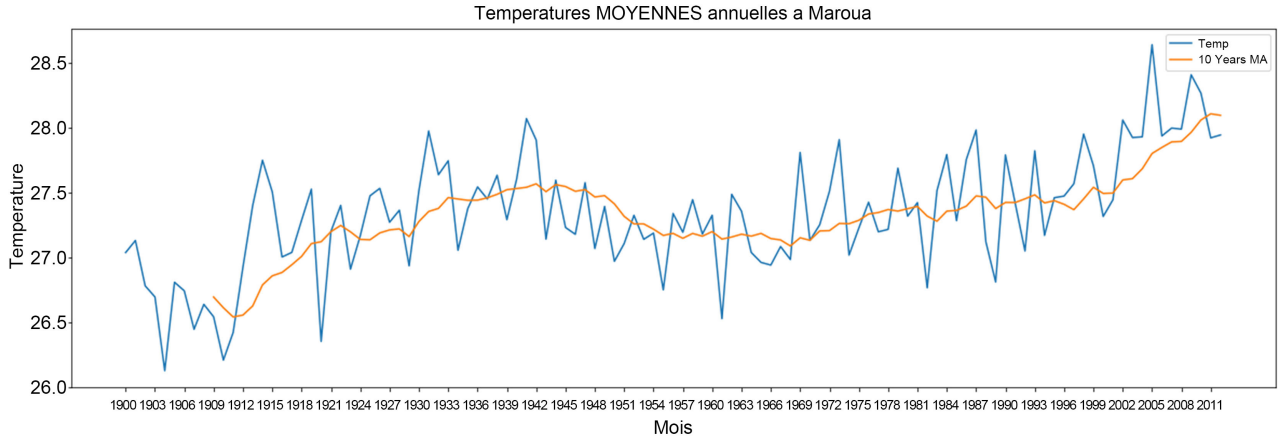

Figure 5. Annual variation of average temperature in Maroua, from 1900 to 2012.

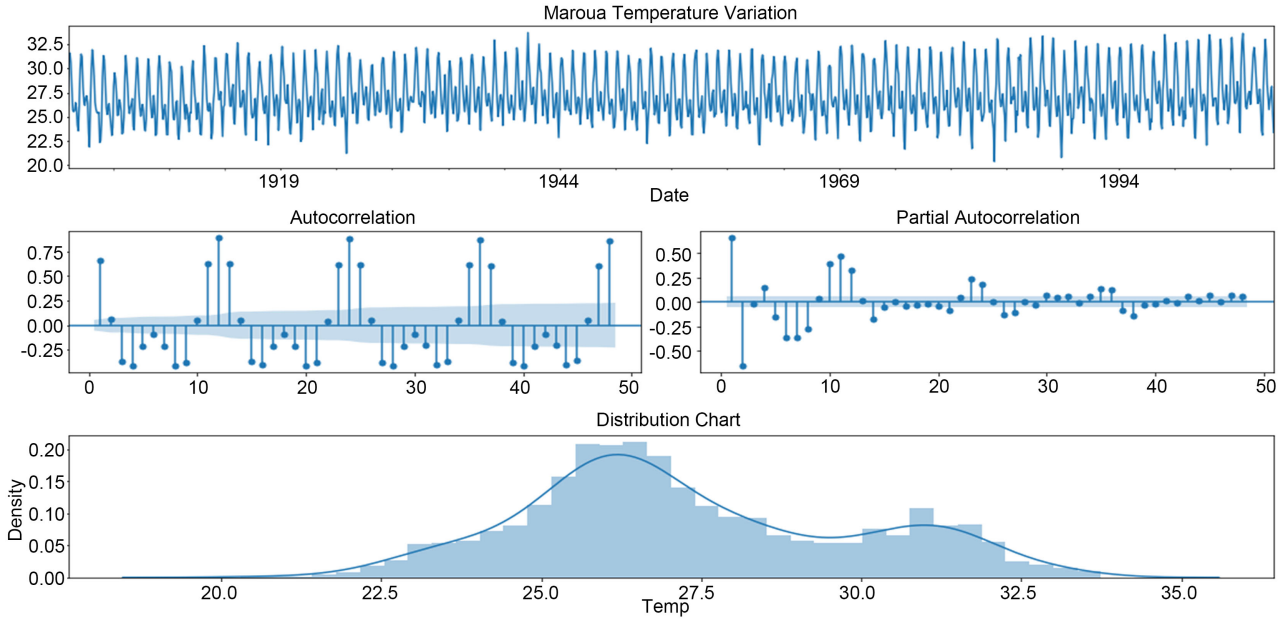

Figure 6. Autocorrelation, partial correlation and distributed graph.

The next month will use the previous month's forecast as a starting point. Negative autocorrelation starts at lag six and occurs once every 12 months. Different seasons have a role to play in this phenomenon.

Therefore, a strong positive autocorrelation is evident, starting at lag 12 and continuing for another 12 lags. The late intervals show a negative PACF. Initially, the PACF shows a positive jump and then declines to a negative PACF. The ACF and PACF diagrams are identical in this situation. An $AR(1)$ model and a first seasonal difference may depend on it ($YtYt12$). The parameters $SAR(P)$ or $SMA(Q)$ may be required, so plot the static function again with the first seasonal difference to confirm (see Figure 7).

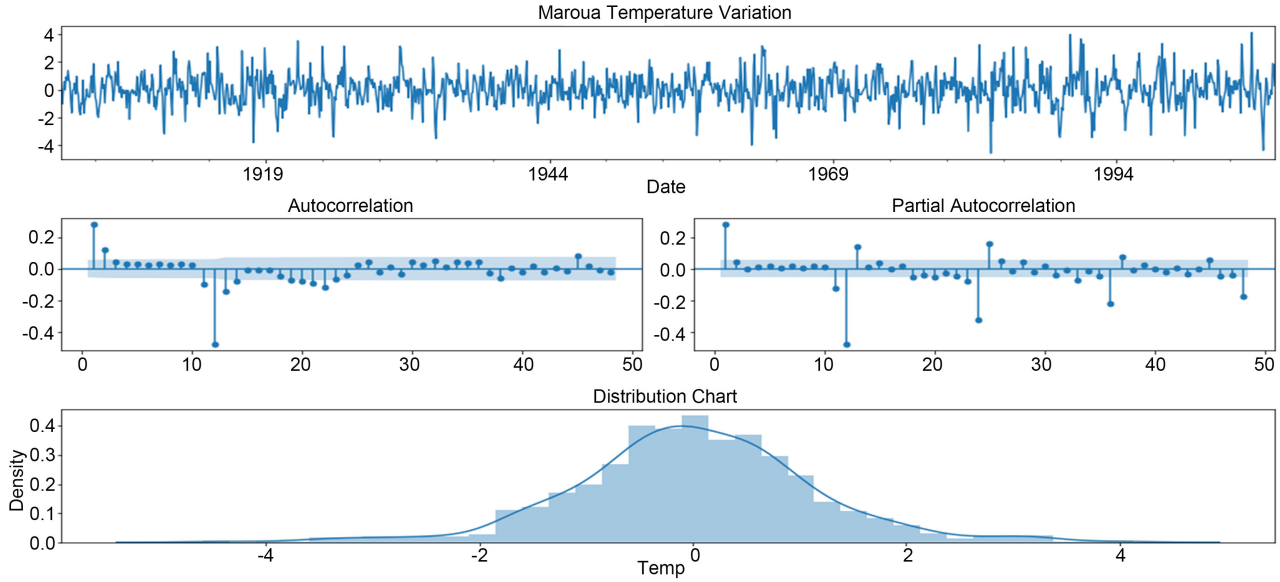

Figure 7. Autocorrelation, Partial Correlation and Climate Change Distributed Graph.

These are $AR(3)$ models, defined as having three parameters. The figures above show that the first ACF lags gradually decrease. The PACF falls below the confidence interval after the third interval. (**Figure 8**) shows the actual, predicted and error values. Both ACF and PACF showed significant decreases at the 12th interval. In other words, it is a $SAR(1)$ with a first difference as the signature SMA includes a parameter of 1 interval. Order 3 and 0 in the model; order 0 and 1 because the series has a clear increasing trend, and order 1 and 2 because the series has a clear decreasing trend. First, create a function that uses one-step predictions from the full validation set to determine how far it is.

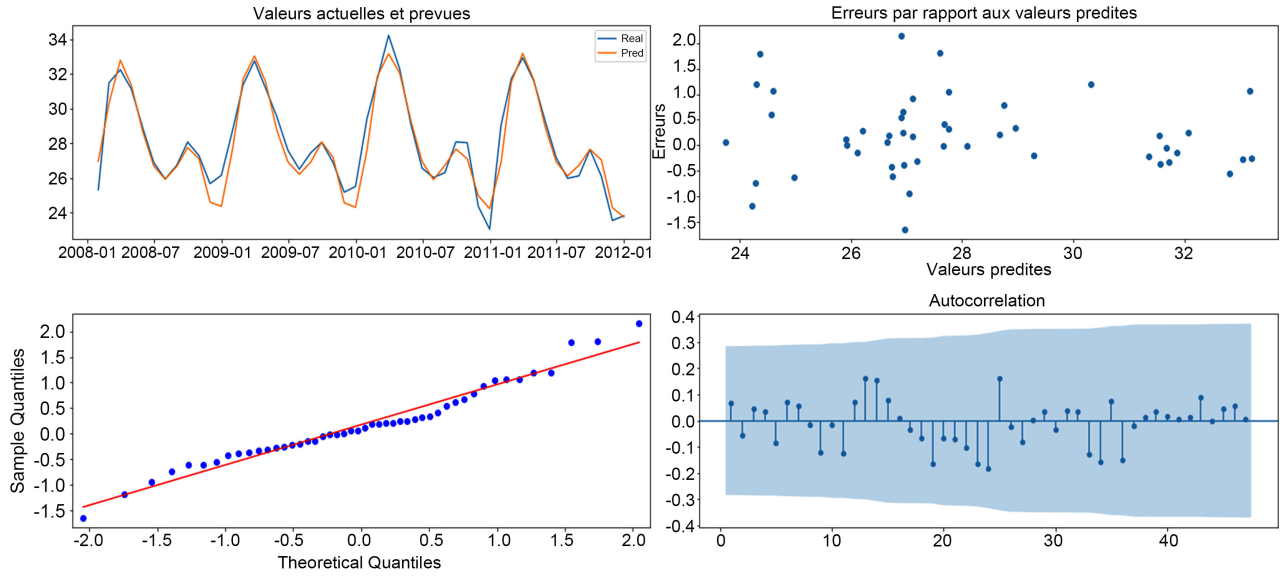

Figure 8. Predicted and actual error values.

The graphs below show historical data as well as projections for the future.

The scattered graph shows the difference between expected and actual values. An autocorrelation plotted graph of the residuals can be used to determine whether there is an association or not.

The errors increase from −1.5 to 1.5 as the temperature increases, as shown in (Figure 9), where the actual values are compared to the extrapolated ones. However, an autocorrelation plot shows a positive peak just above the 2-confidence interval for some of the outliers shown in the **QQ** plot. To assess the predictive accuracy of the test set, we have to consider the 12-month intervals.

3.1. Interannual Evolution of Temperatures (Max, Min and Avg) in Maroua from 1990-2022

Figure 10 shows the variability of maximum and minimum average temperatures in Maroua from 1990 to 2022.

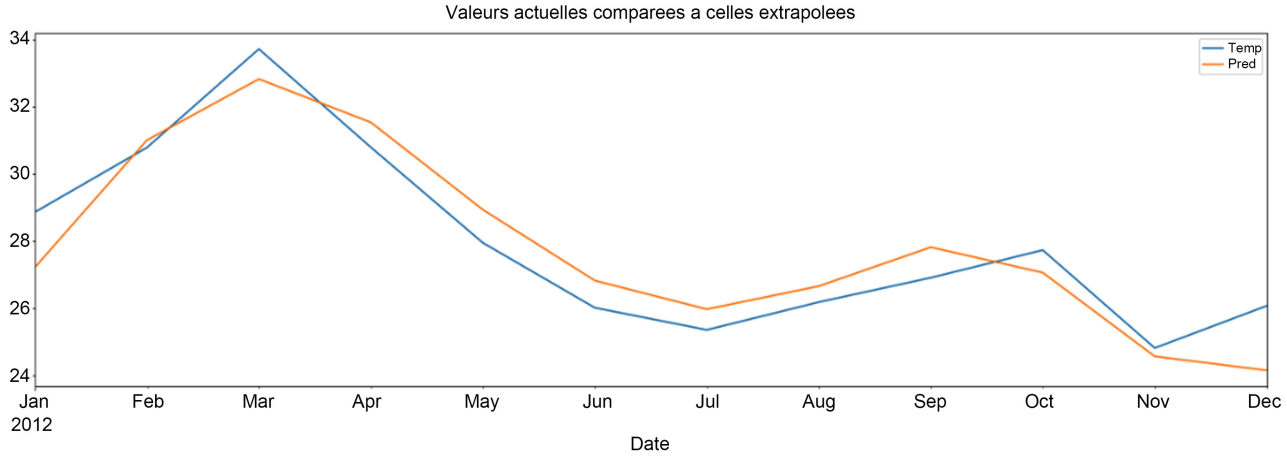

Figure 9. Actual values compared to extrapolated values.

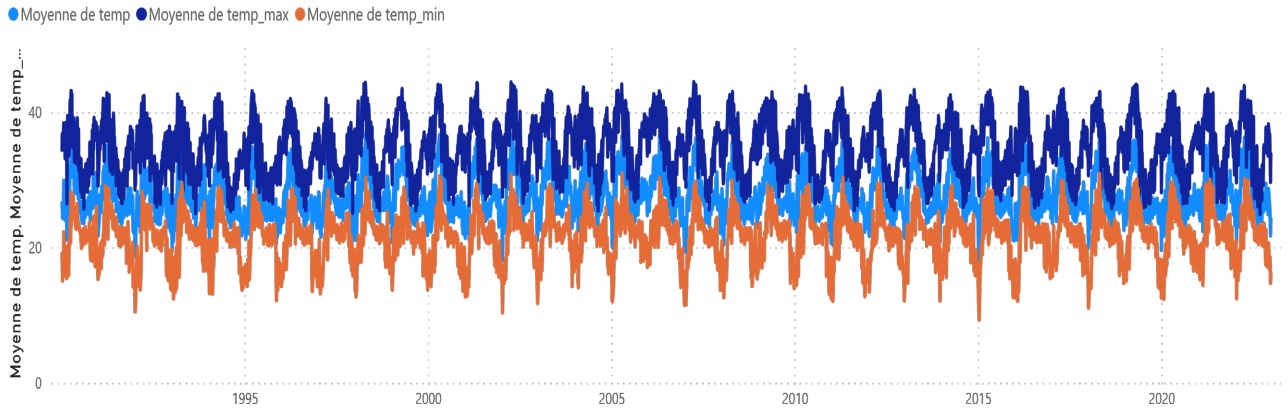

Figure 10. The variability of maximum and minimum average temperatures in Maroua from 1990 to 2022.

3.2. Interannual Rainfall Trends Maroua 1990-2022

Figure 11 shows the variations in rainfall in Maroua from 1990 to 2022. A peak of 115.10 mm is observed on Friday 15 July 2022.

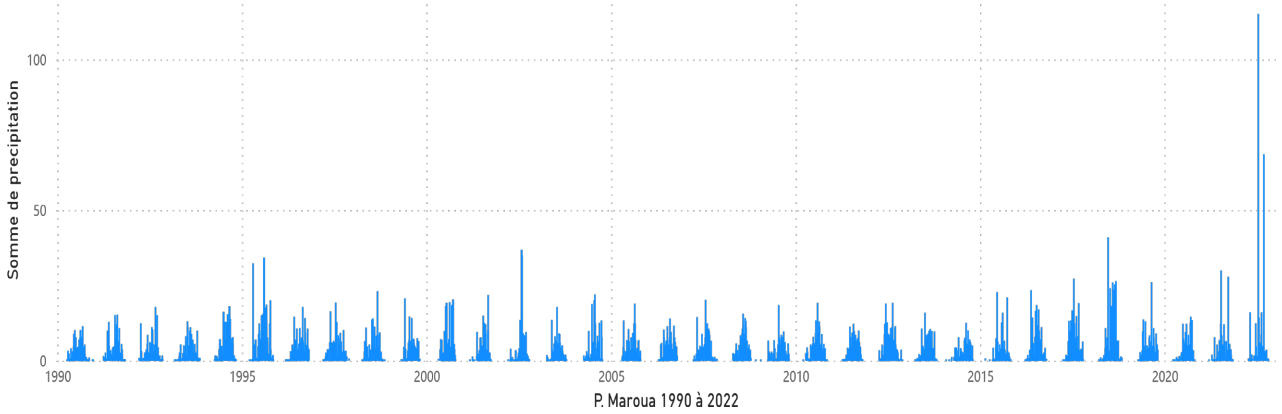

Figure 11. The variations in rainfall in Maroua from 1990 to 2022.

3.3. Temperature Projection in Maroua (January-February-March-April 2023)

The model's simulation of projected temperatures for the periods January, February, March and April 2023 shows a gradual increase in maximum (41.8˚C on 30 April 2023), average (38.7˚C on 20 April 2023) and minimum (36.0˚C on 22 April 2023) temperatures. (**Figure 12**)

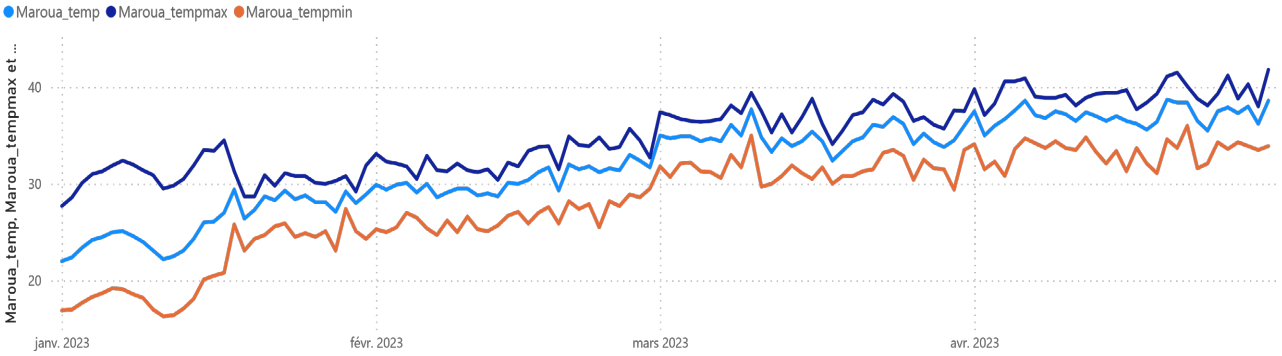

Figure 12. Temperature projection (max, avg and min) Maroua (January-February-March-April 2023).

3.4. Rainfall Projection Maroua (January-February-March-April 2023)

The graph obtained for the rainfall projection in Maroua shows that the area will have no rainfall for the months of January, February, March and April 2023. (**Figure 13**)

3.5. Interannual Evolution of Temperatures (Max, Min and Avg) in Garoua from 1990-2022

(**Figure 14**) shows the variability of maximum and minimum average temperatures in Garoua from 1990 to 2022.

3.6. Interannual Rainfall Trends in Garoua from 1990-2022

Figure 15 shows the rainfall variations in Garoua from 1990 to 2022. A peak of 105.36 mm was observed on Saturday the 8th of September 2018.

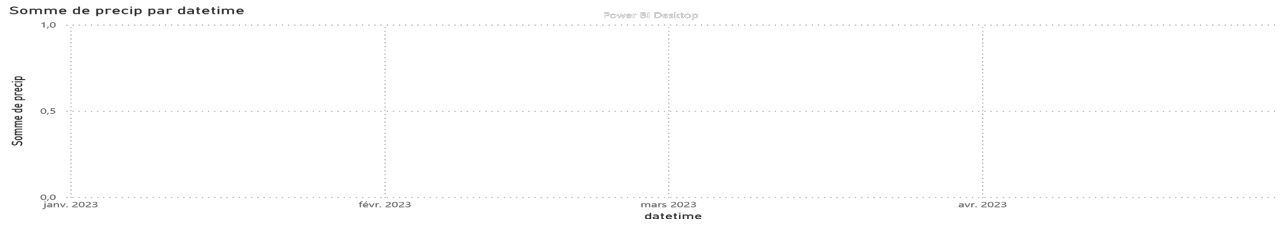

Figure 13. Rainfall projection for Maroua (January-February-March-April 2023).

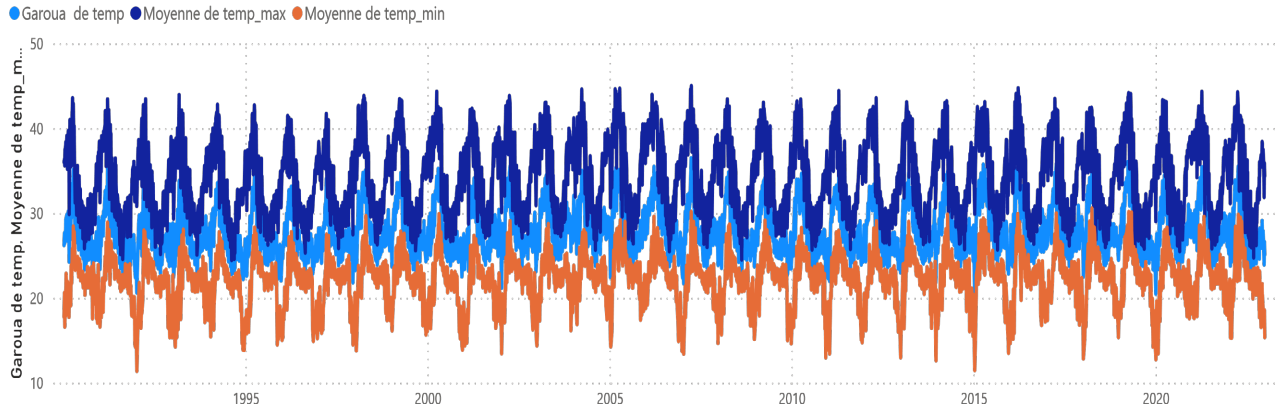

Figure 14. Temperature situation of Garoua from 1990-2022.

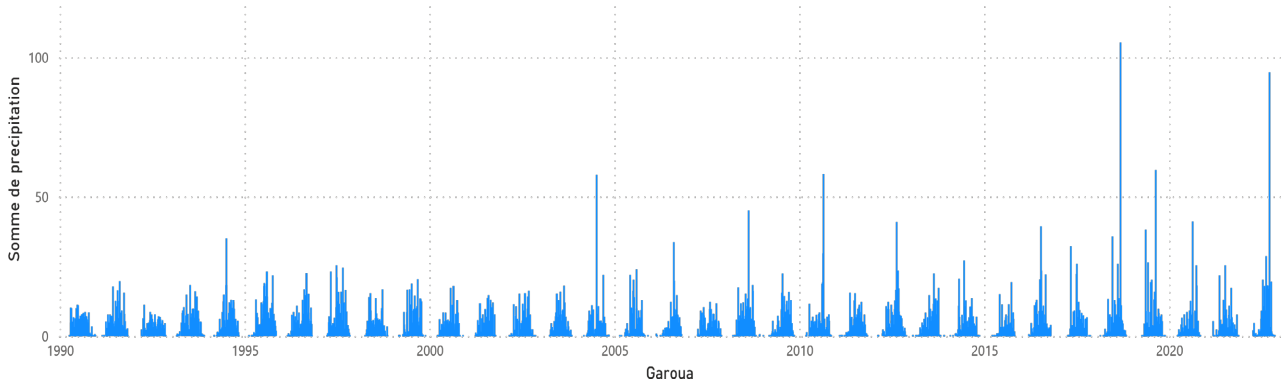

Figure 15. Rainfall situation in Garoua from 1990-2022.

3.7. Temperature Projection (Max, Avg and Min) in Garoua (January-February-March-April 2023)

The model's simulation of projected temperatures for the periods of January, February, March and April 2023 shows a gradual increase in maximum (41.10˚C on 7 April 2023), average (36.20˚C on 7 April 2023) and minimum (29.70˚C on 7 April 2023) temperatures. (**Figure 16**)

3.8. Rainfall Projection Maroua (January-February-March-April 2023)

The graph shows an onset of rainfall in Garoua from 8 March 2023. The graph also shows a peak of rainfall of only 13.40 mm on Thursday 27 April 2023. (**Figure 17**)

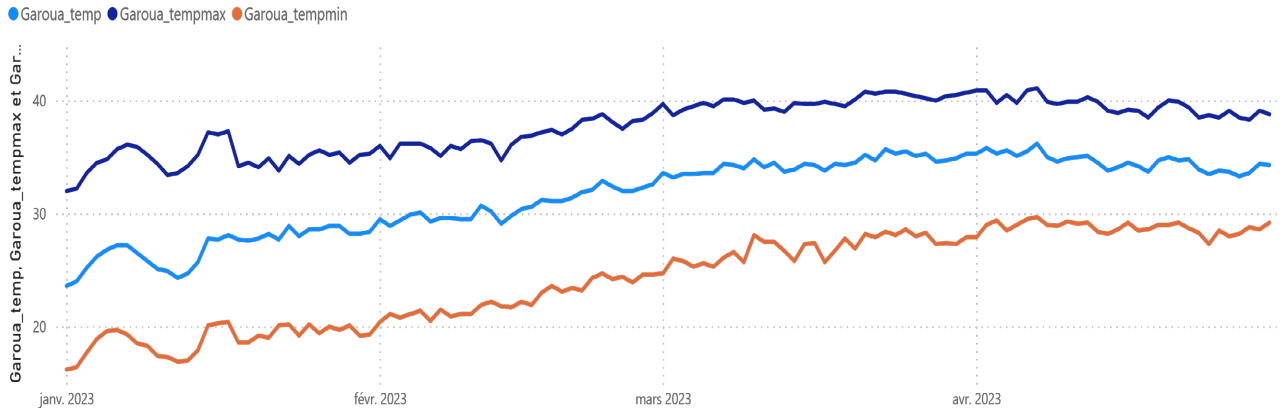

Figure 16. Temperature projection (max, avg and min) in Garoua (January-February-March-April 2023).

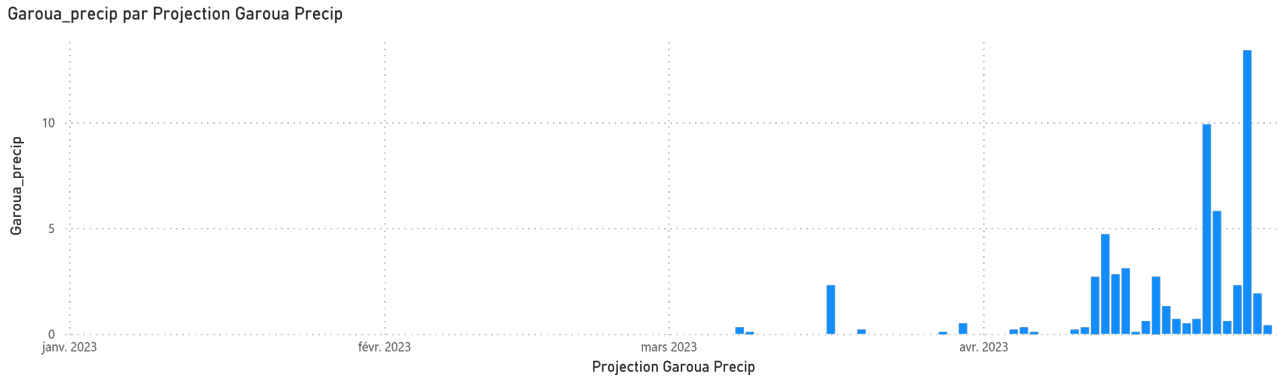

Figure 17. Rainfall projection for Maroua (January-February-March-April 2023).

3.9. Training and Validation of the SARIMA-LSTM Model

PARAMETERS			SSZ	
			Maroua	Garoua
STATISTICAL TESTS			−3.9117	−3.153
P-Value			0.0020	0.0018
Delays used			23.0000	21.0000
Number of observations used			1272.0000	1272.0000
Critical value (1%)			−3.4355	−3.2354
Critical value (5%)			−2.8638	−2.1235
Critical value (10%)			−2.5680	−2.6532
SARIMA model MSE (3, 0, 0), (0, 1, 1, 12)			0.7674	0.6842
Basic RMSE			1.94	1.96
Baseline RMSE for extrapolation			0.97	0.92
Scenarios	Observations	Training period	Train MSE Temp (C)/Precip (mm)	Validation MSE Temp (C)/Precip (mm)
LSTM	1272	2510.91 s	0.22	0.31
SARIMA-LSTM	1272	2493.50 s	0.15	0.19

4. Discussion

Digital forecasts have experienced considerable technological development over the last 30 years, thanks to the use of satellite imagery, which in turn has seen a remarkable explosion in the number of observations and their coverage, but also thanks to the use of digital data assimilation techniques. This progress has been made possible due to the opportunity to place atmospheric observing systems on satellite platforms and the development of digital computers, capable of handling complex non-linear equations. Fundamental knowledge of fluid dynamics and thermodynamics has also improved significantly during this century, putting atmospheric studies on a firm footing as an applied physics problem [13]. These different scientific techniques and approaches have been demonstrated in disciplines such as hydrology, climatology, oceanography, space weather, glaciology, atmospheric physics and chemistry, etc.

The awareness that humans can significantly (if unintentionally) change the global climate is said to be behind the increasing amount of research into monitoring and modelling atmospheric phenomena. Therefore, due to the very large size of the data, the complex and array nature of the system to be simulated, three major forecast steps (observation, data assimilation, model integration) constitute real scientific, numerical and technical challenges. Numerical climate prediction methods have been developed by researchers such as [14] and [15], etc. These authors carried out the disaggregation of the climate scale from global to local. The disaggregation approaches developed on dynamic models have made it possible to pool a multitude of factors. For example, based on the relevant variability regimes for characterizing the state of future climate, the analysis of identified factors, teleconnections and model outputs has enabled the description of future climate prospects.

The SARIMA-LSTM machine learning model, which has some advantages and limitations, fits into the range of local climate sensitivity modelling tools designed to predict the climate of Cameroon. In a context where the impacts of climate change and its effects vary considerably from one climatic zone to another. These impacts would have serious consequences for health, livelihoods and material assets, especially for the urban poor, informal settlements and other vulnerable groups.

The adjusted SARIMA-LSTM model uses disaggregation techniques in its principles, which allow the expression of model outputs at more precise scales. However, despite the quality and accuracy of the results, biases and inaccuracies related to the structure of the models remain, and could be improved by taking into account, practical applications of impact assessment models [13] for example. A major step in this adjustment process will be to incorporate the anthropogenic influences of climate change, in order to formulate strategies to mitigate the socio-economic impacts of changes on the local environment. The present work, which is an initial one to date, intends to extend its analysis methods to the five agro-ecological zones of Cameroon, in order to adjust the climate fore-

casts to local specificities and specific sectors of activity, at different time scales.

However, despite the quality of the results obtained, there are still biases and inaccuracies linked to the structure of the models. There is therefore a need to validate the results of the SARIMA-LSTM digital learning model by comparing the results of the "calculation-testing" approach.

5. Conclusion

A better understanding of complex physical phenomena is possible today thanks to high-performance computer tools, which can even compete with the range of tools of the experimental sciences. Mathematical and computer adjustments must contribute to the construction of efficient digital models, able to provide reliable predictions that accompany the development process. The SARIMA-LSTM digital model was used in this research with the major objective of examining the different factors that allow the development and improvement of a simple machine learning technique to determine the most reliable climate predictions possible. The SARIMA-LSTM digital model has the advantage of an easy formulation and a proven performance, which is not the case for other learning methods. In order to validate the model fitting algorithms, the applications were carried out on the SARIMA-LSTM digital model. The results of the model for the Sudano-Sahelian zone, both for temperatures (maximum, minimum and average) and precipitation, show that the predictions are very satisfactory and allow the application of this digital model to more advanced learning models and to other agro-ecological zones of Cameroon.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Azad, A.S., Sokkalingam, R., Daud, H., Adhikary, S.K., Khurshid, H., Mazlan, S.N.A., *et al.* (2022) Water Level Prediction through Hybrid SARIMA and ANN Models Based on Time Series Analysis: Red Hills Reservoir Case Study. *Sustainability*, **14**, Article 1843. https://doi.org/10.3390/su14031843

[2] Hourdin, F. (2021) Laboratoire de Meteorologie Dynamique, Institut Pierre Simon Laplace. Les principes de la modelisation du climat.

[3] Nie, H.Z., Liu, G.H., Liu, X.M. and Wang, Y. (2012) Hybrid of ARIMA and SVMs for Short-Term Load Forecasting. Elevier.

[4] Hourcade, J.C. (2020) Modelisation integree, evaluation des risques climatiques et des politiques de precaution. https://inis.iaea.org/collection/NCLCollectionStore/_Public/39/075/39075280.pdf

[5] Laurent, F. (2019) Outils de modelisation spatiale pour la gestion integree des ressources en eau: Application aux Schemas d'Amenagement et de Gestion des Eaux. https://core.ac.uk/download/pdf/39852327.pdf

[6] Cho, M., Kim, C., Jung, K. and Jung, H. (2022) Water Level Prediction Model Applying a Long Short-Term Memory (LSTM)-Gated Recurrent Unit (GRU) Method

for Flood Prediction. *Water*, **14**, Article 2221. https://doi.org/10.3390/w14142221

[7] Newbold, P. (1975) The Principles of the Box-Jenkins Approach. *Journal of the Operational Research Society*, **26**, 397-412. https://doi.org/10.1057/jors.1975.88

[8] ONACC (2020-2022) Bulletin des previsions des paramètres climatiques.

[9] Chen, P., Niu, A., Liu, D., Jiang, W. and Ma, B. (2018) Time Series Forecasting of Temperatures Using SARIMA: An Example from Nanjing. *IOP Conference Series: Materials Science and Engineering*, **394**, Article ID: 052024. https://doi.org/10.1088/1757-899x/394/5/052024

[10] PNUD (2016) Renforcer les mesures liees au climat afin de realiser les objectifs de developpement durable.

[11] Sylvie, M. (2018) Les modèles de prevision meteorologique. Encyclopedie de l'Environnement. Centre Europeen de Prevision Meteorologique à Moyen Terme (CEPMMT).

[12] Benestad, R.E. (2001) A Comparison between Two Empirical Downscaling Strategies. *International Journal of Climatology*, **21**, 1645-1668. https://doi.org/10.1002/joc.703

[13] Herrera, E., Ouarda, T.B.M.J. and Bobée, B. (2007) Méthodes de désagrégation appliquées aux Modèles du Climat Global Atmosphère-Océan (MCGAO). *Revue des sciences de l'eau*, **19**, 297-312. https://doi.org/10.7202/014417ar

[14] Tobias, C. and Vanpraet, C. (1980) Notes d'Ecologie Soudano-sahelienne: Quelques relations sols-vegetation dans le parc national de Waza (Nord Cameroun). Pedologue de l'ORSTOM.

[15] Amani, A. and Lebel, T. (1997) Lagrangian Kriging for the Estimation of Sahelian Rainfall at Small Time Steps. *Journal of Hydrology*, **192**, 125-157. https://doi.org/10.1016/s0022-1694(96)03104-6