Scientific Research Publishing

Analysis, Variability and Rainfall Prediction in Sub-Saharan Africa: The Case of the Lake Guiers in Senegal

Abdou Arame Fall[1,2], Saidou Ndao[1,2], Aba Diop[3]

[1]UFR Sciences et Technologies (SET), University Iba Der Thiam of Thies (UIDT), Thies, Senegal
[2]Laboratory of Water and Environment Sciences and Technologies (LaSTEE), Polytechnic School of Thies (EPT), Thies, Senegal
[3]Statistics and Random Models Research Team (ERESMA), University Alione Diop (UAD), Bambey, Senegal
Email: ed2ds@univ-thies.sn, aarame.fall1@univ-thies.sn

Abstract

The aim of this article is to predict the rainfall evolution of a sub-Saharan area in which one of the most important freshwater resources is located: Lake Guiers. Characterized by short seasonal rains of three months, it experienced a long period of drought in the 1970s. We begin by analyzing the temporal distribution of the rainfall including the variability of the data, with a view to predicting a possible return. For this reason, we present here univariate modeling results of rainfall series collected on three stations in the area. The challenge lies in the adequacy of the parameters for the monthly rainfall series, which generates more or less significant forecast errors on the learning bases because of the missing data. This later motivated their conversion to moving average series. On the other hand, the normality of the latter seems to be rejected by the D'Agostino test. Student's and Mann-Whitney's tests confirmed the homogeneity. The autocorlograms show the presence of autoregressive terms in the data. Dickey-Fuller and Mann-Kendall tests reveal both trend and seasonality. The stationarity tests of Dickey-Fuller, Phillips-Perron and KPSS have shown that they are non-stationary. As a result, we did an ARIMA modeling method using the Box-Jenkins [1] method with the R software, which involves estimating model parameters, tests of significance, analysis of residualss, selection according to information criteria and forecasts. The results obtained during the learning-test phase showed a quasi-similarity of the base-tests in all the series except for that of Louga.

Keywords

Time Series, Rain, Trend, Stationarity, Residuals, Model, Forecast

1. Introduction

Statistical analyzes of hydrological time series are generally based on a set of fundamental assumptions: homogeneity, stationarity without overall trends, periodicity and persistence. Homogeneity refers to the belonging of the data in the series to a same population and therefore has a time-invariant mean, in the absence of a change in the method of collecting the data [2]. Stationarity implies that the statistical parameters of the series, calculated at different times, are constant [3]. Trends in hydrological series are introduced by natural changes. Periodicity is usually due to astronomical cycles such as the earth revolution [4].

Time-series studies were conducted to analyze rainfall data [3] [4], particularly in this study area by [5]. Originally, the application of time series analysis was limited to surface water problems, in particular the analysis of extremes (floods and droughts). However, the field has expanded to include groundwater problems [6]. Others were made on flow rate data [7] [8] [9] and surface water quality data [10] [11].

The aim is to detect trends, supplement missing data and develop statistical models that quantitatively describe all the processes underlying a given sequence of observations in order to predict hydrological events on the basis of these data. For example, a design of stormwater collection structures, such as dikes, dams, channels, require measured records of peak runoff, which can occur in shorter periods than the expected return. Thus this field has become a powerful tool for efficient planning and management of water resources.

As a first step, we will conduct a study on homogeneity, trend, stationarity and persistence to determine the lag order for each series by hypothesis testing. Then, once selected d and D, we will return to the identification of a SARIMA. We will push this modeling of the series with and without differentiation, and make a comparison of the qualities of each model according to the information criteria or the predicted values compared on the test bases. The selected models will be presented after a residuals analysis before moving on to an interpretation of the forecasts.

2. Methodology and Data

2.1. Methodology

Consider the series of rain $\{x_t, t \in T\} \subset \mathbb{R}^n$, $T \subset \mathbb{Z}$ as a realization of the possible set value of the monthly rainfall height $\{X_t, t \in T\}$, here taken over 732 months. The SARIMA (Sesonal Autoregressive Integrated Moving Average) model take into account the effect of seasonality and non-stationarity.

$$\phi(B)\Phi(B^S)\Delta^d \Delta_S^D X_t = \theta(B)\Theta(B)\varepsilon_t \tag{1}$$

B, Δ are the "backward shift" and "Differences" operator such as $BX_t = X_{t-1}$, $\Delta X_t = X_t - X_{t-1}$, $\Delta_S = 1 - B^S$, $(\varepsilon_t)_{t \in \mathbb{Z}}$ is a low white noise $BB_f(0, \sigma^2)$ corresponding to unpredictable phenomena causing a disturbance of the series and including an independent residual component $(\eta_t)_{t \in \mathbb{Z}}$ [12]. $\phi(B)$, $\Phi(B)$,

$\theta(B)$ et $\Theta(B)$ being polynomials defining the lags order p of the autoregressive part, P with seasonality, q the moving average part and Q with seasonality. The exponents d, D are the orders of seasonal differentiation and S, the period (or seasonality) of the process. Any time series data can be decomposed into trend, saisonic and residual components.

$$X_t = C_T + S_t + \varepsilon_t \tag{2}$$

A process is said to be stationary in the strict sense when the law of joint probability is invariant for any translation of time [3]. Several tests make it possible to detect it,

- Augmented Dickey-Fuller (ADF) test

$$\begin{cases} H_0 : \phi = 1 \Leftrightarrow \text{unity root existence : the series is not stationary} (d \geq 1) \\ H_1 : \phi \neq 1 \Leftrightarrow \text{no existence of roots unity} (\phi < 1) \end{cases}$$

- Phillips-Perron (pp)/Kwiostowski & al. (KPSS) test

$$\begin{cases} H_0 : d = 1 \ (\text{the series is not stationary}) \\ H_1 : d = 0 \ (\text{the series is stationary}) \end{cases}$$

A non-stationary $(X_t)_{t \in \mathbb{Z}}$ process follows a model SARIMA$(p,d,q)(P,D,Q)[S]$ orders p,d,q, $P,D,Q,S \in \mathbb{N}$, if the process $(\Delta^d \Delta_S^D X_t)_{t \in \mathbb{Z}}$ is SARMA$(p,q)(P,Q)[S]$ and becomes a stationary process.

The partial autocorrelation function $\eta(.)$ takes into account the strength of the bond between X_t et X_{t+h} by removing links induced by intermediate values $X_{t-(h-1)}, \cdots, X_{t-1}$. We note by $X_t^* = \sum_{k=1}^{h-1} \alpha_k X_{t-k}$:

$$\eta(h) = Corr(X_t, X_{t-h} \mid X_{t-1}, \cdots, X_{t-h+1}) = Corr(X_t - X_t^*, X_{t-h} - X_{t-h}^*) \tag{3}$$

The ordinary least squares estimator of σ for $K = p + q + P + Q$ constraints:

$$\hat{\sigma}_{mco}^2 = \frac{1}{n-K} S(\hat{\beta}) \tag{4}$$

The model parameters $\hat{\beta} = (\hat{\phi}, \hat{\Phi}, \hat{\theta}, \hat{\Theta}) \in \mathcal{M}_{1 \times p} \times \mathcal{M}_{1 \times q} \times \mathcal{M}_{1 \times P} \times \mathcal{M}_{1 \times Q}$, are estimated by minimizing the log mean square error and the sum of residual squares (7).

$$\hat{\beta} = \underset{\beta \in \mathbb{R}^4}{\arg \min} \left(\frac{1}{n} \sum_{t=1}^n \log(R_t(\beta)) + \log\left(\frac{1}{n} S(\beta)\right) \right) \tag{5}$$

To test the significance of the parameters, the hypothesis H_0 is assumed that each parameter ϕ_i, $\Phi_{i'}$, θ_j ou $\Theta_{j'}$ is not zero, for $i = 1, \cdots, p$, $i' = 1, \cdots, P$, $j = 1, \cdots, q$, and $j' = 1, \cdots, Q$ and test statistic:

$$T_j = \frac{\hat{\beta}_k}{\hat{\sigma}_{\beta_j}} \xrightarrow[n \to \infty]{\mathcal{L}} t_{n-K} \tag{6}$$

H_0 is accepted at threshold $\alpha \in [0;1]$ if $\left| T_j^{obs} \right| > q_{1-\alpha/2}^{t_{n-K}}$, where T_j^{obs} is the observed value.

Consider $X_t \mid X_{t-1}, \cdots, X_{t-p} \sim \mathcal{N}\left(\hat{X}_t(\beta_0); \hat{\sigma}_0^2(\beta_0)\right)$ and $\hat{\varepsilon}_t = X_t - \hat{X}_t(\beta_0)$ the forecast error where $\hat{X}_t$ is the value of the estimated rainfall height at time

t with the parameters $\beta_0 = (\phi_0, \Phi_0, \theta_0, \Theta_0)$. We note by $R_t(\beta_0)$, the mean square error or forecast error and $S(\beta)$, the sum of the residual squares. The best model is the one that returns the low value of the AIC, BIC information criteria and standard criteria such as MAE, RMSE or MAPE errors and having the most significant parameters.

$$\left(R_t(\beta_0), S(\beta_0)\right) = \left(\frac{1}{\sigma_0}\sum_{t=1}^{n}\left(X_t - \hat{X}_t(\beta_0)\right)^2, \sum_{t=1}^{n}\frac{\left(X_t - \hat{X}_t(\beta_0)\right)^2}{R_t(\beta_0)}\right) \tag{7}$$

Information criteria and standards:

$$\text{AIC} = \log\left(\hat{\sigma}_\varepsilon^2\right) + \frac{2K}{n} \quad \left(\text{Akaike Information Criterion}\right) \tag{8}$$

$$\text{SIC ou BIC ou SBC} = \log\left(\hat{\sigma}_\varepsilon^2\right) + K\frac{\log(n)}{n} \quad \left(\text{Swartchtz Bayesian Criterion}\right) \tag{9}$$

$$\left(\text{ME}, \text{MAE}, \text{MAPE}, \text{RMSE}\right) = \left(\frac{1}{n}\sum_{t=1}^{n}\hat{\varepsilon}_t, \frac{1}{n}\sum_{t=1}^{n}|\hat{\varepsilon}_t|, \frac{100}{n}\sum_{t=1}^{n}\left|\frac{\hat{\varepsilon}_t}{x_t}\right|, \sqrt{\frac{1}{n}\sum_{t=1}^{n}\hat{\varepsilon}_t^2}\right) \tag{10}$$

Lower the value of these criteria, better the model. To study the nullity of the mean of the residualss, we use the classic tests of student (or conformity). Ljung-Box tests can identify processes of white noise. The test statistic being:

$$Q = K\sum_{k=1}^{n-k}\left[\hat{\rho}(k)\right]^2 \xrightarrow[n\to\infty]{\mathcal{L}} \chi_K^2 \tag{11}$$

where $\hat{\rho}(k) = \dfrac{\sum_{t=1}^{n-k}\hat{\varepsilon}_t\hat{\varepsilon}_{t-k}}{\sum_{t=1}^{n}\hat{\varepsilon}_t^2}$. H_0 is accepted at threshold $\alpha \in [0;1]$ if

$\left|Q^{obs}\right| < q_{1-\alpha}^{\chi_K^2}$ where Q^{obs} is the observed value.

The forecast made at time t for the $t+k$ date (k = forecast horizon) is defined by:

$$\hat{X}_{t+k} = \mathbb{E}\left[X_{t+k} \mid I_t\right] \tag{12}$$

where $I_t = \sigma\left(X_1, \cdots, X_t, \varepsilon_1, \cdots, \varepsilon_t\right)$ is a filtration. This predictor represents the best prediction $\left(X_t\right)_{t\in\mathbb{Z}}$ conditionally at the time of the information available at date t, I_t.

2.2. Data

The study focuses on monthly rainfall data over six decades of the three regions that include Lake Guiers. They are collected from the National Agency of Civil Aviation and Meteorology (ANACIM), Dakar (Senegal).

3. Results and Discussions

3.1. Descriptions

We have a series of data on monthly rainfall in Saint Louis, Podor and Louga, three regions in which Lake Guiers is located (16°15'N; 15°48'W), from 1960 to 2020. This zone located in the Sahel, on a strip of 10 to 25 km wide, is characte-

rized by an extreme fragility of ecological balance (Kane A., 2005) [13] due to the alternation of a long dry season (from November to June) and a short wet season (from July to October) [14].

"**Figure 1**" shows the monthly precipitation in each region and shows that there is no trend. The discontinuous lines in blue represent the monthly averages.

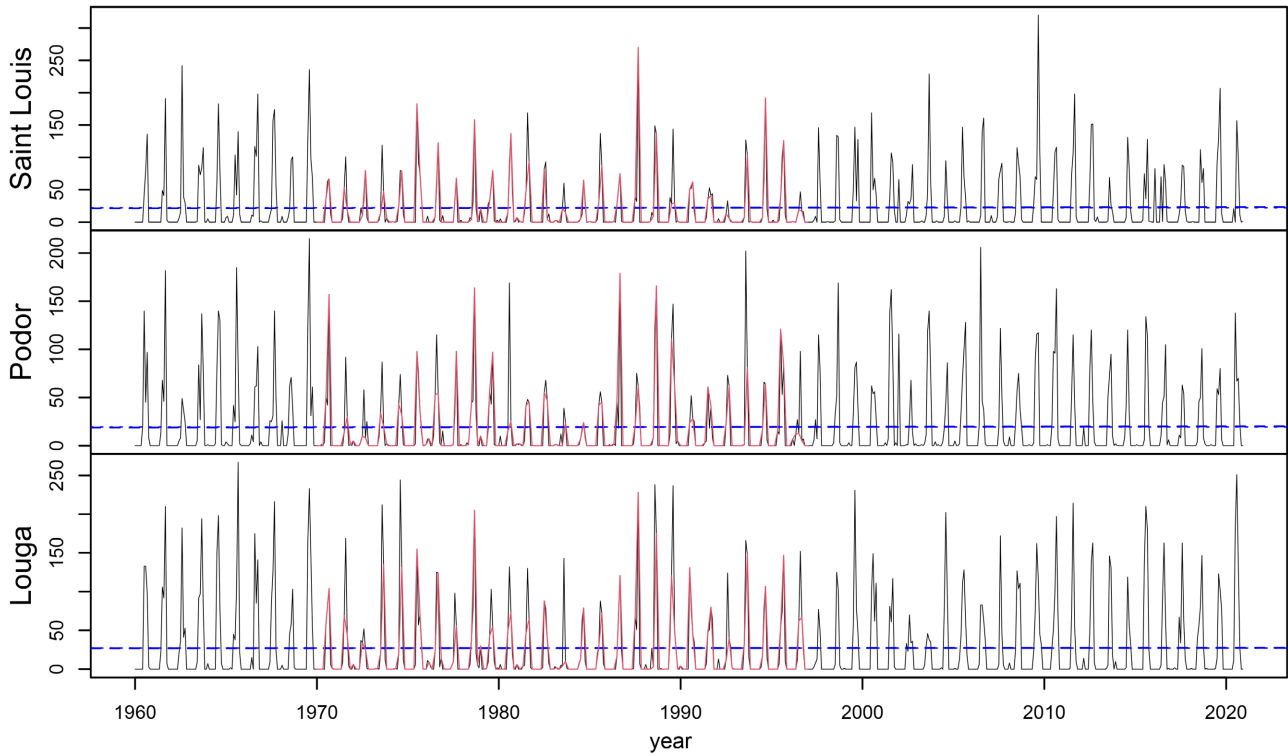The red zones (**Figure 2**) showing changes in the precipitation index correspond

Figure 1. Graphic representation of rainfall data for Saint Louis - Podor - Louga from 1960 to 2020.
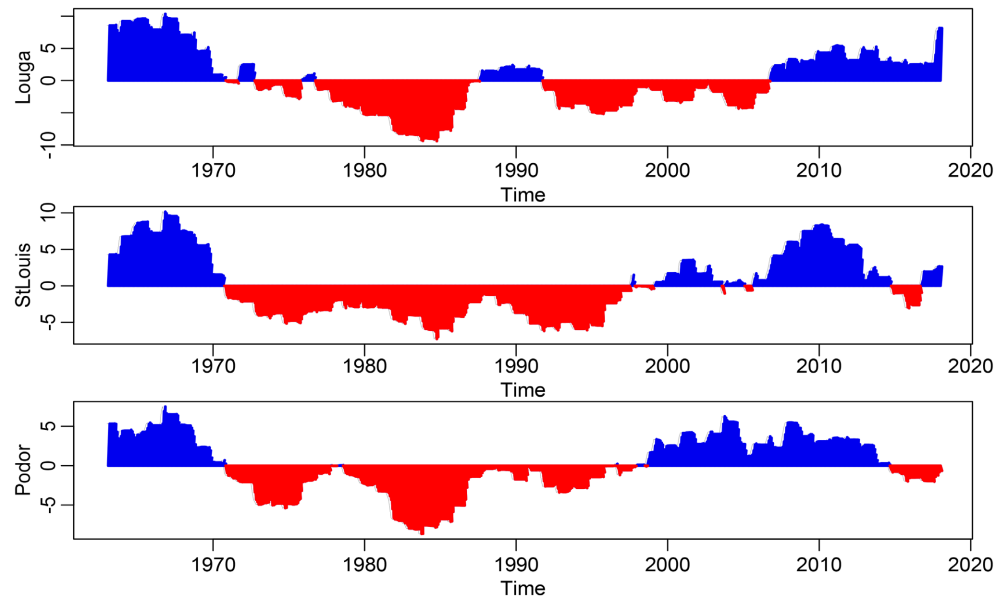

Figure 2. Standardized precipitation index (spi) for Saint Louis - Podor - Louga from 1960 to 2020.

to dry periods in the region. In [15], successive environmental transformations naturally linked to more than two decades of drought (1970-1997).

3.2. Homogeneity

The simple randomness of a chronological sequence is due to the fact that all the observed values come from the same population by random and independent draws [16]. These tests aim at the constancy of the dispersion of the series, that is to say, they study whether the variability of the series is uniform over time and assume that "the series is random" [17]. The parametric test of Student makes it possible to see if the samples share the same mean value while the non-parametric Mann-Whitney [18] one makes it possible to evaluate if the ranks of the values are 2 to 2 homogeneously distributed in the 3 distributions.

With the Student tests, the p-values are <0.05 except for Podor and Saint Louis, where the hypothesis of proximity of rainfall averages is rejected. This may be due to high rainfall variability and its spatial distribution of measuring stations near the desert. As for the p-value obetunes with the Mann-Whitney test are >0.05, we accept the hypothese according to which, the distributions are significantly homogeneous (Table 1).

3.3. Persistence

Persistence or autocorrelation, it represents the short-term dependence of observations. This violates the assumption of independence of observations $(x_t)_t$ or errors $(\varepsilon_t)_t$. Thus, its presence may lead to the rejection of the stationarity hypothesis, that is to say, it may indicate the existence of a trend that, while in reality, there is no [3]. Problems induced by the presence of persistence or seasonality have a related effect.

The autocorrelation function is used to determine the lags order of the terms (Figure 3). In the following, we will consider only the 6-year moving average series of the monthly series. This value, centered of order k at time t being given by: $mm_{k,t} = mm_k(x_t) = \frac{1}{k}\sum_{i=-k}^{k} x_{t-i}$. We get 5 moving series of sizes 660 each.

Table 1. Comparison of Student and Mann-Whitney homogeneity tests.

Statistical tests	(Louga & Podor)	(Louga & Saint Louis)	(Podor & Saint Louis)
mean	(27.32240; 19.41667)	(27.32240; 22.29918)	(19.41667; 22.29918)
Student, H₀: the difference between the 2 means is significant			
test statistic t	3.331	1.9763	1.3145
p-value	0.0009	0.0483	0.1889
Mann-Whitney, H₀: the rank of the measures are distributed homogeneously in the 2 distributions			
test statistic W	277412	277527	267283
p-value	0.1925	0.1856	0.9309

The choice of this period is to remove missing data from the dry season that lasts 9 months each year. This is worth a 3-year loss of data on both sides of the tails.

These diagrams (Figure 4) show the presence of autoregressive terms with lags of order 8.

3.4. Trend

The trend represents the general upward or downward trend over a fairly long period of time.

In "Figure 5", the 6-year moving average rainfall graphs match the trend components from the decomposition of the raw series, with the maximal and
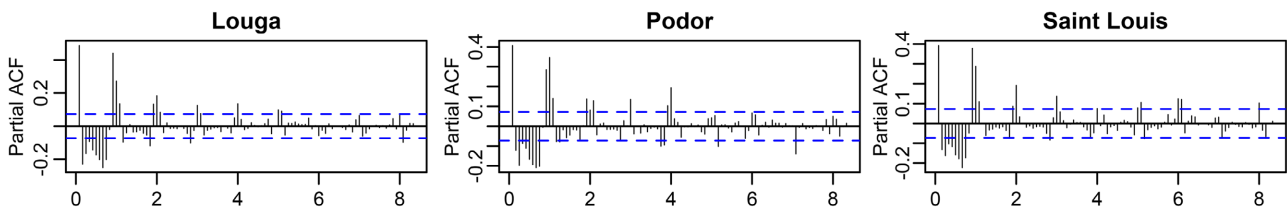

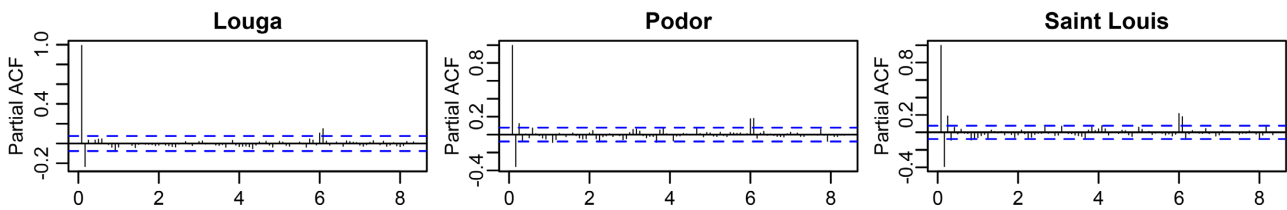**Figure 3.** Partial autocorrelogram of monthly rainfall series.

Figure 4. Partial autocorrelograms of 6-year moving average rain series.
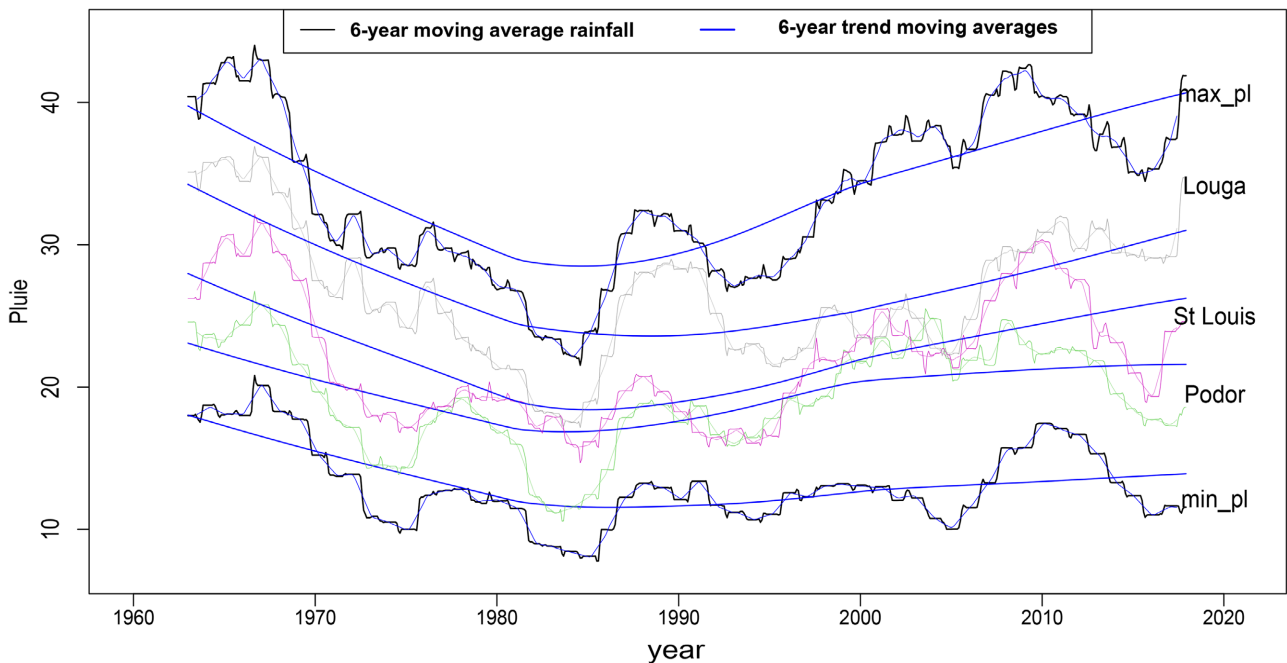

Figure 5. Comparison of trend and 6-year moving average rainfall trends from 1960 to 2020.

minimum rainfall explained respectively to (40.46; 25.49; 34.05)% et (19.49; 47.15; 33.36)% of monthly rainfall respectively on Louga, Podor and Saint Louis.

The nonparametric Mann-Kendall test shows the presence of a trend and determines whether the series includes a seasonal component. The values of the phi3 test statistics are significantly below the critical value 6.25 with the Dickey-Fuller test (Table 2), so the assumption that each series has a unity root and no trend is accepted.

Moreover, the results from the Mann-Kendall test reveal the presence of both trend and seasonality in all series except that of Louga. This can be due to a high spatio-ternporel variability that characterizes this area which has undergone flows of trade winds and monsoons according to the position of the FIT (Intertropical Front) and the note of a decrease in rainfall according to a South-North and West-East gradient [14].

3.5. Stationarity

Three parametric tests (Dickey-Fuller [12], Phillips-Perronet [12] and Kwiatkowski & al. [19]) were used for trend detection.

The p-values in the first two tests are >0.05, so the assumption of non-stationarity H_0 is accepted. This is confirmed by the third test (KPSS), which, for a p-value < 0.05, accepts H_0 as a stationarity hypothesis (Table 3).

The choice of significance level of 5% is arbitrary, but appears to be a reasonable limit for the type of data under study [11]. Thus, given the results of the three tests, it can be concluded that the moving average series are not stationary.

Table 2. Results of three stationarity tests, taking into account the presence of trends and seasonality.

Statistical tests	Louga	Podor	Saint Louis	min rain	max rain
(*Trend*) Mann-Kendall, H_0: *there is no trend*					
test statistic τ	0.009	0.126	0.118	−0.077	0.184
p-value	0.715	0.000	0.000	0.003	0.000
(*ur.df*) Dickey-Fuller/type = "*trend*", H_0: *presence of trend*					
z.diff.lag	9	9	12	8	13
df	638	638	632	640	630
RSE	0.29	0.24	0.28	0.18	0.30
test statistic phi3	2.27	1.62	2.08	1.42	2.48
5% critical value: 6.25					
p-value < 2.2e−16					
(*Seasonality*) Mann-Kendall, H_0: *there is no presence of seasonality*					
test statistic τ	0.011	0.128	0.121	−0.076	0.189
p-value	0.695	0.000	0.000	0.004	0.000

3.6. Normal Moving Average Rainfall Reries

The data can be graphically assimilated as sequences whose evolution is normal, even if the normality test of D'Agostino seems to reject this hypothesis (Figure 6).

3.7. Modeling

With the R software, the commende auto-arima returns the best ARIMA model according to the value of the information criteria.

The values obtained from σ^2 successively for the monthly rain series (Louga, Podor, Saint Louis, rain min, rain max) are 1121; 914; 1144; 413; 1294 as well as for moving average series: 0.0905; 0.0602; 0.0879; 0.0339; 0.1016.

"Table 4" gives the list of the best models obtained with the lowest values of AIC, BIC and criteria of eurreur with the pourtcentage of significant parameters in each model in the 4th column.

Table 3. Comparison of three stationarity tests on monthly series.

Statistical tests	Louga	Podor	Saint Louis	min rain	max rain
(*adf*) *Augmented Dickey-Fuller, H₀: the series is non-stationary*					
test statistic	−1.80	−1.80	−1.61	−1.87	−1.50
ordre du retard = 8					
p-value	0.84	0.66	0.74	0.63	0.79
(*pp*) *Phillips-Perron, H₀: the series is non-stationary*					
test statistic	−4.45	−6.51	−4.88	−6.23	−4.17
TLP = 6					
p-value	0.86	0.74	0.84	0.76	0.88
(*kpss*) *KPSS Test for Level Stationarity, H₀: the series is stationary with a deterministic tendency*					
KPSS Level	1.49	1.17	1.46	1.24	1.63
TLP = 6					
p-value = 0.01					

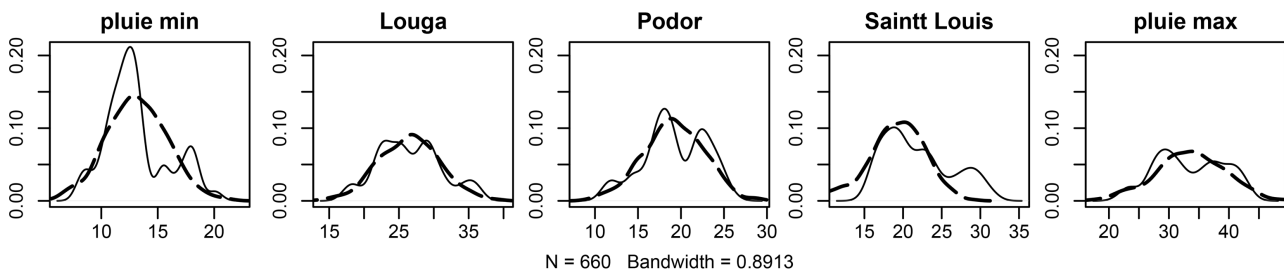Notes: df = degree of freedom; TLP = Truncation lag parameter.

Figure 6. Graphic comparison of the density of the moving average series (continuous lines) and the normal distribution (dotted lines).

3.8. Stochastic Analysis

The hypothesis of nullity of the residualss is accepted by the Student test as well as that of the portmenteau test (Box-Pierce or Box-Ljung), because they follow a white noise, since the p-values are >0.05 (Table 5).

Table 4. Selected models of monthly rainfall series and 6-year moving average.

Models	(p, d, q) (P, D, Q)	σ^2	signif. par.	AIC	BIC	ME	MAE	RMSE
ARIMA (p, q, q) (P, D, Q) [12] *monthly rainfall series*								
Louga	(0, 1, 2) (2, 1, 3)	822.8	6/7	6919.39	6956.01	0.8	15.87	28.28
Podor	(0, 2, 2) (3, 1, 1)	578.6	4/6	6691.12	6723.15	−0.11	13.9	23.72
Saint Louis	(0, 1, 1) (2, 1, 2)	855.3	3/5	6958.13	6985.59	0.57	15.16	28.88
min rain	(3, 0, 1) (2, 1, 0)	413.6	4/6	6396.18	6428.23	−0.12	8.48	20.11
max rain	(4, 1, 0) (5, 1, 4)	1047.0	10/13	7131.79	7195.88	0.00	20.22	31.77
ARIMA (p, d, q) (P, D, Q) [12] *6-year moving average rainfall series*								
Louga	(8, 1, 0) (6, 0, 0)	0.065	11/14	115.6	182.96	−0.003	0.16	0.25
Podor	(8, 2, 0) (6, 0, 1)	0.044	10/15	−156.06	−84.21	−0.005	0.13	0.21
Saint Louis	(8, 1, 0) (6, 1, 1)	0.070	11/15	186.16	257.72	0.009	0.16	0.26
min rain	(7, 1, 0) (4, 0, 4)	0.025	12/15	−476.20	−404.35	−0.012	0.093	0.15
max rain	(7, 0, 3) (1, 0, 2)	0.078	9/13	232.24	295.14	−0.002	0.158	0.28

Note: signif. par. = number of significant parameters.

Table 5. Model residuals on the moving average series.

Statistical Tests	Louga	Podor	Saint Louis	min rain	max rain
Student, H_0: the mean of residuals is zero, df = 659					
mean	−0.003	−0.005	0.009	−0.011	−0.002
conf. int.	[−0.02; 0.02]	[−0.02; 0.01]	[−0.01; 0.03]	[−0.03; 0.002]	[−0.02; 0.02]
test statistic	−0.326	−0.659	0.927	−1.734	t = −0.208
p-value	0.744	0.510	0.3543	0.083	0.835
Box-Pierce/Box-Ljung, H_0: Residuals follow a white noise, df = 1					
test statistic	0.083	0.131	0.049	0.058	0.00032
p-value	0.772	0.717	0.824	0.809	0.9855
D'Agostino, H_0:the Residuals are Gaussian					
Omnibus	83.0	146.21	113.9	141.1	101.60
Skewness ($\kappa = 3$)	3.1	6.4	5.4	1.8	−4.7
Kurtosis ($s = 0$)	8.6	10.2	9.2	11.7	8.9
p-value	(0; 0.001; 0)	(0; 0; 0)	(0; 0; 0)	(0; 0.07; 0)	(0; 0; 0)

Notes: df = degree of freedom, conf. int. = confidence interval.

Gaussity tests appear to reject the H_0 hypothesis because of the high Kurtosis (thick-tailed distribution), as shown in this example of residuals density from the ARIMA (8, 1, 0) (6, 1, 1) [12] the moving average series of the Saint Louis station.

The Quantile-Quantile plot allows to compare two distributions that are considered similar. In abscissa, we have the quantiles of the theoretical distribution of the normorle law adjusted to the mean and slope of the theoretical sample compared to those of the residues from the selected models. Here, the five graphs of the residuals show that most of the points are the theoretical normal line except at the ends which move further and further away, due to heavy tail distributions ($\kappa > 3$ "Table 5") as shown in "Figure 7".

The following table allows the comparison of the results of the estimated values on the test bases.

"Table 6" gives an overview of forecast errors when the model applies on 87% and 76% of the initial bases. And by comparison with the test base respectively 2011-2017 and 2005-2017, gets fairly low errors except the parameters in the Louga series model.

3.9. Forecast

The monthly moving average of the three stations from 1963 to 2017 is represented by the horizontal cyan solid line of ordinates 22.58 on each of the graphs (Figure 8).
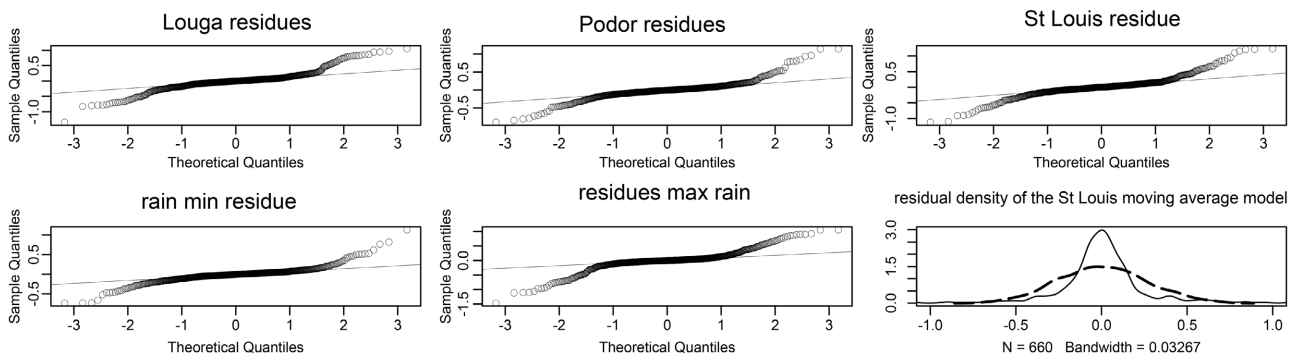

Figure 7. Normality of residualss, example of model residuals density ARIMA (8, 1, 0) (6, 1, 1) [12].

Table 6. Margins of error for learning tests over 7 and 13 years.

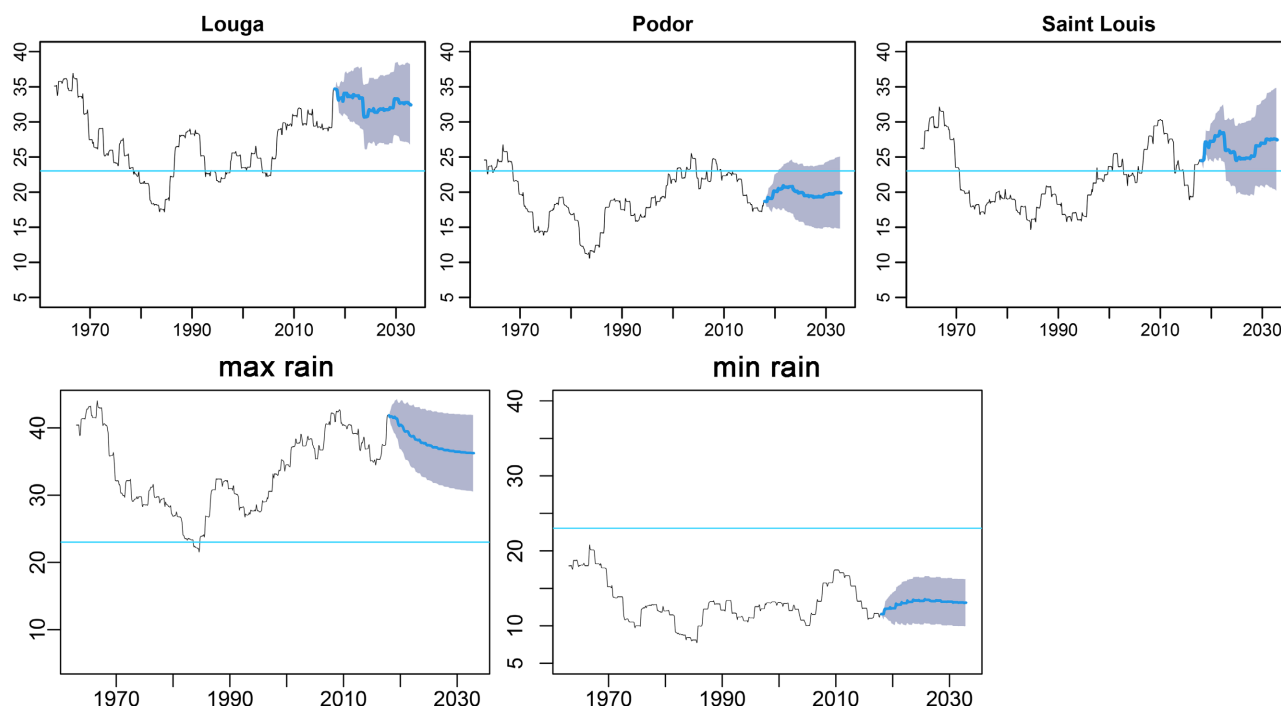| Test learning	errors	Louga	Podor	Saint Louis	min rain	max rain
to 87% (1963-2010)	ME	−1.58	2.86	4.67	1.26	2.30
	MAE	1.80	2.86	4.67	1.81	2.51
	RMSE	2.26	3.44	5.45	2.21	3.04
to 76% (1963-2004)	ME	−6.06	0.58	−0.44	−0.46	−3.69
	MAE	6.06	1.96	3.10	2.51	2.79
	RMSE	6.42	2.43	3.48	2.99	4.42

Figure 8. 15 year forecast.

Table 7. Evolution of the average per section for the periods 1963-1972-1990-2017 in forecasts up to 2032 and comparison with the period 1963-2017.

Periods	(1963-2017)	(1963-1972)	(1973-1990)	(1991-2017)	(2018-2032)
The monthly moving average of the three stations from 1963 *to* 2017 *is* **22.58339**					
Louga	26.56	32.69	↘ 23.78	↗ 26.47	↗ 32.65
Podor	19.24	22.51	↘ 15.85	↗ 20.54	↘ 19.82
Saint Louis	21.95	27.22	↘ 18.41	↗ 22.69	↗ 26.36
min rain	13.10	17.43	↘ 11.21	↗ 12.98	↗ 13.04
max rain	33.44	38.31	↘ 28.29	↗ 35.45	↗ 37.94

A slightly oscillatory downward trend in the previewing results of the Louga series and maximum rainfall from 2023 to 2032 and a constant wave after a short period of climb in Podor and Saint Louis during this same period. As for the minimum rainfall series, it remains moderately stable (Table 7).

4. Conclusions

The study showed that rainfall fell sharply over two decades from 1973 to 1997, after which there was a moderately return. It is followed by an increase from 2005 onwards in Louga and Saint Louis and then a gradual decrease from 2010 onwards. The best models obtained from the evaluation criteria integrated the data. The SARIMA forecasting framework based on the Box-Jenkins method shows that the 6-year average of the maximum rainfall series will continue to decline over the next 15 years. The same remark was made for those of Louga

who represents 40% of the latter. There is also a decrease in the average rainfall in Podor over the next fifteen years compared to the past two decades.

The advantage of the moving average is that it reduces "noise". This facilitates the reading of the evolution and thus makes it possible to quickly detect the trend changes in its evolution. This is important for the design of structures, but at the risk of losing data, especially if they are very volatile. The choice of the 6-year period was justified [3]. The interest was to follow the evolution of the forecasts over a long period of at least 15 years.

However, note that SARIMA processes are additive models and cannot take into account the volatile nature of certain variables. They are missing if the series is nonlinear, nonstationary (especially in variance) or if the variables do not follow the normal distribution following the leptokurtic character (Kurtosis > 3) of their distribution. In the following, we would like to study the Assymptotic behavior of minimum and maximum rainfall heights with the extreme value theory including the GEV (Generalised Extreme Value) and GPD (Generalised of Pareto Distribution) models then make a multivariate analysis by studying the relationship between the rainfall heights of the river flows that feed the Guiers lake via the Taouey channel and by involving certain climatic variables.

Conflicts of Interest

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

References

[1] Pankratz, A. (1983) Forecasting with Univariate Box-Jenkins Models Concepts and Cases. DePauw University, Greencastle. https://doi.org/10.1002/9780470316566

[2] Fernando, D.A.K. and Jayawardena, A.W. (1994) Generation and Forecasting of Monsoon Rainfall Data. *Affordable Water Supply and Sanitation*: *Proceedings of the* 20*th WEDC International Conference*, Colombo, 22-26 August 1994, 310-313.

[3] Ondo, J.C. (2002) Étude comparative des tests de stationnarité. INRS-Eau, Terre et Environnement (ETE), Université du Québec, Québec.

[4] Machiwal, D. and Jha, M.K. (2010) Évaluation comparative de tests statistiques pour l'analyse de séries temporelles: Application à des séries temporelles hydrologiques. *Hydrological Sciences Journal*, **53**, 353-366.

[5] Sambou, S., *et al.* (2006) Critique statistique des séries des pluies annuelles dans le bassin amont du Fleuve Sénégal. Université Cheikh Anta Diop, Dakar.

[6] Rivard, C., *et al.* (2003) Étude de l'impact potentiel des changements climatiques sur les ressources en eau souterraine dans l'Est du Canada. Division Québec, INRS-Eau, Terre et Environnement, Commission géologique du Canada, Québec. https://doi.org/10.4095/214161

[7] Kundzewicz, Z.W., *et al.* (2005) Détection de tendance dans des séries de débit fluvial: 2. Séries d'indices de crue et d'étiage. Centre for Ecology and Hydrology, Crowmarsh Gifford, Wallingford.

[8] Svensson, C., *et al.* (2005) Détection de tendance dans des séries de débit fluvial: 1. Débit maximum annuel. Centre for Ecology and Hydrology, Crowmarsh Gifford, Wallingford.

[9]   Adeloye, A.J.C., *et al.* (2002) Analyses préliminaires des données de débit en vue d'une étude de planification des ressources en eau. Heriot-Walt University, Edinburgh.

[10] Jayawardena, A.W. and Lai, F.Z. (1989) Time Series Analysis of Water Quality Data in Pearl River, China. *Journal of Environmental Engineering*, **115**, Article ID: 590607. https://doi.org/10.1061/(ASCE)0733-9372(1989)115:3(590)

[11] Tidjani, A.E.B., *et al.* (2005) Exploration des séries chronologiques d'analyse de la qualité des eaux de surface dans le bassin de la Tafna (Algérie).

[12] Aragon, Y. (2011) Série temporelle avec R. Méthodes et cas. Département MASS Université Rennes-2-Haute-Bretagne, France.

[13] Kane, A. (2005) Régulation du Fleuve Senegal et flux de matieres particulaire vers l'estuaire depuis la construction du Barrage de Diama.

[14] Mbaye, A.D. (2013) Plan de gestion environnementale et socialedu Projet de restauration des fonctions socio-écologiques du lac de Guiers—PREFELAG. Dakar.

[15] Gac, J.Y., *et al.* (1994) Évaluation des ressources en eau du fleuve Sénégal et Bilan hydrologique du lac de Guiers.

[16] Brunet-Moret, Y. (1977) Présentation d'un test d'homogénéité spécialement conçu pour vérifier d'homogénéité des suites chronologiques de précipitations armuelles dans une zone climatique, en utilisant, si possible, le vecteur régional Hiez. A. Cah. ORSTOM, Sér. Hydrol., Vol. XlV, n° 2.

[17] Lubes, H., *et al.* (1994) Caractérisation de fluctuations dans une série chronologique par applications de tests statistiques Etude Bibliographique. ORSTOM, Montpellier.

[18] Ruxton, G.D. (2006) The Unequal Variance t-Test Is an Underused Alternative to Student's t-Test and the Mann-Whitney U Test. Division of Environmental and Evolutionary Biology, Institute of Biomedical and Life Sciences, Graham Kerr Building, University of Glasgow, Glasgow.

[19] Kokoszka, P., *et al.* (1992) KPSS Test for Functional Time Series. Department of Statistics, Colorado State University, Fort Collins.