International Journal of Innovation and Scientific Research
ISSN 2351-8014 Vol. 11 No. 2 Nov. 2014, pp. 330-338

http://www.ijisr.issr-journals.org/

Modeling social security institution financial resources

***Moulay El Mehdi Falloul*[1] *and Moulay Ali Falloul*[2]**

[1]Ph.D candidate in applied economics and finance,
Hassan II University of Mohammedia,
Mohammedia, Morocco

[2]Ph.D in magnetism and electrical engineering,
Hassan II University of Casablanca,
Casablanca, Morocco

ABSTRACT: This paper focuses on the modeling of financial resources of private social security organism, namely the determination of models or process able to reproduce the evolution of this component. From the optimal model chosen, monthly forecasts are established. Financial resources modeling and forecasts resulting are obtained by application of the univariate analysis of Box-Jenkins method, which is the most suitable for the study of time series like the series of monthly financial resources of social security scheme administered by the institution. This approach is to determine the model that allows more faithfully the reproduction of the evolution of the financial resources of the institution and forecasting on a determined horizon.

KEYWORDS: financial resources modeling, time series, stationary, ARIMA, BOX and JENKINS.

1 INTRODUCTION

The objective of this study is to propose a modeling of social security institution financial resources. To do this, the social security institution financial resources, will be modeled from the method "ARIMA - BOX-JENKINS '.

This note is endeavoring to present the following points:

- A brief survey of the theoretical framework ARIMA models
- An analysis of the social security institution financial resources;
- Research of the most appropriate model for the social security financial resources (identification and validation of models to reproduce the series);
- Determination, from the model chosen the expected values of the financial resources for the prospective period.

2 THEORETICAL FRAMEWORK

2.1 ARIMA MODELS

The ARIMA model, proposed by Box and Jenkins, is generally recommended for model series or stochastic stationarity variables. This methodology is invoking the process ARMA, a noted autoregressive part AR, with a moving average part, noted MA.

In a stationary Autoregressive process of order p, the observation presents Y_t is generated by a weighted average of the past observations until the p^{th} period in the following form:

AR(p) : $Y_t = \sum_{1}^{p} \Phi_i y_{t-i} + \varepsilon_t \quad \forall t \in Z$

Or $Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots\ldots.. + \phi_p Y_{t-p} + \varepsilon_t$

Where ϕ_i are parameters to be estimated positive or negative; ε_t is a hazard.

The identification of the order p is established using the variable correlogram. Indeed, it is shown that the simple correlogram of an AR (p) process is characterized by a geometric decay of its terms and that the partial correlogram has its only p first different terms of 0. The step that follows "identification and estimation" parameters is the test of quality. A good approximation by Autoregressive process AR (p) translates into two types of tests [1]:

- Tests on the parameters: student tests, ϕ is significant at 5% if risk:

$\left| \hat{\Phi} / \hat{\sigma}(\Phi) \right|$ >1,96.

- Tests on white noise: Ljung-box test (Q_{LB}) and Portmanteau test (Q_P))

H0 : ε_t ~BB

$Q_{LB} = [(T+2)\, T \sum_{j=1}^{M} \hat{e}_j / T - j\,] \sim \chi^2_{(M-P)}$ et $Q_P = T \sum_{j=1}^{M} \hat{e}_j \sim \chi^2_M$

Where M =Min[T/2, $3\sqrt{T}$] and P number of parameters to be estimated.

Two statistics Q_{LB} and Q Q_P test the degree of autocorrelations of errors.

Regarding the process of moving average of order q, they allow to estimate each observation y by reference to a weighted average of hazards until the q th period.

MA(q) : $Y_t = \sum_{1}^{q} \theta_i \varepsilon_{t-i} + \varepsilon_t \quad \forall t \in Z$

Or even $Yt = \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \ldots\ldots.. + \theta_q \varepsilon_{t-q} + \varepsilon_t$

In this process, as in Autoregressive model AR, hazards are generated by a process of the white noise type.

The simple correlogram of MA (q) process is characterized by the q first terms that are significantly different from 0. Thus, we can determine the order q through a review of simple correlations (ACF).Estimate quality is assessed against the same tests used at the level of the AR process

ARMA models represent the processes that simultaneously combine values and past mistakes. They are defined by the equation:

ARMA (p,q) : $(1- \phi_1 B- \phi_2 B^2 - \ldots\ldots - \phi_p B^p) Y_t = (1- \theta_1 B- \theta_2 B^2 - \ldots\ldots - \theta_q B^q)\, \varepsilon_t$.

The identification and estimation of this type of models are based on the same combined principles for Autoregressive models AR (p) and medium mobile MA (q) [2].

The ARIMA, autoregressive integrated moving average model is an extension of the simple model ARMA. Generally an ARIMA (p, d, q) model takes the form:

$\phi(B)\, (1-B)^d\, Y_t = \theta\, (B)\, \varepsilon_t + C$

$\phi(B)$, θ (B): polynomials of order p and q respectively. ϕ_i ,θ_i respectively represent the parameters of the Autoregressive and moving average; d is the order of differentiation and B is the delay operator.

The selection of an ARIMA (p, d, q) model is the result of four main steps:

Step 1: identification of the first values of p, d, and q orders is based on the study of simple and partial correlograms.
Step 2: estimation of parameters θi and ϕi is based on a maximization of the likelihood through iterative process functions.

Step 3: the parameters being estimated, the estimation results must be examined by reference to tests on the meaning of the parameters and the quality of residues (absence of autocorrelation). :

Step 4: the choice of the most appropriate model among all the estimated models is carried out on the basis of two criteria: Akaike (AIC) and Schwartz (SC), which measure the quality of approximation of reality by the model chosen.

3 PRELIMINARY ANALYSES OF FINANCIAL RESOURCES

Taking into account the terms of the support of financial resources, the complete situation in respect of a fiscal year is arrested after 24 months of the financial year in question. Therefore, the analysis will focus on the series of monthly values recorded in the declared wage from January 1994 to December 2002 that there is (X_t).

The processed data are from extraction stopped at 31/08/2004; the representation of the raw series of monthly returns for the chosen period is as follows:

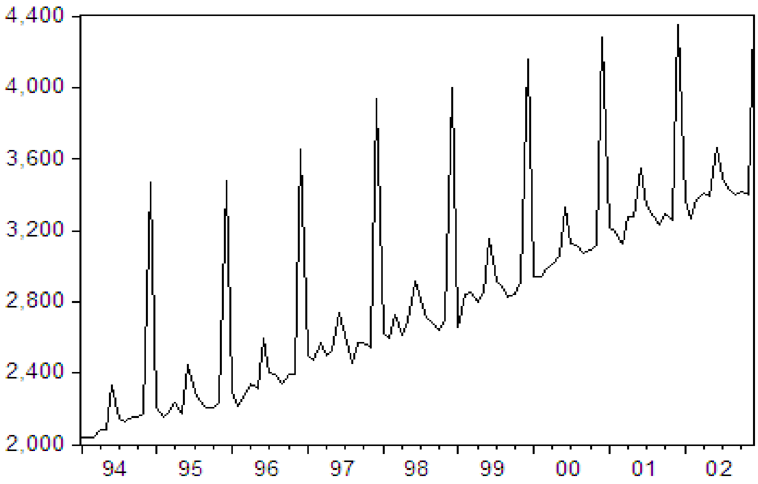


Fig. 1. Evolution of monthly financial resources Jan-94 to Dec-02

After adjustment of outlier values yields the following smoothed series:

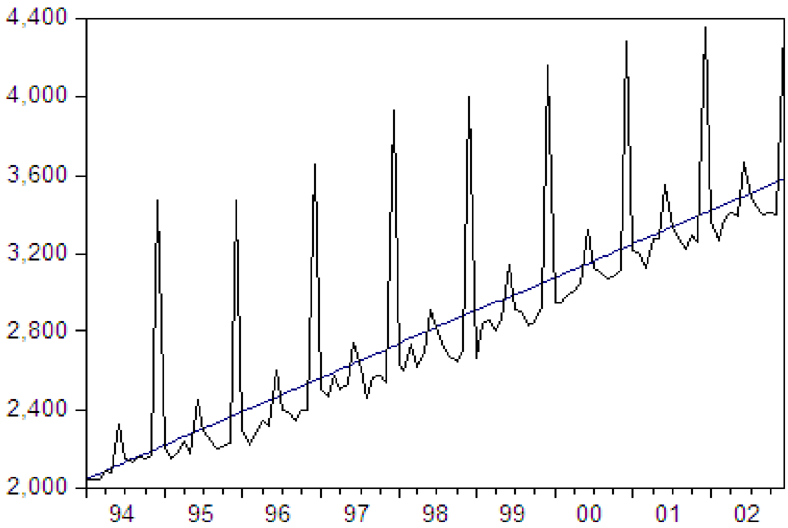


Fig.2. Evolution of monthly financial resources (smoothed series) Jan-94 to Dec-02

The obtained series indicates a growing trend over time and suggests a non-stationary mean. In addition this series presents a seasonality marked by spikes at regular intervals in June and December.

We hereby present the simple correlogram of the corrected series (Xt), which confirms the presence of seasonality:

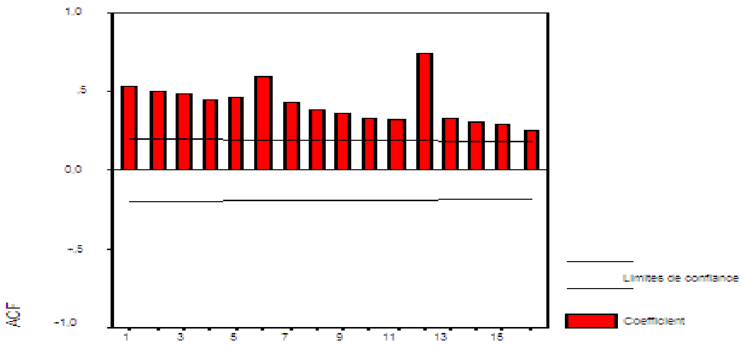


Fig.3. Simple financial resources correlogram X_t

One notes that all simple correlations of the series are significant. The peak in the period K = 12, confirms the presence of seasonality of periodicity equal to 12. What indicates to differentiate the seasonal and non-seasonal part of the series (X_t) to make it stationary.

4 IDENTIFICATION AND VALIDATION OF THE MODEL

To stationnarize series (X_t), is a simple differentiation of order 1 by applying the operator defined for the following series:

$$Y_t = \nabla X_t \text{ où } \nabla X_t = (X_t - X_{t-1})$$

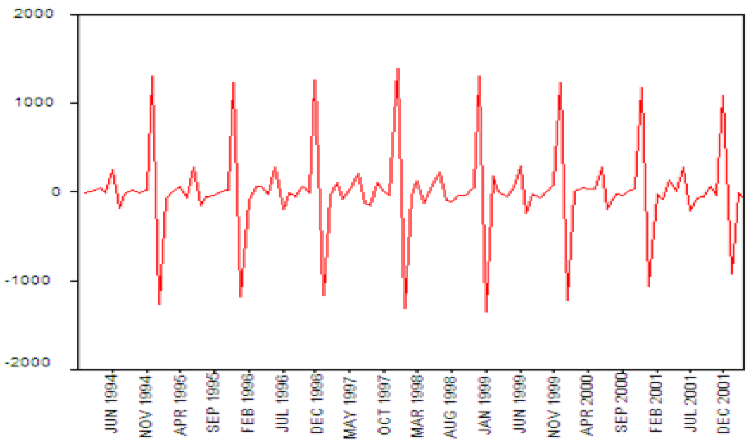


Fig.4. Monthly financial resources X_t 1^{st} differentiation

The new series Y_t has more tendency (the mean and the variance are stable and independent of the time). However, there is always a seasonal behavior that appears clearly in the form of significant peaks to delay k = 12 in the following this simple correlogram.

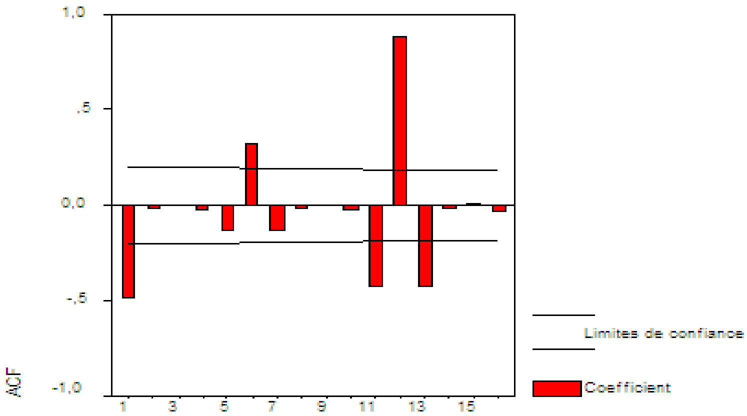


Fig.5. Simple financial resources correlogram y_t

In the step that follows, is the elimination of this seasonality, applying differentiation to the series Y_t of order 1 with a frequency equal to 12 using the operator ∇_{12} [3]*:*

$$w_t = \Delta_{12} Y_t = \Delta\Delta_{12} X_t = \left(X_t - X_{t-1}\right) - \left(X_{t-12} - X_{t-13}\right)$$

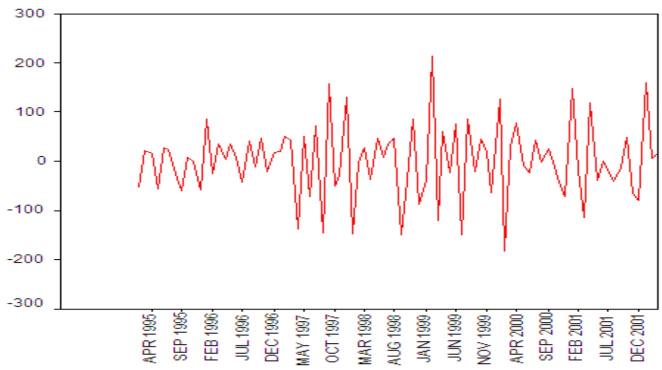


Fig.6. Monthly financial resources X_t 2^{nd} differentiation

The graph of the W_t series highlights the presence of an average zero, a relatively constant variance and the lack of seasonality. Thus, it can be concluded that the new W_t series is a stationary series [4].

The conditions are now fulfilled for research into the family of models ARMA model that best fits the data of the Z_t series.

The determination of such a model requires passing through three phases:

- Model specification;
- Estimation of the parameters of the model;
- Validation of the model.

4.1 ARIMA MODELS

The usual method is to rely on the shape of autocorrelation and partial autocorrelation of the studied series (possibly transformed) functions in order to choose a model ARMA or possibly several models which will be examined:

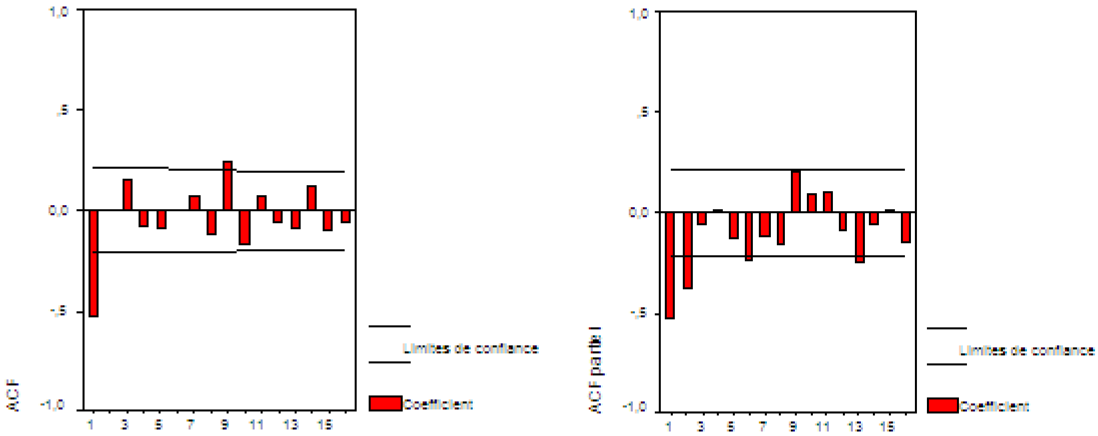


Fig.7. Autocorrelation and partial autocorrelation function time series of financial resources

The simple correlogram has its first significantly different term of 0 (characteristic of the mobile medium order process 1: MA (1)), which leads to analyze this process including the wording is as follows:

$$w_t = \varepsilon_t - \varphi_1 \varepsilon_{t-1} \Leftrightarrow w_t = (1 - \varphi_1 B)\varepsilon_t$$

The partial correlogram has its first two terms significantly non-zero (characteristic of the Autoregressive of order 2 processes: AR (2)), which leads to also consider this process which the mathematical formulation is presented below:

$$w_t = \theta_1 w_{t-1} + \theta_2 w_{t-2} + \varepsilon_t \Leftrightarrow (1 - \theta_1 B - \theta_2 B) w_t = \varepsilon_t$$

A third process to analyze is one that combines the two previous processes (MA (1) and AR (2)) noted ARMA (2.1) and the mathematical formulation is as follows:

$$w_t - \theta_1 w_{t-1} - \theta_2 w_{t-2} = \varepsilon_t - \varphi_1 \varepsilon_{t-1} \Leftrightarrow (1 - \theta_1 B - \theta_2 B) w_t = (1 - \varphi_1 B)\varepsilon_t$$

After identification of the three models, is the estimation of their parameters by the method of maximum likelihood, which is to determine the value of the parameter to estimate (the comments function) which ensures the highest probability of occurrence of these observations.

The results of the parameters estimated for each model are shown below:

Table 1. Estimation of the parameters for each model

	Model MA (1)	Model AR (2)		Model ARMA (2.1)		
	Coefficient	Coefficient	Coefficient	Coefficient	Coefficient	Coefficient
Estimate	0,746	-0,755	-0.342	-0.097	0.0004	0.692
Standard deviation	0.069	0.110	0.100	0.193	0,160	0.161
T statistic	10,542	-7,032	-3,619	-0.502	0.002	4,288

4.2 MODEL VALIDATION

Two types of tests are used for the validation of the models: on the meaning of the parameters of the different models and test the variety of their residues.

4.2.1 TESTS ON PARAMETERS

The coefficients of the first two models (AR (2) and MA (1)) are significantly different from zero at the 5% threshold. In fact their student's T-distribution is higher in absolute terms, the critical value of 1.96 (see table above).

On the other hand, the model coefficients ARMA (2.1) are not all different from zero at the 5% threshold, including coefficients of part autoregressive for which Student's T is lower, in absolute terms, the critical value of 1.96 (see table above). This leads to spread this model of the analysis suite.

4.2.2 TESTS ON RESIDUALS

When the process is well estimated, residues (difference between the observed values and the values estimated by the model) must behave like a white noise. It should be noted later $\hat{\varepsilon}_t$ the residuals of the model estimation.

Test of nullity of the average:

Table 2. Test of nullity of the average

Models	T statistic
Model MA (1)	0,422
Model AR (2)	0.264

The average of the residues of the two models may be considered to be zero at the 95% threshold. In fact their student's T-distribution is less than the critical value of 1.96 in absolute (see table above).

Test of autocorrelation of the residuals:

If residues *(ε_t, $t \in \mathbb{Z}$)) obey a white noise, it must not exist autocorrelation in the series. You can then use the following tests:*

-Study of the simple autocorrelogram (ACF) and the partial autocorrelogram (PACF): it shall be verified that there is no autocorrelation and partial autocorrelation significantly non-zero for the process studied.
- Ljung-Box test

The analysis of the autocorrelograms of both models shows that there is no significant autocorrelation between residues.

Analysis of the residuals of the model series MA (1)

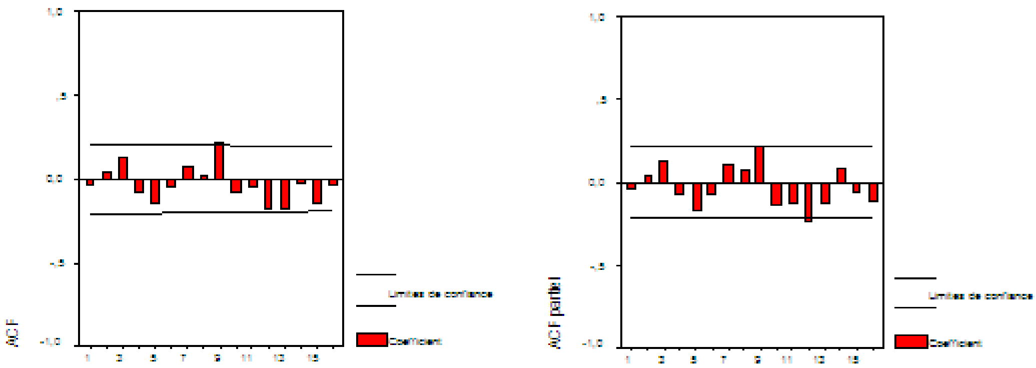
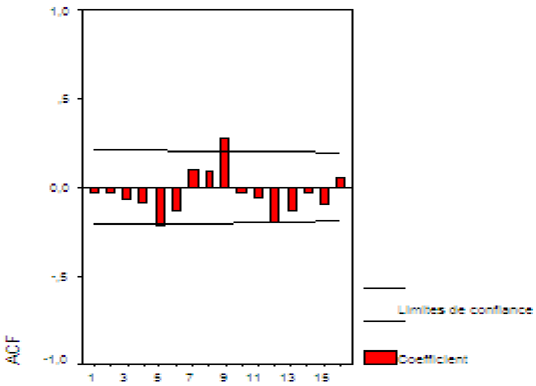


Fig.8. Autocorrelation and partial autocorrelation function time series of financial resources

Analysis of the series of the residuals of the AR (2)



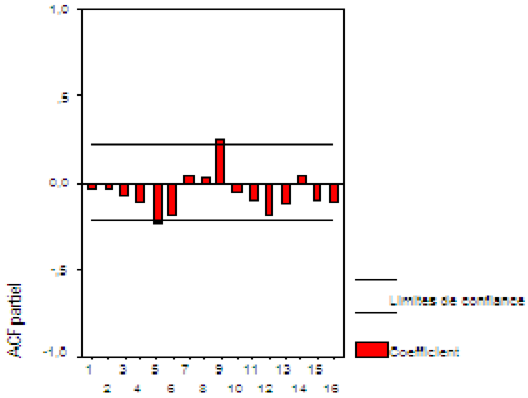


Fig.9. Autocorrelation and partial autocorrelation function time series of financial resources

The global test of Ljung-Box (White noise test) **for** the two models has provided the following table:

Table 3. The global test of Ljung_Box

Delay	Model MA (1)		Model AR (2)	
	Ljung-Box value	Prob. of meaning	Ljung-Box value	Prob. of meaning
1	0.112	0.741	0.074	0.770
2	0.274	0.842	0.144	0.910
3	1.722	0.635	0.644	0.874
4	2,314	0.685	1,441	0.852
5	4,242	0.514	5,752	0.325
6	4,445	0.642	7,441	0.287
7	4,914	0.671	8,442	0.258
8	4,945	0.742	9,112	0.342
9	9,564	0.341	16,752	0.050
10	10,188	0.445	16,963	0.074
11	10.385	0,458	17,241	0.105
12	13.564	0.335	21,163	0.044
13	16,652	0.214	23,153	0.055
14	16,742	0.275	23,244	0.054
15	18,925	0.218	24,242	0.061
16	19,122	0.262	24,442	0.084

The meaning of 'Ljung-Box test probabilities are all higher than 0.05 for both models. Therefore, it can be concluded that residues do not significantly of a white noise process deviate.

At the end of this step, we could consider that the two models estimated AR (2) and MA (1) are satisfactory.

5 Selection Of The Model

It is to choose the best of the two models validated in the previous step on the basis of criteria of information**.** Concretely these criteria are constructed as functions of variance of the model estimated residues and the number of parameters to estimate.

The goal is to find the model that minimizes this function with respect to these two arguments; it retains the criterion of Akaike (AIC) and Bayesian Schwartz (SBC) criterion [5].

The following table shows the calculated value of these two criteria for each model:

Table 4. The global test of Ljung_Box

Criteria	Model MA (1)	Model AR (2)
AIC	943,402	948,852
SBC	944,857	923,745

The model MA (1) (a) of the criteria those are lower than those of the AR (2). Thus, it can be concluded that the best model is the model MA (1).

The equation of the model chosen is as follows:

$$w_t = \varepsilon_t - 0{,}746\,\varepsilon_{t-1} \Leftrightarrow w_t = (1\text{-}0{,}746\mathrm{B})\,\varepsilon_t$$

Where:

$$w_t = \nabla_{12} Y_t = \nabla\nabla_{12} X_t = (X_t - X_{t-1}) - (X_{t-12} - X_{t-13})$$

6 Conclusion

At the end of this step, we could consider that the two models estimated AR (2) and MA (1) are satisfactory and this model allows more faithfully the reproduction of the evolution of the financial resources of the institution and forecasting on a determined horizon.

References

[1] G.E.P Box, and G.M Jenkins: Time Series analysis: forecasting and control, 2nd edition, Holden-Day, 1976.
[2] R. Bourbonnais : Économétrie, Manuel et exercices corrigés, 7e édition. DUNOD, Paris, 2009.
[3] R. Bourbonnais and M. Terraza : Analyse des séries temporelles en économie, 1er édition. PUF, 1998.
[4] C. Brooks: Introductory Econometrics for Finance, 2nd edition. Cambridge University Press, 2008.
[5] J. Hamilton: Time series analysis, 1st edition, Princeton University Press, 1994.