

IDENTIFICATION OF THE COMPONENTS OF THE DRINKING WATER MODEL: THE USE OF ARMA MODELS

N. Loudjani*, A. Messameh, M.T. Bouzaine

Laboratoire de Recherche en Hydraulique Souterraine et de Surface LARHYSS, Université de
Biskra, Algeria

Received: 06 October 2019 / Accepted: 11 July 2020 / Published online: 01 September 2020

ABSTRACT

In this paper, we consider the modeling of time series corresponding to the methodology of Box and Jenkins for building forecasting modeling from ARMA (Auto Regressive Moving Average) processes. This involves the analysis and forecasting of drinking water production data. The purpose of this study is to modeling a time series of drinking water production data according to its past and present values in order to determine the adequate ARMA process by the principle of parsimony. Using the Eviews for Windows software, the results show that the exploration of monthly drinking water consumption data provided by the Biskra Management Authority (A.D.E) during the period from January 2009 to January 2016, Revealed characteristics such as the non-stochastic non-stationarity of the drinking water production series.

Keywords: Time series; Box and Jenkins method; ARMA process; Eviews software.

Author Correspondence, e-mail : nloudjani@gmail.com

doi: <http://dx.doi.org/10.4314/jfas.v12i3.6>



1. INTRODUCTION

L'analyse des séries chronologiques est un outil couramment utilisé de nos jours pour la prédiction de données futures [1]. La prévision à court terme de la consommation d'eau potable joue un rôle essentiel dans la gestion efficace des ressources allouées à la production d'eau potable. Dans ce travail, nous nous concentrons sur la consommation mensuelle d'eau potable gérée par l'Algérienne Des Eaux (ADE) de Biskra durant la période du Janvier 2009 au Janvier 2016.

La plupart des logiciels de prévision, comme les logiciels statistiques et économétriques, permettent rapidement de réaliser ces analyses sur micro-ordinateur mais les concepts de base des méthodes les plus avancées restent complexes à aborder. Les fonctionnalités du module Séries chronologiques d'EViews pour Windows [2] nous permettent de construire une modélisation des variations mensuelles des données de production d'eau potable considérée comme un processus temporel soumis à des perturbations aléatoires. Pour ces raisons, nous utiliserons le logiciel Eviews afin d'obtenir des résultats fiables.

L'objet de cette étude est de modéliser une série temporelle des données de production d'eau potable de la ville de Biskra pour tester les modèles de prévision de cette consommation à court terme afin de déterminer le processus ARMA adéquat par principe de parcimonie.

2 DONNEES ET METHODES

2.1 Données

C'est une phase d'analyse descriptive des données brutes et l'analyse et évolution des séries chronologiques de consommation d'eau potable dans la zone d'étude. La caractérisation des composantes d'une série : tendance, effet saisonnier et erreurs par l'analyse graphique et du corrélogramme [3]. Modéliser les mouvements de la série temporelle des données de la production d'eau potable à partir des valeurs présentes et passées d'un bruit blanc avec un modèle linéaire. La modélisation permettant d'effectuer des prévisions. Nous nous sommes concentrés sur la prévision à court terme. Dans notre contexte, lorsqu'un modèle purement $AR(p)$ (Autoregressive) est utilisé pour modéliser le profil temporel de consommation d'eau potable [4] les données considérées à cet effet sont mensuelles et la période retenue pour

l'étude va du janvier au 2009 janvier 2016 (85 observations). Pour construire le modèle de prévision, nous avons adopté la méthodologie de Box-Jenkins.

2.2 Méthodes

Dans toute approche de prévision, le choix de la méthode dépend de l'objectif de l'étude.

Nous utiliserons le logiciel «Eviews» afin d'obtenir des résultats fiables à propos de modéliser les mouvements de la série temporelle à partir des valeurs présentes et passés d'un bruit blanc avec un modèle linéaire [5].

La méthodologie de Box et Jenkins permet de déterminer le processus ARMA (Auto Regressive Moving Average) adéquat pour la modélisation d'une chronique. La méthodologie B & J suggère quatre étapes, à savoir : l'identification, l'estimation, la validation et la prévision [6]. La méthodologie de Box & Jenkins vise à formuler un modèle permettant de représenter une chronique avec comme finalité de prévoir des valeurs futures. Ils se réfèrent à deux types de modèle : des processus moyennes mobiles, des modèles autorégressifs, ou à une combinaison des deux. Les modèles ARMA permettent de représenter la plupart des processus aléatoires stationnaires [7].

3. RÉSULTATS ET DISCUSSION

Dans cette étude, nous avons analysé une série des données mensuelles de l'eau Potable de la ville de Biskra selon la méthodologie synthétisée par Box et Jenkins pour mieux caractériser le comportement d'une série chronologique à court terme [8].

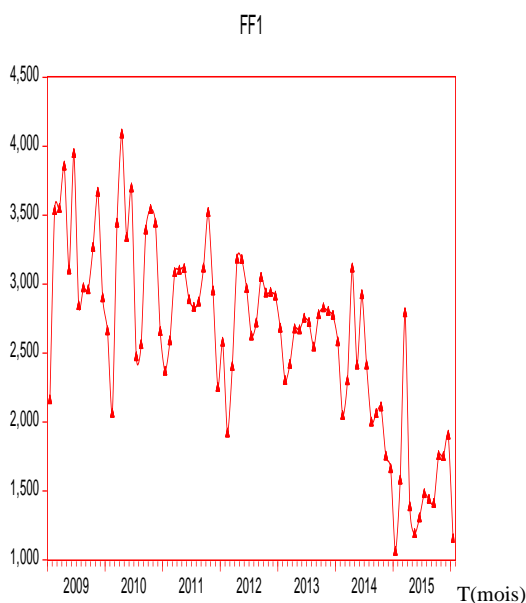
Tableau 1. Les données de production d'eau potable mensuelles Vp (volume produit) en (m³/j) de la ville de Biskra pour la période de Janvier 2009 à Janvier 2016

Année \ Mois	2009	2010	2011	2012	2013	2014	2015	2016
Janvier	2160	2660	2367	2576	2678	2581	1306	1153
Février	3532	2061	2587	1918	2300	2045	1577	
Mars	3547	3440	3082	2400	2416	2298	2791	
Avril	3854	4087	3099	3181	2674	3114	1384	
Mai	3100	3337	3111	3176	2666	2411	1189	
Juin	3945	3696	2885	2966	2751	2922	1302	
Juillet	2843	2472	2830	2621	2719	2411	1478	
Aout	2972	2560	2868	2716	2542	1998	1437	
Septembre	2957	3390	3115	3048	2777	2060	1408	
Octobre	3264	3540	3518	2933	2827	2108	1754	
Novembre	3664	3440	2950	2938	2800	2750	1747	
Décembre	2900	2654	2250	2911	2771	2660	1903	

A. Analyse descriptive

Tableau 2 : Corrélogramme de la série FF1

Vp (M³/j)



Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
1	0.437	0.437	14.536	0.000	
2	0.149	-0.052	16.257	0.000	
3	-0.034	-0.100	16.349	0.001	
4	0.035	0.121	16.454	0.002	
5	0.305	0.328	23.943	0.000	
6	0.380	0.149	35.755	0.000	
7	0.275	0.022	42.040	0.000	
8	0.004	-0.141	42.042	0.000	
9	-0.111	-0.053	43.089	0.000	
10	0.029	0.102	43.152	0.000	
11	0.278	0.185	50.004	0.000	
12	0.430	0.177	66.582	0.000	
13	0.300	0.038	74.788	0.000	
14	0.095	0.019	75.619	0.000	
15	-0.059	-0.045	75.943	0.000	
16	0.049	-0.001	75.978	0.000	
17	0.109	-0.138	77.145	0.000	
18	0.233	0.010	82.546	0.000	
19	0.198	0.052	86.501	0.000	
20	-0.091	-0.188	87.347	0.000	
21	-0.178	-0.067	90.697	0.000	
22	-0.126	-0.015	92.394	0.000	
23	0.055	-0.012	92.728	0.000	
24	0.237	0.044	98.009	0.000	
25	0.145	-0.023	100.40	0.000	
26	0.020	0.023	100.45	0.000	
27	-0.091	0.050	101.42	0.000	
28	-0.149	-0.139	104.13	0.000	
29	0.022	0.019	104.19	0.000	
30	0.113	-0.010	105.81	0.000	
31	0.103	0.017	107.20	0.000	
32	-0.056	-0.012	107.62	0.000	

Fig.1. L'évolution mensuelle de la série FF1 (Forage F1)

Le graphe de la figure 1 représente l'évolution mensuelle de la production d'eau potable de la ville de Biskra. Il laisse apparaître une tendance déterministe sur toute la durée ainsi que des fluctuations irrégulières qui traduisent la saisonnalité ; ceci est confirmé par la fonction d'autocorrélation. Il existe des pics très importants aux niveaux de $K = 1$, $K = 6$, $K = 12$. Ce qui montre que la série est non stationnaire. Nous avons indiqué que l'analyse du corrélogramme simple et partiel nous montre une non stationnarité de la série. Il a donc été prolongé par le test de Dickey et Fuller augmenté (ou test ADF) afin de détecter la présence d'une racine unitaire pour les processus de type AR(p) [9].

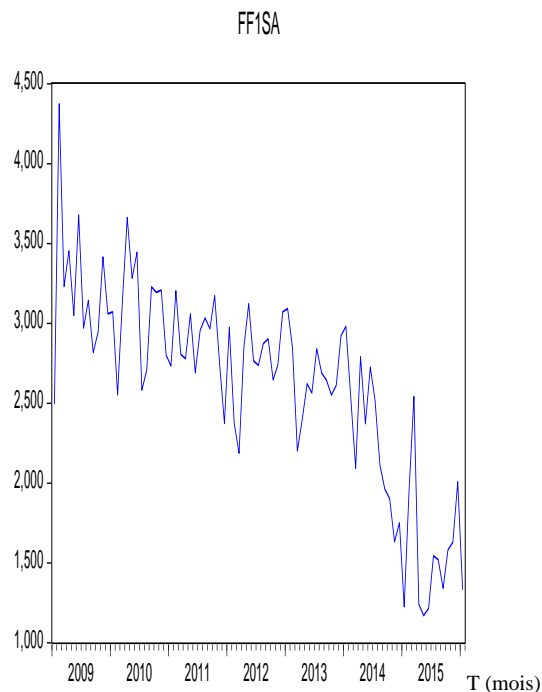
Le corrélogramme (Tableau 2) représentant l'analyse de la série FF1 est calculé à l'aide du logiciel EVIEWS sur 32 retards. Son examen présente une décroissance de ses retards (ce qui indique l'existence du facteur tendanciel). Le premier terme du corrélogramme simple et partiel est très important (0.437) ceci nous amène à tester l'existence d'une tendance déterministe ou stochastique par l'emploi du test de racine unitaire. Les autocorrélations s'annulent très lentement. Donc la série brute est affectée de la saisonnalité de la tendance, ce qui nous amène à appliquer le test de DICKEY-FULLER pour confirmer la non stationnarité de la série FF1SA.

B. Etude de la stationnarité de la série désaisonnalisée (FF1SA)

Avant d'utiliser la procédure ARMA, il faut tout d'abord examiner la série et vérifier sa stationnarité avec un graphique temporel [10]. On va désaisonnaliser la série FF1SA à l'aide de coefficients saisonniers (CVS) et la méthode de moyenne mobiles (Moving Average). Désaisonnaliser une chronique c'est éliminer cette saisonnalité sans modifier les autres.

Tableau 3. Corrélogramme de la série FF1SA

Vp (M³/j)



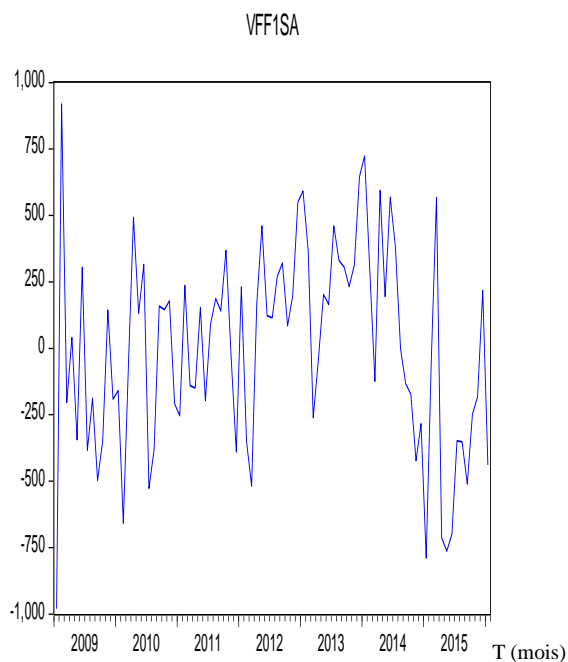
Date: 03/07/17 Time: 08:51
Sample: 2009M01 2016M01
Included observations: 85

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
1	0.718	0.718	45.364	0.000	
2	0.654	0.287	83.507	0.000	
3	0.618	0.173	117.92	0.000	
4	0.626	0.202	153.75	0.000	
5	0.604	0.100	187.47	0.000	
6	0.538	-0.042	214.57	0.000	
7	0.487	-0.049	237.02	0.000	
8	0.464	-0.003	257.69	0.000	
9	0.418	-0.062	274.70	0.000	
10	0.344	-0.125	286.36	0.000	
11	0.358	0.095	299.15	0.000	
12	0.300	-0.059	308.25	0.000	
13	0.267	-0.028	315.55	0.000	
14	0.269	0.102	323.07	0.000	
15	0.214	-0.058	327.91	0.000	
16	0.183	-0.047	331.48	0.000	
17	0.105	-0.133	332.68	0.000	
18	0.107	0.028	333.94	0.000	
19	0.125	0.076	335.70	0.000	
20	0.104	0.009	336.94	0.000	
21	0.053	-0.011	337.27	0.000	
22	0.043	0.008	337.49	0.000	
23	0.024	-0.033	337.56	0.000	
24	0.038	0.056	337.74	0.000	
25	0.035	0.018	337.89	0.000	
26	0.059	0.112	338.33	0.000	
27	0.057	0.000	338.74	0.000	
28	-0.015	-0.171	338.78	0.000	
29	0.004	0.041	338.78	0.000	
30	0.001	-0.047	338.78	0.000	
31	0.003	-0.005	338.78	0.000	
32	0.002	0.044	338.78	0.000	

Fig.2. Profil de la série désaisonnalisé FF1SA

Tableau 4. Corrélogramme de la série VFF1SA

Vp (M³/j)



Date: 03/07/17 Time: 09:08
Sample: 2009M01 2016M01
Included observations: 85

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
1	0.268	0.268	6.3464	0.012	
2	0.206	0.144	10.110	0.006	
3	0.165	0.088	12.578	0.006	
4	0.268	0.203	19.120	0.001	
5	0.278	0.168	26.288	0.000	
6	0.207	0.063	30.317	0.000	
7	0.128	-0.012	31.859	0.000	
8	0.163	0.045	34.395	0.000	
9	0.119	-0.029	35.780	0.000	
10	0.007	-0.146	35.784	0.000	
11	0.076	0.014	36.354	0.000	
12	-0.030	-0.118	36.445	0.000	
13	-0.015	-0.066	36.469	0.001	
14	0.024	0.046	36.528	0.001	
15	-0.038	-0.033	36.680	0.001	
16	-0.017	0.024	36.710	0.002	
17	-0.146	-0.121	39.042	0.002	
18	-0.060	0.035	39.444	0.002	
19	-0.014	0.051	39.466	0.004	
20	-0.079	-0.063	40.180	0.005	
21	-0.163	-0.079	43.241	0.003	
22	-0.167	-0.095	46.521	0.002	
23	-0.143	-0.058	48.959	0.001	
24	-0.088	-0.002	49.895	0.001	
25	-0.113	-0.013	51.474	0.001	
26	-0.029	0.155	51.581	0.002	
27	-0.019	0.076	51.626	0.003	
28	-0.193	-0.141	56.463	0.001	
29	-0.108	0.013	58.002	0.001	
30	-0.111	-0.072	59.661	0.001	
31	-0.088	-0.064	60.730	0.001	
32	-0.052	-0.003	61.103	0.001	

Fig.3. Profil de la Série stationnaire VFF1SA

On remarque du corrélogramme de la fonction d'autocorrélation simple (ACF) trois pic important pour le terme MA (1.4.5) sont significativement différentes de zéro et le corrélogramme de la fonction partielle (PACF) est un pic important pour le terme AR(1) et d'après la combinaison des deux fonctions il ya cinq modèle : **AR (1), MA(1), ARMA (1.1), ARMA (1.4), ARMA (1.5).**

A partir de ce graphe, on peut dire que la série VFF1SA est stationnaire, on obtient VFF1SA est une série stationnaire autour de sa moyenne et la fonction d'autocorrélation converge rapidement vers le zéro et n'est pas des effets saisonniers.

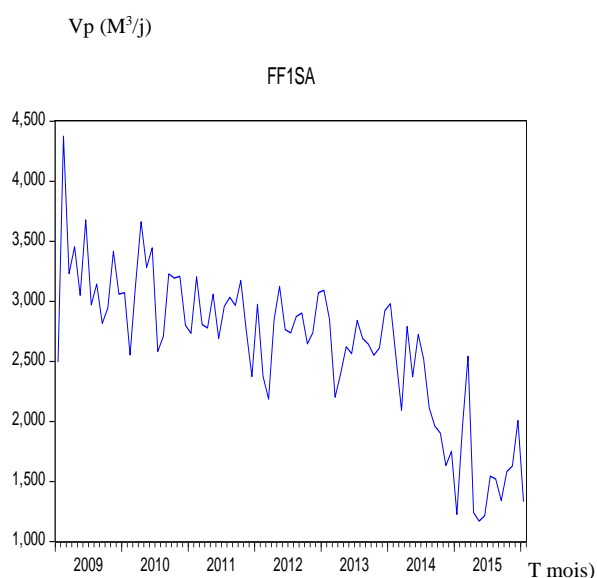


Fig.4. Série non stationnaire FF1SA

(Série observée)

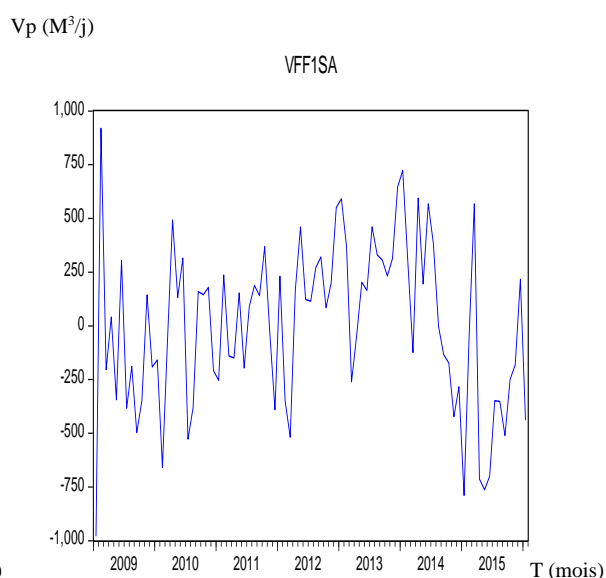


Fig.5. Série stationnaire VFFSA

(Série Calculée)

Test de Dickey Fuller Augmenté (ADF):

La série est stationnaire mais on doit confirmer par test de Dikey- Fuller [11].

Tableau 5 : Test de stationnarité de la série des données de production d’eau potable (Modèle avec tendance et avec constante)

Null Hypothesis: VFF1SA has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 1 (Fixed)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-4.467740	0.0030
Test critical values:		
1% level	-4.072415	
5% level	-3.464865	
10% level	-3.158974	

*MacKinnon (1996) one-sided p-values.

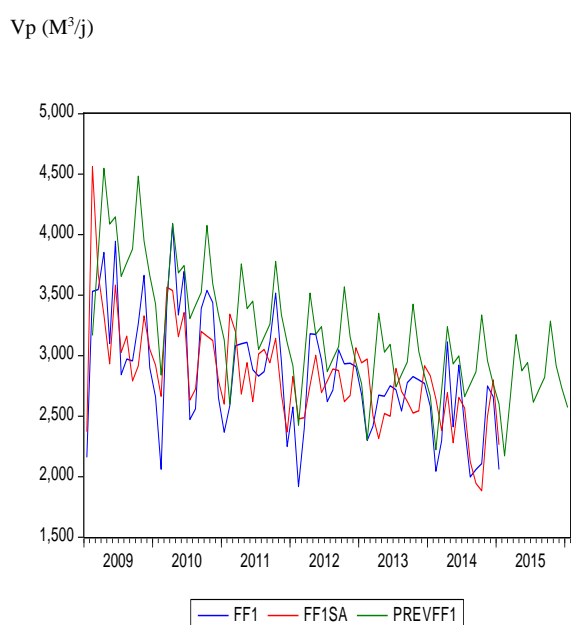
Augmented Dickey-Fuller Test Equation
Dependent Variable: D(VFF1SA)
Method: Least Squares
Date: 02/15/18 Time: 10:28
Sample (adjusted): 2009M03 2016M01
Included observations: 83 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
VFF1SA(-1)	-0.545860	0.122178	-4.467740	0.0000
D(VFF1SA(-1))	-0.113792	0.100664	-1.130418	0.2617
C	-8.224961	76.48300	-0.107540	0.9146
@TREND(2009M01)	0.065668	1.552609	0.042296	0.9664

R-squared	0.337731	Mean dependent var	-16.37420
Adjusted R-squared	0.312581	S.D. dependent var	408.4401
S.E. of regression	338.6407	Akaike info criterion	14.53475
Sum squared resid	9059523.	Schwarz criterion	14.65132
Log likelihood	-599.1921	Hannan-Quinn criter.	14.58158
F-statistic	13.42894	Durbin-Watson stat	1.866171
Prob(F-statistic)	0.000000		

Pour La confirmation par le test de Dickey-FULLER, la série n’admet pas de racine unitaire (la série est stationnaire) ($0.003 < 0.05$) et la série n’admet pas de tendance ($0.9664 > 0.05$). On peut conclure que la série VFF1SA est stationnaire [12].

Tableau 6. Corrélogramme des résidus carrés



Date: 01/29/18 Time: 10:59
Sample: 2009M02 2016M01
Included observations: 84
Q-statistic probabilities adjusted for 1 ARMA term(s)

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 0.067	0.067	0.3897	
		2 0.076	0.072	0.8966	0.344
		3 0.067	0.058	1.2980	0.523
		4 0.164	0.153	3.7258	0.293
		5 0.240	0.222	8.9870	0.061
		6 0.091	0.056	9.7506	0.083
		7 0.035	-0.011	9.8647	0.130
		8 0.088	0.038	10.608	0.157
		9 0.080	0.002	11.224	0.189
		10 -0.026	-0.118	11.292	0.256
		11 0.072	0.031	11.798	0.299
		12 -0.072	-0.109	12.324	0.340
		13 -0.071	-0.124	12.834	0.381
		14 0.055	0.060	13.146	0.437
		15 -0.004	0.022	13.148	0.515
		16 0.034	0.048	13.271	0.581
		17 -0.138	-0.091	15.330	0.501
		18 -0.085	-0.040	16.125	0.515
		19 -0.003	-0.004	16.126	0.584

T(Mois)

Fig.6. Profil des données d’observation et de prévision de la production d’eau potable avec le modèle ARMA

Observation

- 1)- La constante n'est pas significative car la probabilité =0.7959>0.05
- 2)- Le coefficient de AR(1) significative car la probabilité=0.0096 <0.05

Interprétation

D'après le corrélogramme des résidus, nous constatons que la probabilité très supérieure à 0.05 car tous les pics à l'intérieure de l'intervalle de confiance, le résidu peut être assimilé à un bruit blanc, donc le modèle est validé. On remarque que tous les pics sont à l'intérieur de l'intervalle de confiance ce qui confirme l'absence de l'effet **ARCH**. C'est-à-dire **la variance des résidus est homogène**.

L'observation des graphiques montre bien que les variables de la valeur observée (données de production d'eau potable) sont collées avec celle de la variable projetée (prédite) (PREVFF1) et que le résidu se comporte maintenant comme un bruit blanc.

4. CONCLUSION

Dans cette étude, nous avons utilisé les données mensuelles de la production d'eau potable de la ville de Biskra pour tester les modèles de prévision de cette consommation à court terme. L'exploration de ces données nous a révélé des caractéristiques telles que la non-stationnarité stochastique de la série de production d'eau potable. Les modèles estimés doivent cependant être améliorés. Une piste évidente à poursuivre est de prendre en compte les cycles mensuels et saisonniers (printemps, été, automne et hiver). Les modèles de prévision fournissent une aide précieuse à la prise de décision dans le domaine de la gestion des ressources en eau, et demeurent fiables dans la prévision du comportement du phénomène étudié malgré la présence d'un bruit blanc. Ainsi on recommande de faire une étude d'analyse de données de consommation en regression non linéaire par la methode de réseaux de neurones.

5. REFERENCES BIBLIOGRAPHIQUES

- [1] Michel.P et Christine.K et Raphael.R (2010-2011) séries Chronologiques, cours et exercices, Université Paris-Sud.
- [2] Bourbonnais, R. et Terraza, M. (2004). *Analyse des séries temporelles*, Dunod, Paris. Economica, Paris.
- [3] Martino, Joseph P. (1983). *Technological Forecasting for Decision Making*, Elsevier, New
- [4] Hamilton, J. (1994). *Time Series Analysis*, Princeton University Press, Princeton.
- [5] Box, G. E., & Jenkins, G. M. (1970). *Time Series Analysis: Forecasting and Control*. Holden-D. San Francisco.
- [6] Gorioux, C. et Monfort, A. (1990). *Séries temporelles et modèles dynamiques*, Economica, Paris
- [7] Coutrot, B. et Dreesbeke, J.-J. (1990). *Les méthodes de prévision*, Que Sais-je ? n°2157, Presses Universitaires de France, Paris (2e éd.).
- [8] Bisgaard, S. and M. Kulachi (2011). *Time's series analysis and forecasting by example*. John Wiley & sons.
- [9] David M. et Michaud J.C. (1989) *La prévision, approche empirique d'une méthode statistique*,
- [10] Lejeune M. 1997, *Statistique, cours B7: séries chronologiques*, Collection des cours du CNAM, 63 p.
- [11] Bisgaard, S. and M. Kulachi (2011). *Time's series analysis and forecasting by example*. John Wiley & sons.
- [12] DICKEY DAVID A. et WAYNE A. FULLER, (1979), "Distribution of the Estimators for Autoregressive Time Series with a Unit Root", *Journal of the American Statistical Association*, 74, 427 – 431.

6. LISTE DES ABREVIATIONS

ACF : Fonction d'autocorrélation

PACF : Fonction d'autocorrélation partielle

AR(p) : Processus autorégressif AR(p) de l'ordre p

MA(q) : Processus moyenne mobile de l'ordre q

How to cite this article:

Loudjani N, Messameh A, Bouzaine M T, Identification of the components of the drinking water model: The use of ARMA models. *J. Fundam. Appl. Sci.*, 2020, *12(3)*, 1090-1100.