ISSN: 2278-3369

International Journal of Advances in Management and Economics

Available online at www.managementjournal.info

CASE STUDY

Water Demand, Distribution and Consumption Forecasting: Case of Tunisia

Imen Gam[1]*, Rabiaa Ben Aïcha[1], Jaleleddine Ben Rejeb[2]

[1]*Law, Economics and Politics of Sousse University.*

[2]*High Institute of Management of Sousse University.*

*Corresponding Author: Email: ***frtngamimen@yahoo.fr***

Abstract

Researches, nowadays, led in hydraulic domain concern, principally, the management of water, especially in countries endowed with limited water resources. We analyzed the current situation of water in Tunisia as well as its state of shortage aggravated more and more in the future by envisaging it by the methodology of Box and Jenkins allowing, suchlike, a manager and a treasurer to work out plans of actions to meet growing needs of the populations and development for today and tomorrow. Results reveal an overall upward trend in the production, distribution and consumption of water in Tunisia and a bullish gap reflecting losses on adduction and on distribution.

Keywords: *Box and Jenkins, Forecasting, Scarceness of water, Sustainable development.*
JEL Classification: Q25, Q21, C53

Introduction

Water is the first element of life. It constitutes an economic, social and environmental good and as such must be managed with the objective to protect a common heritage in the interest of the entire community.

In a few decades, the population growth as well as the development of industries and agriculture multiplied the uses of water doing so explode demand water. And faced with limited water resources and unequally distributed in space and in time, a problem of water shortage appears for many countries of the world having a situation of in equation, translated by excess water demands compared to the available resources. More specifically, and in the Mediterranean basin, Tunisia counts among the least provided countries in means in water. To be stricter, we try to analyze actual situation and trend some water in Tunisia by referring to predictions performed from series of real data. We begin by informing about the data on which we were based to perform prediction, in a second stage, of the series of production, of distribution as well as of consumption of water in Tunisia using the methodology of Box and Jenkins. On the other hand, we trace the evolution of the availability of water per capita as well as domestic consumption in order to characterize the water situation in Tunisia. We finish our work with a conclusion including recommendations applicable in this topical issue.

Theoretical Basis

The review of past studies shows an important number of works in different categories of water use with the dominance of modeling residential water demand. Residential water demand has been extensively analyzed during the last decades. The mainly aim of those analyses is to design the most effective water demand policy otherwise to determine the policy that allows the most efficient water allocation by detecting the variables that underlie water demand, In other words, to quantify their impact on water demand. The founding articles of this literature dated back to the late sixties and focused on area of the USA [1,3] Later, researchers have used European data to decrypt the increase of water demand. Those studies used time series, cross sectional data or panel data and their main aim was to estimate elasticities of residential water demand with respect to price, income, population and household characteristics, precipitation moreover other variables both on the long and short-term.

Attention has lately shifted to forecast water demand. Researchers have underlined that an accurately and adequately forecasted water demand either in short-term, medium-term or long term is very relevant for operational management of a water network (capacity planning, scheduling maintenance…) and financial planning (reviewing prices…). Otherwise forecasting is of key importance in order to optimize all the operations of the water system and the making of strategic decision In other words, to ensure the adequate, efficient and sustainable water management.

In order to trace adequately the evolution of water demand in the future, most of the previous studies used the artificial neural network such as Bougadis and al. [4] and Ghiassi and al. [5]. From their part, Nasseri and al. [6], Yi Wu and Yan [7] applied the genetic algorithm to achieve the same goal. Moreover, to predict the monthly water consumption for the city of Tehran, Nasseri and al. [6] applied also the kalman filter. Using the same technique, Jacobi and al. [8] tried to predict water demand for the city of Beijing.

The box and Jenkins methodology known by its simplicity and effectiveness has also been applied in a large package of work such as that of Maidment [9]. Based on monthly data from January 1978 to December 1987, Al Dhowalia [10] tried to predict the future water demand for the city of Riadh referring to this methodology. By applying the same procedure the aim of Lawgali's [11] paper was to forecast the water demand for agricultural, industrial and domestic use in Libya.

Data

A source data contains annual water production; distribution and consumption in the state of Tunisia were obtained from statistical reports of SONEDE (National Water Distribution Utility), Statistical Division over the years 1971-2010 (A total of 40 observations by series).

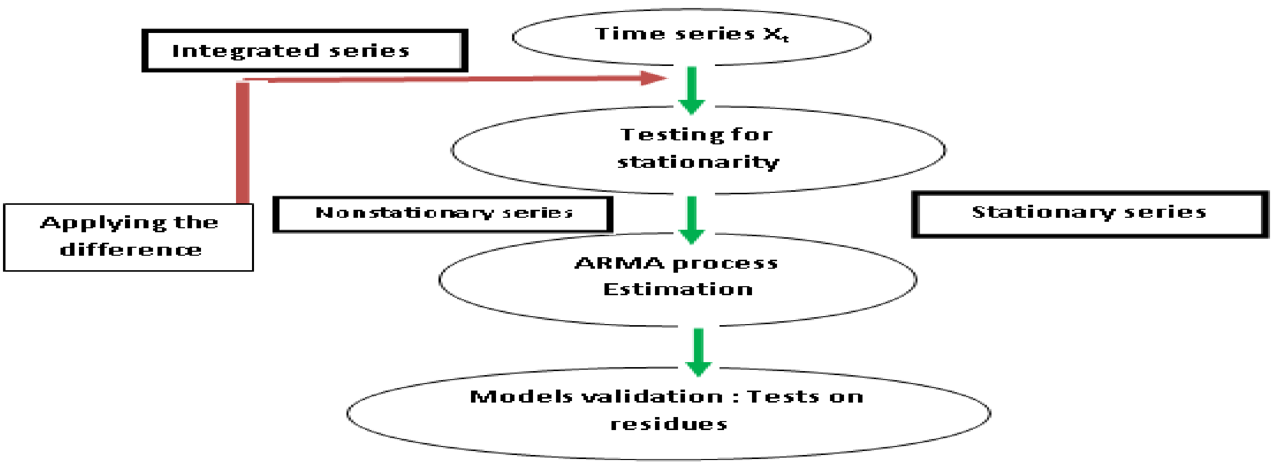
Fig.1 reveals the dramatic growth in water production in Tunisia. In 1971, production was approximately 100 million m^3; by 2010 it had grown five times to reach over 500 million m^3. Fig. 2 shows that in 1971, the volume of annual water distribution in Tunisia was about 98 million m^3. This volume was grown by nearly five times in the same period. At that time, annual water consumption grew by about 5 times from 74 million m^3 to 387.7 million m^3 Fig. 3. graphics superposition shows the existence of an important volume of water waste due to the difference between water production and water distribution from the first hand moreover water distribution and water consumption from the second hand.

Methodology and Empirical Results

Referring to the Box-Jenkins methodology, we adopt a systematic study for each of the three series (production, distribution and consumption of water in Tunisia) basing on their characteristics. This study aims to determine the most appropriate ARIMA (Autoregressive Integrated Moving Average) representation. The ARIMA models provide a useful framework to understand how the water time series is generated subsequently, we can pass to trace the most accurately and adequately evolution of water demand, distribution and consumption in the future. In other words, we seek to establish the closest prediction to the reality. This approach is based on four basic steps namely:

- Identification of parameters
- Estimation
- Validation
- Prediction

The procedure of modeling Time Series by an ARMA model is reported in scheme.



Modeling Time Series by an ARIMA model

Testing for Stationarity

A time series X_t (t =1,…, T) is defined usually as a collection of observations sampled at equally-spaced and ordered time points. Prior to undertaking estimation and forecasting, it was essential to study the characteristics of the time series. In other words, we should analyze its mean and its variance. Indeed if those characteristics vary over the time, the series is said to be nonstationary; otherwise and in presence of an invariant stochastic process, X_t is considered stationary.

One of the major criticisms of non-stationary series, also called integrated series or series with presence of unit root, is that we can study its comportment only in the concerned period and we cannot generalize to other periods. Thus, for a forecasting purpose, non-stationary series have a very limited value.On a formalized way, the stochastic process X_t is stationary only if those three conditions are satisfied:

$$E(Xt)=E(Xt+m)=\mu;\quad \forall t\, et\, \forall m \qquad (1)$$
$$Var(Xt)<\infty;\quad \forall t \qquad (2)$$
$$Cov(Xt,Xt+h)=E[(Xt-\mu)(Xt+h-\mu)]=\gamma_h \qquad (3)$$

There are two general approaches to testing for stationarity, formal and informal.

Informal Tests

Informal tests are based essentially on the visual analysis of the graph and the correlogram of the time series.

Graphic Analysis

A first indication about the stationarity of our series can be provided by the graph analysis. In fact, visually, the different series (Fig. 1, 2, 3) exhibit an overall upward trend and even the application of the logarithm transformation to our series (Fig. 4, 5, 6) does not solve this problem and no significant changes on those trends are detected. This observation leads us to suspect the presumption of non-stationarity of our time series. This intuition can be reinforced by the analysis of correlograms.

A brief overview of the different correlograms (Fig. 7, 8, 9) confirms and strengthens the presumption of non-stationarity. Indeed, the autocorrelations are significantly different from zero and decreases very slowly. In addition, only the first partial autocorrelation for each of the three correlograms is significantly different from zero.

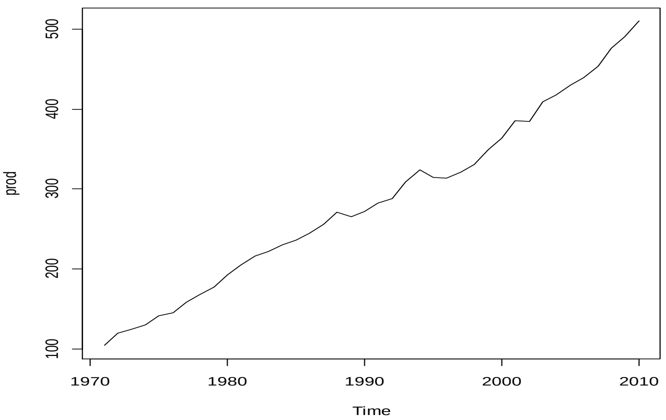


Fig. 1: Water production in Tunisia

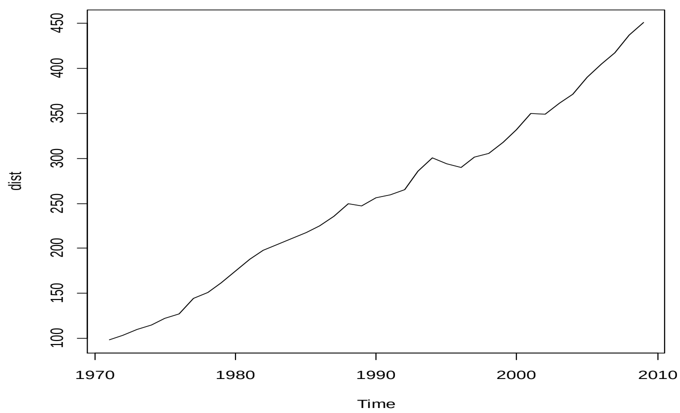


Fig. 2: Water distribution in Tunisia

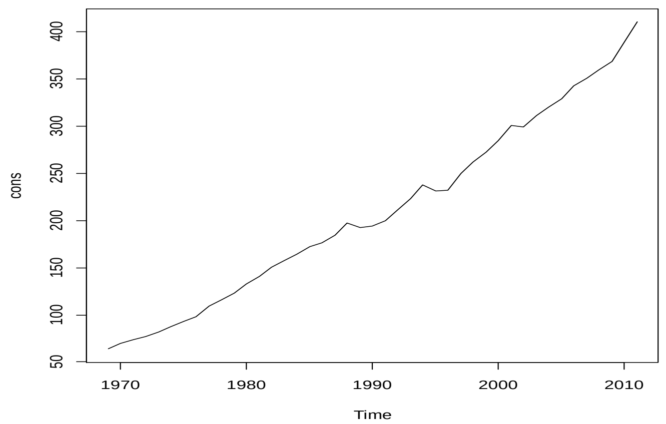


Fig. 3: Water consumption in Tunisia

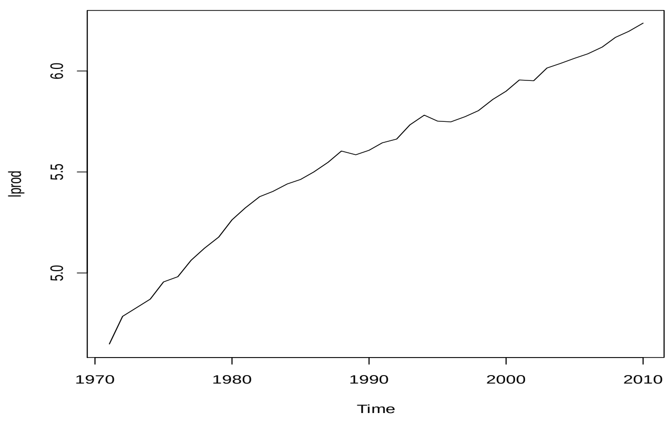


Fig. 4 : Logarithmique transformation of the production series

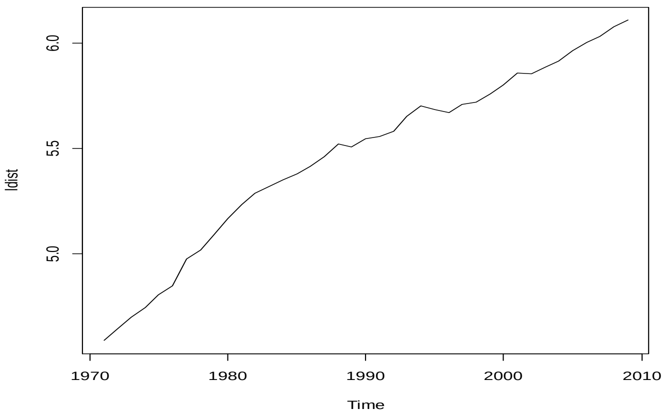


Fig. 5: Logarithmique transformation of the distribution series

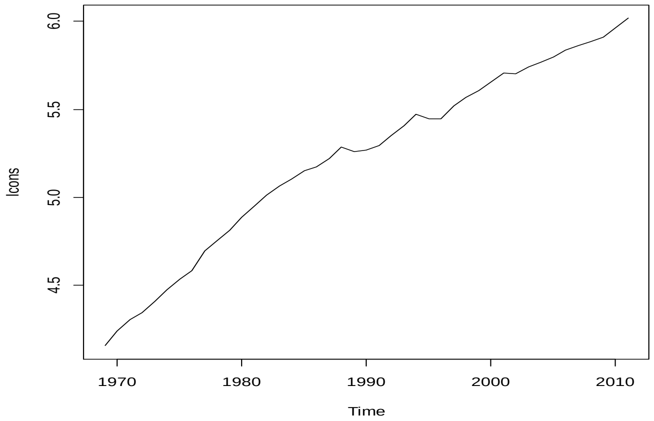


Fig. 6: Logarithmique transformation of the consumption series

Correlogram Analysis

Then, it becomes necessary to verify this intuition by applying more efficient and reliable statistical tests called formal tests.

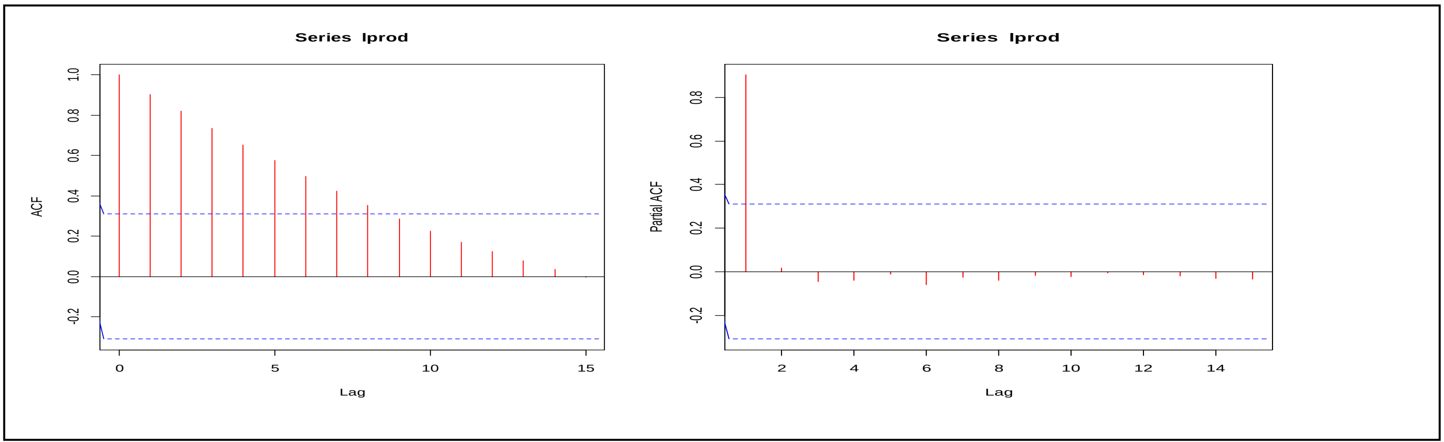


Fig. 7: lprod series correlogram

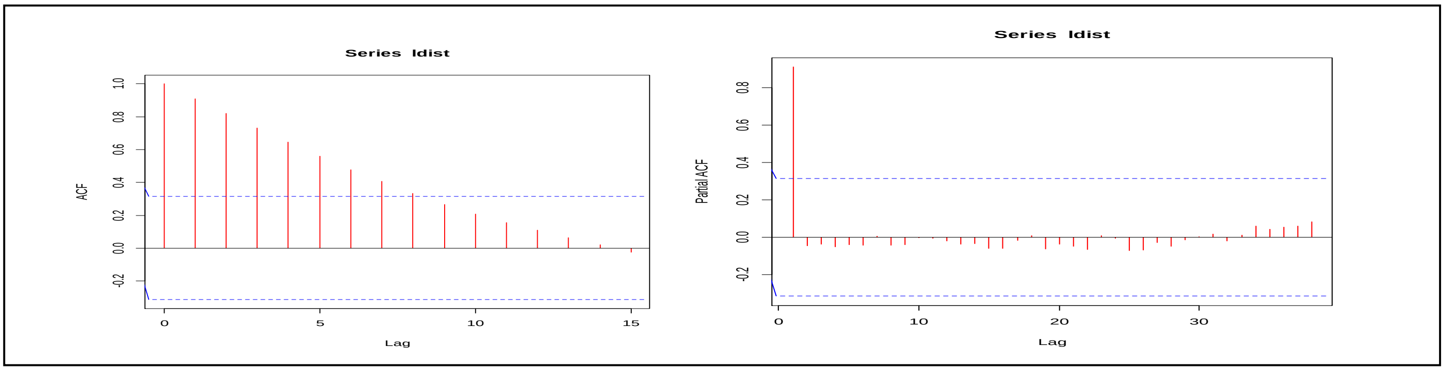


Fig. 8: ldist series correlogram

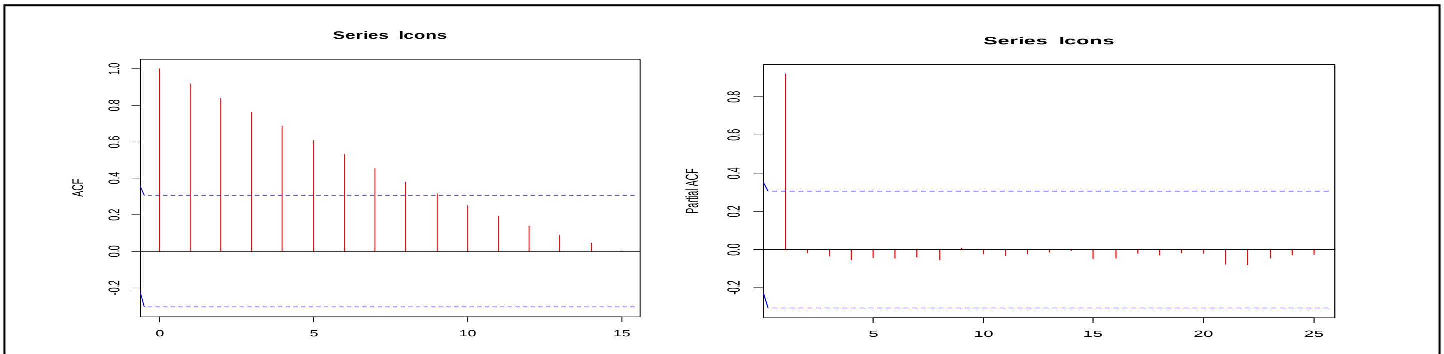


Fig. 9: lcons series correlogram

Formal Tests

Several stationarity tests are developed in the literature such as Dickey-Fuller test (DF), Augmented Dickey Fuller test (ADF), Philips-Perron test, KPSS test …Generally those tests can be classified into two groups. The first group is based on the null hypothesis of presence of unit root (non-stationary) wich belongs the ADF test and another one based on the null hypothesis of no unit root wich belongs the KPSS test (1992). In our analysis we apply three tests (ADF, PP, and KPSS) to test the stationarity of our series. Table 1 summarizes the results of those tests.

Table 1 : Stationarity tests

	ADF		PP		KPSS	
Lcons	-2.4412	(0.3994)	-3.7587	(0.8959)	2.1644	(0.01)
Dlcons	-2.8528	(0.2375)	-36.0947	(0.01)	0.7658	(0.01)
Ddlcons	-5.088	(0.01)*	-45.4091	(0.01)*	0.0465	(0.1)*
Lprod	-2.8848	(0.2262)	-6.8142	(0.7013)	2.0084	(0.01)
Dlprod	-3.2845	(0.08884)	-40.3974	(0.01)	0.7265	(0.01114)
Ddlprod	-5.0815	(0.01)*	-13.8731	(0.01)*	0.1143	(0.1)*
Ldist	-2.7804	(0.2678)	-3.6216	(0.9029)	1.9398	(0.01)
Dldist	-2.4062	(0.4146)	-34.3627	(0.01)	0.6784	(0.01551)
ddldist	-4.3404	(0.01)*	-44.9768	(0.01)*	0.0286	(0.1)*

(.)p-value

*Series is stationary in the level of 5%

The examination of this table reveals that none of our variables are stationary in levels. Therefore, we proceed of the differentiation of all variables. Again, the variables in the first difference are non-stationary and we move to the second difference. Thus, it appears that in the level of 5% all variables are integrated of order 2.Fig. 10, 11, 12 of the series in the second difference confirm clearly the results found.

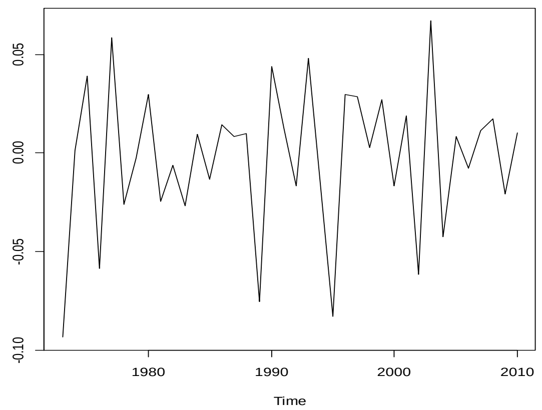


Fig. 10: Evolution of the series ddlprod

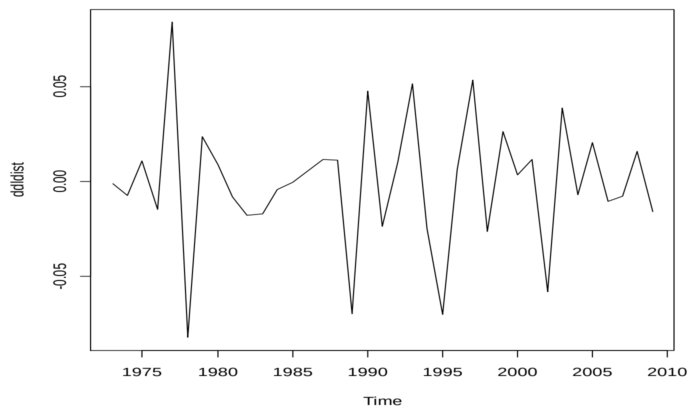


Fig. 11: Evolution of the series ddldist

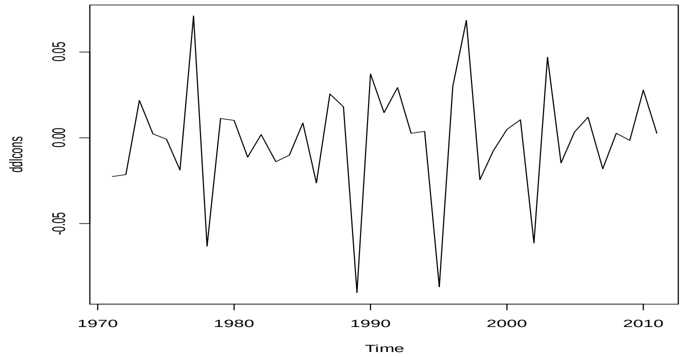


Fig. 12: Evolution of the series ddlcons

Parameters Identification

After obtaining stationary series we can proceed immediately to the orders identification of the sought process by examining the correlograms of each series. In other words, we must determine the order p and q of the ARMA process.The examination of the correlogram of the series ddlprod (Fig. 13) shows that only the first partial autocorrelation is significantly different from zero then we can deduce that p = 1. For the simple autocorrelations, the two first out from the confidence interval so we can conclude that q = 2. After this step, we identify five processes: AR(1), MA(1), ARMA (1,2), ARMA (1,1) and MA (2). Similarly, the examination of the correlogram of the series ddldist (Fig. 14) shows that only the first partial autocorrelation is significantly different from zero then we can deduce that p = 1. For the simple autocorrelations, the two first out from the confidence interval so we can conclude that q = 2. After this step, we identify five processes: AR(1), MA(1), ARMA (1,2), ARMA (1,1) and MA (2). The examination of the correlogram of the series ddlcons (Fig. 15**)** shows that the first two partial autocorrelation are significantly different from zero and we can deduce that p = 2. For the simple autocorrelations, also the two first out from the confidence interval so we can conclude that q = 2. After this step, we identify eight processes: AR(1), MA(1), ARMA (2,2), ARMA (1,1), MA (2), AR(2), ARMA (1,2) and ARMA (2,1). that will be estimated in the next step.

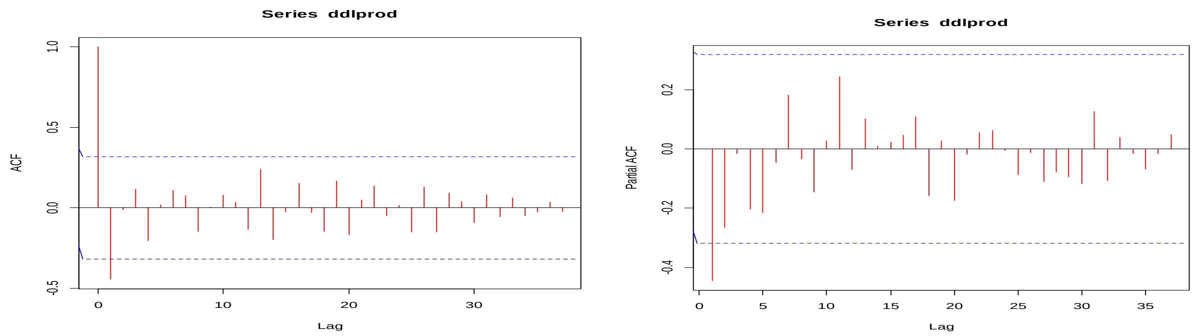


Fig. 13 : Corrélogramme de la série ddlprod

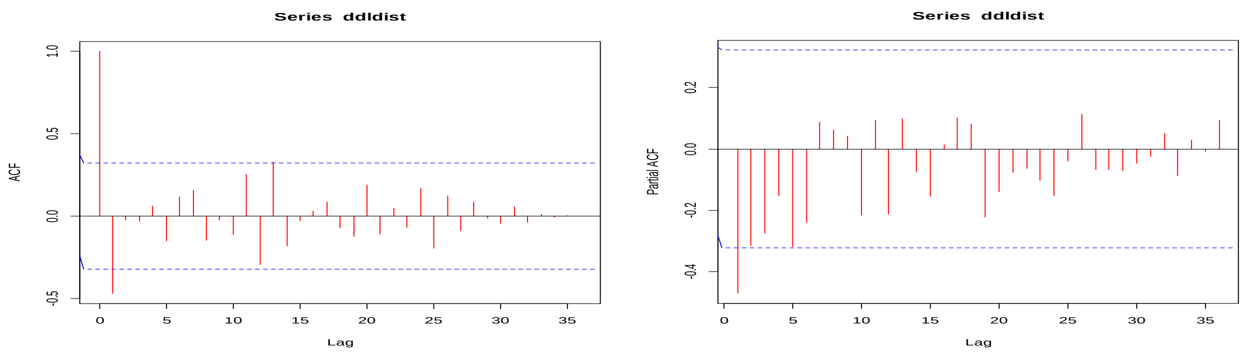


Fig.14 : Corrélogramme de la série ddldist

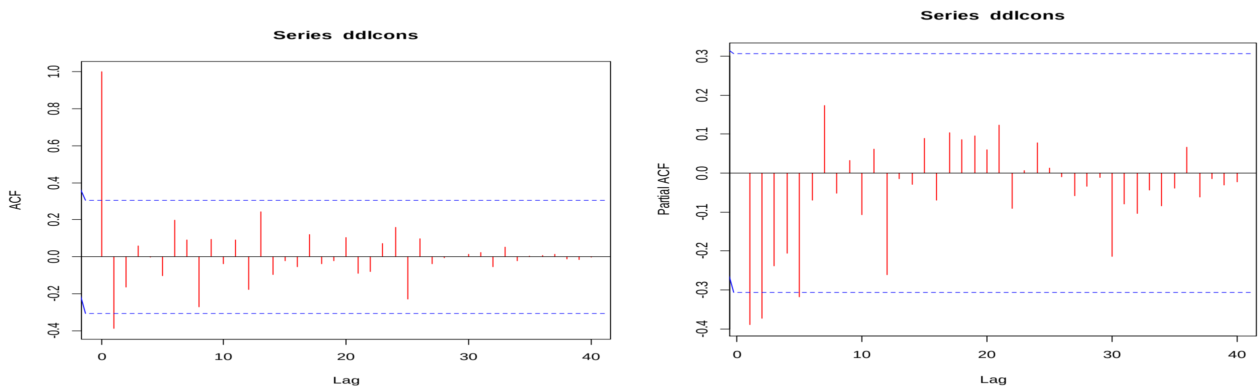


Fig. 15 : Corrélogramme de la série ddlcons

Validation and Estimation of Models

To choose the suitable orders of p and q, we proceed to estimate all possible combinations of ARMA models and then we retain the model that minimizes the criterion information (in our case we use the Akaike criterion "AIC"). Tables 2, 3 and 4 depict the various levels of AIC for each ARMA model.

At the end of this procedure we choose a MA(1) representation to model the production series (ddlprod). The model MA(1) is also valid for modeling the distribution series (ddldist) and the consumption series (ddlcons).Thereafter, we pass to test the residuals of those models to ensure their adequacy.

Table 2: Production series (ddlprod)

Model	AR (1)	MA (2)	MA (1)	ARMA (1,2)	ARMA(1,1)
AIC	-148.23	-155.15	-156.7	-156.2	-155.08

Table 3: Distribution series (ddldist)

Model	AR (1)	MA (2)	MA (1)	ARMA (1,2)	ARMA(1,1)
AIC	-150.76	-157.88	-159.88	-155.88	-157.88

Table 4: Consumption series (ddlcons)

Model	AR(2)	AR(1)	MA(2)	MA(1)	ARMA(2,2)	ARMA(1,2)	ARMA(2,1)	ARMA(1,1)
AIC	-169.8	-165.3	-175.85	-177.77	-175.7	-176.14	-175.29	-175.82

Residuals Tests

To assess the good quality of the estimation, residuals tests are often used. To diagonise the various abnormalities of residues we proceed to test the absence of autocorrelation and normality tests. Residues of the adequate process present several qualities such as independence and normality.

Test of Absence of Autocorrelation

By definition, the residuals autocorrelation problem arises when residuals covariance is different from zero. Thus, in the presence of this problem, the variance-covariance matrix of residuals, the error terms of different observations are not independent and the estimators obtained by applying the Ordinary Least Square method despite being unbiased, they don't have a minimum variance.In general, a well estimated process is characterized by the estimated residuals $e_t = \frac{\hat{\Phi}(L)}{\hat{\Theta}(L)}$ that behave like a white noise. In other words, by a test of absence of autocorrelation, we seek to test the following two hypotheses:

H_0: e_t *is a white noise*

H_1: e_t *is not a white noise*

The Ljung-Box Statistic which can be written in the following form:

$$Q(k) = n(n+2)\sum_{j=1}^{k}\frac{\hat{\rho}^2(j)}{(n-j)}$$

(with n is the total number of period) is often used to test the hypothesis of serial independence. This statistic is distributed according to a Chi-Squared distribution with k degree of freedom. The application of this test allows us to accept the null hypothesis (H_0) at the 5% level if $Q(k)$ is less than the 0.95 quantile of the corresponding chi-squared (the critical probability " p-value" must be high and above the fixed level. The results of this test for our different series are represented in the table 5.

Table 5: Ljung-Box test

	MA(1) (ddlprod)	MA(1) (ddldist)	MA(1) (ddlcons)
p-value	0.796	0.6597	0.4741

The examination of the table 5 reveals that the residuals are not correlated (the p-value is above the 5% level). Thus, the process chosen have a good quality. Following the test of absence of autocorrelation we pass to test the normality of residuals.

Test of Normality of Residuals

Residues of an ideal model must possess the properties of a normal distribution. The literature provides us an important package of normality tests. We apply in these work three tests:

Shapiro-Wilk test: this test proposed in 1965 by Samuel Shapiro and Martin Wilk is based on the following statistic:

$$W = \frac{\left(\sum_{i=1}^{n} a_i x_i\right)^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

We reject the hypothesis of normality if this statistic is very small. In other words, if the p-value is less than the fixed level, we reject the hypothesis of normality.Jarque-Bera test: In 1980 Carlos Jarque and Anil Bera proposed the statistic below to test the residuals normality

$$JB = \frac{n}{6}\left(S^2 + \frac{(K-3)^2}{4}\right)$$

With S represent the empirical skewness

$$S = \frac{\mu_3}{\sigma^3} = \frac{\mu_3}{(\sigma^2)^{3/2}} = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^3}{\left(\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2\right)^{3/2}}$$

K is the empirical kurtosis

$$k = \frac{\mu_4}{\sigma^4} = \frac{\mu_4}{(\sigma^2)^2} = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^4}{\left(\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2\right)^2}$$

This statistic follows a Chi-Squared distribution with 2 degrees of freedom.
Anderson-Darling test: Theodore Anderson and Donald Darling this test in 1952.

Referring to the table 6 reported, the Anderson-Darling test has shown that the residues of the production and the distribution series follow a normal distribution at the statistical level of 5%. The same test confirms that the residues of consumption series follow the normal distribution only at the level of 1%.

Table 6 : Normality test

	MA(1) (ddlprod)	MA(1) (ddldist)	MA(1) (ddlcons)
Anderson-Darling	0.05556	0.1315	0.04799
Shapiro-Wilk	0.04498	0.1517	0.0549
Jarque-Bera	0.1698	0.1753	0.03553

The Shapiro-Wilk test noted that the residues of the distribution and the consumption series follow the normal distribution at the statistical level of 5%, while the residues of the production series follow a normal distribution at the level of 1%.
The Jarque-Bera test showed that at the statistical level of 5%, only the residues of the production and distribution series follow a normal distribution. Against the residues of the consumption series follow a normal distribution at the 1% level.

In view of all the foregoing, we can conclude that the residues of all the series follow a normal distribution at the statistical level of 1% and thereafter are estimated process are of a good quality and can be validated. We can then proceed to the forecast step.

Forecasting

The prediction step takes according to Dominique Geoffray major challenge. The reasons are twofold:

- For the planner interested by the mobilization of water resources and its investment program: this one seek to anticipate and realize in advance extensions for infrastructure mobilization, supply, storage and distribution to cope with the socio-economic development planned for urban areas in order to not fall into a water deficit situation.

- For the manager of water distribution service: without an efficient and adequate prediction of water demand, he wouldn't be able to project into the future its operating accounts and the sales prices of water.

The prediction of the three series studied over a horizon of twenty years (2010-2030) is shown in the fig. 16.

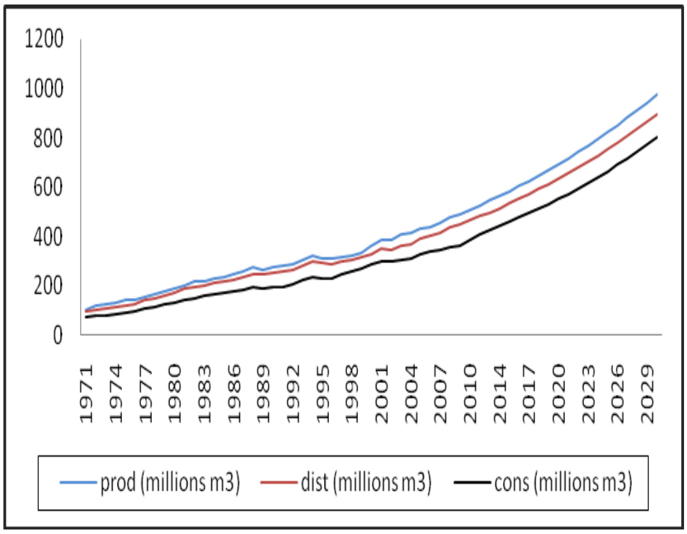


Fig. 16: Evolution of annual production, distribution and consumption of fresh water

Visual analysis of this figure shows that the three series of production, distribution and consumption of water continue to increase in the future. The difference between the production and the distribution series represents the volume of adduction waste. While the difference between the distribution and the consumption volume represents the volume of distribution losses.

This evolution, which seems, at first glance, easeful and not worrying since the consumption level is below the production and the distribution levels, is rather a sign of an alarming and critical situation in several levels.

Firstly, the level of losses increases more and more. The reasons to explain this situation are twofold. From the first hand, the aging of infrastructure can be a direct cause. On the second hand, the very limited infrastructure capacity can lead directly to a considerable loss of water.

Secondly, the relevant development in the production and in the consumption may not lead to serious problems for countries reach with water resources but this is not the case for Tunisia Known by the avarice of its water resources. At this level, the Fig. 17 shows that the availability of m³/year/habitant is below 500 (the international standard of a critical situation).

Thereby, and instead of increasing the volume of production to catch the consumption evolution, it becomes more aware to streamline the latter especially if we notice the existence of wasteful and unnecessary use of water. The figure 18 outlines this problem and shows that, despite the avarice of resources, water consumption in liters/day/habitant for the domestic sector exceeds the international standard (50 l/d/h). Hence, it becomes essential to find solutions and effective measures to conserve this precious resource and reduce losses in the domestic sector.

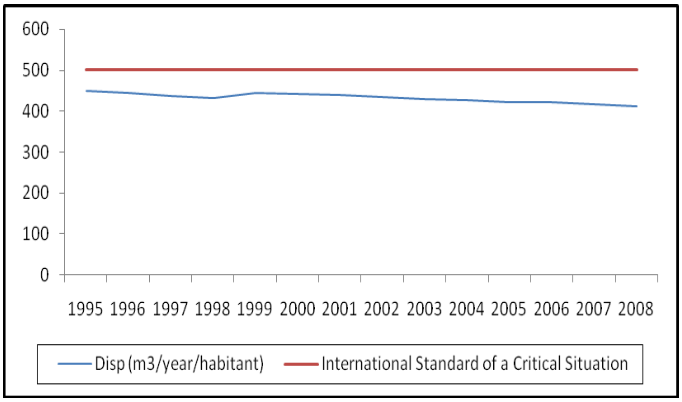


Fig. 17: Evolution of water availability (m³/year/habitant)

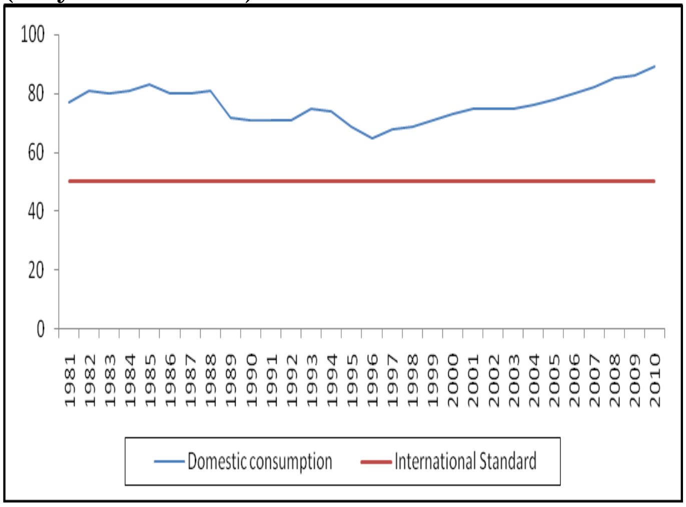


Fig. 18: Evolution of domestic consumption (l/day/habitant)

Conclusion

Without pretending the exhaustiveness, the issue of the situation of water, although it is very interesting and treated by several researchers, we treat it a bit in another way by using the prediction of Box and Jenkins that reveals an overall upward trend in the production, distribution and consumption of water in Tunisia and a bullish gap reflecting losses on adduction and on distribution. This clearly shows that the Tunisian State water policies are still too focused on offer to satisfy the increasing water demand, which leads to overexploitation of scarce resources. Do not forget that this requires improving the efficiency of the water supply network and reduce, thus, water losses (leaks on network, wastage among users). Clearly said, Tunisia needs to develop a more sustainable development scenario which can be achieved only gradually through the necessary reforms working on two axes: economy of water and integrated management of networks and water resources.

At this stage, and in a context of structural shortage ((less than 500 m3/year per capita) and face increasing uncertainties related to climate change with population growth, the development of tourism, industry and irrigated land, it is clear that the growth of water demand will remain strong. Hence, it would be better that the Tunisian state work out plans of action allowing to use resources in a more efficient and optimal way, to manage better different manners and encourage subscribers to rationalize their domestic usage.

References

1. Howe CW, Line weaver FP (1967) The impact of price on residential water demand and its relation to system design and price structure. Water Resources Research, 3:13-32.
2. Gibbs KC (1978) Price variable in residential water demand models. Water Resources Research, 14:15-18.
3. Foster HS, Beattie BR, (1979) Urban Residential Demand for water in the United States. Land Economics 55.1.
4. Bougadis J (2005) Short-term municipal water demand forecasting, Hydrological Processes, Volume 19, 137–148
5. Ghiassi M, Zimbra D, Saidane H (2008) Urban Water demand forecasting with a dynamic artificial neural network model. J. Water Resources Planning and Management, 134:138-146.
6. Nasseri M (2011) Forecasting monthly urban water demand using Extended Kalman Filter and Genetic Programming. Expert Systems with Applications, 38:7387-7395.
7. Yi Wu Z, Yan X, (2010) Applying genetic programming approaches to short-term water demand forecast for district water system. Water Distribution Systems Analysis, 1498-1506.
8. Jacobi M. et al (2007) Water demand forecasting using kalman filtering. Applied Simulation and Modeling, 199-202.
9. Maidment DR, et al (1985) Transfer Function Models of Daily Urban Water Use. Water Resources Research 21: 425-432.
10. Al-Dhowalia KH (1996) Modeling municipal water demand using box-jenkins technique. Jkau:eng. sct., 8:61-71
11. Lawgali FF (2008) Forecasting water demand for agricultural, industrial and domestic use in Libya. International Review of Business Research Papers, 4:231-248

-