

Comparison of Parametric and Nonparametric Techniques for Water Consumption Forecasting

Bisher M. Iqelan

Abstract— Forecasting water consumption is an important foundation for the development of water demand management strategy. In this paper, a parametric time series exponential smoothing state space (ETS) method was employed to monthly water consumption for Gaza City/Palestine. It is compared with nonparametric Reweighted Nadaraya-Watson (RNW) method. The root mean square error (RMSE) and mean absolute scaled error (MASE) are used for comparing the forecasting accuracy. This paper finds statistically significant evidence showing that the ETS state space model outperforms RNW approach at forecasting water consumption.

Index Terms— Exponential smoothing, Nadaraya-Watson estimator, Nonparametric, Parametric, Reweighted, State space models, Water consumption forecast.

1 INTRODUCTION

With the rapid growth of world population, problems of water consumption are threatening and involving the attention of every human being. The implementation of long-term water consumption forecasting stands as one solution to the problem of water consumption. A good path to water management and design is through water prediction and water consumption forecasting. They can give crucial water data required in future development and predict other types of water problems. As a result, water consumption forecasting holds a significant theoretical and practical value that raises the predictive accuracy of water consumption. There are different conventional approaches to water consumption forecasting including parametric and nonparametric techniques.

Located along the Mediterranean Sea, the Gaza Strip is a very small Palestinian territory. It is 362-kilometer square and is home to a population of more than 2 million people. Therefore, it ranks as one of the most densely populated places around the world. As the siege continues to escalate, normal life would not be viable since the standards of living are declined. This paper focuses on the city of Gaza as it contains most of the population. The water crisis in Gaza is resulted from over-pumping of the aquifer, water loss, and salt accumulation. Water consumption forecasts can provide crucial insights into the future needs of Gaza's life requirement. This paper attempts to provide a coherent water consumption forecasting model for Gaza policy makers as well. The objective of this study is to evaluate the application of two different forecasting techniques for Gaza City. These techniques include both parametric exponential smoothing state space (ETS) and nonparametric Reweighted Nadaraya-Watson (RNW).

The literature on forecasting water consumption is rich since it is a very important field of study. Maidment [20] applied short-term Box and Jenkins models on daily municipal water use study. These ARIMA time series models were used as a function of rainfall and air temperature.

The technique of artificial neural networks (ANN) was used in several civil engineering applications. In addition, [27] conducted a study to forecast daily municipal water demand using time series models. Jowitt and Xu [17] used exponential smoothing and autoregression to forecast water demand. Shvarster et al. [26] provided a model for hourly water use forecast based on time series analyses and pattern recognition. Moreover, [15] presented a technique to forecast daily water use. This technique includes two steps; the first step uses an exponential smoothing algorithm to forecast monthly average water use. The second step contains the monthly average forecast that is employed to obtain a daily forecast. Zhou [33] developed time series models for daily water consumption in Melbourne, Australia. Khan and Coulibaly [19] provided a comparative study between time series models in the forecasting of the water level of a lake. Horielova and Zadachyn [10] concluded that ARIMA models and ANN models are most adequate in forecasting water consumption of large cities.

Data Source: We use the data set of water consumption in Gaza City / Gaza Strip/ Palestinian Territories from Municipality of Gaza <https://www.gaza-city.org>. We considered monthly data sets for water consumption from January 1990 to July 2016.

The section above contains a summary of water consumption problem with an application on Gaza City. In addition, an overview of previous literature is also included. The remainder of the paper is organized as follows. Section 2 presents the parametric exponential smoothing state space models and the nonparametric Reweighted Nadaraya-Watson methods used in this paper. Section 3 discusses the methodology used in the selection of parametric and nonparametric techniques and forecasting evaluation. Data description and empirical results are provided in section 4. Section 5 presents the summary and conclusion of this paper.

2 MATERIALS AND METHODS

2.1 The Reweighted Nadaraya-Watson Estimator

Let $(X_t, Y_t), t = 1, 2, \dots, n$ be observations of a bivariate random variable (X, Y) and assume that all are continuously

• Bisher M. Iqelan is currently one of the staff members at Department of Mathematics, College of Science, The Islamic University of Gaza, Gaza Strip, Palestine. E-mail address: biqelan@iugaza.edu.ps

distributed with joint probability density function $f(x, y)$. This section presents a kernel regression estimator known as the reweighted Nadaraya-Watson estimator.

Let $f(y|x) = \frac{f(x, y)}{g(x)}$ be the conditional density of Y_t given X_t , where $g(x) = \int f(x, y)dy$ is the marginal density of X_t . The simple nonparametric regression function for Y_t on X_t is defined as

$$m(x) = E(Y_t | X_t = x) = \int \frac{yf(x, y)}{g(x)} dy$$

The basic idea is to estimate this non-parametrically, with minimal assumptions about $m(x)$. Suppose the relationship between the dependent variable Y and the independent variable X , the regression equation can be written as,

$$Y_t = m(X_t) + \varepsilon_t, \quad t = 1, 2, \dots, n$$

Where $\varepsilon_t, t = 1, 2, \dots, n$ are called the measurement errors such that $E(\varepsilon_t | X_t) = 0$ and $E(\varepsilon_t^2 | X_t = x) = \sigma^2(x)$. Using the kernel estimation, the regression mean function $m(x)$

is estimated by $\hat{m}(x)$, where

$$\hat{m}(x) = \int y \hat{f}(y|x) dy$$

A class of kernel-type estimator is called the Nadaraya-Watson (NW) estimator which is one of the most widely known and used for estimating $f(y|x)$. The NW estimator has been suggested independently by [29] and [22].

Let $\hat{f}_{NW}(y|x)$ stands for Nadaraya-Watson kernel estimation and defined as,

$$\hat{f}_{NW}(y|x) = \frac{\sum_{t=1}^n K_h(x - X_t) Y_t}{\sum_{t=1}^n K_h(x - X_t)},$$

Where $k(\cdot)$ is a kernel function, $K_h(\cdot) = \frac{1}{h} K\left(\frac{\cdot}{h}\right)$ and

$h = h_n > 0$ is the bandwidth. Consequently, the NW estimator, $\hat{m}(x)$, of the regression mean function $m(x)$ can be defined as

$$\hat{m}_{NW}(x) = \frac{\sum_{i=1}^n K_h(x - X_i) Y_i}{\sum_{i=1}^n K_h(x - X_i)}.$$

It is worth noting that there is some weakness in applying the NW estimator especially when the estimation is done near the boundary points or when the design density is highly clustered. For more details see [8], [4], [5], and [32].

To overcome the weakness of the NW estimator, a new class of the kernel estimation of the conditional density function and regression mean function has been proposed by [8]. The new estimator is called the Reweighted Nadaraya-Watson (RNW) estimator, see [4], [6]. The RNW estimator of $f(y|x)$ is denoted

by $\hat{f}_{RNW}(y|x)$ and defined as

$$\hat{f}_{RNW}(y|x) = \frac{\sum_{i=1}^n \tau_i(x) K_h(x - X_i) K_h(y - Y_i)}{\sum_{i=1}^n \tau_i(x) K_h(x - X_i)},$$

where $\tau_t(x)$ are probability weights functions satisfying the following properties

1. $\tau_t(x) \geq 0$,
2. $\sum_{t=1}^n \tau_t(x) = 1$,
3. $\sum_{t=1}^n \tau_t(x) (X_t - x) K_h(x - X_t) = 1$.

$$\tau_t(x) = \frac{1}{n(\lambda(X_t - x) K_h(x - X_t) + 1)}$$

Where λ is the unique minimizer of $L(\lambda)$, Where $L(\lambda)$ is defined by

$$L(\lambda) = \sum_{t=1}^n \log(1 + \lambda(X_t - x) K_h(x - X_t)).$$

Moreover, the Reweighted Nadaraya-Watson estimator, $\hat{m}_{RNW}(x)$, of $m(x)$ is given by

$$\hat{m}_{RNW}(x) = \frac{\sum_{i=1}^n \tau_i(x) K_h(x - X_i) Y_i}{\sum_{i=1}^n \tau_i(x) K_h(x - X_i)}.$$

Recently, [25] studied the RNW and proved the joint asymptotic normality of it estimated at a finite number of conditional points.

2.2 Exponential Smoothing State Space Models

A common class of forecasting models are exponential smoothing techniques. It's simple but very helpful of adjusting time series forecasting. R. G. Brown [3], [9], and [31] have initially introduced these methods in their early works. The idea behind forecasting using exponential smoothing is to assign exponentially declining weights to observations as they go back in age, i.e., recent observations have a larger weight than the old ones.

Pegels [24] was the first to classify exponential smoothing techniques and propose a taxonomy of the trend component and seasonal component. Pegels' taxonomy was later extended by [7], who added damped trend to the classification. This extension is then modified by [15], before the final extension is proposed by [28] who extended the classification to include damped multiplicative trends.

TABLE 1
THE FIFTEEN EXPONENTIAL SMOOTHING METHOD

Trend Component		Seasonal Component		
		N (None)	A (Additive)	M (Mult.)
N	(None)	N,N	N,A	N,M
A	(Additive)	A,N	A,A	A,M
A _d	(Addi. damped)	A _d	A _d	A _d
M	(Multiplicative)	M,N	M,A	M,M
M _d	(Mult. damped)	M _d	M _d	M _d

The classification of exponential smoothing methods shows a total of 15 different methods regarding trend and seasonal components [12], [11]. These methods can be seen in Table 1. Each method is labelled by ordered pair of letters (T,S) showing the type of 'trend' and 'seasonal' components. However, some of the methods are better known with other names. For instance, cell (N,N) is the simple exponential smoothing (or SES). Similarly, cell (A,N) describes Holt's linear method, and the damped trend method corresponds to cell (A_d,N). Also, cell (A,A) and (A,M) are the additive and the multiplicative Holt-Winters' methods, respectively [12]. The following component is to illustrate the additive Holt-winter method (A,A) for an observed time series {y_t}. The following notation follows that in [12],

$$\begin{aligned}
 \text{Level: } \ell_t &= \alpha(y_t - s_{t-m}) + (1-\alpha)(\ell_{t-1} + b_{t-1}) \\
 \text{Growth: } b_t &= \beta^*(\ell_t - \ell_{t-1}) + (1-\beta^*)b_{t-1} \\
 \text{Seasonal: } s_t &= \gamma(y_t - \ell_{t-1} - b_{t-1}) + (1-\gamma)s_{t-m} \\
 \text{Forecast: } \hat{y}_{t+h|t} &= \ell_t + b_t h + s_{t-m+h_m^+}
 \end{aligned}$$

Where *m* is the number of seasons in a year (e.g. for quarterly data *m* = 4, and for monthly data *m* = 12). ℓ_t is the level, b_t is growth, s_t is the seasonal component, $\hat{y}_{t+h|t}$ denotes the *h*-step-ahead forecast and $h_m^+ = \lfloor (h-1) \text{ mod } m \rfloor + 1$ which establishes that the estimates of the seasonal indices for forecasting come from the last year of the sample. (The notation $\lfloor u \rfloor$ means the largest integer not greater than *u*.)

The smoothing parameters, α , β^* and γ , show how fast the

level, trend and seasonality, respectively, fit to new information. For more details regarding the component for all fifteen exponential smoothing methods, see [12] pp.18.

In the state space approach for the exponential smoothing methods, the error terms are smoothed. The corresponding models are called innovation state space models. Note that it is an innovation because all equations in this type use the same error term ϵ_t . For more details, see [1], [2], and [11]. Hyndman [15], [23] and [28] showed that there are two possible underlying state space models, for each of the 15 methods in Table 1. There is one with additive errors and one with multiplication errors. A list of all the 30 possible specifications are described in [12]. An extra letter for error is added in the method notation. Each stated space model is labelled as ETS(·,·,·) for (error, trend, seasonal). So ETS(M, A, A) refers to a model with multiplicative error, additive trend and additive seasonality. To illustrate the idea of the multiplicative error model ETS(M, A, A), let first $\mu_t = \hat{y}_t = \ell_{t-1} + s_{t-m}$ denote the one-step forecast of y_t supposing that all parameters are known. Also, let $\epsilon_t = (y_t - \mu_t) / \mu_t$, so that ϵ_t is a relative error. The general form of the resulting state equations is (See [12] pp. (19-22)

$$\begin{aligned}
 y_t &= (\ell_{t-1} + b_{t-1} + s_{t-m})(1 + \epsilon_t), \\
 \ell_t &= \ell_{t-1} + b_{t-1} + \alpha(\ell_{t-1} + b_{t-1} + s_{t-m})\epsilon_t, \\
 b_t &= b_{t-1} + \beta(\ell_{t-1} + b_{t-1} + s_{t-m})\epsilon_t, \\
 s_t &= s_{t-m} + \gamma(\ell_{t-1} + b_{t-1} + s_{t-m})\epsilon_t,
 \end{aligned} \tag{2.1}$$

By defining the state space vector as $x_t = (\ell_t, b_t, s_t, s_{t-1}, \dots, s_{t-m+1})'$, the system (2.1) can be written in standard state space notation:

$$\begin{aligned}
 y_t &= [1 \ 1 \ 0 \ \dots \ 0 \ 1]x_{t-1}(1 + \epsilon_t), \\
 x_t &= \begin{bmatrix} 1 & 1 & 0 & \dots & 0 & 1 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 \end{bmatrix} x_{t-1} \\
 &+ [1 \ 1 \ 0 \ \dots \ 0 \ 1]x_{t-1} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} \epsilon_t
 \end{aligned} \tag{2.2}$$

where x_{t-1} , is the state space vector such that $x_{t-1} = (\ell_{t-1}, b_{t-1}, s_{t-1}, s_{t-2}, \dots, s_{t-m})'$. Usually the assumption of the error term ϵ_t is that $\epsilon_t \sim \text{NID}(0, \sigma^2)$.

3 MODEL SELECTION

3.1 Bandwidth Selection

The RNW estimator, $\widehat{m}_{RNW}(x)$ is a nonparametric approach which has no need for a distributional assumption. The goodness of the kernel regression estimate depends decisively on the selection of bandwidth, h . When the bandwidth is very small, the estimate will be very close to the original data. If the bandwidth is very large, the estimate will be very smooth, lying close to the mean of the data.

Many suggestions for bandwidth selection have been taken by [30]. Formally, the suitable bandwidth is the one minimizing the mean square error (MSE) of the RNW estimator,

$$MSE(\widehat{m}_{RNW}(x)).$$

3.2 Automatic Forecasting

The parametric ETS method is an automatic forecasting method which is applied to perform forecasting using `ets()` function via the forecast package in R software. The suitable exponential smoothing model is selected among 30 ETS models using maximum likelihood estimator (MLE) and selecting the best model based on information criteria AIC, AICc, and BIC. The model which minimizes the criteria is chosen as preferable for the data. For more details, see [15], [12], [11].

3.3 Forecast Evaluation

There are many different accuracy measures which can be used to select the forecasting model perform carefully. Denote the actual observation at time t by y_t and its forecasted value by \hat{y}_t . The following accuracy measures are considered in Table 2.

TABLE 2
ACCURACY MEASURES

Acronyms	Definition	Formula
MSE	Mean square error	$\sum_{i=1}^n e_i^2 / n$
MAE	Mean absolute error	$\sum_{i=1}^n e_i / n$
RMSE	Root mean square error	\sqrt{MSE}
MAPE	Mean absolute percentage error	$\sum_{i=1}^n 100 \frac{ e_i / y_i }{n}$

where n is the number of observations in the sample. It is worth noting that MSE, MAE and RMSE are scale dependent measures and because of that they should not be used in comparing models estimated from different data sets. Moreover, MAPE measure is based on percentage error which has the benefit of being scale independent but is not valid for series that are very sparse and scattered, as they are undefined or infinite for $y_t = 0$. This makes MAPE inappropriate in this application.

The accuracy measure, suggested by [13], proposes comparing the errors given by the forecasting models to the errors from naïve

models, i.e., a method where the forecast for the observation time t is simply the observation at time $t - 1$. The scaled error would then be defined as,

$$q_t = \frac{e_t}{\frac{1}{n-1} \sum_{t=2}^n |y_t - y_{t-1}|}$$

Measures based on scaled errors have benefit of being neither scale dependent nor implicated by series fluctuating around zero. The mean absolute scaled error (MASE) is given by

$$MASE = \frac{\sum_{t=1}^n |q_t|}{n},$$

Moreover, the MASE is very flexible and has a simple interpretation. If the value is less than one, it means the forecast is more accurate than the naïve forecast.

4 APPLICATION TO REAL DATA

4.1 Data

We have analyzed the monthly water consumption series in Gaza city/Gaza strip/ Palestinian territories from January 1990 to July 2016. We have transformed the data by dividing it by 1000000. Table 3 presents some descriptive statistics for this data set.

TABLE 3
DESCRIPTIVE STATISTICS FOR MONTHLY WATER CONSUMPTION
DATA 1990:1-2016:7

n	Min.	1 st Qu	Mean	Std.dev	3 rd Qu	Max
319	1.379	1.880	2.262	0.528	2.519	3.906

Figure 1 displays a graphical representation of the decomposed series of water consumption. Clearly by studying the plot of the water consumption in the first panel of Figure 1, it is easy to realize that the series does not have a constant mean. This reflects that the water consumption is a non-stationary series. It gives us an indication of the seasonality of water consumption. The series displays periodic attitude every 12 months. Moreover, the third panel in Figure 1 displays the trend of the series. It is the same series, but with seasonality extracted.

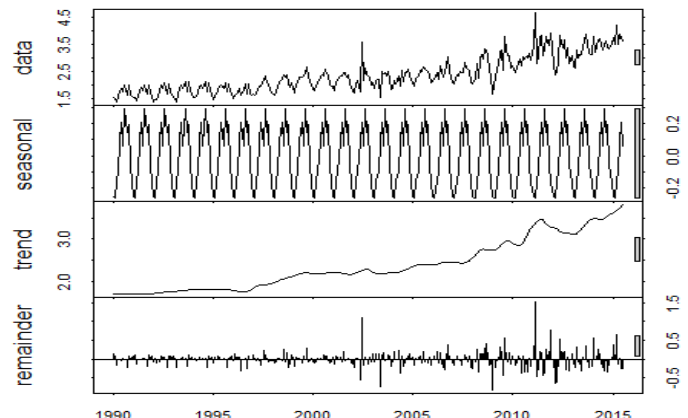


Fig. 1. Decomposition of the water consumption series

We use the first 307 observations from January 1990 to July 2015 as training sample for model prediction, and the remaining 12 observations from August 2015 to July 2016 as post-sample for forecast evaluation.

4.2 Empirical Study

This section establishes the empirical results on the training sample for fitting models of water consumption in Gaza City by using the two different methods, the nonparametric RNW kernel estimation method and the parametric ETS method. R statistical software was used for fitting the models and to forecast water consumption in Gaza City for 12 months from August 2015 to July 2016. The procedure for the nonparametric RNW is based on selecting the suitable bandwidth, h , which minimizes $MSE(\hat{m}_{RNW}(x))$. Table 4 shows different values of bandwidth and the corresponding $MSE(\hat{m}_{RNW}(x))$. Comparing the values to obtain that the minimum value of $MSE(\hat{m}_{RNW}(x))$ is equal to 0.1898 corresponding to the bandwidth $h = 0.4939$.

TABLE 4
COMPARISON OF THE BANDWIDTH WITH CORRESPONDING VALUE OF $MSE(\hat{m}_{RNW}(x))$

$MSE(\hat{m}_{RNW}(x))$	The bandwidth h			
	0.10974	0.21949	0.49385	0.54872
0.36144	0.26491	0.18975	0.99956	

When fitting an exponential smoothing model with a state space approach, the ets() function in forecast package was used. It is used to choose the suitable model automatically on the basis of maximum likelihood method (MLE) and then to calculate the point forecasts for water consumption in Gaza City for the 12 months from August 2015 to July 2016. When applying the function, the results showed that the best performing model was ETS(M, A, A), that is a model with multiplicative error, additive trend and additive seasonality. The general equation of this model was explained in section 2.2 by the system 1.2.

4.3 Forecasting Models

Table 5 compares the actual with the point forecast for the 12 months in the period from August 2015 to July 2016 for water consumption in Gaza City in the Palestinian Territories based on ETS and RNW methods.

TABLE 5
ACTUAL AND PREDECTION FOR WATER CONSUMPTION FOR GAZA CITY IN 2015:8-2016:7

Month	Actual Data	Prediction	
		ETS(M,A,A)	RNW method
Aug 2015	3798939	3973588	3298502
Sep 2015	3994249	3821226	3700635
Oct 2015	3842813	3922643	3085275
Nov 2015	3361297	3678210	3243449
Dec 2015	3956254	3583795	3850790
Jan 2016	3589814	3489380	3505749
Feb 2016	3657156	3516415	3492082
Mar 2016	4229867	3607954	3699882
Apr 2016	3795809	3694175	3266582
May 2016	4151937	3865150	4255222
Jun 2016	3972153	4019894	3544637
Jul 2016	3837802	3917415	3094285

It is quite clear that the forecast error of ETS(M, A, A) technique is much less than that of the RNW technique. This outcome points out that ETS(M, A, A) method fits water consumption data much better than the RNW kernel estimator method. It is interesting to note that the parametric exponential smoothing model outperforms the nonparametric Re-weighted Nadaraya-Watson kernel estimator method. Comparing the results of the measures of forecast accuracy in Table 6, the ETS(M, A, A) model still performs best based on all the measures values.

The predicted values using ETS(M, A, A) is highly accurate as RMSE = 0.2620172 and MASE = 0.613292; it has the lowest RMSE and MASE estimate.

TABLE 6
FORECAST ACCURACY FOR THE ETS AND RNW TECHNIQUES

	MAE	RMSE	MASE
RNW estimator	0.3631308	0.4356086	1.070811
ETS(M,A,A)	0.2079781	0.2620172	0.613292

The smoothing parameters of the exponential smoothing state space model ETS(M, A, A) that adequately fit the water consumption data are listed as follows:

$$\alpha = 0.1502, \beta = 0.0026, \lambda = 0.0001$$

with initial states (the components of the state space vector x_{t-1}) are listed as follows:

$$\ell_{t-1} = 1.7337, b_{t-1} = -0.0031$$

$$s = -0.1573, -0.0519, 0.2035, 0.1129, 0.2763, 0.0998, 0.2131, 0.0693, -0.0907, -0.166, -0.2465, -0.2626$$

where s above refers to the values of the seasonal initials

$S_{t-1}, S_{t-2}, \dots, S_{t-m}$. Putting all these values in the system (2.2) above to obtain the standard state space equations of ETS(M, A, A) model.

Figure 2 shows the forecasts from ETS(M, A, A) model. The thin blue indicates the fitted values, while the forecasts are shown as the dashed thicker red line. The prediction intervals are clarified by the light 95% and dark 80% grey confidence bands.

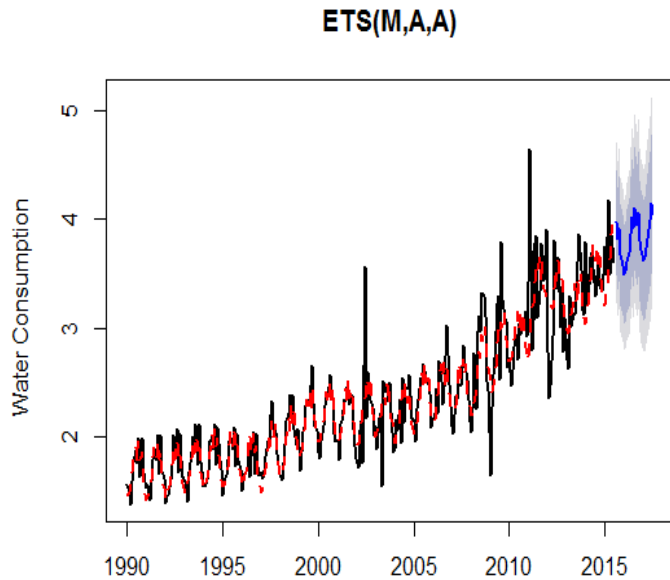


Fig. 2. Forecast for h=12 months ahead with ETS (M,A,A) model.

However, to make prediction intervals using exponential smoothing techniques, the prediction intervals require that the forecast errors should be uncorrelated and normally distributed with mean zero and constant variance.

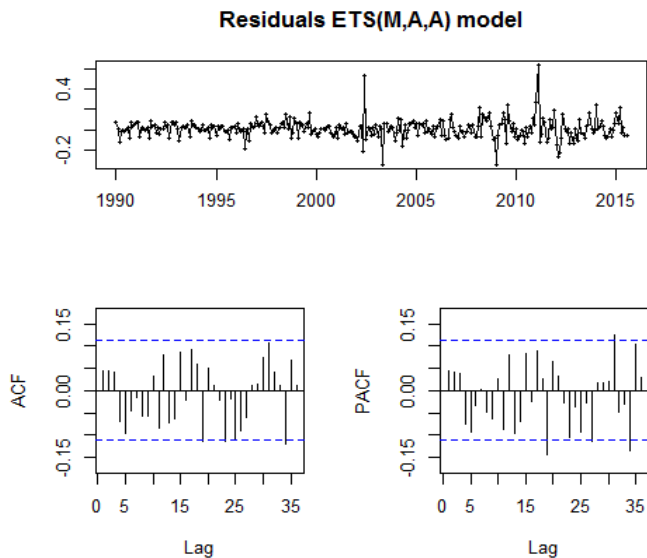


Fig. 3. Residual diagnostics for ETS(M,A,A) model.

The sample correlation in Figure 3 displays that the most sample autocorrelation coefficients of the residuals are within the confidence limits, consequently the residuals are white noise reflecting that ETS(M, A, A) is adequate.

To confirm the evidence of autocorrelations, the Box-Ljung test p-value result in Table 7 displays that there is evidence of no autocorrelations in the forecasts errors. Furthermore, the Jarque Bera test p-value shows that there is a strong evidence for normality of forecasts errors. Moreover, the test of homoscedasticity p-value proves the evidence that the variance is constant.

TABLE 7
RESIDUAL DIAGNOSTICS TESTS

Residual Test	p-values
Residual Autocorrelation test Box-Ljung test	$7.695 * 10^{-05}$
Residual Normality test Jarque Bera Test	$< 2.2 * 10^{-16}$
Homoscedasticity test Box-Ljung test (Squared Residuals)	0.01018

5 CONCLUSION REMARKS

The main objective of this paper is to forecast water consumption in Gaza City/ Palestinian Territories from January 1990 to July 2016. Two approaches are used attempting to obtain a suitable fit. They are a parametric exponential smoothing state space (ETS) model and a nonparametric Reweighted Nadaraya-Watson (RNW) kernel estimator method.

The analysis recognizes that the ETS state space model has good prediction results than the RNW approach based on both mean absolute error (MAE), root mean square error (RMSE) and mean absolute scaled error (MASE). However, we can note that although ETS state space models are not widely applied in forecasting of water consumption, the empirical results assert the importance of their application

ACKNOWLEDGMENT

I would like to thank Raed B. Salha for the very useful discussion and suggestions he provided in this study. I would also like to thank Municipality of Gaza for providing the water consumption data.

REFERENCES

- [1] Anderson, B.D.O and Moore, J.B. "Optimal Filtering". Prentice-Hall, Englewood Cliffs. 1979.
- [2] Aoki, M. "State Space Modeling of Time Series". Springer-Verlag, Berlin. 1987.
- [3] Brown, R. G. "Statistical Forecasting for Inventory Control". McGraw-Hill, New York. 1959.
- [4] Cai, Z. "Regression Quantiles for Time Series Data". Econometric Theory, 18, 169-192. 2002.
- [5] Cai, Z. "Weighted Nadaraya-Watson Regression Estimation". Statistics and

- Probability Letters, 51, 307-318. 2001.
- [6] DeGooiger, J., & Zerom, D. "On Conditional Density Estimation". *Statistical Neerlandica*, 57, 159-176. 2003.
- [7] Gardner, Jr, E. S. "Exponential Smoothing: The State of the Art", *Journal of Forecasting* 4, 1-28. 1985.
- [8] Hall, P., Wolff, R. and Yao, Q. "Methods for Estimating a Conditional Distribution". *Journal of the American Statistical Association*, 94, 154-163. 1999.
- [9] Holt, C. C. "Forecasting Trends and Seasonals by Exponentially Weighted Averages". O.N.R. Memorandum 52/1957, Carnegie Institute of Technology. 1957.
- [10] Horielova, K. and Zadachyn V. "Planning of City Water Supply System Modernization Based on Water Consumption Forecast". *Information Technologies in the Economy, Ecology, Medicine, and Education*. 2016.
- [11] Hyndman, R. J., & Athanasopoulos, G. "Forecasting: Principles and Practice". OTexts. 2014.
- [12] Hyndman, R. & Khandakar, Y. "Automatic Time Series Forecasting: The Forecast Package for R". *Journal of Statistical Software* 26(3). 2008.
- [13] Hyndman, R. J. and A. B. Koehler "Another Look at Measures of Forecast Accuracy". *International Journal of Forecasting*, 22, 679-688. 2006.
- [14] Hyndman, R. J., A. B. Koehler, J. K. Ord and R. D. Snyder. "Forecasting With Exponential Smoothing: The State Space Approach". Springer-Verlag, Berlin. 2008.
- [15] Hyndman, R. J., A. B. Koehler, R. D. Snyder and S. Grose. "A State Space Framework for Automatic Forecasting Using Exponential Smoothing Methods". *International Journal of Forecasting*, 18(3), 439-454. 2002.
- [16] Jones, M. C., Marron, J. S. and Sheather, S. J. "A Brief Survey of Bandwidth Selection for Density Estimation". *Journal of the American Statistical Association* 91, 401-7. 1996.
- [17] Jowitt, P. W., and Xu, C. "Demand Forecasting for Water Distribution Systems." *Civ. Eng. Syst.*, 9, 105-121. 1992.
- [18] Kenward, T. C., and Howard, C. D. D. "Forecasting for Urban Water Demand Management." *Proc., 26th Annual ASCE Conference on Water Resources Planning and Management*, Reston, Va., Chap. 9A188. 1998.
- [19] Khan, M., Coulibaly, P. "Application of Support Vector Machine in Lake Water Level Prediction". *Journal of Hydrological Engineering*, 11 (3), 199-205. 2006.
- [20] Maidment, D., Miaou, S., Crawford, M. "Transfer Function Models of Daily Urban Water Use". *Water Resources Research* 21 (4), 425-432. 1985.
- [21] Marron, J. S. and Wand, M. P. "Exact Mean Integrated Squared Error". *Ann. Statist.* 20, 712-36. 1992.
- [22] Nadaraya, E. A. "One Estimating Regression". *Theory of the Probability of its Applications*, 9, 141-142. 1964.
- [23] Ord, J. K., A. B. Koehler and R. D. Snyder. "Estimation and Prediction for a Class of Dynamic Nonlinear Statistical Models". *Journal of the American Statistical Association*, 92, 1621-1629. 1997.
- [24] Pegels, C. C. "Exponential Forecasting: Some New Variations". *Management Science* 15(5), 311-315. 1969.
- [25] Salha, R. and EL Shekh Ahmed, H. "Reweighted Nadaraya-Watson Estimator of Regression Mean". *International Journal of Statistics and Probability*, 4(1), 138-147. 2015.
- [26] Shvarster, L., Shamir, U., and Feldman, M. "Forecasting Hourly Water Demands by Pattern Recognition Approach." *J. Water Resour. Plann. Manage.*, 1196, 611-627. 1993.
- [27] Smith, J. "A Model of Daily Municipal Water Use for Short-Term Forecasting". *Water Resources Research* 24 (2), 201-206. 1988.
- [28] Taylor, J. W. "Exponential Smoothing with a Damped Multiplicative Trend". *International Journal of Forecasting*, 19, 715-725. 2003.
- [29] Watson, G.S. "Smooth Regression Analysis". *Sankhya Series A*, 26, 359-372. 1964.
- [30] Wand, M. P. and Jones, M. C. "Kernel Smoothing". Chapman and Hall, London. 1995.
- [31] Winters, P. R. "Forecasting Sales by Exponentially Weighted Moving Averages". *Management Science*, 6, 324-342. 1960.
- [32] Yu, K and Stander, J. "Quantile Regression: Applications and Current Research Areas". *The Statistician*, 52 (3), 331-350. 2003.
- [33] Zhou, S.L., McMahon, T.A., Walton, A., Lewis, J. "Forecasting Operational Demand for an Urban Water Supply Zone". *Journal of Hydrology* 259, 189-202. 2002.