RESEARCH ARTICLE                                                                OPEN ACCESS

# A Singular Spectrum Analysis Technique to Electricity Consumption Forecasting

Bisher M. Iqelan
*(Department of Mathematics, the Islamic University of Gaza, Gaza Strip, Palestine)*

**ABSTRACT**
Singular Spectrum Analysis (SSA) is a relatively new and powerful nonparametric tool for analyzing and forecasting economic data. SSA is capable of decomposing the main time series into independent components like trends, oscillatory manner and noise. This paper focuses on employing the performance of SSA approach to the monthly electricity consumption of the Middle Province in Gaza Strip\Palestine. The forecasting results are compared with the results of exponential smoothing state space (ETS) and ARIMA models. The three techniques do similarly well in forecasting process. However, SSA outperforms the ETS and ARIMA techniques according to forecasting error accuracy measures.
*Keywords*: SSA, ETS state space, ARIMA, Forecasting, Electricity consumption time series

## I. INTRODUCTION

Nowadays, life is impossible without electricity. Electricity provides homes and public places with lights and heat. Without it, health, education, finance, technology and other critical services collapse. Therefore, electricity consumption is without a doubt an important issue that has been of interest over the past few years.

It is always more challenging to examine densely populated places. For instance, Gaza Strip is a very small Palestinian territory that is home to a population of more than 2 million people. As a result, it suffers from a chronic crisis in the electricity supply for many years mainly because of the Israeli siege. This paper focuses on the Middle Province of Gaza Strip or known as Central Gaza Strip which consists of the refugee camps of Bureij, al-Maghazi, and al-Nussairat, and the city of Deir al-Balah. It attempts to provide a coherent electricity consumption forecasting for the Middle Province of Gaza Strip using Singular Spectrum Analysis SSA.

Although it is considered as a new nonparametric tool, there exists an extensive literature on Singular Spectrum Analysis (SSA). To begin with, [1] assert that the role of Singular Spectrum Analysis (SSA) is to provide estimates of the statistical dimension. SSA also aims to describe the main physical phenomena reflected by the data. It gives adaptive spectral filters connected with the dominant oscillations of the system, and clarifies the data's noise characteristics. Moreover, [2] add that SSA is both a linear analysis and prediction method. It is superior to the other classical spectral methods because of the data-adaptive character of the eigenelements it is based on. It can also use concepts from nonlinear dynamics. Adding to that, [3] asserts that SSA is a powerful tool used in time series

analysis. It gives a much more accurate forecast results than other methods when applied to many practical problems. Its aim is to decompose original series into small number of independent and interpretable components like oscillatory components, a structureless noise, and a slowly varying trend. Moreover, [4] also illustrate that SSA technique performs four steps. First, computing the trajectory matrix. Second, constructing a matrix for applying SVD. Third, grouping and corresponding to splitting the matrices that were computed at the SVD step. Finally, reconstructing the one-dimensional series. In addition, [5] add that SSA has attained successful application in different branches such as meteorological, biomechanical, hydrological, physical sciences, economics and others. The aim of SSA is to look for nonlinear, non–stationary, and intermittent or transient behavior in an observed time series. Furthermore, [6] employ SSA to decompose the original electricity price series into trend, periodic and noisy components. This approach is evaluated by analysing and forecasting the day ahead electricity prices in the Australian and Spanish electricity markets. The forecasting results assert the dominance of the SSA approach compared with the other forecasting techniques. In addition, [7] assert that SSA is a powerful and well-developed tool of time series analysis and forecasting. SSA can be applied to a wide number of time series analysis problems like exploratory analysis for data-mining and parameter estimation in signal processing. Adding to that [8] also assumes that SSA is a model-free tool that can be applied to all types of series. It comprises time series analysis tools, multivariate statistics tools, dynamical systems and signal processing tools. Finally, [9] investigate the use of SSA in mid-term forecasting of the monthly

electricity consumption for the residential class in Brazil. The results show that the SSA method with graphical analysis of singular vectors presented the more accurate forecasts.

The section above contains a summary of electricity consumption with an application on the Middle Province of Gaza Strip. Moreover, an overview of previous literature is also added. The remainder of the paper is organized as follows. Section 2 discusses the methodology. Section 3 contains the application to real data. Section 4 consists of comparison between SSA approach and other well-known models. Section 5 presents the conclusion of this paper.

## II. METHODOLOGY

Consider a real-valued nonzero time series $Y_N = (y_1, y_2, \cdots, y_N)$ of length $N$. The main point of SSA is to make a decomposition of the original time series into the sum of independent fundamental component parts such as slowly varying trend, oscillatory component and noise.

The SSA technique consists of two stages: decomposition and reconstruction and both of which contain two distinct steps. A brief description of each stage and a discussion on the methodology of the SAA technique will be presented in the following sections following [10] and [11].

### 2.1 Decomposition

This stage includes two steps: embedding and Singular Value Decomposition.

2.1.1. Step 1: Embedding:
Embedding can be considered as a procedure that takes a univariate time series $Y_N = (y_1, y_2, \cdots, y_N)$ and makes it a multivariate set of observations $X_1, X_2, \cdots, X_K$ where the lagged vector $X_i$, $i = 1, 2, \cdots, K$ is defined as $X_i = (y_i, y_{i+1}, \cdots, y_{i+L-1})^T \in R^L$. Let $L$ $(1 < L < N)$ be some integer called window length (the single parameter of the embedding step) and $K = N - L + 1$. As a result, the trajectory matrix can be constructed as shown below in equation (1):

$$X = \left[ X_1 \middle| X_2 \middle| \cdots \middle| X_K \right] = \begin{bmatrix} y_1 & y_2 & \cdots & y_K \\ y_2 & y_3 & \cdots & y_{K+1} \\ \vdots & \vdots & \cdots & \vdots \\ y_L & y_{L+1} & \cdots & y_N \end{bmatrix} \quad (1)$$

The trajectory matrix $X$ is a Hankel matrix that means all the elements on the off diagonals ($i + j = \text{constant}$) are equals [3]. For example, consider the observations

$Y = (10, 20, 30, 40, 50, 60)$, $N = 6$. Choose window length $L = 4$ then $K = 6 - 4 + 1 = 3$, and the trajectory matrix is:

$$X = \begin{bmatrix} 10 & 20 & 30 \\ 20 & 30 & 40 \\ 30 & 40 & 50 \\ 40 & 50 & 60 \end{bmatrix} \quad (2)$$

2.1.2. Step 2: Singular Value Decomposition (SVD):
From the trajectory matrix $X$, define the covariance Matrix $X X^T$. Let $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_L$ denote the eigenvalues of $X X^T$ and the corresponding eigenvectors $U_1, U_2, \cdots, U_L$ in SSA literature are donated by (the empirical orthogonal functions). Let d be the number of nonzero eigenvalues (that is the rank of $X X^T$) and set the principal components $V_i = X^T U_i / \sqrt{\lambda_i}$, then the SVD of the trajectory matrix $X$ can be expressed as a sum of matrices:

$$X = E_1 + E_2 + \cdots + E_d \quad (3)$$

where $E_i = \sqrt{\lambda_i} U_i V_i^T$, $i = 1, 2, \cdots, d$. The gathering $\left( \sqrt{\lambda_i}, U_i, V_i \right)$ is referred to the $i^{th}$ eigentriple (square root of eigenvalue, eigenvector, factor vector) of the matrix $X$. The ratio $\lambda_i \Big/ \sum_{i=1}^{d} \lambda_i$ is the share of the trajectory matrix $X$ explained by the sum in (1). For more details see [12].

### 2.2 Reconstruction

This stage includes two steps: grouping and averaging.

2.2.1. Grouping:
The grouping procedure divides the set of elementary matrices indices $\{1, \cdots, d\}$ into $m$ disjoint subsets $I_1, \cdots, I_m$. Let $I_k = \{i_{k,1}, \cdots, i_{k,p}\}$ be a group of indices $i_{k,1}, \cdots, i_{k,p}$, $k = 1, \cdots, m$. Then the matrix $X_{I_k}$ corresponding to the group $I_k$ is defined as $X_{I_k} = E_{i_{k,1}} + E_{i_{k,2}} + \cdots + E_{i_{k,p}}$. Then the corresponding decomposition is written as:

$$X = X_{I_1} + \cdots + X_{I_m} \quad (4)$$

For a given group $I_k$, $k = 1, \cdots, m$ the share of the component $X_{I_k}$ into the equation (3) is measured by the contribution of the corresponding eigenvalues: $\sum_{i \in I_k} \lambda_i \Big/ \sum_{i=1}^{d} \lambda_i$. For more details see [13]. Note that in the previous numeric example, there are only two nonzero components and at most two groups that can be defined as $I_1 = \{1\}, I_2 = \{2\}$.

### 2.2.2. Step 4: Diagonal averaging:

The group of $m$ components chosen in the above step is used to rebuild the deterministic components of the original time series $Y_N$. The aim of diagonal averaging is to transform all matrix components $X_{I_1}, \cdots, X_{I_m}$ in equation (4) above to Hankel matrices $\tilde{X}_{I_1}, \cdots, \tilde{X}_{I_m}$ where $\tilde{X}_{I_j}$ is the Hankelization of $X_{I_j}$ for $j = 1, \cdots, m$. The expression (4) is transformed to the Hankelized form:

$$X = \tilde{X}_{I_1} + \cdots + \tilde{X}_{I_m} \qquad (5)$$

If $w_{ij}$ stands for an element of a matrix $W$, then the $k^{th}$ element of the produced time series is obtained by averaging $w_{ij}$ over all $i$, $j$ such that $i + j = k + 1$. (For example, the $3^{rd}$ element is obtained by $(w_{1,3} + w_{2,2} + w_{3,1})/3$ and so on).

Let $\{\tilde{y}_{j1}, \tilde{y}_{j2}, \cdots, \tilde{y}_{jN}\}$ denote the reconstructed components of the original time series corresponding to the Hankel matrix $X_{I_j}$ for $j = 1, \cdots, m$. Therefore, the original time series will be decomposed into the sum of $m$ series such that:

$$y_k = \sum_{j=1}^{m} \tilde{y}_{jk}, \quad k = 1, 2, \cdots, N \qquad (6)$$

Forming diagonal averaging on the components of the previous example, the Hankelized expansion becomes:

$$X = \begin{bmatrix} 10 & 20 & 30 \\ 20 & 30 & 40 \\ 30 & 40 & 50 \\ 40 & 50 & 60 \end{bmatrix} = \tilde{X}_1 + \tilde{X}_2$$

$$= \begin{bmatrix} 15.38 & 21.63 & 28.70 \\ 21.63 & 28.70 & 38.27 \\ 28.70 & 38.27 & 49.91 \\ 38.27 & 49.91 & 62.39 \end{bmatrix} + \begin{bmatrix} -5.38 & -1.63 & 1.30 \\ -1.63 & 1.30 & 1.73 \\ 1.30 & 1.73 & 0.08 \\ 1.73 & 0.08 & -2.39 \end{bmatrix} \qquad (7)$$

Consequently, the reconstructed time series components are:

$$Y = \begin{bmatrix} 10 \\ 20 \\ 30 \\ 40 \\ 50 \\ 60 \end{bmatrix} = \begin{bmatrix} 15.38 \\ 21.63 \\ 28.70 \\ 38.27 \\ 49.91 \\ 62.39 \end{bmatrix} + \begin{bmatrix} -5.38 \\ -1.63 \\ 1.30 \\ 1.73 \\ 0.08 \\ -2.39 \end{bmatrix} \qquad (8)$$

### 2.3 Forecasting

The basic SSA forecasting method is known as the recurrent method (R-forecasting). It uses the concept of Linear Recurrent Formula (LRF). A time series $Y_N = (y_1, \cdots, y_N)$ satisfies the LRF and can be successfully forecasted via the SSA. $Y_N$ recognizes the LRF of order $d$ if:

$$y_t = \alpha_1 y_{t-1} + \cdots + \alpha_d y_{t-d}, t = d + 1, \cdots, N \qquad (9)$$

where $\alpha_1, \alpha_2, \cdots, \alpha_d$ are the coefficients of the LRF of order $d$. Wide class of time series is ruled by LRF such as harmonic, polynomial and exponential series. Subsequently, there is a summarized description of the so called SSA recurrent forecasting algorithm (for more details see [10]).

The LRF coefficients $\alpha_1, \alpha_2, \cdots, \alpha_d$ in the linear combination (9) can be obtained by using the eigenvectors acquired in the SVD step of the SSA algorithm. Let $U_j^{\triangledown}$ denote the vector of the 1st $L - 1$ components of the eigenvector $U_j$ and $\beta_j$ is the last component of $U_j$ (for $j = 1, \cdots, r, r < L$). The LRF coefficients can be calculated as follows:

$$\Lambda = (\alpha_1, \alpha_2, \cdots, \alpha_{L-1})^T = \frac{1}{1 - \upsilon^2} \sum_{j=1}^{r} \beta_j U_j^{\triangledown} \qquad (10)$$

where $\upsilon^2 = \sum_{j=1}^{r} \beta_j^2$.

Considering these notations, the time series $\hat{Y}_{N+h} = (\hat{y}_1, \cdots, \hat{y}_{N+h})$ can be defined by the following formula:

$$\hat{y}_i = \begin{cases} \tilde{y}_i & \text{for } i = 1, \cdots, N \\ \sum_{j=1}^{L-1} \alpha_j y_{i-j} & \text{for } i = N+1, \cdots, N+h \end{cases} \quad (11)$$

where $\hat{y}_i$ ($i = 1, \cdots, N$) are the reconstructed time series. Then $\hat{y}_{N+1}, \cdots, \hat{y}_{N+h}$ are the $h$-step-ahead recurrent forecast. Regarding the previous numerical example, let $r = 3 < L$. Then the LRF coefficients can be calculated and have the values:

$$\Lambda = (\alpha_1, \alpha_2, \alpha_3)^T = (-1.50, 2.00, 0.50)^T \quad (12)$$

and the 1st-step-ahead forecast is then:

$$\hat{y}_7 = 0.50 * 62.39 + 2.00 * 49.91 \\ + (-1.50) * 38.27 = 73.61. \quad (13)$$

### 2.4 Parameters Selection

As explained in the previous sections, SSA approach requires two important parameters: the window length $L$ (in the embedding step) and the grouping effect parameter $r$ (in the reconstruction stage). Certainly, there are great efforts and suggestions of various methods of selecting the suitable value of $L$ and $r$ (see [10], [14], [15] and [16]). Taking into account theoretical results for structure of the trajectory matrix and seprability, it seems better to select window length less than $N/2$ or $N/3$ as other recommended values (see [17]). Moreover, values for $L$ and $r$ be selected depend on both information provided by time series under study and analysis that need to be performed. In this paper, SSA is used as a technique of forecasting. Therefore, the implemented criteria are based on the forecasting errors. However, different error measures are used to find the error forecast. These include Mean absolute error (MAE), Mean square error (MSE), Root mean square error (RMSE) and Mean absolute percentage error (MAPE). These accuracy measures are defined in Table 1 below.

**Table 1** Accuracy Measures

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

where $n$ is the number of observations in the sample. Note that the selection of $L$ and $r$ is obtained by minimizing the forecasting error using these accuracy measures (see [16] and [18]).
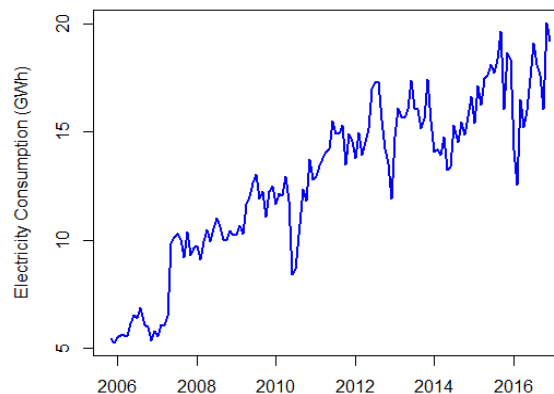


**Figure 1.** Monthly electricity consumption in the middle area of Gaza Strip (2005:11-2016:12)

## III. APPLICATION TO REAL DATA

SSA approach is used to decompose and forecast the monthly electricity consumption series (in GWh) in the Middle Province of Gaza Strip. The time series contains 134 observations from November 2005 to December 2016. We have transformed the data by dividing it by 1000000. Table 2 shows some descriptive statistics for this data set. On the other hand, Fig. 1 displays the electricity consumption over the period (2005:11 - 2016:12).

**Table 2.** Descriptive Statistics for Monthly Electricity Consumption (in GWh) for the Middle Area of Gaza 2005:11-2016:12

| $n$ | Min. | 1st Qu | Median | Mean | 3rd Qu | Max. |
|---|---|---|---|---|---|---|
| 134 | 5.26 | 10.22 | 13.45 | 12.81 | 15.56 | 20.02 |

Visual analysis of the drawn time series in Fig. 1 specifies that it has a trend and this trend can be resembled either by exponentially increasing function or linear function. Moreover, it seems that the seasonal component has a complicated and changeable attitude. Fig. 2 displays the periodogram of electricity consumption series which emphasizes this assumption. The periodogram is established for purifying and revising the representation of the time series.
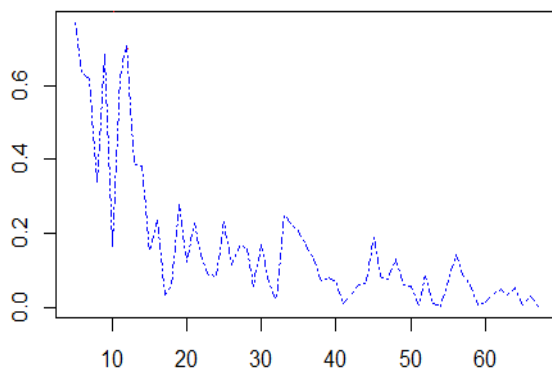
**Figure 2.** Periodogram of series electricity consumption (the first 70 points)

Moreover, the first 122 observations from November 2005 to December 2015 are used as a training sample for model prediction and the remaining 12 observations from January 2016 to December 2016 are used as a testing data set to assess the electricity consumption forecasts. Results from SSA technique are compared with other models' results, like Box-Jenkins (ARIMA) model and Exponential Smoothing (ETS) model. SSA computations are implemented by using the RSSA package in R (see [19]), while the other two models performed using the forecast package [20]. For each approach, the following performance indexes related to the forecasting errors are offered: MAE, MSE, RMSE and MAPE, and residuals test statistics.

### 3.1. Electricity consumption by SSA

As explained in the previous section, it is required as a first step to select the proper parameters, i.e., the window length $L$ and the group indices $r$ . As mentioned before, the suitable values of $L$ and $r$ are that allowing the minimal values of the accuracy measures of Table 1.

The process of selection $L$ and $r$ carried out through two steps. The first step is fixing $L$ , $2 \leq L \leq N / 2$ and changing the value of $r$ , $1 \leq r \leq L - 1$ until reaching the choice of the pair $(L , r)$ which minimizes the accuracy measures. Repeat this step many times with different choices of $L$ . The second step is choosing the optimal pair $(L , r)$ , throughout all pairs of the first step, which provide minimum accuracy measures. Table 3 shows different choices of $(L , r)$ and the corresponding values of the accuracy measures (all values are rounded).

**Table 3** Different selections of $(L , r)$ with corresponding values of accuracy measures

| $(L , r)$ | MAE | MSE | RMSE | MAPE |
|-----------|------|------|------|------|
| (37,3) | 1.568 | 3.952 | 1.988 | 0.102 |
| (38,9) | 1.516 | 3.914 | 1.978 | 0.101 |
| (39,7) | 1.501 | 3.622 | 1.903 | 0.097 |
| (40,7) | 1.416 | 3.560 | 1.887 | 0.094 |
| (41,7) | 1.454 | 4.014 | 2.004 | 0.097 |
| (42,12) | 1.445 | 3.417 | 1.848 | 0.095 |
| (55,4) | 1.528 | 4.134 | 2.033 | 0.102 |
| (56,4) | 1.488 | 3.886 | 1.971 | 0.099 |

Note that many other selections and their corresponding values of accuracy measures are also computed but they are improper so ignored and not reported. Comparing the values in Table 3 concludes that the best selection of SSA parameters of electricity consumption time series is that $L = 40$ and $r = 7$ . Fig. 3 illustrates eigenvectors from the first 12 eigentriples. It shows the participation of each eigenvalue after the SVD stage using $L = 40$ as window length required for the embedding step.
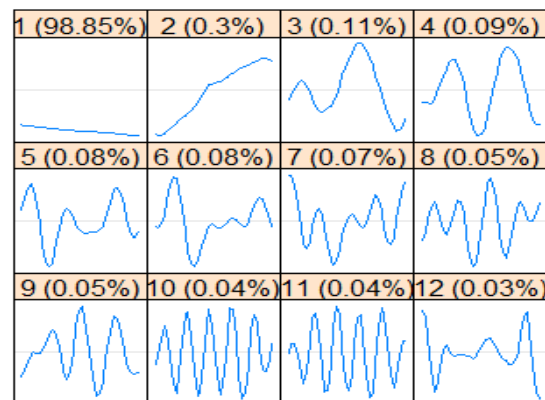


**Figure 3.** The first 12 principal components plotted as time series.

Fig. 4 displays the plot of the reconstructed components. It describes the 12 most significant initial reconstructed components of the original electricity consumption time series. After taking a quick look, the first two reconstructed components are related to slow motion components (i.e., the trend behavior), while the remainder of the reconstructed components are connected to fluctuating components.
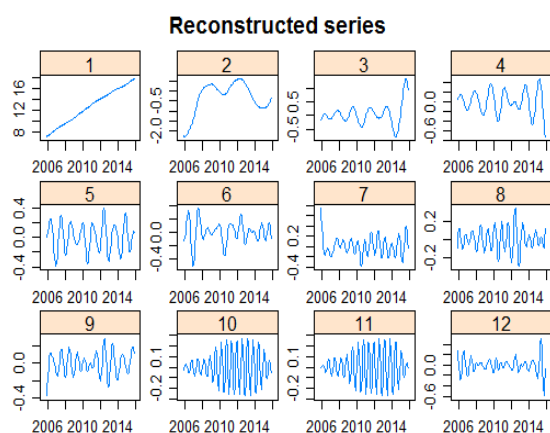
**Figure 4.** Reconstructed components related to the first 12 eigentriples

Furthermore, for the sake of evaluating the separability of the various eigentriples, Fig. 5 displays a graphical representation of the w-correlation matrix. The correlation between components $m$ and $n$ can be represented as a cell $(F_m, F_n)$. It has a color scale from black to white equivalent to 1 to 0, respectively. Notice that cell $(F_1, F_2)$ is white (w-correlation=0), so components 1 and 2 are clearly separable whereas cell $(F_3, F_4)$ are grey color (w-correlation=0.25), so components 3 and 4 are almost separable. Components 5-6 and 7-8 have high w-correlations so they could be grouped. Also, the first two eigentriples in Fig. 5 correspond to the trend components while the next eigentriples produce harmonic (oscillatory) components. The large shining cells on the top-right shows noise components. Therefore, the first 7 eigentriples will be used in the reconstruction of the electricity consumption series.

## IV. COMPARISON

In this section, the SSA approach is compared with several well-known models, the Box-Jenkins model namely, Autoregressive Integrated Moving Average (ARIMA) models and the Exponential Smoothing (ETS) models.
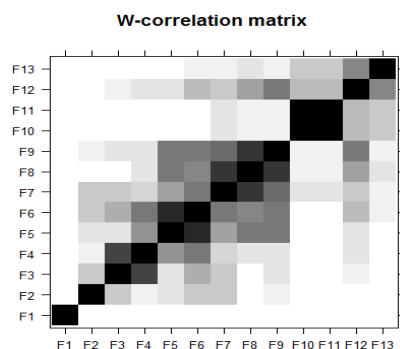


**Figure 5.** A graphical representation of the w-correlation matrix.

### 4.1. ARIMA Model

A non-stationary time series $\{X_t\}$ is said to follow a non-stationary autoregressive integrated moving average (ARIMA) denoted by $ARIMA(p,d,q)$ if it is expressed as:

$$\Phi_p(\mathbf{B})\nabla^d Y_t = \mu + \Theta_q(\mathbf{B})\varepsilon_t \qquad (14)$$

where $\varepsilon_t$ are identically and independently distributed as $N(0,\sigma^2)$, $t = 1,2,\cdots,N$ and $N$ is the number of observations, $d$ is the order of non-seasonal differences and $\nabla$ is the non-seasonal differencing operator, $\nabla = 1 - \mathbf{B}$. $\mu$ is the mean of a series assuming that after differencing it is stationary.

As mentioned above, $\mathbf{B}$ is the backshift operator, used to simplify the representation of lag values, by $\mathbf{B}X_t = X_{t-1}$. Also, $\Phi_p(\mathbf{B})$ and $\Theta_q(\mathbf{B})$ are the autoregressive polynomial of $\mathbf{B}$ for order $p$ and the moving average polynomial of $\mathbf{B}$ for order $q$ respectively where:

$$\Phi_p(\mathbf{B}) = 1 - \phi_1\mathbf{B} - \phi_2\mathbf{B}^2 - \cdots - \phi_p\mathbf{B}^p \qquad (15)$$

and

$$\Theta_q(\mathbf{B}) = 1 - \theta_1\mathbf{B} - \theta_2\mathbf{B}^2 - \cdots - \theta_q\mathbf{B}^q \qquad (16)$$

More details can be found in ([21], [22], [23], [24], and [25]).

### 4.1.1. Parameters Estimation

The estimation of parameters for ARIMA model is a nonlinear problem that requires some special processes such as the maximum likelihood method or nonlinear least-squares estimation. At this stage of model building, the estimated parameter values should minimize the sum of squared residuals. For this purpose, many software packages are applicable for fitting ARIMA models. In this current study, the forecast package in R software will be used. To choose the best ARIMA model based on observation data, we use the corrected Akaike Information Criterion (AICc). For more details see [21].

Based on the methodology summarized above, to calculate the point forecasts for electricity consumption in the Middle Province of Gaza Strip for the 12 months from January 2016 to December 2016, the results showed that the best ARIMA model was $ARIMA(1,1,4)$:

$$(1 - 1.201\mathbf{B} + 0.201\mathbf{B}^2)Y_t = 0.028 + \Theta_4(\mathbf{B})\varepsilon_t \qquad (17)$$

where

$$\Theta_4(\mathbf{B}) = (1 - 0.202\mathbf{B} - 0.102\mathbf{B}^2 - 0.116\mathbf{B}^3 - 0.100\mathbf{B}^4) \qquad (18)$$

### 4.2. Exponential Smoothing State Space Model

Another popular type of forecasting models are exponential smoothing techniques. It's simple but very helpful of adjusting time series forecasting. [26], [27], and [28] 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 [29] 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 [30], who added damped trend to the classification. This extension is then modified by [31], before the final extension is proposed by [32] who extended the classification to include damped multiplicative trends. For more details, see [33].

When fitting an exponential smoothing model with a state space approach, the ets() function in forecast package [20] and [34] was used. It is utilized to choose the suitable model automatically on the basis of maximum likelihood method (MLE) and then to calculate the point forecasts for electricity consumption in the Middle Area of Gaza Strip for the 12 months from January 2016 to December 2016. When applying the function, the results showed that the best performing model was $ETS(M,A,N)$, that is a model with multiplicative error, additive trend and no seasonality. The exponential smoothing state space $ETS(M,A,N)$ model is:

$$y_t = \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 4.689 \\ 0.118 \end{bmatrix} (1 + \varepsilon_t),$$

$$x_t = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 4.689 \\ 0.118 \end{bmatrix} + \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 4.689 \\ 0.118 \end{bmatrix} \begin{bmatrix} 0.908 \\ 0.0001 \end{bmatrix} \varepsilon_t$$

(19)

where the state vector $x_t = (\ell_t, b_t)^T$, includes the level and growth components respectively. For more details about the structures and notations of exponential smoothing state space models, see [33].

### 4.3. Forecasting Results

Table 4 presents the results for several techniques for the forecasting of the 12-step-ahead data points for electricity consumption in the Middle Area of Gaza Strip. The scores show that the SSA forecasts are comparable with the forecasts acquired from the Box-Jenkins $ARIMA(1,1,4)$ and the exponential smoothing state space $ETS(M,A,N)$ models. Moreover, the SSA forecasts outperform the forecasts produced by the ARIMA and exponential

smoothing models and consequently the forecasted values for electricity consumption of the Middle Province in Gaza Strip are very close to the original data for the SSA technique. Adding to that, SSA is able to provide further details about the decomposition of the time series.

Comparing the results of the measures of forecast accuracy in Table 5, the SSA technique still performs best based on all the measures values. The predicted values using SSA are highly accurate as MSE=3.5604 is approximately 3 times less than the other models. RMSE = 1.8869 is 1.5 times less than the other models and MAPE = 0.0938 is approximately 2 times less than the other models.

**Table 4.** Actual and Prediction for Electricity Consumption (in GWh) for the Middle Province of Gaza Strip in 2.16:1-2016:12

| Month | Actual Data | Prediction | | |
|---|---|---|---|---|
| | | SSA | ARIMA (1,1,4) | ETS(M, A,N) |
| 01/2016 | 14.1886 | 17.2325 | 18.4250 | 18.4873 |
| 02/2016 | 12.5564 | 16.9441 | 19.1469 | 18.6053 |
| 03/2016 | 16.4811 | 16.8670 | 18.1730 | 18.7232 |
| 04/2016 | 15.1924 | 16.9982 | 18.7956 | 18.8412 |
| 05/2016 | 16.0113 | 17.2629 | 18.5982 | 18.9592 |
| 06/2016 | 17.4678 | 17.5599 | 18.8822 | 19.0771 |
| 07/2016 | 19.0960 | 17.8049 | 18.8836 | 19.1951 |
| 08/2016 | 18.1212 | 17.9669 | 19.0509 | 19.3131 |
| 09/2016 | 17.6333 | 18.0759 | 19.1208 | 19.4310 |
| 10/2016 | 16.0417 | 18.2045 | 19.2478 | 19.5490 |
| 11/2016 | 20.0189 | 18.4360 | 19.3413 | 19.6670 |
| 12/2016 | 19.2120 | 18.8230 | 19.4545 | 19.7849 |

By all the odds, the SSA technique has the lowest forecasting accuracy measures.

**Table 5** Forecast Accuracy for the SSA, ARIMA and Exponential Smoothing Techniques

| | MAE | MSE | RMSE | MAPE |
|---|---|---|---|---|
| SSA | 1.4158 | 3.5604 | 1.8869 | 0.0938 |
| ARIMA(1,1,4) | 2.2399 | 8.3199 | 2.8844 | 0.1499 |
| ETS(M,A,N) | 2.3597 | 8.5087 | 2.9170 | 0.1563 |

Fig. 5 exhibits the actual and forecasted electricity consumption in the Middle Province of Gaza Strip, as well as 80% and 95% prediction intervals. Therefore, a more detailed picture of the SSA technique can be drawn by diving into the numbers of the previous tables.
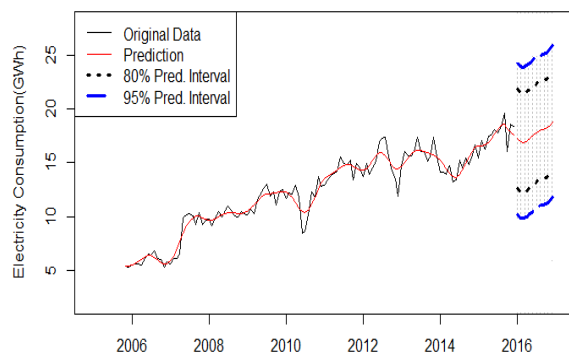
**Figure 5**. Actual and forecasted (red line) electricity consumption in the Middle Province of Gaza Strip, 80% and 95% (blue) prediction intervals.

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.

The sample correlation in Fig. 6 displays that most of sample autocorrelation coefficients of the residuals are within the confidence limits, consequently the residuals are white noise reflecting that SSA technique is adequate.



**Figure 6.** Residual diagnostics for SSA technique.

To confirm the evidence of autocorrelations, the Box-Ljung test p-value result in Table 6 displays that there is an 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 6.** Residual Diagnostics Tests

| Residual Test | p-values |
|---|---|
| Residual Autocorrelation test Box-Ljung test | $8.243 * 10^{-08}$ |
| Residual Normality test Jarque Bera Test | 0.01426 |
| Homoscedasticity test Box-Ljung test (Squared Residuals) | 0.04635 |

## V. CONCLUSION

The main objective of this research is to clarify the methodology of Singular Spectrum Analysis and explain that SSA can be successfully utilized to analyze and predict the monthly electricity consumption in the Middle Province of Gaza Strip\Palestine. This paper has explained that the SSA technique is a very powerful appliance for decomposing and forecasting a non- linear and/or a non-stationary time series into a collection of independent components. In the given example, regarding the monthly electricity consumption, the performance of SSA is the leading, compared with other well-known forecasting models, namely Box-Jenkins ARIMA model and Exponential Smoothing State Space ETS model. The numerical results obtained using R packages declared that the error came by the SSA technique were smaller than those obtained by the ARIMA and ETS state space models according to mean absolute error (MAE), mean square error (MSE), root mean square error (RMSE) and mean absolute percentage error (MAPE). However, the SSA outperforms the ARIMA and ETS state space models. Adding to that, the obtained numerical results assert the potentiality of the SSA technique for electricity consumption predicting implementations.

## REFERENCES

[1] R. Vautard and M. Ghil, Singular Spectrum Analysis in Nonlinear Dynamics, With Applications to Paleoclimatic Time Series, *Physica D, 35,* 1989, 395-424.

[2] R. Vautard, P. Yiou and M. Ghil, Singular-Spectrum Analysis: A Toolkit for Short, Noisy Chaotic Signals, *Physica D, 58,* 1992, 95-126.

[3] H. Hassani, Singular Spectrum Analysis: Methodology and Comparison, *Journal of Data Science*, 5(2), 2007, 239–257.

[4] N. Simões, L. Wang, S. Ochoa1, J. P. Leitão, R. Pina, C. Onof, A. Sá Marques, Č. Maksimović, R. Carvalho, L. David, A Coupled SSA-SVM Technique for Stochastic Short-Term Rainfall Forecasting, *12nd International Conference on Urban Drainage*, Porto Alegre, 2011, 1-8.

[5] M. Zokaei, R. Mahmoudvand, N. Najari, Comparison of Singular Spectrum Analysis and ARIMA, *Proc. 58th World Statistical Congress*, Dublin, 2011, 3991- 3996.

[6] A. Miranian1, M. Abdollahzade, H. Hassani, Day-Ahead Electricity Price Analysis and Forecasting by Singular Spectrum Analysis, *IET Generation, Transmission & Distribution, 2012, 7(4),* 2012, 337–346.

[7] N. Golyandina and A. Korobeynikov, Basic Singular Spectrum Analysis and forecasting

with R, *Computational Statistics and Data Analysis, 71*, 2013, 934–954.

[8]  C. Deng, *Time Series Decomposition Using Singular Spectrum Analysis*, Master Thesis, the Department of Mathematics, East Tennessee State University, 2014.

[9]  M. Lima de Menezes, R. Castro Souza and J. Francisco Moreira Pessanha, Electricity Consumption Forecasting Using Singular Spectrum Analysis, *Universidad Nacional de Colombia, 82 (190)*, 2015, 138-146.

[10]  N. Golyandina, V. Nekrutkin, and A. Zhigljavski, *Analysis of Time Series Structure: SSA and Related Techniques* (Boca Raton: CRC Press, 2001).

[11]  N. Golyandina, A. Zhigljavsky, *Singular Spectrum Analysis for Time Series* (Springer Science & Business Media, 2013).

[12]  H. Hassani, S. Heravi, and A. Zhigljavsky, Forecasting European Industrial Production with Singular Spectrum Analysis, *International Journal of Forecasting*, 25, 2009, 103–118.

[13]  S. Rocco, Singular Spectrum Analysis and Forecasting of Failure Time Series, *Reliability Engineering and System Safety, 114*, 2013, 126–136.

[14]  H. Hassani, R. Mahmoudvand and M. Zokaei, Separability and Window Length in Singular Spectrum Analysis, *C. R. Acad. Sci. Paris, Ser. I, 349*, 2011, 987–990.

[15]  R. Mahmoudvand and M. Zokaei, On the Singular Values of the Hankel Matrix with Application in Singular Spectrum Analysis, *Chilean Journal of Statistics 3*, 2012, 43-56.

[16]  R. Mahmoudvand, F. Alehosseini, M. Zokaei, Feasibility of Singular Spectrum Analysis in the Field of Forecasting Mortality Rate, *Journal of Data Science, 11*, 2013, 851-866.

[17]  N. Golyandina, A. Korobeynikov, Basic Singular Spectrum Analysis and forecasting with R, *Computational Statistics & Data Analysis, 71*, 2014, 934-954.

[18]  M. Ghodsi and M. Yarmohammadi, Exchange Rate Forecasting With Optimum Singular Spectrum Analysis, *Journal of Systems Science and Complexity – Springer, 27*, 2014, 47–55.

[19]  A. Korobeynikov, Computation- and space-efficient implementation of SSA, *Statistics and Its Interface, 3(3)*, 2010, 257-368.

[20]  Hyndman RJ. Forecast: Forecasting Functions for Time Series and Linear Models. *R package version 8.0*. 2017.

[21]  B. M. Iqelan, Time Series Modeling of Monthly Temperature Data of Jerusalem/Palestine, MATEMATIKA, 31(2), 2015, 159–176.

[22]  J. D. Cryer and K. Chan, *Time Series Analysis with Applications in R.* (Springer, Second edition, 2008).

[23]  D.S. Shumway and D. S. Stoffer, *Time Series Analysis and Its Applications: With R Examples* (Springer, third edition, 2011).

[24]  G. Box and G. Jenkins, *Time Series Analysis: Forecasting and Control* (Holden-Day, the University of Michigan, 1976).

[25]  P. Brockwell and R. Davis, *Time Series: Theory and Methods* (Springer; 2nd ed. 1991).

[26]  R. G. Brown, *Statistical Forecasting for Inventory Control* (McGraw-Hill, New York. 1959).

[27]  C. C. Holt, *Forecasting Trends and Seasonals by Exponentially Weighted Averages* (O.N.R. Memorandum 52/1957, Carnegie Institute of Technology. 1957).

[28]  P. R. Winters, Forecasting Sales by Exponentially Weighted Moving Averages, *Management Science, 6,* 1960, 324–342.

[29]  C. C. Pegels, Exponential Forecasting: Some New Variations. *Management Science 15(5),* 1969, 311–315.

[30]  Jr. Gardner, Exponential Smoothing: The State of the Art, *Journal of Forecasting 4,* 1985, 1–28.

[31]  R. J. Hyndman, 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),* 2002, 439–454.

[32]  J. W. Taylor, Exponential Smoothing with a Damped Multiplicative Rend, *International Journal of Forecasting, 19,* 2003, 715–725.

[33]  B. M. Iqelan, Comparison of Parametric and Nonparametric Techniques for Water Consumption Forecasting, *International Journal of Scientific & Engineering Research, 8(1),* 2017, 1530-1536.

[34]  R.J. Hyndman and Y. Khandakar, Automatic Time Series Forecasting: The Forecast Package for R, *Journal of Statistical Software, 26(3),* 2008, 1–22