Journal of Statistics Applications & Probability
An International Journal

http://dx.doi.org/10.18576/jsap/080207

Statistical Modeling and forecasting of weather Data Distribution Using Improved Time Series Analysis

*Afrah Al-Bossly**

Department of Mathematics, College of Science and Humanities in Al-Kharj, Prince Sattam Bin Abdulaziz University, Al-kharj 11942, Saudi Arabia.

Received: 3 Aug. 2018, Revised: 12 Oct. 2018, Accepted: 19 Jan. 2019.
Published online: 1 Jul. 2019.

Abstract: The current study is intended to investigate the applicability of a special class of time series models; autoregressive integrated moving-average (ARMIA) for the estimation of temperature distribution forecast model. Different transformations of ARMIA models such as differencing and smoothing are investigated, in addition to study the effect of each model parameters on the accuracy of the derived model. This study is applied at a temperature time series data of Riyadh city in KSA. By investigating a number of smoothing techniques, simple exponential smoothing (with = 0.2) is found to be the most adequate forecasting model for the case under study as it yields highest correlation factor (R2= 0.9337).

Keywords: Time series analysis, ARIMA, Transformation, Smoothing.

1 Introduction

Quantitative models are very useful in forecasting and have become essential in many applications. As the findings of several types of research pointed to the fact that integrating different models can improve their predictive performance. Khashei et al.[1] presented a hybrid model of "ARIMA" and "Probabilistic Neural Network" (PNN), in order to reach more exact results. In their proposed model, the authors modified ARIMA model's estimated values using ARIMA residuals trend, which are respectively obtained from PNN and optimum step length.

In a study to forecast the capacity of electricity generation in Malaysia, Haigeset at [2] deployed ARIMA approach to model and forecast about 50 years data [2]. The authors evaluated different models using the Schwarz Bayesian Criterion (SBC). Models accuracy was calculated using Root Mean Square Error (RMSE) and Mean Absolute Error (MAE). Degerine and Lambert-Lacroix [3], characterized the partial autocorrelation function of a non-stationary time series using the auto-covariance function. Kumar and Jain [4] also used ARIMA to model the time series of traffic noise [4]. A new forecast methodology that combines Gray Model (GM(1,1)) and (ARIMA) was developed by Jia et al [5] to predict UT1-UTC earth orientation parameters after the removal of the leap second and Earth's zonal harmonic tidal.

Vector Autoregressive Moving- Average (VARMA) is one of the main conventional multivariate time series forecasting models. A new framework for predicting the online VARMA time series was proposed [5]. The results of experimental work validated the efficacy and of the proposed algorithms [6].

2 Time-Series Models

One of the most significant quantitative models is time-series forecasting. Time-series analysis includes methods to analyze data series in order to reach significant statistics and other data characteristics.

Methods for time-series analysis may be categorized as either frequency-domain methods; which comprise spectral and wavelet analyses, or time-domain methods that include autocorrelation and cross-correlation analysis.

*Corresponding author e-mail: a.basli@psau.edu.sa

The first method refers to analyzing mathematical functions with respect to frequency, instead of time. A function can be converted from one domain to the other by means of mathematical operators transforms. One of these transforms is "Fourier transform", which change a time function into a sum of different frequency sine waves, each of which represents a frequency component [7].

Time-series techniques may also be categorized as parametric and non-parametric methods. In parametric methods, the stationary process could be described by a small number of parameters using "autoregressive or moving average" model.

Chris Chatfield and Mohammed Yar [8] suggested an approach for predicting intervals for the "additive Holt-Winters" forecasting procedure extending the results to the "multiplicative seasonal" case [8]. The multiplicative prediction showed a contrasting behavior to the "additive" case.

2.1 Auto-Regressive Moving Average Models (ARMA):

AR models; first introduced by Yule [9] were complemented by MA schemes. The combined ARMA processes can be used to model all stationary time series if the suitable number of AR and MA terms (p, q) are appropriately specified [10]. This means that any series $x_t$can be modeled as a group of precedent x_t values and/or precedent e_t errors, as in equation (1).

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q} \quad (1)$$

Where $\emptyset_1, \emptyset_2, \dots \emptyset_p$are the AR coefficients and $\theta_1, \theta_2, \dots \theta_q$are the MA coefficients. The value of these parameters are estimated using some non-linear optimization procedure that minimizes the sum of square errors or some other appropriate loss function [11].

A general "ARIMA model" of order (p, d, and q) symbolizing the time series as:

$$\phi(B)\nabla^d x_t = \theta(B)e_t \quad (2)$$

Where: "x_t and e_t" represent variable and random error at time t respectively. B is a backward operator defined by;

$$Bx_t = x_{t-1} \quad (3)$$

$$\nabla = 1 - B, \nabla^d = (1-B)^d; \text{ d is the order of differencing} \quad (4)$$

$$\phi(B) \text{and } \theta(B) \quad (5)$$

and …B† are autoregressive (AR) and moving averages (MA) operators of p and q orders, respectively, given by:

$$\phi(B) = 1 - \phi_1 B - \phi_2 B - \cdots - \phi_p B^p \quad (6)$$

and

$$\theta(B) = 1 - \theta_1 B - \theta_2 B - \cdots - \theta_q B^q \quad (7)$$

2.1.1 Steps of ARIMA Time Series Modeling

ARIMA model is prevalent for their statistical properties and due to the modeling process using Box–Jenkins methodology [11]. In addition, ARIMA models can employ a number of exponential smoothing models.

Once the series patterns, trends, cycles, and seasonality, are visualized, to use Box & Jenkins methodology for ARIMA modeling, it is essential firstly to determine whether the time series is stationary or non- stationary. Dickey-Fuller is one of the tests used to check stationarity [12].If the series is nonstationary; it should be stationarized (transformed) using one of the transformation techniques.

Three techniques are normally used to transform a time series [13]:

- **Detrending**
- **Differencing**: through modeling the differences between series data points
- **Seasonality** which is easy to incorporate in the "ARIMA model"

Following the transformation step, optimal parameters of the resulting "stationary time series" are to be defined. The parameters p,d,q can be found using autocorrelation function (ACF) and partial autocorrelation function (PACF) plots. If both "ACF and PACF" decrease regularly, this is an indication that the time series should be stationarized introducing value to "d". Once the optimal parameters are determined ARIMA model is built.

2.1.2 Autocorrelation and Partial Autocorrelation Function

There are several autocorrelation coefficients, corresponding to each panel in the lag plot. For example, r_1 and r_2 measure the relationships between (y_t and y_{t-1}), and (y_t y_{t-2}), respectively and so on[14].The value of r_k can be written as:

$$\boldsymbol{r_k} = \frac{\sum_{t=k+1}^{T}(y_t - \bar{y})(y_{t-k} - \bar{y})}{\sum_{t=1}^{T}(y_t - \bar{y})^2} \quad (8)$$

Where:

r_k is the autocorrelation coefficient of lag k, $k<\frac{T}{4}$, T is the sample size

Time series that show no autocorrelation is called white noise.

Partial correlation was presented long time ago [14, 15], but the parameterization of a stationary time-series through the partial autocorrelation function (PACF) is relatively recent. This was recognized by Nielsen and Schou [16] for "autoregressive processes" and by Ramsey [17] for the "general stationary" case. The partial autocorrelation coefficients are [11]

$$g_{kk} = \{\frac{r_k}{1-\sum_{j=1}^{k-1} g_{k-1,j}} \frac{\sum_{j=1}^{k-1} g_{k-1,j} r_{k-j}}{r_j}\}_{k=2,3,\ldots} \quad (9)$$

$$g_{kj} = \{g_{k-1,j} - g_{kk} g_{k-1,k-j}\}_{j=1,2,\ldots,k-1} \quad (10)$$

Where:
g_{kk} is the k^{th} PAC in an autoregressive process of order k

g_{kj} is the j^{th} recursive computational coefficient in the autoregressive process

2.2 Time-Series Transformation

Transformations are used to improve time series statistical analysis, through determining a right scale for a model (of known class) having the best performance.

2.2.1 Box-Cox Transformation

An important group of proper transformations for time-series measured was proposed by Tukey [18],to achieve a simple structured model with normal errors, and constant error variance. This was then modified by Box and Cox [19].This transformation method is regarded as parametric pre-processing technique intended to make the distribution of a set of data approximately Gaussian [19]. Bicego and Baldo indicated that the technique is also useful in Pattern Classification, in case that Gaussianity of datasets is not so critical. They also showed that; probably as a result nonlinear nature of the Box-Cox transformation, class separability can be improved.

The Box-Cox transformation has the following mathematical form [20]

$$Y = (X +) \quad (11)$$

Where λ is the exponent (power) and δ is a shift amount that is added when X is zero or negative. When λ is zero, the above definition is replaced by:

$$Y = \ln(X + \delta) \tag{12}$$

Usually, the standard λ values of -2, -1.5, -1, -0.5, 0, 0.5, 1, 1.5, and 2 are investigated to determine which, if any, is most suitable.

Decompositions "additive and multiplicative" can be reached by means of "Box-Cox transformation" of data with (0 <λ< 1). A value of λ = 0 *corresponds to "multiplicative decomposition" while λ=1 is equivalent to an "additive decomposition".*

Box-Cox transformation which is used to improve time series normality is defined by the equation:

$$Y_t = \begin{cases} \frac{X_t^{\lambda}-1}{\lambda}, & (X_t > 0, \lambda \neq 0) \text{or} (X_t \geq 0, \lambda > 0) \\ \ln(X_t), & X_t > 0, \lambda = 0 \end{cases} \tag{13}$$

2.2.2 Differencing

Differencing is used to remove non-stationarity by modeling the differences. For instance,

$$x(t) - x(t-1) = ARMA\ (p,\ q) \tag{14}$$

The differencing equation is:

$$Y_t = (1-B)^d (1-B^s)^D X_t \tag{15}$$

where d is the order of the first differencing component, s is the seasonal component period, is the order of the seasonal component, and B is the lag operator defined by:

$$BX_t = X_{t-1}$$

Typical values of (d, D, s) are (1,1,s), (2,1,s). s equals (12) for monthly data with an annual seasonality and it equals 0 in case of no seasonality".

This differencing is called as the integration part in ARIMA. Hence, there are three parameters: p which expresses the AR factor, d is the differencing (I) and q is the moving average factor MA.

2.2.3 Seasonal Decomposition

As mentioned above, there are two forms of classical decomposition: "additive and multiplicative decomposition". In additive decomposition, the seasonal component is assumed to be constant from year to year. For "multiplicative seasonality", “seasonal indices” are the m values that form the seasonal component.
Detrending and deseasonalizing using the classical decomposition model could be written as:

$$X_t = m_t + s_t + \epsilon_t \tag{16}$$

where mt is the trend component and st the seasonal component, and et is a N(0,1) white noise component. Using XLSTAT software allows to fit this model in two separate and/or successive steps:

$$X_t = m_t + \epsilon_t = \sum_{i=0}^{k} a_i t^i + \epsilon_t \tag{17}$$

1 – Detrending model:

$$Y_t = \epsilon_t = X_t - \sum_{i=0}^{p} a_i t^i \tag{18}$$

where k is "the polynomial degree". The a_i parameters are obtained by fitting a"linear model to the data. The "transformed time series writes [21]:

$$X_t = s_t + \epsilon_t = \mu + b_i + \epsilon_t,\ i = t \bmod p \tag{19}$$

1 – Deseasonalizing model:

$$X_t = m_t + s_{t\,mod\,p} + \in_t \quad (20)$$

where p is the period. The bi parameters are obtained byfitting the data using a linear regression model too. The transformed time series is:

$$m_t = \sum_{i=-P/2}^{P/2} w_i X_{t+i} \quad (21)$$

where P /2 is the integer division of P by 2 and the coefficients (wi) are :

$$w_i \begin{cases} \frac{1}{2P} si|i = P/2 \\ \frac{1}{P} otherwise \end{cases} \quad (22)$$

Each seasonal index s_i is calculated from $s_t = X_t - m_t$ as the average of the elements of st for which t mod P = i. Their values are then centered as:

$$\widehat{S_\iota} = \widehat{s_\iota} - \frac{1}{p}\sum_{j=1}^{p} \widehat{S_J} \quad (23)$$

The random component is estimated as :

$$\widehat{\epsilon_\iota} = X_t - \widehat{m}_t - \hat{S}_{tmodP} \quad (24)$$

If the "multiplicative type of decomposition" is chosen, the model is given by:

$$X_t = m_t \times s_{tmodp} \times \in_t \quad (25)$$

The trend component is estimated in the same way as indicated for the "additive decomposition".

The seasonal indices si are computed as the average of the elements of $s_t = X_t / m_t$ for which (tmod P= i).

The normalized form is:

$$\hat{S}_i = \hat{s}_i \times \left(\prod_{j=1}^{p} \hat{s}_j\right)^{-1/P} \quad (26)$$

2.3 Smoothing

If {Y t}, (t=1,…,n), the time series of interest, is defined by P_tY_{t+h} the forecaster of Y_{t+h} with minimum "mean square error", and e_t a N(01) white noise. The smoothing methods are as follows**.**

2.3.1Moving Average Smoothing

The first step in a classical "time series decomposition" is to use a "moving average method" to estimate the trend-cycle. A "moving average" of order m can be written as [22]:

$$T_t = \frac{1}{m}\sum_{j=-k}^{k} y_{t+j} \quad (27)$$

Where: m=2k+1, which means that the "estimate of the trend-cycle" is obtained by taking the average of the time series for k periods of t. Consequently, averaging reduces data randomness, leading to a smooth trend-cycle component. This is called an m-**MA**, i.e. a moving average of order m.

2.3.2 Simple Exponential Smoothing (SES)

This is the simplest of the exponentially smoothing methods and it is appropriate if the forecasted data has no obvious trend or seasonal pattern. Sometimes this model is referred to as brown's simple exponential smoothing, or "exponentially weighted moving average model", as the weights decrease exponentially thus, the smallest weights are assigned to the oldest observations. The equations of the model are written as [23]:

Forecast equation $\hat{y}_{t+h|t} = \ell_t$ (28)

Smoothing equation $\ell_t = \alpha y_t + (1-\alpha)\ell_{t-1},$ (29)

Where l_t is the smoothed series value; level, at a time. Substituting h=1provides the fitted values, but if t is set to be equal to T, true forecasts beyond the training data are obtained.

To estimate unknown parameters and initial values of any exponential smoothing method SSE are minimized. The residuals are defined as $e_t = y_t - \hat{y}_{t|t-1}$ for t=1,…, T. Thus, minimize:

$$\textbf{SSE}=\sum_{t-1}^{T}\left(y_t - \hat{y}_{t|t-1}\right)^2 = \sum_{t=1}^{T} e_t^2 \quad (30)$$

As mentioned above, higher weight is given to the more recent observation [21].

2.3.3 Double Exponential Smoothing (Holt's linear trend method)

This model is also called "Brown's Linear Exponential Smoothing" and the method extended "simple exponential smoothing" to permit forecasting data with a trend. The method employs a forecast equation in addition to a smoothing equation for the level and another one for the trend; taking into consideration its variation with time [23].

Forecast equation: $y_{t+h|t} = l_t + hb_t$ (31)

Level equation: $l_t = \alpha y_t + (1-\alpha)(l_{t-1} + b_{t-1})$ (32)

Trend equation: $b_t = \beta^*(l_t - l_{t-1}) + (1-\beta^*)b_{t-1}$ (33)

where ℓ_t represents series level estimate at time t, b_t indicates the series trend estimate at time t, α is the "smoothing parameter", $(0 \le \alpha \le 1)$, and β^* is the "trend smoothing parameter", $(0 \le \beta^* \le 1)$.

Similar to "simple exponential smoothing", equation (32) shows that (ℓ_t) is observation (y_t) "weighted average" and the training forecast one-step-forward of time t is given by $(\ell_{t-1}+b_{t-1})$. Equation (33) indicates that b_t is the "estimated trend weighted average" at time t based on $(\ell_t-\ell_{t-1})$ and b_{t-1}, the previously-estimated trend.

As the h-step-ahead forecast is equal to the last estimated level plus *h* times the last estimated trend value, the forecasts are linear functions of *h*.

2.3.3.1 Holt-Winters Seasonal Additive Model

This method permits the predictions to take into considerations both "trend" and "seasonality". The model was called "additive" due to the stable nature of the seasonality effect. The additive method is formulated as [24]:

$$y_{t+h|t} = l_t + hb_t + s_{t+h-m(k+1)} \quad (34)$$

$$l_t = \alpha(y_t - s_{t-m}) + (1-\alpha)(l_{t-1} + b_{t-1}) \quad (35)$$

$$b_t = \beta^*(l_t - l_{t-1}) + (1-\beta^*)b_{t-1} \quad (36)$$

$$s_t = \gamma(y_t - l_{t-1} - b_{t-1}) + (1-\gamma)s_{t-m} \quad (37)$$

Where: k is the integer part of (h−1)/m. l_t equation showing a "weighted average" between the "seasonally adjusted observation" $(y_t - s_t - m)$ and the "non-seasonal forecast" $(\ell_{t-1}+b_{t-1})$ for time t. Trend equation is the same as "Holt's linear method". Similarly, "seasonal equation" shows a "weighted average" between the "current seasonal index", $(y_t - \ell_{t-1} - b_{t-1})$, and the "seasonal index" of last year's same season.

The seasonal component equation is written as:

$$s_t = \gamma*(y_t - \ell_t) + (1-\gamma*)\, s_{t-m} \quad (38)$$

Substituting ℓ_t from equation (29), then:

$$s_t = \gamma*(1-\alpha)(y_t-\ell_{t-1}-b_{t-1})+[1-\gamma*(1-\alpha)]\, s_{t-m} \quad (39)$$

which is the same as the seasonal component smoothing' equation with $\gamma=\gamma*(1-\alpha)$.

2.3.3.2 Holt-Winters Seasonal Multiplicative Model

This method considers the case of a trend; varying with time, and a seasonal component with a period p, which is the reason of name "multiplicative model". As the discrepancies between the observations increase, the seasonal component also increases. The model equations are as follows:

The component form for the multiplicative method is [23]:

$$y_{t+h|t} = l_t + hb_t + s_{t+h-m(k+1)} \quad (40)$$
$$l_t = \alpha \frac{y_t}{(s_{t-m})} + (1-\alpha)(l_{t-1} + b_{t-1}) \quad (41)$$
$$b_t = \beta^*(l_t - l_{t-1}) + (1-\beta^*)b_{t-1} \quad (42)$$
$$s_t = \gamma \frac{y_t}{(l_{t-1} - b_{t-1})} + (1-\gamma)s_{t-m} \quad (43)$$

3 Results and Discussion

Three years data sets of Riyadh, KSA temperature (°F) obtained from the website https://www.wunderground.com/ are used to define the time series model that best fit the data using XLSTAT software package. ARIMA modeling of the time-temperature series is shown in Figure 1. The series ACF (a) and PAC (b) plots are shown in Figure.2.

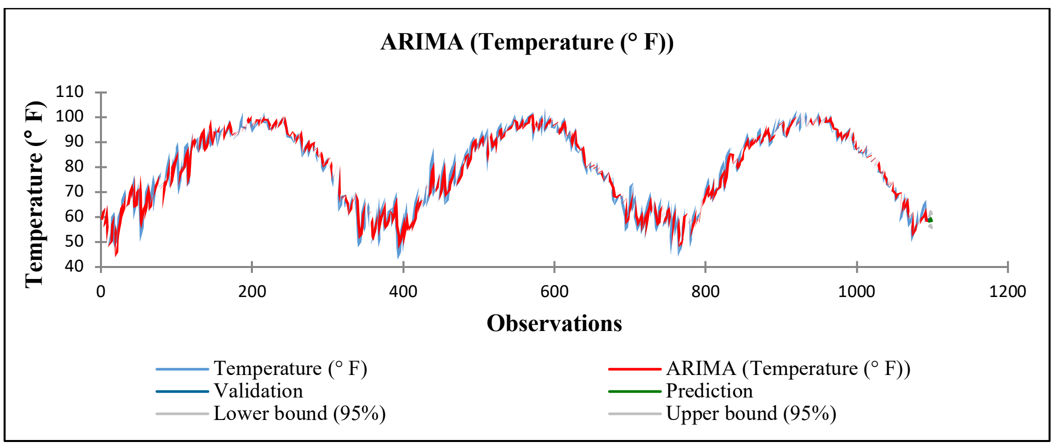
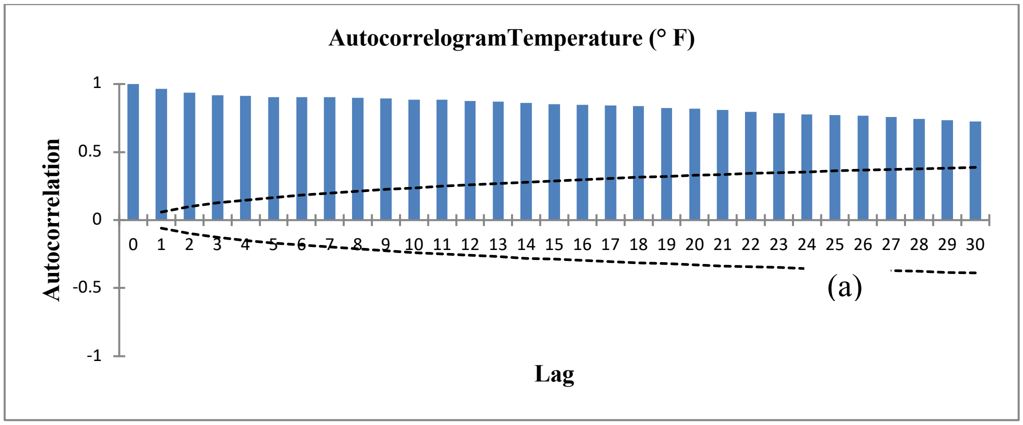


Fig. 1: Temperature ARIMA model of Riyadh, KSA.



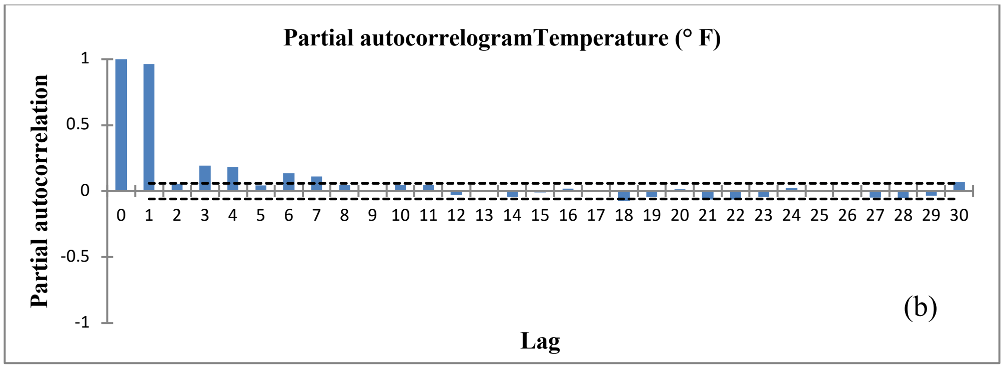


Fig. 2: Autocorrelation (a) and partial autocorrelation (b) of temperature time series.

ACF and PACF plots are used to identify the proper ARMA model by comparison with the identification characteristics method.

Different investigated smoothed models are tested for Goodness of fit calculating the coefficients, SSE, MSE, and $R^2$where:

SSE is the Sum of Squares of Errors: This statistic is minimized if the Least Squares option has been selected for the optimization.

MSE is the Mean Square Error and R^2 is the correlation factor. Data summary statistics are shown in Table 1.

Table 1: Data summary statistics.

Variable	Observations	Minimum	Maximum	Mean	Std. deviation
Series1	1096	43.000	104.000	79.375	15.244

3.1 Moving Average Results

Table 2: exhibit the results of using the moving average method.

Table 2: Moving average results.

Model parameter k (Temp. ° F)	DF	SSE	MSE	RMSE	MAPE	MPE	MAE	R^2
2	1096	8147.7	7.434058	2.726547	2.774814	-0.195	1.999103	0.96798
4	1096	13235	12.07591	3.475041	3.602905	-0.292	2.560795	0.94798
6	1096	15907	14.51375	3.809692	3.977352	-0.349	2.811272	0.93748

8	1096	16959	15.47348	3.933635	4.145084	-0.370	2.924862	0.93335

From Table 2 it clear that Goodness of Fit test results indicate that minimum errors (SSE, MSE, and RMSE) obtained for the parameter value (k =2). The correlation factor (R^2) in this case is 0.96798.

3.2 Exponential Smoothing Results

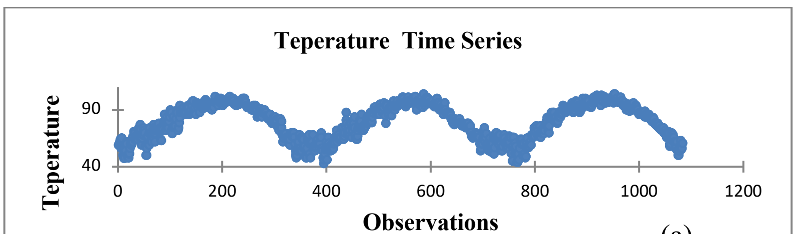
Table 3, exhibits the results of using simple and double exponential smoothing for the 1096 observations time series, for different values of α (0 <α < 1).

The temperature data series is first testes for stationarity and appeared to be nonstationary. Hence, it is stationarized investigating different transformation techniques; differencing, detrending and seasonality. Figure 3(a, b, c, d, e) illustrates the original time series and the data series transformation using each of these methods.

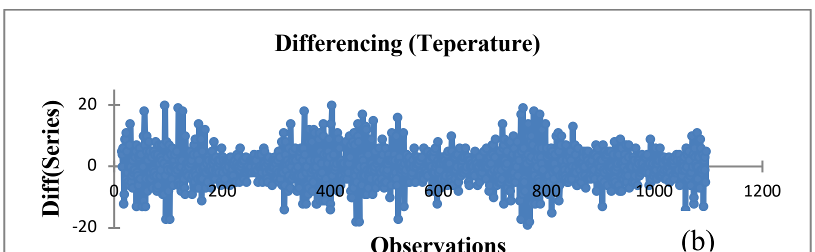
Table 3: Exponential smoothing results

Smoothing Method	Model parameters α Temperature (° F)	DF	SSE	MSE	RMSE	R^2
Simple exponential	0.2	21115	19.28304	4.391246	4.68667	0.916881
Simple exponential	0.4	18387	16.79186	4.097788	4.30566	0.927619
Simple exponential	0.6	17304	15.80287	3.975282	4.14126	0.931882
Simple exponential	0.8	16850	15.38858	3.922827	4.03124	0.933668
Double exponential	0.2	175993	160.7245	12.67772	6.28982	0.307204
Double exponential	0.4	33206	30.3249	5.506805	4.90302	0.869286
Double exponential	0.6	24895	22.73479	4.768101	4.82119	0.902003
Double exponential	0.8	27699	25.29593	5.029506	5.15803	0.890963

From Table 3 it could be seen that the goodness of fit test results indicate that minimum errors (SSE, MSE, and RMSE) are obtained for α = 0.8 using simple exponential smoothing. The correlation factor (R^2) in this case is 0.9337.



(a)



(b)

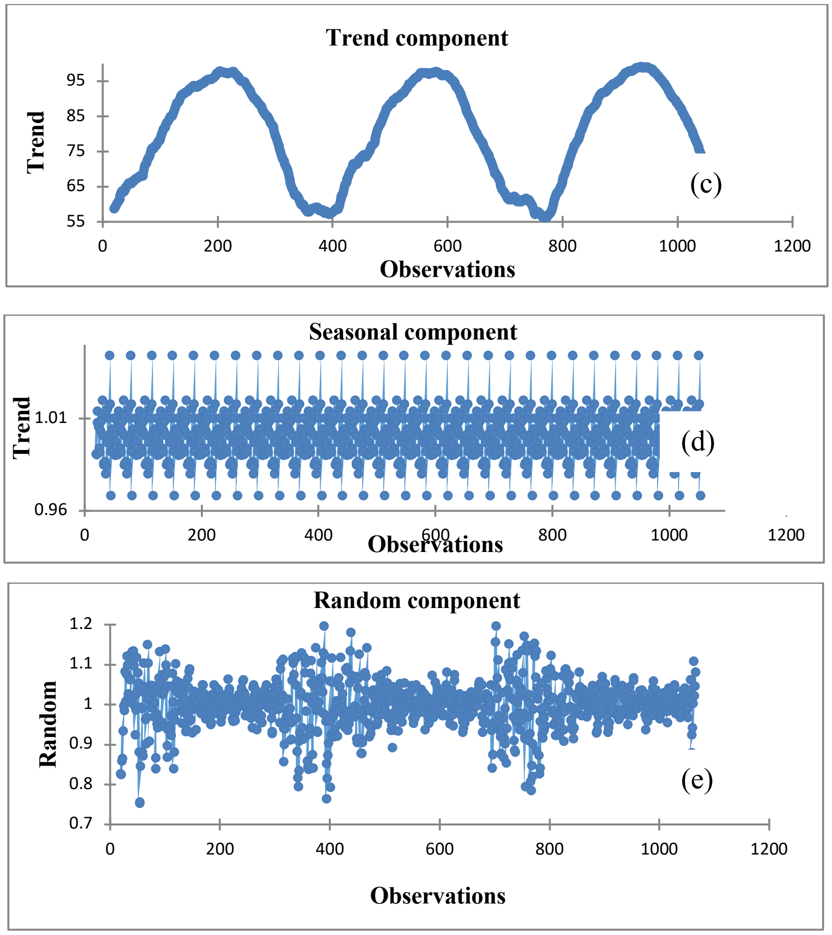


Fig. 3: Original data (a), Differencing (b), Trend(c), Seasonal (d), and Random component (e) of the fitted series

3.3 Double Exponential Smoothing with Linear Trend (Holt-Winters)

Using smoothing with trend adjustment (Holt) method with different values of (α) and (β), the output results are shown in Table 4.

Table 4: Double exponential smoothing for different values of (α) and (β),

Model parameters (Temp.° F))		DF	SSE	MSE	RMSE	MAPE	R^2
α	β						
0.2	0.2	1092	23828	21.82023	4.671212	4.907857	0.906063
0.4	0.2	1092	21845	20.00448	4.472637	4.63389	0.91388
0.6	0.2	1092	20500	18.77306	4.332789	4.480055	0.919181
0.8	0.2	1092	19886	18.21059	4.267387	4.353577	0.921603
0.2	0.4	1092	28217	25.83949	5.083256	5.334534	0.88876
0.2	0.6	1092	32776	30.01486	5.478582	5.701223	0.870785
0.2	0.8	1092	36944	33.83151	5.816486	6.048816	0.854354

From table 4, it is clear that minimum errors are obtained for smoothing parameter **(α = 0.8**) and trend parameter **(β =0.2)**. The correlation factor in this case (R^2 = 0.9216).

3.3.1 Results of Holt-Winters Seasonal Multiplicative Model

Using smoothing while taking into account both trend and the seasonality; Holt seasonal additive method, with different values of the model parameters (α,β, and γ) the output results are shown in Table 5.

Table 5: Results of Holt-Winters seasonal multiplicative model.

Model parameters			DF	SSE	MSE	RMSE	MAPE	MPE	R^2
α	β	γ							
0.2	0.2	0.2	1080	27763	25.7068	5.070187	5.240485	-0.349	0.887509
0.2	0.2	0.4	1080	30521	28.25985	5.315999	5.4728	-0.397	0.876337
0.2	0.2	0.6	1080	35231	32.62124	5.711501	5.885435	-0.471	0.857252
0.2	0.2	0.8	1080	42900	39.72219	6.302554	6.492011	-0.582	0.826178

From Table 5 it is clear that minimum errors are realized for the model parameters **(α,β, and)** each equals **0.2**. In this case $R^2 = 0.8875$

3.3.2 Holt-Winters Additive Model

Table 6: Results of Holt-Winters seasonal additive model.

Model parameters			DF	SSE	MSE	RMSE	MAPE	MPE	R^2
α	β	γ							
0.2	0.2	0.2	1080	26983	24.98385	4.998384	5.161299	-0.269	0.890672
0.2	0.2	0.4	1080	30003	27.78016	5.270689	5.430159	-0.278	0.878436
0.2	0.2	0.6	1080	34457	31.90504	5.648455	5.826373	-0.285	0.860386
0.2	0.2	0.8	1080	41182	38.13172	6.175088	6.366483	-0.291	0.833138

From table 6 it is clear that minimum errors are realized for the model parameters **(α,β, and γ)** each equals **0.2**. In this case $R^2 = 0.89067$

To sum the above-mentioned results, the obtained parameter values for each smoothing method and the correlation factor are exhibited in Table 7.

Table 7: Summary of the results

Model	**Parameters**			$\mathbf{R^2}$
	α	β	γ	
Simple exponential smoothing	0.8	--	--	0.9337
Double exponential smoothing	0.8	0.2	--	0.9216
Holt-Winters seasonal multiplicative model	0.2	0.2	0.2	0.8875
Holt-Winters seasonal additive model	0.2	0.2	0.2	89067

It is clear that simple exponential smoothing is the best forecasting model in the case under study as it results in the highest correlation factor ($R^2 = 0.9337$).

Figure 4 presents the exponentially-smoothed time series simple exponential (a) and the residuals (b) for the best-obtained fit with α =0.2.

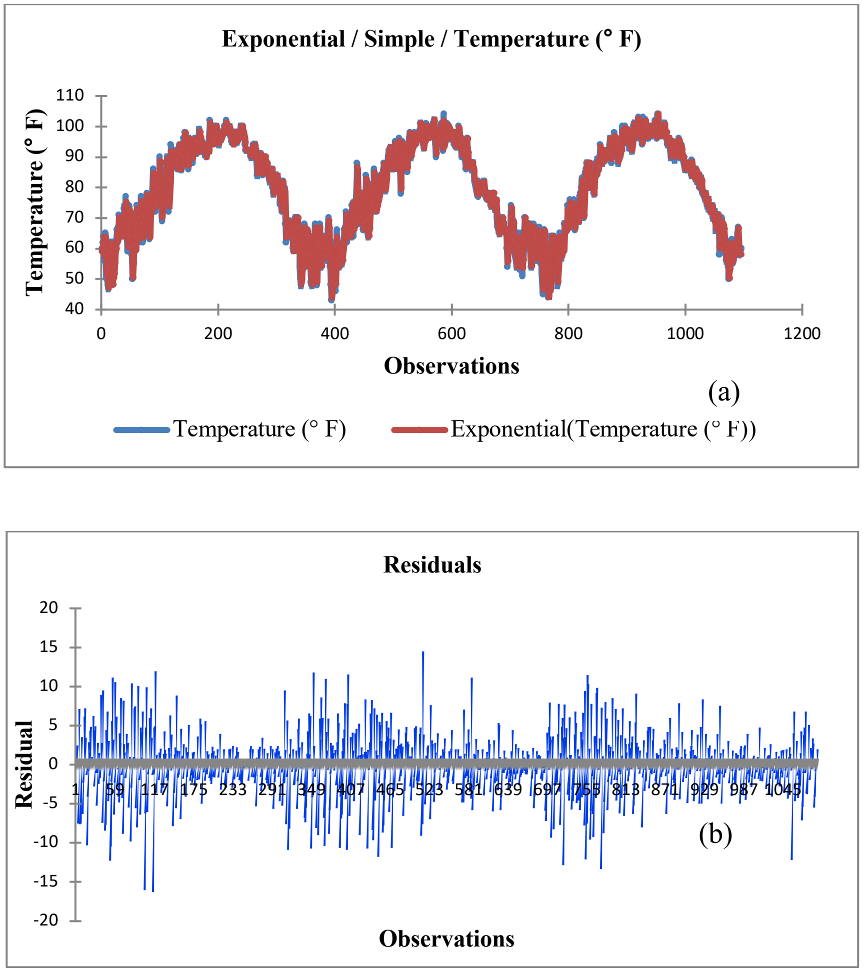


Fig. 4: Simple exponentially smoothed time series (a) and residuals (b)
(For the best fit with α =0.2)

From Figure 4 it could be determined that the fit is adequate as the residual errors are randomly-distributed with zero means.

4 Conclusions

Autoregressive Integrated Moving-Average" (ARMIA) models are applied for the estimation of temperature distribution forecast model.

The smoothing models used to fit the time series are adequate as the residual errors are randomly distributed with almost zero mean.

Based on the value of the correlation factor R^2, simple exponential smoothing is the most adequate forecasting model; following the moving average, for the case under study (forecasting temperature distribution for Riyadh, KSA), as for this model R^2 is found to be equal to (0.9337) which is higher than double exponential, Holt-Winters seasonal multiplicative and additive models value of the investigated cases.

The models of autoregressive integrated moving-average (ARMIA) may be applied on some new subjects in engineering and applied mathematics (see Refs. [25]-[32]).

Acknowledgment

I would like to thank Department of Mathematics, College of Science and Humanities in Al-Kharj, Prince Sattam Bin Abdulaziz University, Al-kharj 11942, Saudi Arabia for their kind supporting my research work, so I am feeling confident to be a faculty member in this brilliant institution.

References

[1] M. Khashei, M. Bijari, and G. A. Raissi Ardali, "Hybridization of autoregressive integrated moving average (ARIMA) with probabilistic neural networks (PNNs)," Computers & Industrial Engineering., **63**, 1, 37–45, 2012.

[2] R. Haiges, Y. D. Wang, A. Ghoshray, and A. P. Roskilly, "Forecasting Electricity Generation Capacity in Malaysia: An Auto Regressive Integrated Moving Average Approach," Energy Procedia., **105**, 3471–3478, 2017.

[3] S. Dégerine and S. Lambert-Lacroix, "Characterization of the partial autocorrelation function of nonstationary time series," Journal of Multivariate Analysis., **87**(1), 46–59, 2003.

[4] K. Kumar and V. K. Jain, "Autoregressive integrated moving averages (ARIMA) modeling of a traffic noise time series," Applied Acoustics., 58(3), 283–294, 1999.

[5] S. Jia, T. Xu, Z. Sun, and J. Li, "Middle and long-term prediction of UT1-UTC based on the combination of Gray Model and Autoregressive Integrated Moving Average," Advances in Space Research., **59(3)**, 888–894, 2017.

[6] H. Yang, Z. Pan, Q. Tao, and J. Qiu, "Online learning for vector autoregressive moving-average time series prediction," Neurocomputing., **315**, 9–17, 2018.

[7] Broughton, S.A.; Bryan, K. "Discrete Fourier Analysis and Wavelets: Applications to Signal and Image Processing". New York: Wiley., **72**, 2008.

[8] C. Chatfield and M. Yar, "Prediction intervals for multiplicative Holt-Winters," International Journal of Forecasting., **7(1)**, 31–37,1991.

[9]Yule, G.U. "Why Do We Sometimes Get Nonsense-Correlations between Time Series: A Study in Sampling and the Nature of Time Series". Journal of Royal Statistical Society., **89**, 1-64, 1926.

[10] Wold, H. "A Study in the Analysis of Stationary Time Series". Stockholm: Almgrist &Wiksell.,1938.

[11] Box, G. E. P and Jenkins, G.M., "Time series analysis: Forecasting and control". Holden-Day, San Francisco., 1976.

[12] A. Steland, "Weighted Dickey-Fuller processes for detecting stationarity," Journal of Statistical Planning and Inference., **137(12)**, 4011–4030, 2007.

[13] Ahmad Rizal Alias, Noor Yasmin Zainun and Ismail Abdul Rahman" Comparison between ARIMA and DES Methods of Forecasting Population for Housing Demand in Johor" MATEC Web of Conferences 81 07002 ICTTE ., 2016.

[14] Jozef Barunik" Autocorrelation and Partial Autocorrelation Functions of AR(1) Process" http://demonstrations.wolfram.com/AutocorrelationAndPartialAutocorrelationFunctionsOfAR1Proces/WolframnDemonstrationsProject , March **7**, 2011.

[15] T. Etchison, S. G. Pantula, and C. Brownie, "Partial autocorrelation function for spatial processes," Statistics & Probability Letters, **21(1)**, 9–19,. 1994.

[16] O. Barndorff-Nielsen, G. Schou, "On the parameterization of autoregressive models by partial autocorrelations". J. Multivariate Anal., **3**, 408–419, 1973.

[17] F.L. Ramsey, Characterization of the partial autocorrelation function, Ann. Statist., **2**, 1296–1301, 1974.

[18] Tukey, J. W.,"On the comparative anatomy of transformations", Annals of Mathematical Statistics., 28, 602-632, 1957.

[19]Box G. E. P. and Jenkins G. M. "Time Series Analysis: Forecasting and Control", 3rd edition. Pearson Education, Upper Saddle River,1984.

[20] R. M. SAKIA, "The Box-Cox transformation technique: a review", The Statistician., **41**, 169-178, 1992.

[21] E. B. Dagum, & S. Bianconcini, "Seasonal adjustment methods and real-time trend-cycle estimation". Springer., 2016.

[22] I. T. Christou, Quantitative Methods in Supply Chain Management, DOI: 10.1007/978-0-85729-766-2_2, Springer-Verlag London Limited , 2012.

[23] Hyndman, R. J., Koehler, A. B., Ord, J. K., & Snyder, R. D. (2008). "Forecasting with exponential smoothing: The state space approach". Berlin: Springer-Verlag.http://www.exponentialsmoothing.net, 2008.

[24] Xiaochen Li, "Comparison and Analysis between Holt Exponential Smoothing and Brown Exponential Smoothing Used for Freight Turnover Forecasts", Third International Conference on Intelligent System Design and Engineering Applications, 2013.

[25] Abd-alla A.N., Hamdan A.M., Giorgio I. and Del Vescovo D., The mathematical model of reflection and refraction of longitudinal waves in thermo-piezoelectric materials, Archive of Applied Mechanics., **84**, 1229-1248, 2014.

[26] Abd-alla A.N., Giorgio I., Galantucci L., Hamdan A.M. and del Vescovo D., Wave reflection at a free interface in an anisotropic pyroelectric medium with nonclassical thermoelasticity, Continuum Mechanics and thermodynamics., **28**, 67-84, 2016.

[27] Abd-alla A.N., Alshaikh, F., Giorgio, I., Della Corte, A., A mathematical model for longitudinal wave propagation in a magnetoelastic hollow circular cylinder of anisotropic material under the influence of initial hydrostatic stress, Mathematics and Mechanics of Solids., **21(1)**, 104-118, 2016.

[28] Abd-alla A.N., Raizah, A., Placidi, L., The influence of hydrostatic stress on the frequency equation of flexural waves in a magnetoelastic transversely isotropic circular cylinder, ZAMM Zeitschrift fur Angewandte Mathematik und Mechanik., **96(1)**, 53-66, 2016.

[29] Abbas I.A., Abd-alla A.N., Alzahrani F.A. and Spagnuolo M., Wave propagation in a generalized thermoelastic plate by using eigenvalue approach, Journal of Thermal Stresses., **39(11)**,1367-1377, 2016.

[30] Abd-alla A.N. and Asker N., Numerical simulations for the phase velocities and the electromechanical coupling factor of the Bleustein-Gulyaev waves in some piezoelectric smart materials, Mathematics and Mechanics of Solids., **21(5)**, 539-551, 2016.
[31] Abd-alla A.N., Alshaikh F., Mechai I. and Abass I.A., Influence of initial stresses and piezoelectric constants on the propagation bulk acoustic waves in an anisotropic smart material (Aluminum Nitrite), Journal of Journal of Computational and Theoretical Nanoscience., **13**, 1-7, 2016.
[32] Abd-alla A.N., Hasbullah N.F. and Hossen H.M., The frequency equations for shear horizontal waves in semiconductor/ piezoelectric structures under the influence of initial stress, Journal of Journal of Computational and Theoretical Nanoscience., **13**, 7-14, 2016.

Afrah Al-Bossly is an assistant professor of mathematics, "Prince Sattam Bin Abdulaziz University", KSA. She awarded her Ph.D. degree in: "A Study of Some Selected Bivariate Exponential with Applications in Reliability". Her interests are in the areas of applied statistics including engineering systems (methods and models). She has published some research articles in international journals of statistical and engineering sciences. She had held various managerial positions with good experience.