Archives of Current Research International

Volume 25, Issue 4, Page 120-131, 2025; Article no.ACRI.126836
ISSN: 2454-7077

On the Effects of Missing Values on Estimates of Trend Parameters and Seasonal Indices in Descriptive Time Series Analysis

Lawrence C. Kiwu [a*], Eleazar C. Nwogu [a], Chukwudi J. Ogbonna [a], Hycinth C. Iwu [a] and Iheanyi S. Iwueze [a]

[a] *Department of Statistics, Federal University of Technology, Owerri, Nigeria.*

Authors' contributions

This work was carried out in collaboration among all authors. All authors read and approved the final manuscript.

Article Information

DOI: https://doi.org/10.9734/acri/2025/v25i41142

Open Peer Review History:
This journal follows the Advanced Open Peer Review policy. Identity of the Reviewers, Editor(s) and additional Reviewers, peer review comments, different versions of the manuscript, comments of the editors, etc are available here: https://pr.sdiarticle5.com/review-history/126836

Original Research Article

Received: 15/09/2024
Accepted: 18/11/2024
Published: 26/03/2025

ABSTRACT

In literature, when missing values are observed in a series, emphasis has always been on obtaining the estimates of the missing values while little or no attention has been given to assessing the effects of the missing values on parameter estimates. The aim of this study therefore is to at assess the effect of missing values on the estimates of trend parameters and seasonal indices in descriptive time series analysis when trending curve is linear and the decomposition model is additive. Estimates of trend parameters and seasonal indices were obtained using descriptive time

---

**Corresponding author: Email: lawrence.kiwu@futo.edu.ng;*

Cite as: *Kiwu, Lawrence C., Eleazar C. Nwogu, Chukwudi J. Ogbonna, Hycinth C. Iwu, and Iheanyi S. Iwueze. 2025. "On the Effects of Missing Values on Estimates of Trend Parameters and Seasonal Indices in Descriptive Time Series Analysis". Archives of Current Research International 25 (4):120-31. https://doi.org/10.9734/acri/2025/v25i41142.*

series methods, while the performances of the estimates in the presence and absence of missing values are assessed using the values of summary statistics (MSE, RMSE and MAE), based on deviation of estimates from the actual parameters used in simulation. The results show that estimates of the trend parameter, error mean and standard deviation appear not to be affected when the number of missing values is less than ten. Estimates of seasonal indices also seem to be affected only slightly. For ten or more missing values, the summary statistics are only slightly higher especially when the missing follow consecutives. Specifically, the study found that trend parameters, error mean and error standard deviation used in simulation are recovered better than seasonal indices. It has therefore been recommended that when using the descriptive time series methods to obtain estimates of trend parameters and seasonal indices in presence of missing values there is no need to obtain estimates of the missing values first.

Keywords: Trend parameters; seasonal indices; consecutive; separated; accuracy measures.

1. INTRODUCTION

In time series analysis, data is required to be complete, made sequentially in time and equally spaced. In other words, it makes no provision for missing observations. However, in real life missing values are unavoidable for some reasons. Most time series data such as financial time series does not require the data to have missing observations over a long period of time. This is because the statistical properties of the series are preserved by its sequence using such a complete data. Schmitt et al., (2015) confirmed that missing data does not only affect properties of statistical estimations, it also introduces some element of ambiguity in the data analysis and result in misleading inferences and conclusions. Similarly, the correlation structure of a dataset may not be captured if the decision is to replace missing values with zeros, Kerkri et al., (2015). Thus, the need to address the problem of missing values before proceeding with the analysis (Sohae, 2015). Some conventional methods used in handling missing values include ignoring or deleting the missing observation, Arslan & Aydilek, (2013). This is in contrast to one of the assumptions of the Box-Jenkins method which entails that the series be equally spaced over time and that there are no lost values in the series Yaffee & McGee, (1999). Each record in time series is unique; hence, ignoring missing observation may result to having a series that is unusable for many purposes with truncated time series plot Tusell, (2005). When analyst encounter missing observations in time series data, one of the remedial measures is to replace the missing observation by its estimate (David, 2006 and Howell, 2007). Imputation of missing value is an enormous field of study, where a lot of research has already been conducted. Popular techniques include the works of Rubin, (1987), Dempster, Laird, Rubin, (1977) and Vacek & Ashikaga, (1980) on Multiple Imputation, Expectation-Maximization and Nearest Neighbor respectively. However, different methods of missing values imputation have been observed to give inconsistent estimates. This led to the comparison of estimates using different summary statistics by Iwueze et al., (2018). Adejumo et al., (2021) used similar summary statistics to compare different imputation methods on estimates of missing values. The study observed that the Kalman filter algorithm (KAL) produced the best missing value estimates. However, the KAL involves complex computational algorithm, a major limitation to the use of the method. In other to generalize on the best method of estimating missing values, Afkanpour et al., (2024) carried out a review on various imputation methods in the healthcare field. The study classified the missing value imputation methods into; conventional statistical, machine learning, deep learning, and applied hybrid imputation methods. The study found that greater percentage of researcher use the conventional statistical methods followed by machine learning methods of missing values imputation. From the foregoing, most studies have concentrated efforts on obtaining the estimates of missing values without considering the effects of the missing values on the estimates of the parameters of the models. This study is therefore aimed at assessing the effect of missing values on the estimates of trend parameters and seasonal indices in descriptive time series, when trending curve is linear and decomposition method is additive.

2. MATERIALS AND METHODS

The traditional method of time series (also known as the descriptive time series analysis) is adopted for the decomposition of the observed series. The models are commonly used for the decomposition are:

The Additive model,

$$X_t = T_t + S_t + C_t + e_t, \quad t = 1, 2, ..., n \quad (2.1)$$

the Multiplicative model,

$$X_t = T_t \times S_t \times C_t \times e_t, \quad t = 1, 2, ..., n \quad (2.2)$$

and the Mixed Model.

$$X_t = T_t \times S_t \times C_t + e_t, \quad t = 1, 2, ..., n \quad (2.3)$$

where for time t, T_t is the trend, S_t is the seasonal effect, C_t is cyclical and e_t is irregular or error term. According to Chatfield (2004) the cyclical component is superimposed into the trend when a short period of time is involved, when this happens, the trend-cycle component is given by M_t. This adjustment reduces equations (2.1) – (2.3) to;

The Additive,

$$X_t = M_t + S_t + e_t, \quad t = 1, 2, ..., n \quad (2.4)$$

the Multiplicative model,

$$X_t = M_t \times S_t \times e_t, \quad t = 1, 2, ..., n \quad (2.5)$$

and the Mixed Model.

$$X_t = M_t \times S_t + e_t, \quad t = 1, 2, ..., n \quad (2.6)$$

where,

$M_t = a + bt$ is the trending curve, a and b are the trend parameters, S_j is the seasonal indices and e_t is the error.

This study is limited to the model given in equation (2.4) with following assumptions,

(a) The irregular component e_t is normal with mean zero and constant variance, that is,

$$e_t \sim N\left(0, \sigma^2\right) \quad (2.7)$$

(b) The seasonal effect, when it exists, has periods; it repeats after s time periods. That is,

$$S_{t+s} = S_t, \text{ for all } t \quad (2.8)$$

subject to;

$$\sum_{j=1}^{s} S_{t+j} = 0 \quad (2.8)$$

The estimates seasonal indices ($\hat{S}_j$) in Equation (2.9) were calculated by subtracting the overall average of the detrended time series from the seasonal averages; Box et al., (2016).

$$\bar{S}_j = \frac{1}{m}\sum_{i=1}^{m}\left(X_{ij} - T_{ij}\right) \quad (2.9)$$

where, i is the period, j represent the season and m is the number of observations for that season. The seasonal index in Equation (3.0) is adjusted to meet the condition in Equation (2.8).

$$\hat{S}_j = \text{Adjusted}\left(\bar{S}_j\right) \quad (2.10)$$

In traditional method of decomposition, the trend component M_t in the observed series is first identified and isolated either by subtraction for Equation (2.4). The de-trended series is obtained as $X_t - \hat{M}_t$ for Equation (2.4). From the detrended series, the seasonal component is estimated and isolated by estimating the average of the de-trended series at each season. The de-trended, de-seasonalized series is obtained as $X_t - \hat{M}_t - \hat{S}_t$ for Equation (2.4). This gives the irregular component which may or may not be random is left.

3. PARAMETER ESTIMATION

According to Agung et al., (2020), the estimates of trend parameters is given by,

$$\hat{b} = \frac{\sum_{i=1}^{n}\left(t_i - \bar{t}\right)\sum_{i=1}^{n}\left(x_i - \bar{x}\right)}{\sum_{i=1}^{n}\left(t_i - \bar{t}\right)^2} \quad (3.1)$$

$$\hat{a} = \frac{1}{n}\sum_{i=1}^{n} x_i - \hat{b}\left(\frac{1}{n}\sum_{i=1}^{n} t_i\right) \quad (3.2)$$

4. ACCURACY MEASURE ASSESSMENT

The effect of missing values on the model parameters will be assessed using the accuracy measures. The accuracy measures; Mean square error (MSE), Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) are based on deviation of the parameter estimates from the parameters. When ϕ denotes the parameters used in simulation while $\hat{\underline{\phi}}$ denotes the corresponding estimates in the presence of the missing values. That is,

$$\hat{e} = \underline{\phi} - \hat{\underline{\phi}} \quad (3.3)$$

where,

$$\underline{\phi}' = [a, b, s_1, s_2, s_3, s_4, s_5, s_6, s_7, s_8, s_9, s_{10}, s_{11}, s_{12}, \mu_e, \sigma_e]$$

and

$$\hat{\underline{\phi}}' = \left[\hat{a}, \hat{b}, \hat{s}_1, \hat{s}_2, \hat{s}_3, \hat{s}_4, \hat{s}_5, \hat{s}_6, \hat{s}_7, \hat{s}_8, \hat{s}_9, \hat{s}_{10}, \hat{s}_{11}, \hat{s}_{12}, \hat{\mu}_e, \hat{\sigma}_e\right]$$

The accuracy measures are defined as:

$$MSE = \frac{1}{k}\sum_{i=1}^{k} e_i^2 \quad (3.4)$$

$$RMSE = \sqrt{\frac{1}{k}\sum_{i=1}^{k} e_i^2} \quad (3.5)$$

$$MAE = \left[\frac{1}{k}\sum_{i=1}^{k} |e_i|\right] \quad (3.6)$$

$i = 1, 2, ..., k$; t is the number of parameters

5. EMPIRICAL RESULTS

The empirical examples for this study consist of 100 series of 120 observations each simulated from the Additive Model: $X_t = M_t + S_j + e_t$ with $M_t = a + bt$, $a = 1.0$ and $b = 2.0$, e_t (the error component) assumed, $e_t \sim N(0,1)$ and seasonal indices, $S_j, j = 1, 2, ..., 12$ shown in Table 1. The simulations were done with Minitab Statistical Software, Version 20. For want of space, results for only six series, one each for none, one, two, five, ten and fifteen, missing values are shown in Table 2 through 4, while the corresponding graphs are shown in Figs. 1 through 6.

As Fig. 1 shows, the plot is the normal plot of a series without missing value moving in an upward direction in a linear form. Fig. 2 is the time plot of series moving upward in a linear form but with a missing value at position 3.

Figs. 3(a) and (b) are the time plots of series moving upward in a linear form but with two missing values at consecutive and separated positions respectively.

Similarly, Figs. 4(a) and (b) show the time plots of series with five missing values at consecutive and separated positions respectively. The time plots for series with ten missing values at consecutive and separated positions respectively are shown in Figs. 5(a) and (b).

Furthermore, Figs. 6(a) and (b) are the time plots of series with fifteen consecutive and separated missing values.

Furthermore, Figs. 6(a) and (b) are the time plots of series with fifteen missing values at consecutive and separated positions respectively. In all the plots, the shape of the plots remained the same but with gabs in the position of the missing values.

6. ESTIMATES OF TREND PARAMETERS AND SEASONAL INDICES

Table 2 presents the estimate of trend parameters and seasonal indices for series with no missing value and series with one missing value at different positions. Estimates of trend parameters are derived using Equations (2.7) and (2.8), while estimates of seasonal indices are derived using Equation (3.0). The errors $(\hat{e})$, that is, the difference between the actual value and estimated values of the parameters and summary statistics for the series are also shown in Table 2.

Table 1. The seasonal indices for the simulated series

j	1	2	3	4	5	6	7	8	9	10	11	12
S_j	-1.5	2.5	3.5	-4.5	-1.5	2.5	3.5	-4.5	-1.5	2.5	3.5	-4.5

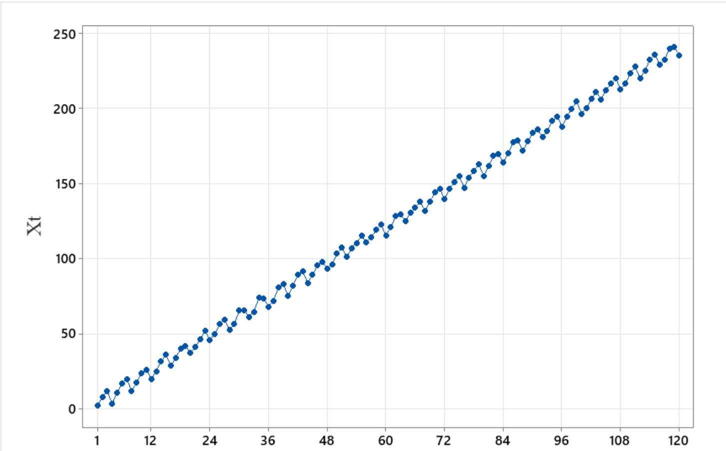

Fig. 1. Time plot of simulated series without missing value

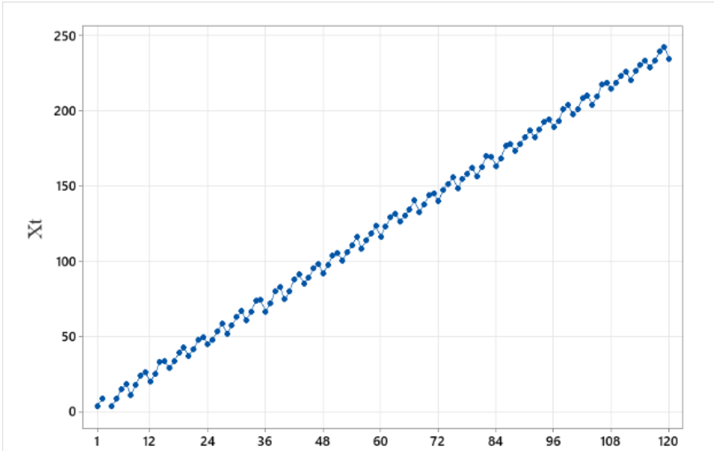

Fig. 2. Time plot of simulated series with one missing value at position 3

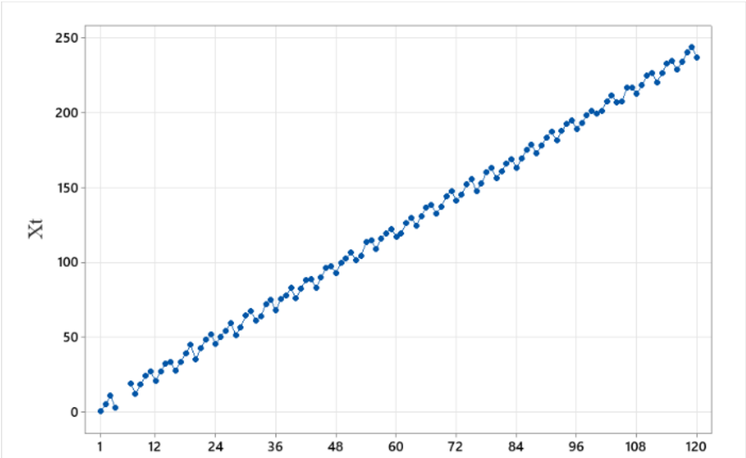

Fig. 3(a). Time plot of simulated series with two consecutive missing values

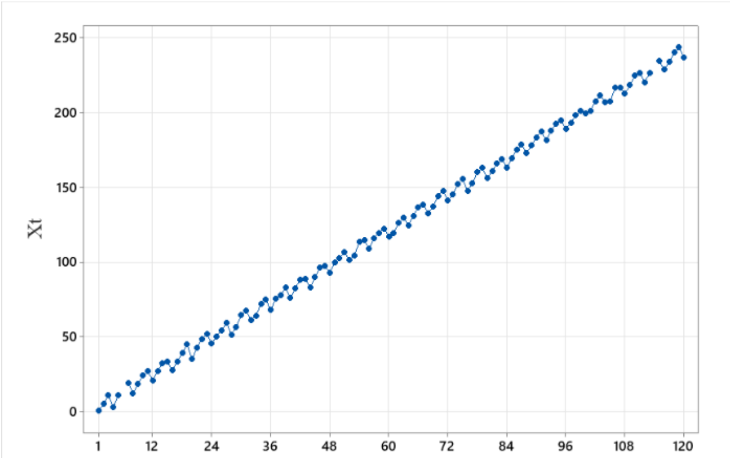

Fig. 3(b). Time plot of simulated series with two separated missing values

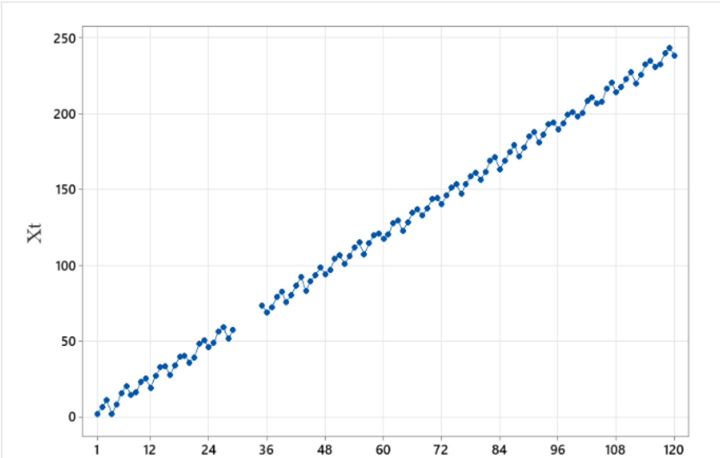

Fig. 4(a). Time plot of simulated series with five consecutive missing values

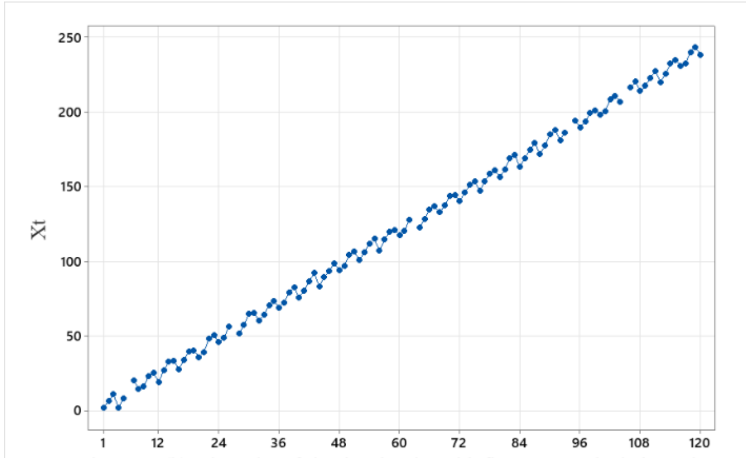

Fig. 4(b). Time plot of simulated series with five separated missing values

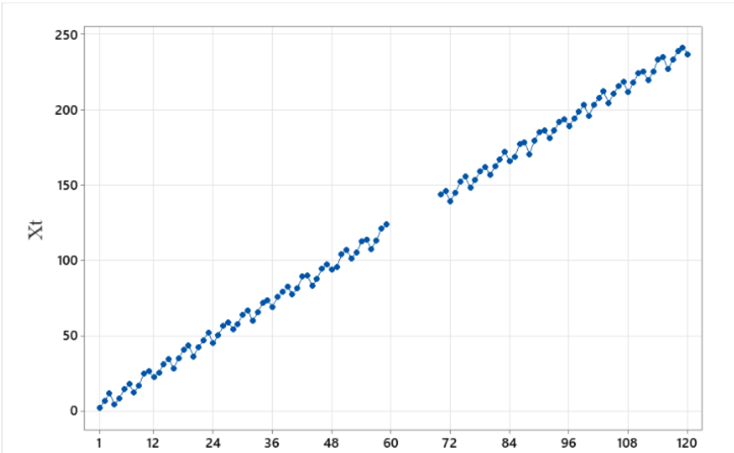

Fig. 5(a). Time plot of simulated series with ten consecutive missing values

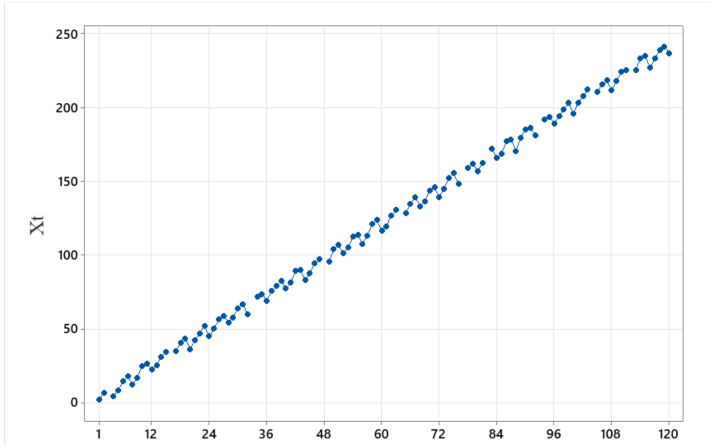

Fig. 5(b). Time plot of simulated series with ten separated missing values

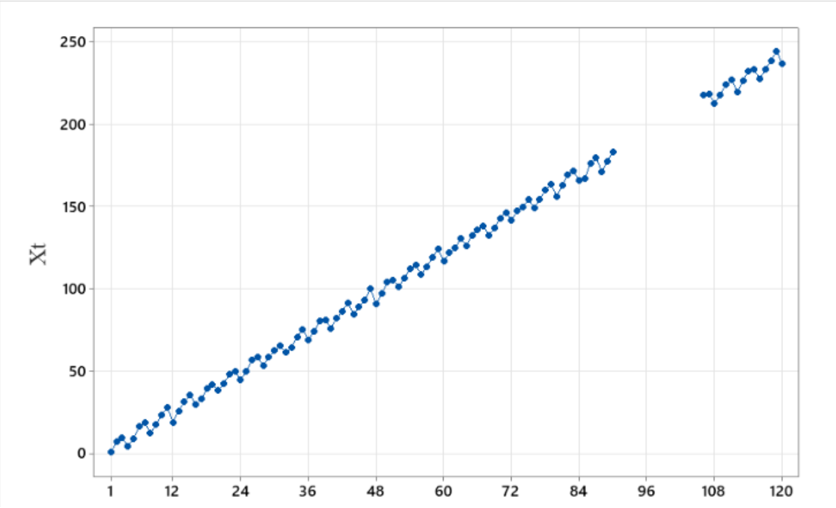

Fig. 6(a). Time plot of simulated series with fifteen consecutive missing values

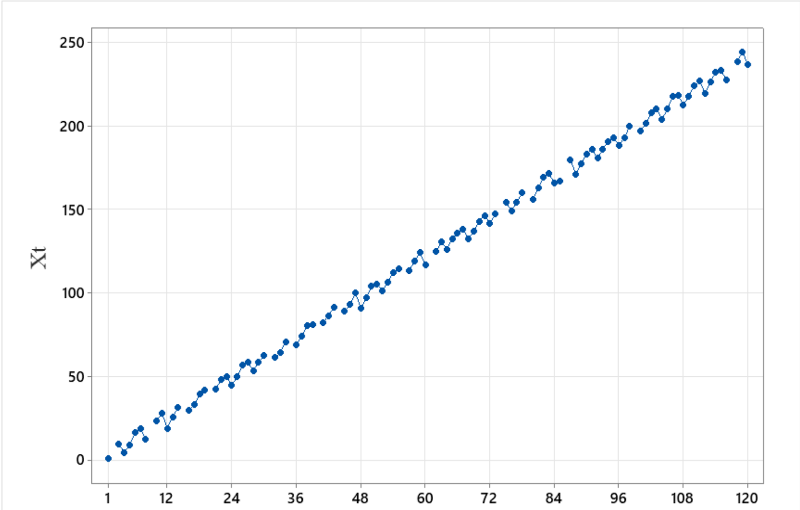

Fig. 6(b). Time plot of simulated series with fifteen separated missing values

Table 2. Estimates of trend parameter and seasonal indices for series with no/one missing value at different Position in the data

		Position/location of one missing value																		
		NONE		**1**		**2**		**5**		**10**		**20**		**50**		**100**		**120**		
Parameter	**Actual** (ϕ)	**Eat,** $(\hat{\phi})$	**Error** (e)	**Eat,** $(\hat{\phi})$	**Error** (e)	**Eat,** $(\hat{\phi})$	**Error** (e)	**Eat,** $(\hat{\phi})$	**Error** (e)	**Eat,** $(\hat{\phi})$	**Error** (e)	**Eat,** $(\hat{\phi})$	**Error** (e)	**Eat,** $(\hat{\phi})$	**Error** (e)	**Eat,** $(\hat{\phi})$	**Error** (e)	**Eat,** $(\hat{\phi})$	**Error** (e)	$\lvert error \rvert$ **Range**
a	1	1.09	-0.09	1.06	-0.06	1.10	-0.10	1.07	-0.07	1.10	-0.10	1.08	-0.08	1.11	-0.11	1.10	-0.10	1.08	-0.08	0.06 - 0.11
b	2	2.00	0.00	2.00	0.00	2.00	0.00	2.00	0.00	2.00	0.00	2.00	0.00	2.00	0.00	2.00	0.00	2.00	0.00	0
S_1	-1.5	-2.04	0.54	-2.04	0.54	-2.04	0.54	-2.03	0.53	-2.07	0.57	-2.05	0.55	-2.08	0.58	-1.96	0.46	-2.04	0.54	0.46 - 0.58
S_2	2.5	2.68	-0.18	2.68	-0.18	2.68	-0.18	2.69	-0.19	2.65	-0.15	2.70	-0.20	2.87	-0.37	2.70	-0.20	2.68	-0.18	0.15 - 0.37
S_3	3.5	4.16	-0.66	4.16	-0.66	4.16	-0.66	4.18	-0.68	4.13	-0.63	4.18	-0.68	4.12	-0.62	4.24	-0.74	4.16	-0.66	0.62 - 0.76
S_4	-4.5	-5.06	0.56	-5.06	0.56	-5.06	0.56	-5.05	0.55	-5.09	0.59	-5.04	0.54	-5.10	0.60	-5.11	0.61	-5.06	0.56	0.54 - 0.61
S_5	-1.5	-1.02	-0.48	-1.02	-0.48	-1.02	-0.48	-1.01	-0.49	-1.05	-0.45	-0.81	-0.69	-1.06	-0.44	-0.94	-0.56	-1.02	-0.48	0.44 - 0.69
S_6	2.5	2.55	-0.05	2.55	-0.05	2.55	-0.05	2.56	-0.06	2.52	-0.02	2.57	-0.06	2.51	-0.01	2.63	-0.13	2.55	-0.05	0.01 - 0.13
S_7	3.5	3.61	-0.11	3.61	-0.11	3.61	-0.11	3.62	-0.12	3.58	-0.08	3.63	-0.13	3.20	0.30	3.32	0.18	3.61	-0.11	0.08 - 0.3
S_8	-4.5	-4.18	-0.32	-4.18	-0.32	-4.18	-0.32	-4.16	-0.34	-4.21	-0.30	-4.52	0.02	-4.22	-0.28	-4.47	-0.03	-4.18	-0.32	0.02 - 0.34
S_9	-1.5	-1.58	0.08	-1.58	0.08	-1.58	0.08	-1.57	0.07	-1.39	-0.11	-1.56	0.06	-1.40	-0.10	-1.50	0.00	-1.58	0.08	0 - 0.11
S_{10}	2.5	2.59	-0.09	2.59	-0.09	2.59	-0.09	2.46	0.04	2.69	-0.19	2.61	-0.11	2.96	-0.46	2.65	-0.15	2.59	-0.09	0.04 - 0.46
S_{11}	3.5	3.38	0.13	3.38	0.13	3.38	0.13	3.39	0.11	3.35	0.15	3.39	0.11	3.33	0.17	3.46	0.04	3.38	0.13	0.04 - 0.17
S_{12}	-4.5	-5.09	0.59	-5.09	0.59	-5.09	0.59	-5.08	0.58	-5.12	0.62	-5.07	0.57	-5.13	0.63	-5.01	0.51	-5.09	0.59	0.51 - 0.63
μ_e	0	0.07	-0.07	0.07	-0.07	0.07	-0.07	0.07	-0.07	0.07	-0.07	0.07	-0.07	0.08	-0.08	0.07	-0.07	0.08	-0.08	0.07 - 0.08
σ_e^2	1	0.95	0.05	0.96	0.04	0.96	0.04	0.96	0.04	0.96	0.04	0.96	0.04	0.96	0.04	0.96	0.04	0.95	0.05	0.04 - 0.05
MSE			**0.11**		**0.11**		**0.11**		**0.11**		**0.11**		**0.12**		**0.13**		**0.11**		**0.11**	
RMSE			**0.33**		**0.33**		**0.33**		**0.33**		**0.33**		**0.34**		**0.36**		**0.33**		**0.33**	
MAE			**0.24**		**0.24**		**0.24**		**0.23**		**0.24**		**0.23**		**0.28**		**0.23**		**0.24**	

Table 3. Estimates of trend parameter and seasonal indices when the number of missing values is consecutive

		Number of consecutive missing values								
		Two		Five		Ten		Fifteen		
Parameter	**Actual** (ϕ)	**Eat,** $(\hat{\phi})$	**Error** (e)	**Eat,** $(\hat{\phi})$	**Error** (e)	**Eat,** $(\hat{\phi})$	**Error** (e)	**Eat,** $(\hat{\phi})$	**Error** (e)	$\lvert error \rvert$ **Range**
a	1	1.07	-0.07	1.04	-0.04	1.12	-0.12	1.11	-0.11	0.04 - 0.12
b	2	2.00	0.00	2.00	0.00	2.00	0.00	2.00	0.00	0
S_1	-1.5	-2.01	0.51	-2.01	0.51	-2.34	0.84	-2.07	0.57	0.51 - 0.84
S_2	2.5	2.71	-0.21	2.62	-0.12	2.69	-0.19	2.85	-0.35	0.12 - 0.35
S_3	3.5	4.19	-0.69	4.25	-0.75	4.22	-0.72	4.15	-0.65	0.65 - 0.75
S_4	-4.5	-5.03	0.53	-4.94	0.44	-5.13	0.63	-4.66	0.16	0.16 - 0.63
S_5	-1.5	-0.99	-0.51	-0.79	-0.71	-1.04	-0.46	-0.84	-0.67	0.46 - 0.71
S_6	2.5	2.58	-0.08	2.22	0.28	2.71	-0.21	3.17	-0.67	0.08 - 0.67
S_7	3.5	3.27	0.23	3.83	-0.33	3.87	-0.37	3.95	-0.45	0.23 - 0.45
S_8	-4.5	-4.15	-0.35	-4.63	0.13	-3.98	-0.52	-5.18	0.68	0.13 - 0.68
S_9	-1.5	-1.55	0.05	-1.43	-0.07	-1.62	0.12	-1.59	0.09	0.05 - 0.12
S_{10}	2.5	2.62	-0.12	2.56	-0.06	2.65	-0.15	2.51	-0.01	0.01 - 0.15
S_{11}	3.5	3.41	0.09	3.41	0.09	2.97	0.53	2.88	0.62	0.09 - 0.62
S_{12}	-4.5	-5.06	0.56	-5.09	0.59	-5.02	0.52	-5.20	0.70	0.52 - 0.7
μ_e	0	0.07	-0.07	0.05	-0.05	0.11	-0.11	0.05	-0.05	0.05 - 0.11
σ_e^2	1	0.97	0.03	0.91	0.09	0.96	0.04	0.95	0.05	0.03 - 0.09
MSE			**0.11**		**0.12**		**0.17**		**0.20**	
RMSE			**0.33**		**0.35**		**0.42**		**0.45**	
MAE			**0.24**		**0.25**		**0.33**		**0.35**	

Table 4. Estimates of trend parameter and seasonal indices when the number of missing values is separated

		Number of separated missing values								
		Two		Five		Ten		Fifteen		
Parameter	**Actual** (ϕ)	**Eat,** $(\hat{\phi})$	**Error** (e)	**Eat,** $(\hat{\phi})$	**Error** (e)	**Eat,** $(\hat{\phi})$	**Error** (e)	**Eat,** $(\hat{\phi})$	**Error** (e)	$\lvert error \rvert$ **Range**
a	1	1.05	-0.05	1.09	-0.09	1.06	-0.06	1.12	-0.12	0.05 - 0.12
b	2	2.00	0.00	2.00	0.00	2.00	0.00	2.00	0.00	0
S_1	-1.5	-2.04	0.54	-1.66	0.16	-1.72	0.22	-1.90	0.40	0.16 - 0.54
S_2	2.5	2.68	-0.18	2.40	0.10	2.80	-0.30	2.66	-0.16	0.1 - 0.3
S_3	3.5	4.16	-0.66	3.98	-0.48	4.03	-0.53	4.50	-1.00	0.48 - 1
S_4	-4.5	-5.06	0.56	-4.89	0.39	-4.76	0.26	-4.59	0.09	0.09 - 0.56
S_5	-1.5	-1.02	-0.48	-1.01	-0.49	-1.40	-0.10	-1.40	-0.10	0.1 - 0.46
S_6	2.5	2.55	-0.05	2.56	-0.06	2.82	-0.32	2.32	0.18	0.05 - 0.32
S_7	3.5	3.61	-0.11	3.78	-0.28	3.62	-0.12	2.79	0.72	0.11 - 0.72
S_8	-4.5	-4.18	-0.32	-4.47	-0.03	-4.97	0.47	-4.49	-0.01	0.01 - 0.47
S_9	-1.5	-1.58	0.08	-1.67	0.17	-1.58	0.08	-1.55	0.05	0.05 - 0.17
S_{10}	2.5	2.59	-0.09	2.75	-0.25	2.58	-0.08	3.21	-0.71	0.08 - 0.71
S_{11}	3.5	3.38	0.13	3.49	0.01	3.55	-0.05	3.39	0.11	0.01 - 0.13
S_{12}	-4.5	-5.09	0.59	-5.27	0.77	-4.99	0.49	-4.94	0.44	0.44 - 0.77
μ_e	0	0.08	-0.08	0.05	-0.05	0.05	-0.05	0.11	-0.11	0.05 - 0.11
σ_e^2	1	0.96	0.04	0.90	0.10	0.98	0.02	0.97	0.03	0.02 - 0.1
MSE			**0.11**		**0.09**		**0.07**		**0.15**	
RMSE			**0.33**		**0.29**		**0.26**		**0.38**	
MAE			**0.24**		**0.20**		**0.19**		**0.25**	

The results in Table 2 show that the value of the slope (b) was recovered almost perfectly well, while the value of the intercept (a) was recovered, but not as well as the slope (b). The value of a was recovered with absolute errors ranging from 0.06 to 0.11. On the other hand, the seasonal indices (S_j) were recovered with greater absolute error. The poor recovery of the seasonal indices is the same for series with and without missing values. This is true when two or more missing values are located at consecutive or different positions in the series. The estimate of trend parameters and seasonal indices for series with two or more missing values, the errors, and summary statistics are presented in Tables 3 and 4. Table 3 is for missing values that are located at consecutive positions while Table 4 is for missing values that are located at separated positions. The results show that the value of the slope (b) was recovered almost perfectly well for series in Table 3 and Table 4, while the value of the intercept (a) was recovered but not as well as the slope (b). The intercept was recovered with absolute error ranging from 0.04 to 0.12 for Table 3 (when the missing values are located at consecutive position) and 0.05 to 0.12 for Table 4 (when the missing are separated). The seasonal indices were recovered with greater absolute error. Similarly, the poor recovery of the seasonal indices is the same for all number of missing values in both consecutive and separated position.

7. SUMMARY AND CONCLUSION

This study has investigated the effect of missing values on the estimates of trend parameter and seasonal indices on descriptive time series analysis. The aim is to see if it is necessary to first obtain estimate of the missing values before further analysis. The methods of descriptive time series analysis were used to obtain estimates of trend parameters and seasonal indices when trending curve is linear and decomposition model is additive. Summary statistics were used to compare estimates of the parameters in absence and presence of one or more missing values.

The results of the analysis show that the trend parameters, error means and standard deviation are recovered almost perfectly while the seasonal indices are recovered less precisely. These implies that obtaining the estimates of the missing values before continuing with the analysis may not be necessary when dealing with time series data with linear trend and additive model. It has therefore been recommended that when analyzing a time series data with missing value using descriptive method it may not be necessary to obtain the estimates of the missing values before continuing with the analysis.

DISCLAIMER (ARTIFICIAL INTELLIGENCE)

Author(s) hereby declare that NO generative AI technologies such as Large Language Models (ChatGPT, COPILOT, etc.) and text-to-image generators have been used during the writing or editing of this manuscript.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

Adejumo, O. A., Onifade, O. C., & Albert, S. (2021). Kalman filter algorithm versus other methods of estimating missing values: Time series evidence. *African Journal of Mathematics and Statistics Studies, 4*(2), 1-9. https://doi.org/10.52589/AJMSS-VFVNMQLX

Afkanpour, M., Hosseinzadeh, E., & Tabesh, H. (2024). Identify the most appropriate imputation method for handling missing values in clinical structured datasets: A systematic review. *BMC Medical Research Methodology, 24*, 188. https://doi.org/10.1186/s12874-024-02310-6

Agung, P., Agus, S., Agustini, T., Mustafa, M., Sukono, & Ruly, B. (2020). A new method to estimate parameters in the simple regression linear equation. *Mathematics and Statistics, 8*(2), 75-81. https://doi.org/10.13189/ms.2020.080201

Arslan, A., & Aydilek, I. B. (2013). A hybrid method for imputation of missing values using optimized fuzzy c-means with support vector regression and a genetic algorithm. *Information Sciences, 233*, 25-35. https://doi.org/10.1016/j.ins.2013.01.021

Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (2016). *Time series analysis: Forecasting and control*. John Wiley & Sons.

Chatfield, C. (2004). *The analysis of time series: An introduction*. Chapman and Hall/CRC Press.

David, S. C. F. (2006). *Methods for the estimation of missing values in time series*. Cowan University Press, Western Australia. https://ro.ecu.edu.au/theses/63

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B (Methodological), 39*(1), 1-38. http://links.jstor.org/sici?sici=0035-9246%281977%2939%3A1%3C1%3AMLFIDV%3E2.0.CO%3B2-Z

Howell, D. C. (2007). The analysis of missing data. In W. Outhwaite & S. Turner (Eds.), *Handbook of social science methodology* (pp. 231-249). Sage. https://doi.org/10.4135/9781848607958.n11

Iwueze, I. S., Nwogu, E. C., Nlebedim, V. U., Nwosu, U. I., & Chinyem, U. E. (2018). Comparison of methods of estimating missing values in time series. *Open Journal of Statistics, 8*, 390-399. https://doi.org/10.4236/ojs.2018.82025

Kerkri, A., Zarrouk, Z., & Allal, J. (2015). A comparison of NIPALS algorithm with two other missing data treatment methods in a principal component analysis. http://paperssondages14.sfds.asso.fr/submission_41.pdf

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. John Wiley & Sons Inc. http://dx.doi.org/10.1002/9780470316696

Schmitt, P., Mandel, J., & Guedj, M. (2015). A comparison of six methods for missing data imputation. *Journal of Biometrics & Biostatistics, 6*(224), 1-6. https://doi.org/10.472/2155-6180.1000224

Sohae, O. (2015). Multiple imputation in missing values in time series data (Master's thesis). Duke University. https://dukespace.lib.duke.edu/server/api/core/bitstreams/027b3bc4-6c6b-40e4-a8ff-bef39fc59785/content

Tusell, F. P. (2005). Multiple imputation of time series with an application to the construction of historical price indices. *University of the Basque Country Go*. https://addi.ehu.es/handle/10810/5663

Vacek, P., & Ashikaga, T. (1980). An examination of the nearest neighbor rule for imputing missing values. *ASA Proceedings of the Statistical Computing Section*, 326-331. https://hellanicus.lib.aegean.gr/bitstream/handle/11610/17202/Missing_Data_in_Time_Series_and_Imputation_methods.pdf?sequence=1

Yaffee, R., & McGee, M. (1999). *Introduction to time series analysis and forecasting: With applications of SAS and SPSS*. Academic Press. https://www.researchgate.net/publication/234789625

Peer-review history:
The peer review history for this paper can be accessed here:
https://pr.sdiarticle5.com/review-history/126836