Scholars Journal of Physics, Mathematics and Statistics

OPEN ACCESS

Abbreviated Key Title: Sch J Phys Math Stat
ISSN 2393-8056 (Print) | ISSN 2393-8064 (Online)
Journal homepage: https://saspublishers.com

The Effect of Over-Differencing on Model Validity

Nasir Elmesmari[1*], Rajab Suliman[2], Mohammed Elnazzal[3]

[1]Department of Statistics, Faculty of arts and Science, Al Maraj, University of Benghazi, Benghazi, Libya
[2]Information System and Supply Chain Management Department, John B. and Lillian E. Neff College of Business and Innovation, University of Toledo, USA
[3]Department of Statistics, Faculty of Science, University of Benghazi, Benghazi, Libya

DOI: 10.36347/sjpms.2022.v09i08.001 | **Received:** 11.10.2022 | **Accepted:** 19.11.2022 | **Published:** 24.11.2022

***Corresponding author:** Nasir Elmesmari
Department of Statistics, Faculty of arts and Science, Al Maraj, University of Benghazi, Benghazi, Libya

Abstract | **Original Research Article**

We examine the effect of unnecessary differencing (over-differencing) on the appropriateness of the proposed models (Autoregressive of order one AR(1), Autoregressive of order two AR(2), and Moving Average of order two MA(2)). Our interest arises from the fact that in practical applications the fitted model due to inappropriately differenced data can still suitably describe the data sample based on the goodness of fit test using residual analysis. Given that we use simulation study to detect the consequences of unnecessary differencing on the fitted model. While the simulation study can be controlled using different scenarios, it becomes more challenging when dealing with real data. The validity and performance of the fitted models was checked by observing the changes in the estimated coefficients, the associated standard errors (SE), the residual variance, and Akaike information (AIC) by comparing them with the true parameters of the system (true model). The uniqueness of this paper is to examine how the fitted model is sensitive (valid) to the over-differencing.

Keywords: Time series, over-differencing, AR(1), AR(2), MA(2), simulation, sensitivity.

1. INTRODUCTION

Data mining has recently found several research and development initiatives driven by time series. Time series is an important class of behavioral data that can be measured from scientific and financial applications (Montgomery and Johnson, 1976). It may be an hourly record of temperature at a given place or a quarterly record of gross national product, monthly sales totals, prices of stocks and mutual funds. A time series is number of observations are collected over excessive period. The time series data by its nature is huge in the data sample, highly dimensional, and collected continually. Additionally, time series data is distinguished by its numerical nature and continuously measured. There are various of time series data studies such as finding similar time series (Agrawal *et al.*, 1993; Berndt and Clifford, 1996; and Fu, 1999), time series sequence searching (Faloutsos *et al.*, 1994), dimensionality reduction (Keogh, 1997; Keogh *et al.*, 2000) and segmentation (Abonyi *et al.*, 2005). Based on those variety of time series domains, studies have received extensive attention from the database and pattern recognition communities (Keogh and Kasetty, 2002). The literature has a variety of mining tasks based on time series, these tasks can be generally classified into four categories, rule discovery and summarization, pattern discovery, clustering, and classification. While some studies focus on one of these categories, the others may focus on more than one of the processes described above (Shukla, Amit K., Manvendra Janmaijaya, Ajith Abraham, and Pranab K. Muhuri, 2019) [1]. Over-differencing a series will produce a loss of the performance of forecasting one-step-ahead [2]. Also, John Cochrane, (2012) illustrates the dangers of over-differencing. The paper's contribution is an assessment of the sensitivity (validity) of the fitted model to over-differencing.

2. THE DIFFERENCING METHOD

Box and Jenkins (1976) argue that homogeneous nonstationary sequences X_t can be transformed into stationary sequences by taking successive differences of the series; that is, by considering the series $\nabla X_t, \nabla^2 X_t, \ldots$ where ∇ is the

[1]For details on the four categories, see Shukla *et al.*, (2019).
[2]For details on Over-differencing, see De Gooijer *et al.*, (2006).

backward difference operator $\nabla = 1 - B$, and B is the backward shift operator which is defined by $Bx_t = x_t - 1$. Deterministic polynomial trends are ones where the polynomial's coefficients remain constant over time. The coefficient in stochastic trends is subject to random variation; thus, the trend fluctuates stochastically corresponding to random shocks that enter the system. By considering the following three linear trend models, the distinction may be explained in more detail. The first model (M1) is the deterministic linear trend model $X_t = \beta_0 + \beta_1 t + \varepsilon_t$ where ε_t is a white-noise sequence or more generally, a zero-mean stationary stochastic process. In M1 the level $\mu_t = E(X_t) = \beta_0 + \beta_1 t$ grows in a deterministic linear fashion with respect to time. The second model (M2) entails the first differences of the series and a constant trend parameter β_1. It is given by $(1-B)X_t = \beta_1 + \varepsilon_t$ or $X_t = X_{t-1} + \beta_1 + \varepsilon_t$. This model also indicates to a linear trend. The level of the series at t (or, more properly, the conditional expectation of X_t given $X_{t-1}, X_{t-2}, ...$) is $\mu_t = X_{t-1} + \beta_1$. The level is determined based on the previous observation. Since X_{t-1} is subject to random shocks, the trend changes stochastically. The third model (M3) requires the second differences of the data $(1-B)^2 X_t = \varepsilon_t$ or $X_2 = 2X_{t-1} - X_{t-2} + \varepsilon_t$. Since X_{t-1} and X_{t-2} in the level $\mu_t = 2X_{t-1} + X_{t-2}$ are influenced by random shocks, both the intercept and the slope of the linear trend, which passes through X_{t-1} and X_{t-2} alter stochastically (Abraham and Ledolter, 1983).

3. Identifying The Order of Differencing:

We begin with the first step, which is a crucial step in fitting a time series model. This step is the determination of the order of differencing needed in the series. The purpose of this step is to stationaries the series. Generally, the lowest order of differencing that produces a time series fluctuates around a well-defined mean value and whose autocorrelation function (ACF) plot decline rapidly to zero, either from above or below. If the series still reveals a long-term trend, or otherwise lacks a tendency to return to its mean value, or if its autocorrelations are positive out to a high number of lags (i.e., 10 or more), then it requires a higher order of differencing. Next, we discuss the rule of differencing.

Rule 1: A higher order of differencing is likely necessary if the series exhibits positive autocorrelations out to a high number of lags. Differencing tends to produce negative correlation: if the series primarily shows strong positive autocorrelation, then a nonseasonal difference will reduce the autocorrelation and the lag-1 autocorrelation to a negative value. In case of applying a second nonseasonal difference (which is occasionally necessary), the lag-1 autocorrelation will be led even further negative direction. If the lag-1 autocorrelation is zero or even negative, then the series does not need additional differencing. The argue of differencing is due to a random pattern in the autocorrelations. One of the most common errors in time series modeling is to "over-difference" the series that result in adding extra Autoregressive (AR) or Moving Average (MA) terms to solve the damage. If the magnitude of lag-1 autocorrelation is more negative than -0.5 (and theoretically a negative lag-1 autocorrelation should never be greater than 0.5), which means the series has been over-differenced.

Rule 2: If the lag-1 autocorrelation is zero or negative, or are all small and has random pattern, then the series does not need a higher order of differencing. If the lag-1 autocorrelation is -0.5 or smaller, the series may be over-differenced. Also, a higher standard deviation is an indication of possible over-differencing rather than a reduction when the order of differencing is increased.

Rule 3: Often, the optimal order of differencing is the order of differencing at which the standard deviation is lowest.

Rule 4: A model with no orders of differencing assumes that the original series is stationary (mean-reverting). A model with one order of differencing assumes a constant average trend for the original series. A model with two orders of total differencing assumes that the original series has a time-varying trend.

Rule 5: A model with no orders of differencing normally contains a constant term (which represents the mean of the series). A model with two orders of total differencing typically does not include a constant term. In a model with one order of total differencing, a constant term should be included if the series has a non-zero average trend (Beusekkom, 2003).

4. Risk of Over-Differencing:

Although further differences of stationary series will again be stationary, over-differencing can lead to serious difficulties. It can unnecessarily confound the autocorrelation structure and produce higher variance for the over-differencing series (Fuller, W.A., 1976). These difficulties are illustrated in the following subsections.

4.1. Effect of over-differencing on the autocorrelation structure

Consider the stationary MA(1) process $X_t = (1 - b_1 B)\varepsilon_t$

And autocorrelation function of this process is given by indictor function as follows:

$$\rho_k = \begin{cases} \frac{-b_1}{1+b_1^2} & if \quad k = 1 \\ 0 & if \quad k > 1 \end{cases} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \ (1)$$

The first difference of the process is

$$(1-B)X_t = [1 - (1+b_1)B + b_1 B^2]\varepsilon_t \ \ldots\ldots\ldots\ (2)$$

And the autocorrelations of this difference are given by

$$\rho_k = \begin{cases} \frac{-(1+b_1)^2}{2(1+b_1+b_1^2)} & if \quad k = 1 \\ \frac{b_1}{2(1+b_1+b_1^2)} & if \quad k = 2 \\ 0 & if \quad k > 2 \end{cases} \ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \ (3)$$

Eq.1 through Eq. 3 shows that structure of the ACF of the over-differenced series is more complicated than that of the original process. This leads to a non-parsimonious representation, since it requires the estimation of two parameters as compared to one in the original MA(1) model. Furthermore, over-differencing will force the moving average operator to be of noninvertible form (Anderson, T.W., 1971).

4.2. Effect of Over-Differencing on the variance

The variance of the MA(1) process is given by $\gamma_0(X) = (1 + b_1^2)\sigma^2$. The variance of the over-differenced series $\gamma_0(Z) = 2(1 + b_1 + b_1^2)\sigma^2$. Hence, which shows that the variance of the over-differenced process will always be larger than that of the original process.

Next, consider the stationary AR(1) process $(1 - aB)X_t = \varepsilon_t$ with variance $\gamma_0(Z) = \sigma^2/(1 - a^2)$, so the first difference Z_t follows the Autoregressive Moving Average process (ARMA(1,1)): $(1 - aB)Z_t = (1 - B)\varepsilon_t$.

The variance of this process is given by $\gamma_0(Z) = \frac{2(1-a)\sigma^2}{1-a^2}$ From $\gamma_0(Z) - \gamma_0(X) = \frac{(1-a)\sigma^2}{1-a^2}$; we find that for $a < 0.5$ over-differencing will increase the variance. Indeed, changing in the sample variances of successive differences help in determining the appropriate degree of differencing. For nonstationary sequences the sample variances will be higher, since the squared deviations are taken from its mean. The bias will be produced associated with nonstationary sequences that do not have a fixed level, therefore, Over-differencing will produce more variation and the sample variances will become higher (Abraham and Ledolter, 1983).

5. DATA AND SAMPLE SIZE

In this study we use data given by Priestley, M.B., (1981) [3] which generated from the following stationary models:

- **AR(1) model:** $X_t - 0.6X_{t-1} = \varepsilon_t$
- **AR(2) model:** $X_t - 0.4X_{t-1} + 0.7X_{t-2} = \varepsilon_t$
- **MA(2) model:** $X_t = \varepsilon_t + 1.1\varepsilon_{t-1} + 0.2\varepsilon_{t-2}$

Where ε_t in each model is a normal white noise (i.e. $\varepsilon_t \sim N(0,1)$).

The main analysis is based on 500 observations simulated from models AR (1), AR (2) and MA (2), Also the procedure of each model is based on three steps.

i. A model with no orders of a nonseasonal difference (d = 0).
ii. A model with one order of a nonseasonal difference (d = 1).
iii. A model with two orders of a nonseasonal difference (d = 2).

6. ANALYSIS AND RESULTS

We present our analysis in next section, and the result of the effect of Over-differencing on AR(1) model are reported in Table 1.

Table 1: The result of the effect of Over-differencing on AR(1) model in which the model with no order, one order, and two orders of a nonseasonal difference

Estimates	Order of Difference		
	D = 0	**D = 1**	**D = 2**
$\hat{a}$	0.55	-0.21	-0.54
SE $\hat{a}$	0.037	0.043	0.037
AIC	1410.01	1509.53	1800.13
Residual Variance	0.97	1.20	2.16

Table 1 shows that the Over-differencing has significant impact on the estimated coefficients and the associated standard errors. Table 1 also shows that there is a significant change in the sign of the estimated coefficients as well as the increase in the residual variance and Akaike information (AIC). Therefore, the result show that over-differencing is completely produce a different model. The following Figures show the time series plots, ACF, PACF, residuals, box plots, and histogram for the fitted models of AR(1) calculated respect to the differenced series.

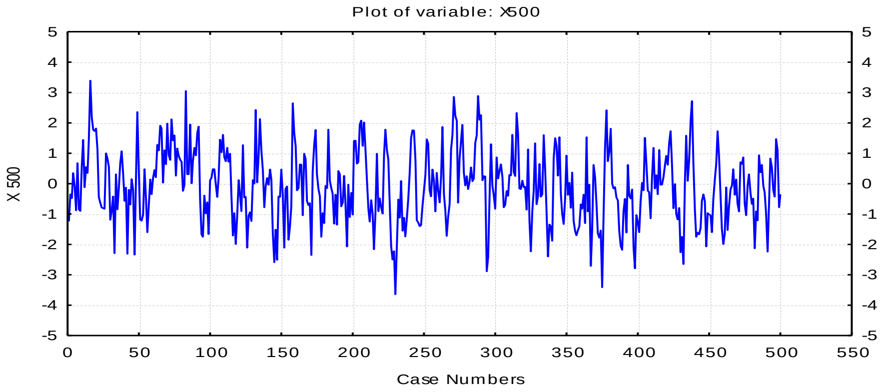

Figure 1.1: Plot of AR (1) model with no orders of difference

[3]For details on the data sample, see Priestly, M.B., (1981).

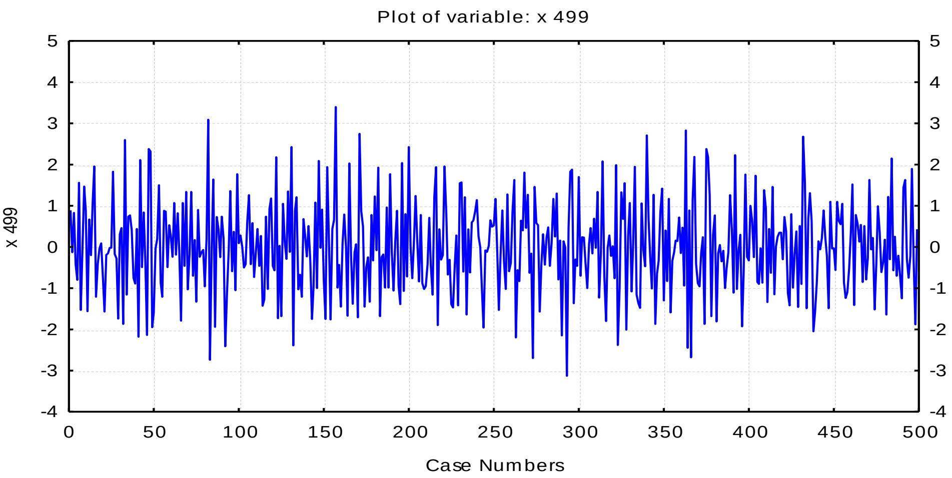

Figure 1.2: Plot of AR (1) model with one order of difference

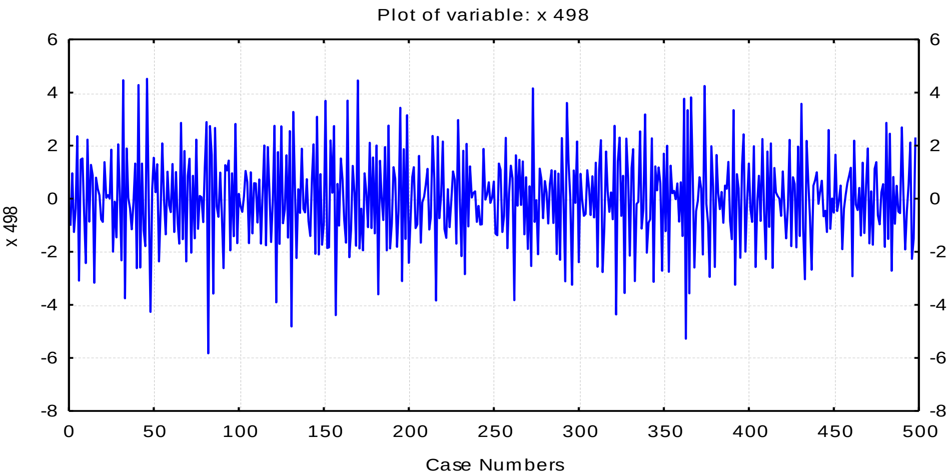

Figure 1.3: Plot of AR (1) model with two orders of difference

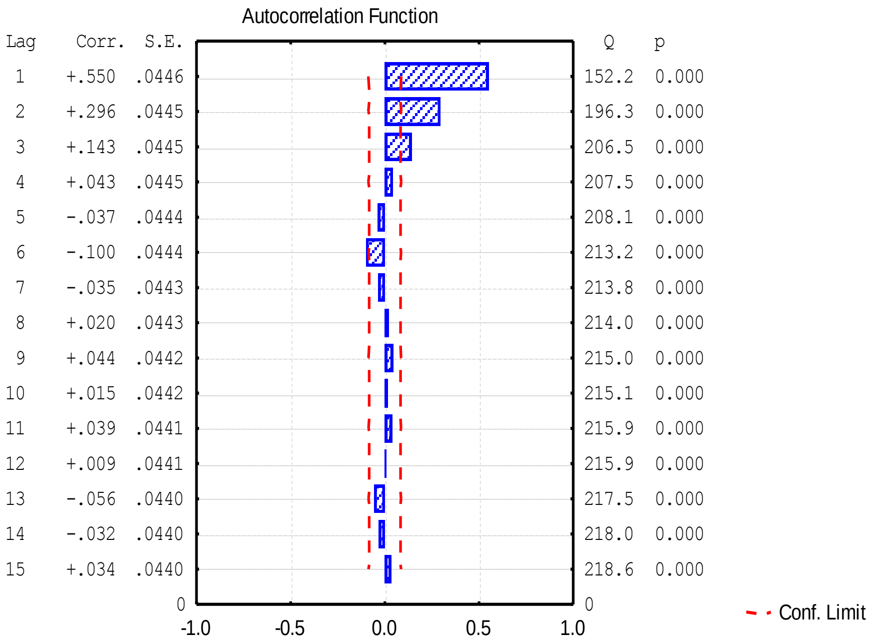

Figure 1.4: The ACF for AR (1) model with no order of difference

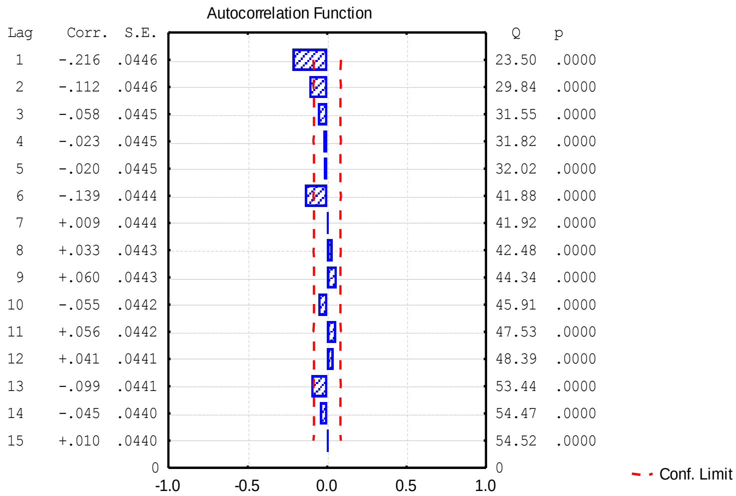

Figure 1.5: The ACF for AR (1) model with one order of difference

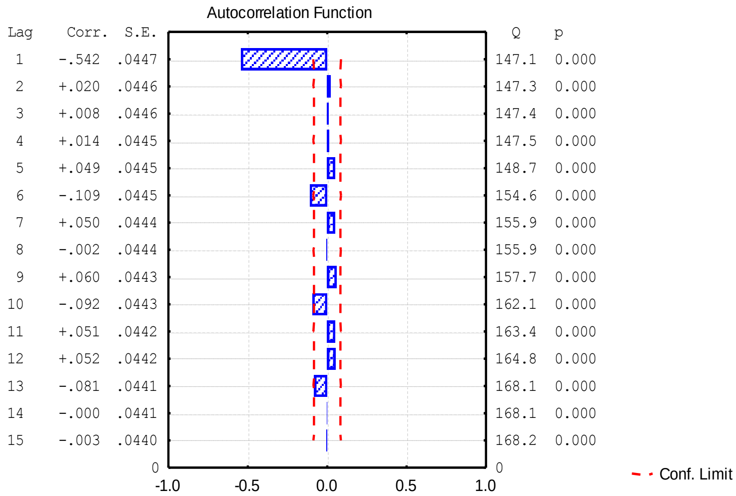

Figure 1.6: The ACF for AR (1) model with two orders of difference

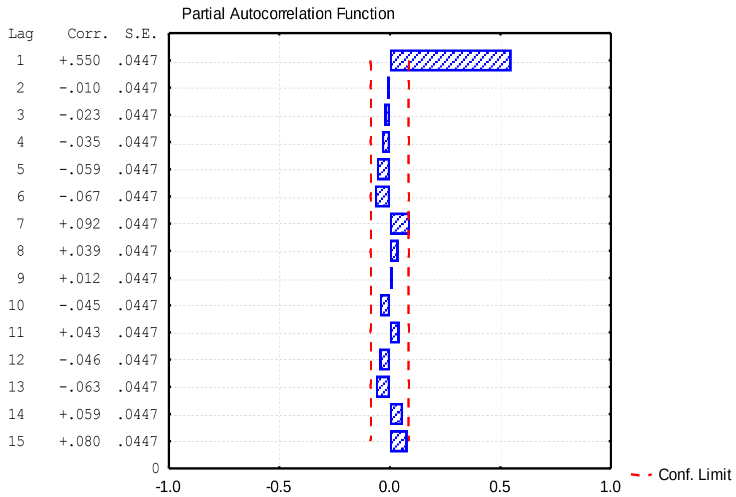

Figure 1.7: The PACF for AR (1) model with no order of difference

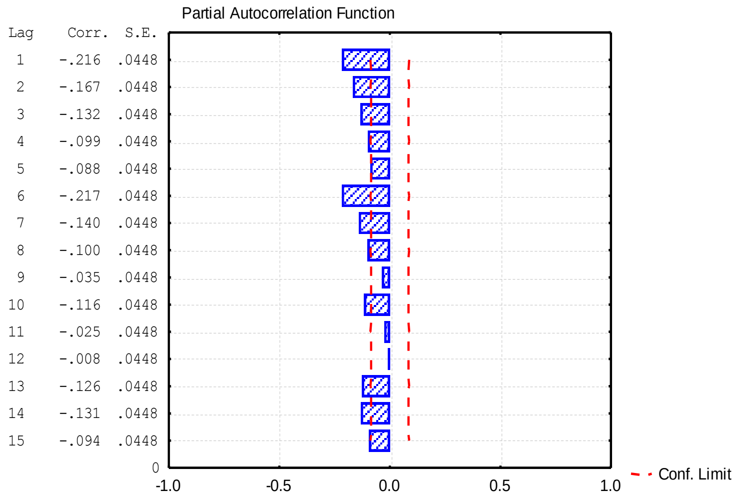

Figure 1.8: The PACF for AR (1) model with one order of difference

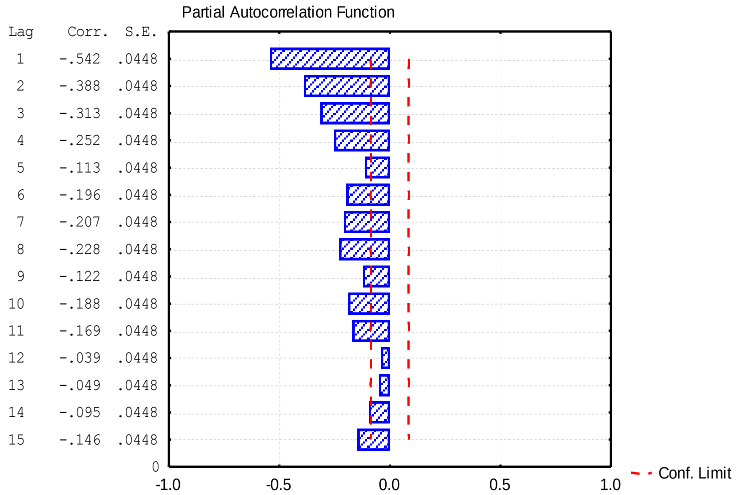

Figure 1.9: The PACF for AR (1) model with two orders of difference

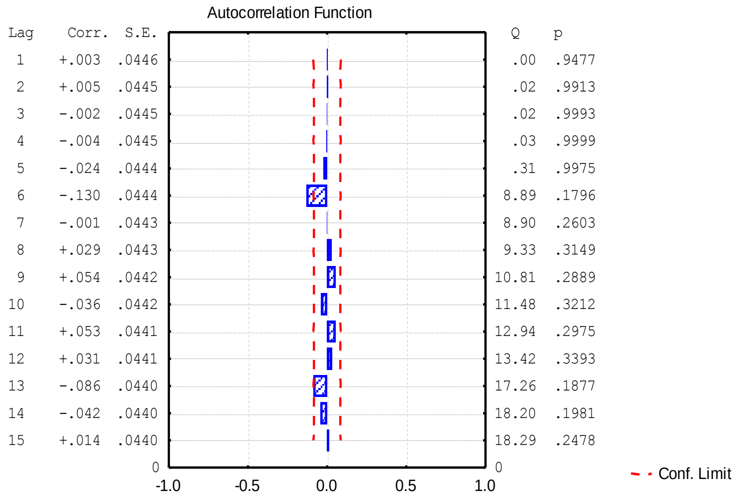

Figure 1.10: The ACF for AR (1) model residuals with no order of difference

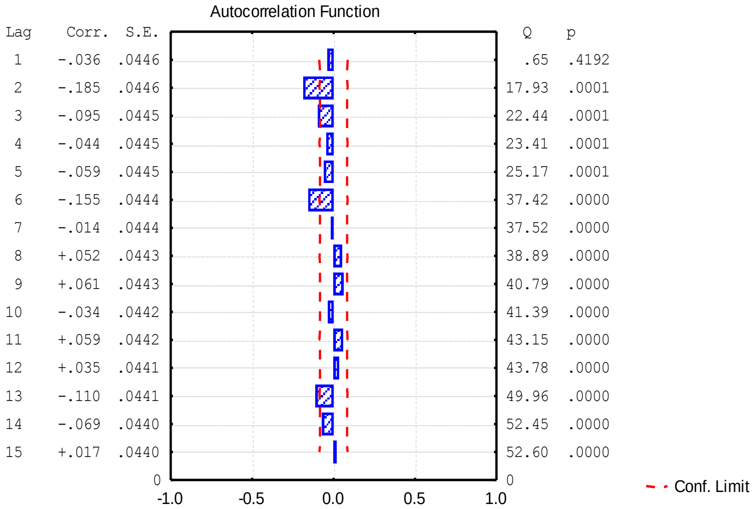

Figure 1.11: The ACF for AR (1) model residuals with one order of difference

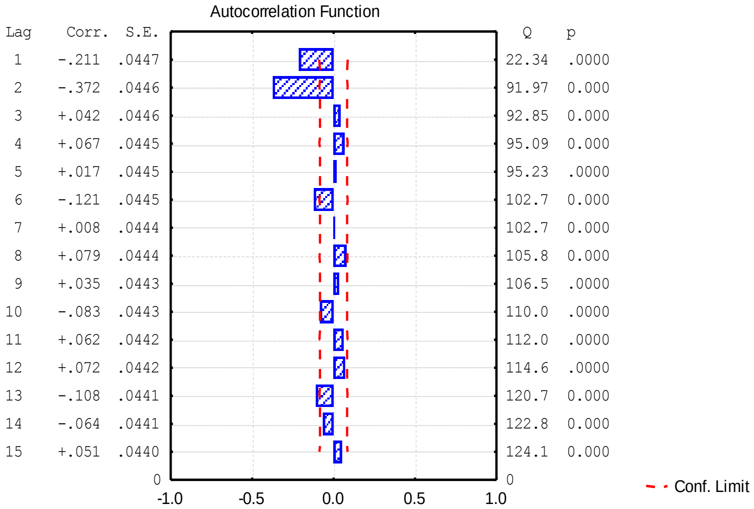

Figure 1.12: The ACF for AR (1) model residuals with two orders of difference

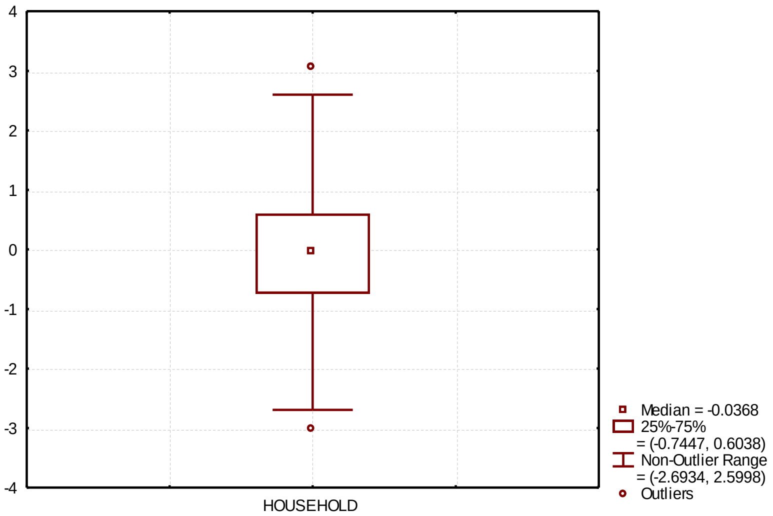

Figure 1.13: The box plot for AR (1) model residuals with no order of difference

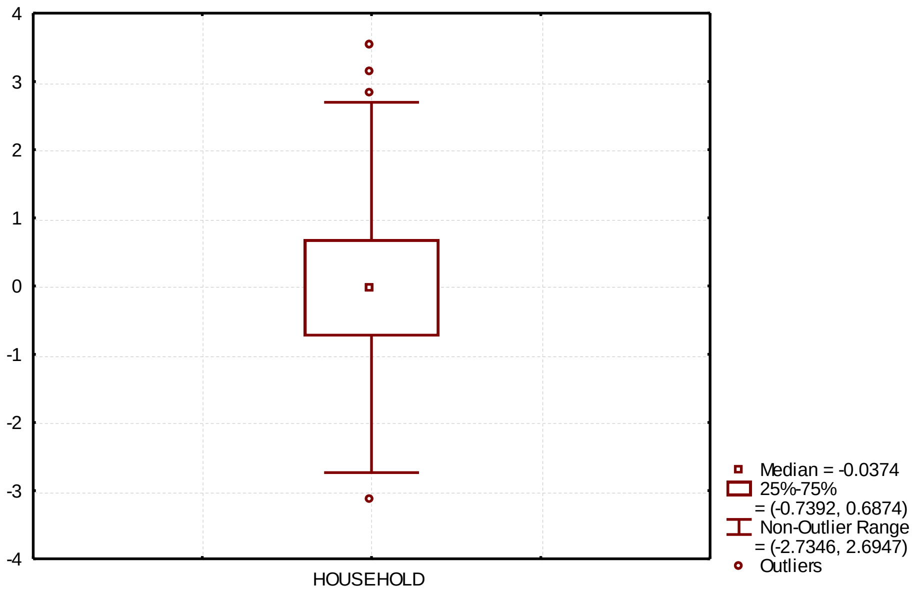

Figure 1.14: The box plot for AR (1) model residuals with one order of difference

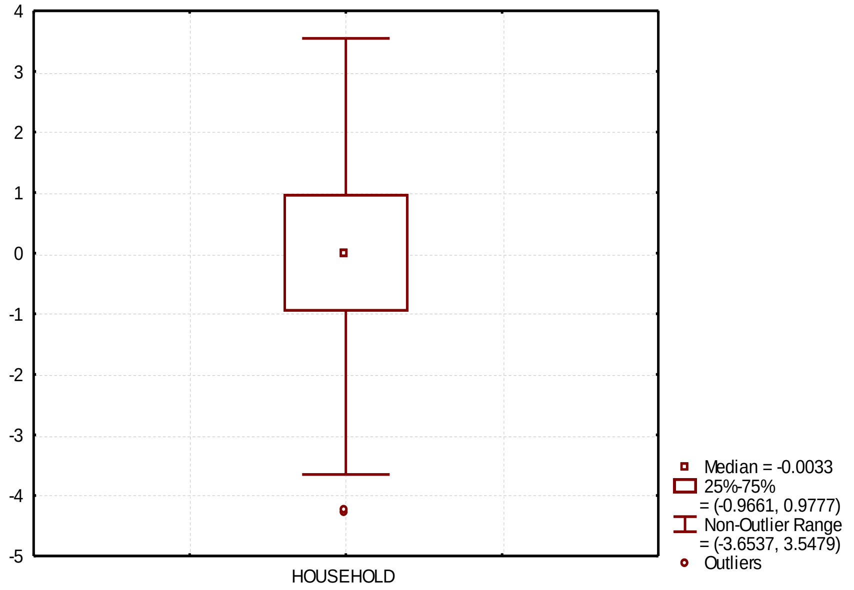

Figure 1.16: The box plot for AR (1) model residuals with two orders of difference

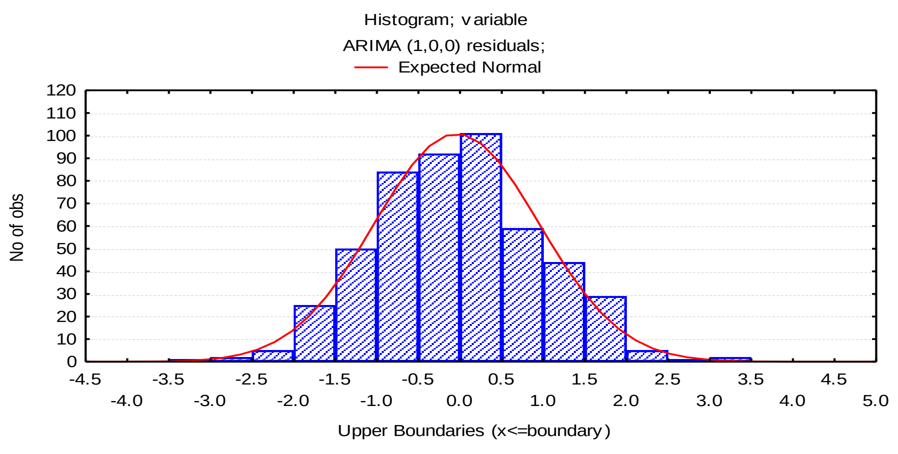

Figure 1.16: The histogram of AR (1) model residuals with no order of difference

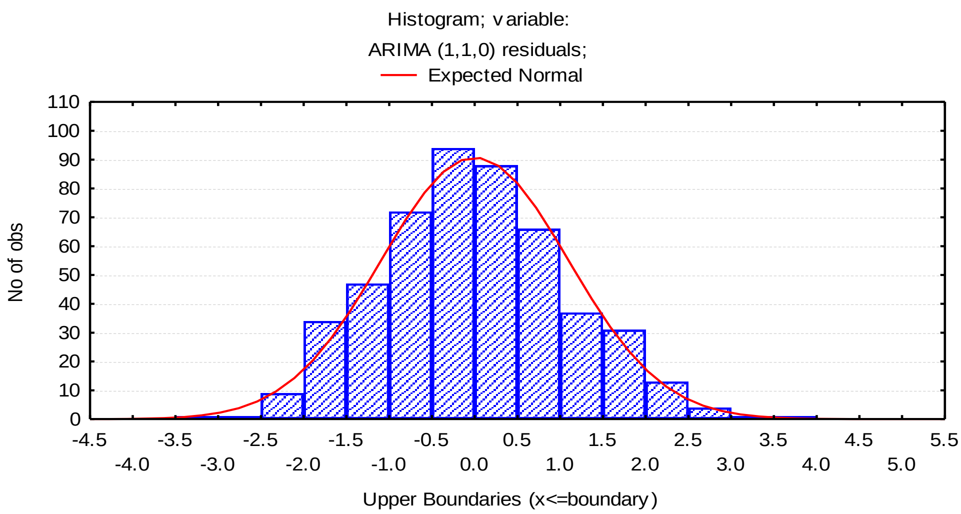

Figure 1.17: The histogram of AR (1) model residuals with one order of difference

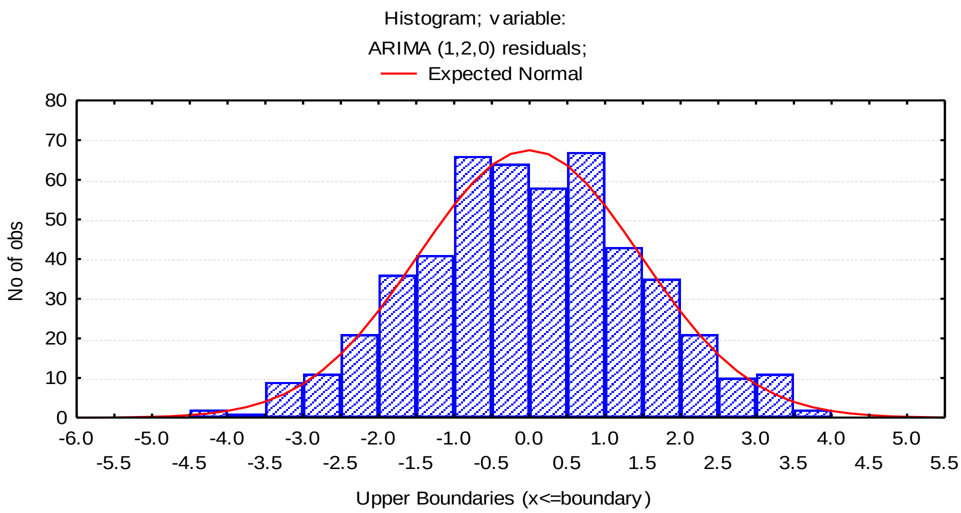

Figure 1.18: The histogram of AR (1) model residuals with two orders of difference

Figure 1.1 through Figure 1.18 show the time series plots of the differenced series, ACF, PACF as well as the ACF, the box plots, and the histograms for the residuals calculated from the fitted models AR(1) respect to over-differencing, these results reveal the following two important points 1) the plots of the differenced series still stationary, and 2) the histograms of the residuals look nearly symmetrical and the corresponding box plots have extreme symmetry of the central portion where the median is equidistant from the lower and upper limits, suggesting that the residuals computed from the fitted models to improperly differenced data are probably an approximate white noise which means that over-differencing can not be completely detected by just only examining the residuals which is a very important conclusion to draw from this study.

Next, we report the result of the effect of Over-differencing on AR(2) that are reported in Table 2.

Table 2: The result of the effect of Over-differencing on AR(2) model in which the model with no order, one order, and two orders of a nonseasonal difference.

	Order of Difference		
Estimates	**D = 0**	**D = 1**	**D = 2**
$\hat{a}_1$	0.35	0.41	- 0.21
$\hat{a}_2$	- 0.69	- 0.69	- 0.61
$SE\ \hat{a}_1$	0.066	0.032	0.035
$SE\ \hat{a}_2$	0.067	0.032	0.035
AIC	1411.56	1667.25	2072.76
Residual variance	0.97	1.64	3.73

Table 2 shows that the estimated coefficient $\hat{a}_1$ is decreasing as the order differencing increases, more specifically, when D = 2, and the magnitude of $\hat{a}_1$ is largely affected when D = 0 and D =1 relative to D = 2 (Positive effect 0.35, 0.41 to negative effect - 0.21). Conversely, Table 2 shows that the estimated coefficient $\hat{a}_2$ is increasing as the order differencing increases, especially when D = 2. Similarly, the ***AIC*** and residual variance measures are higher for higher order differencing. Taken all together, the results of over differencing for AR(2) indicate that over differencing lead to entirely different models. Figures 2.1 through 2.18 show the time series plots, ACF, PACF, residuals, box plots, and histogram for the fitted models of AR(2) calculated respect to the differenced series.

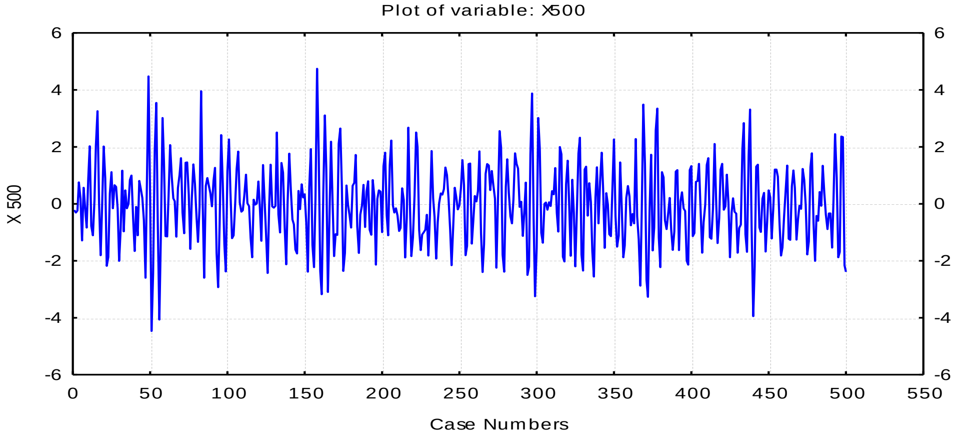

Figure 2.1: Plot of AR (2) model with no order of difference

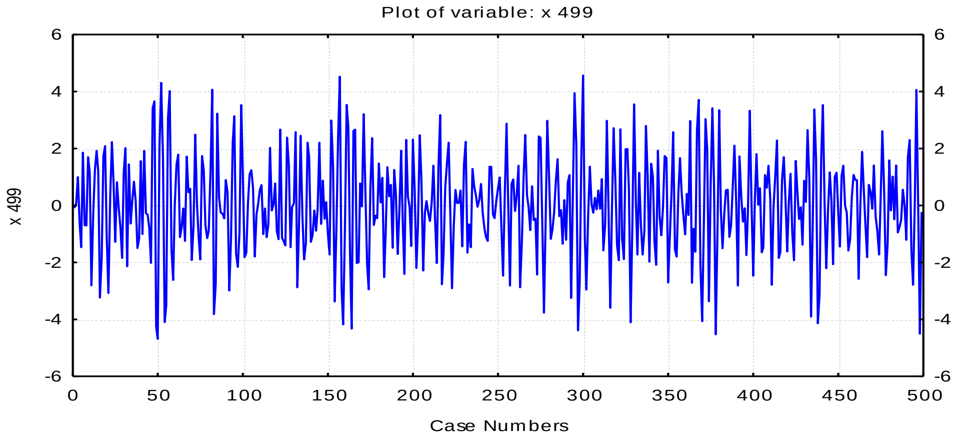

Figure 2.2: Plot of AR (2) model with one order of difference

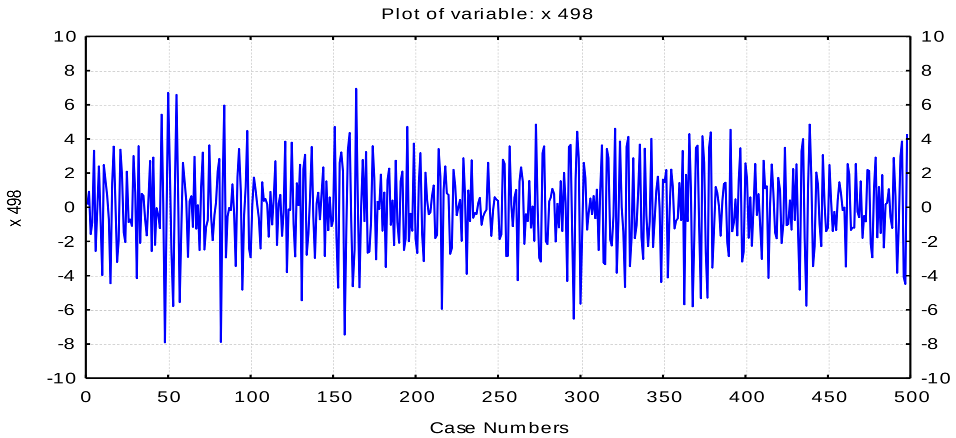

Figure 2.3: Plot of AR (2) model with two orders of difference

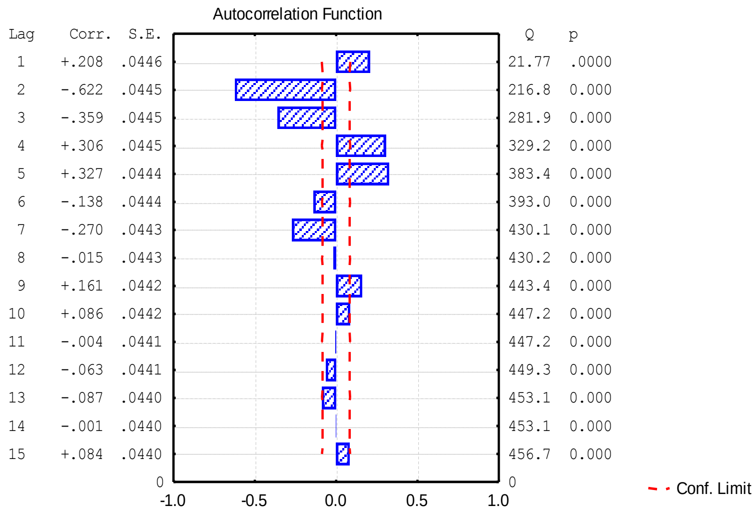

Figure 2.4: The ACF for AR (2) model with no order of difference

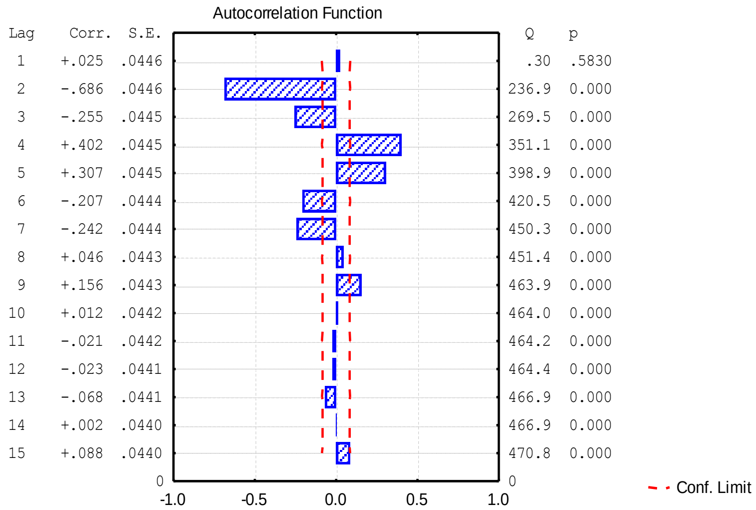

Figure 2.5: The ACF for AR (2) model with one order of difference

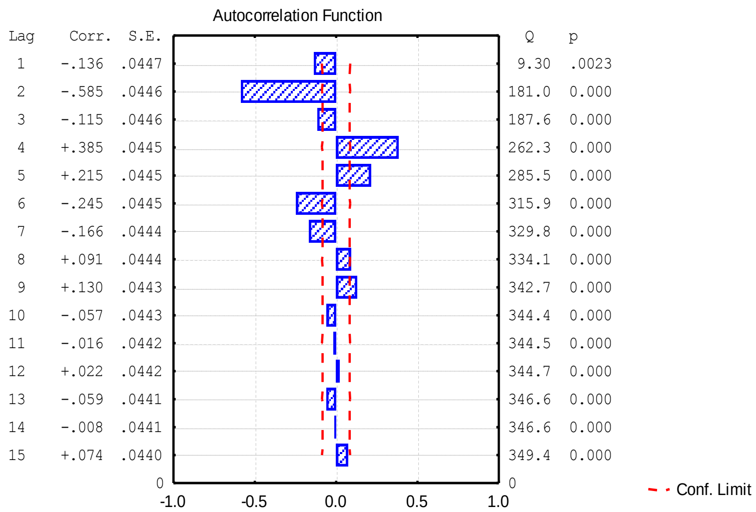

Figure 2.6: The ACF for AR (2) model with two orders of difference

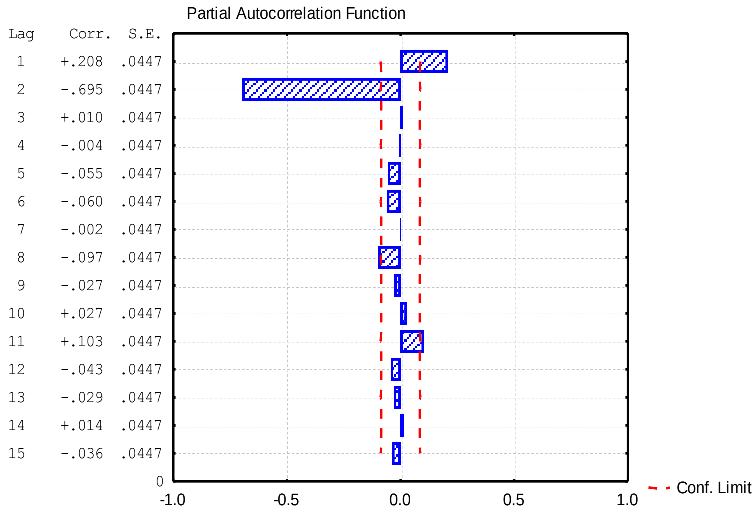

Figure 2.7: The PACF for AR (2) model with no order of difference

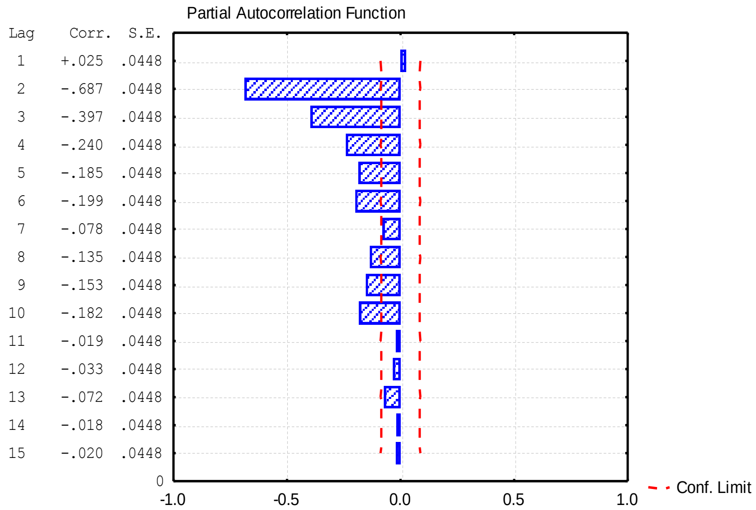

Figure 2.8: The PACF for AR (2) model with one order of difference

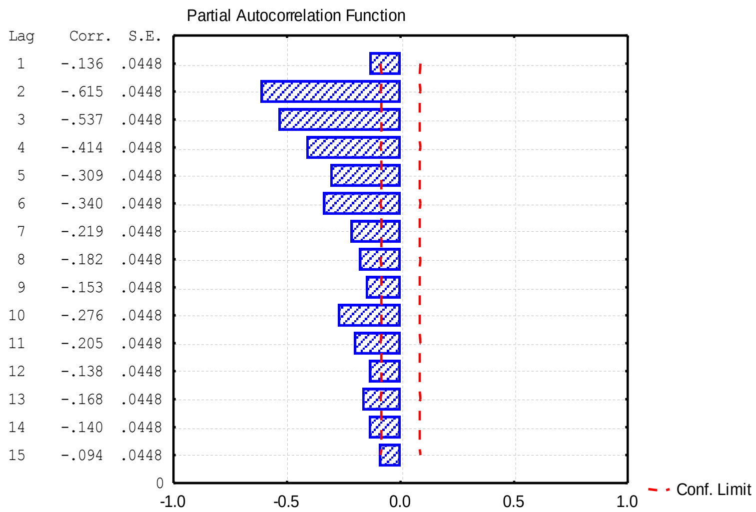

Figure 2.9: The PACF for AR (2) model with two orders of difference

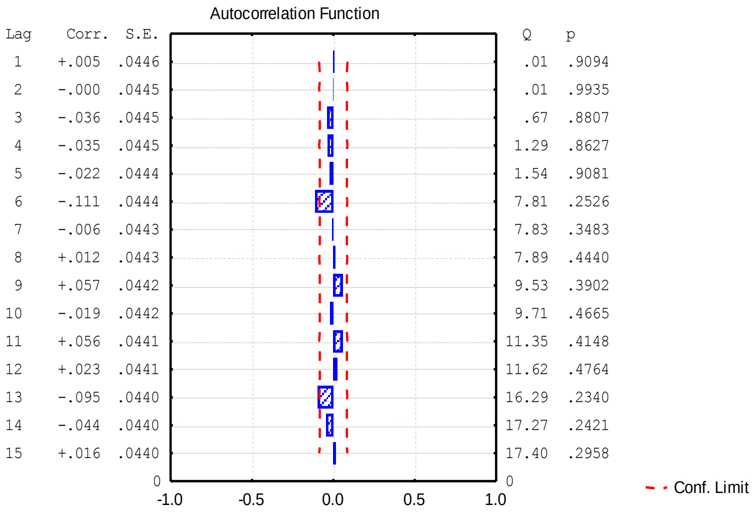

Figure 2.10: The ACF for AR (2) model residuals with no order of difference

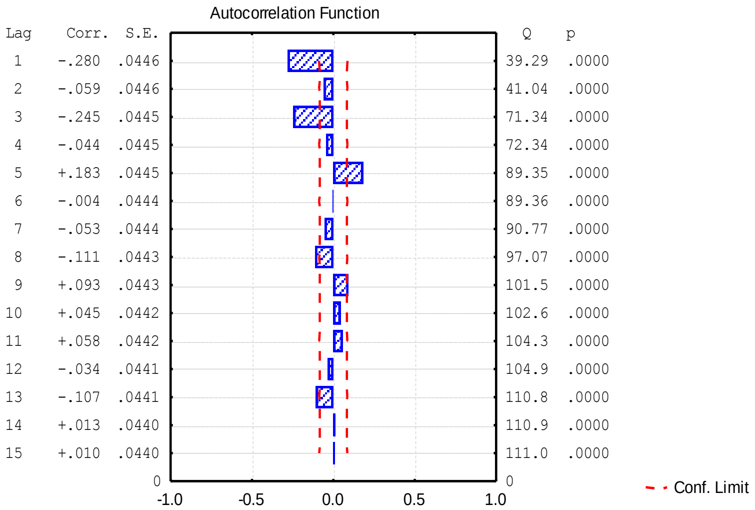

Figure 2.11: The ACF for AR (2) model residuals with one order of difference

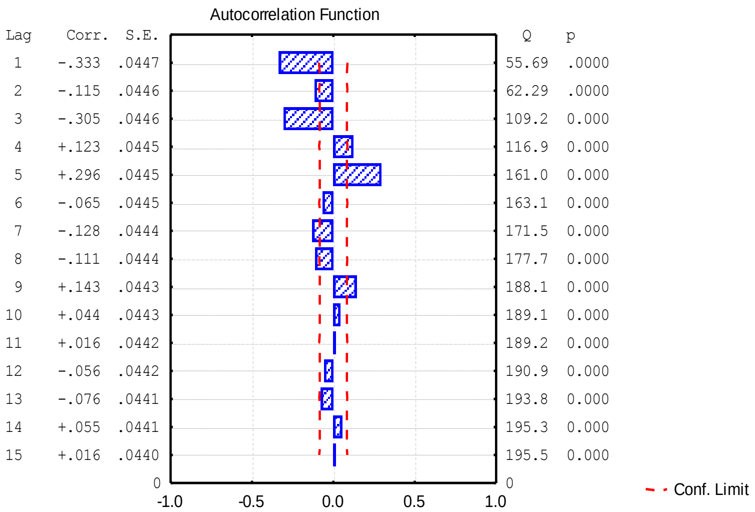

Figure 2.12: The ACF for AR (2) model residuals with two orders of difference

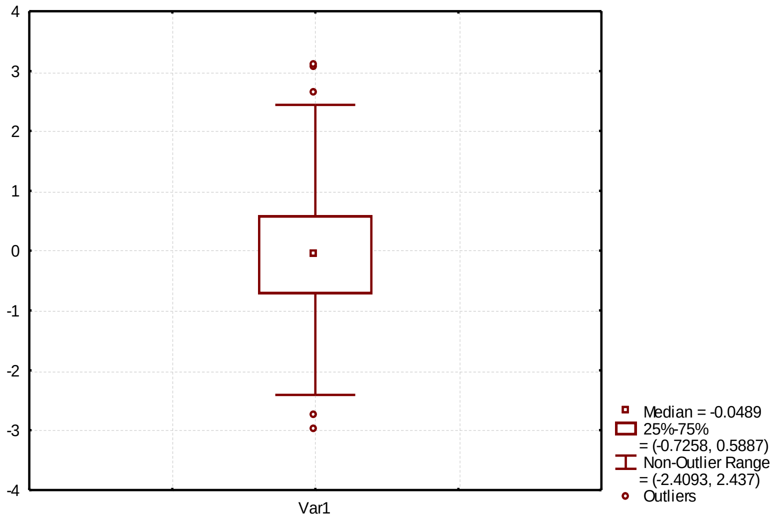

Figure 2.13: The box plot for AR (2) model residuals with no order of difference

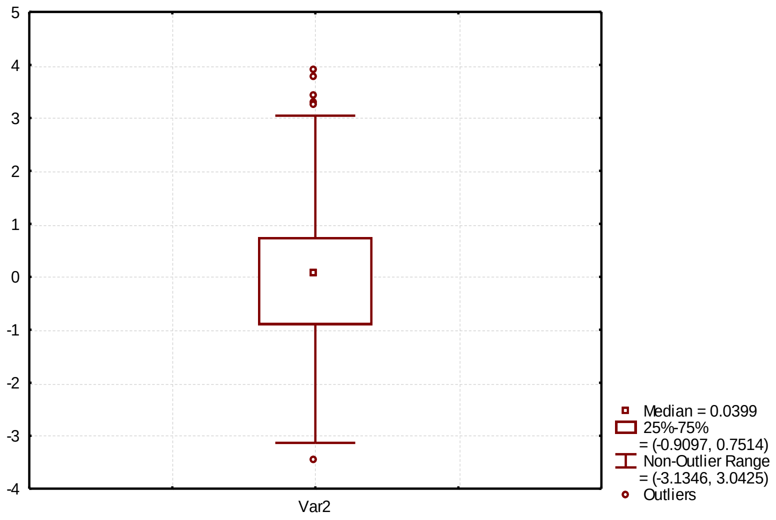

Figure 2.14: The box plot for AR (2) model residuals with one order of difference

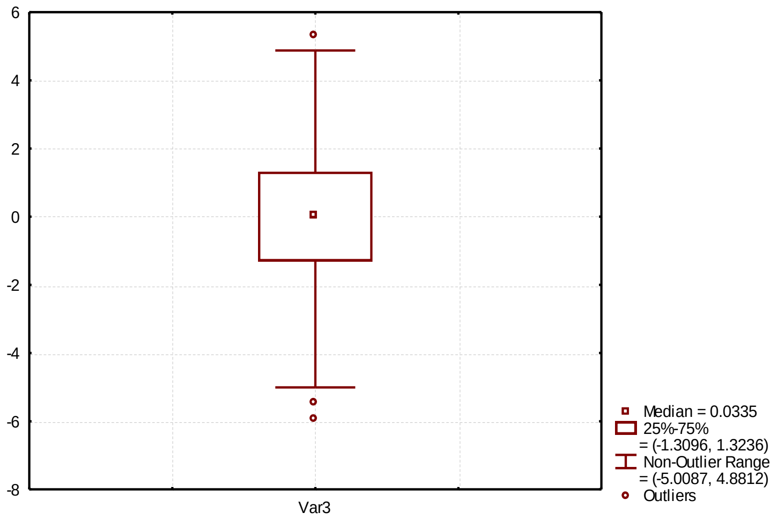

Figure 2.15: The box plot for AR (2) model residuals with two orders of difference

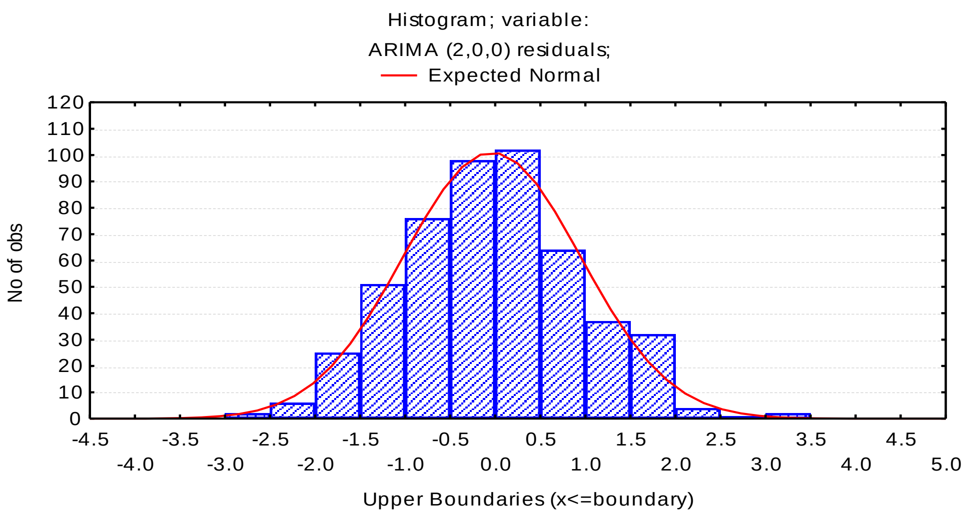

Figure 2.16: The histogram of AR (2) model residuals with no order of difference

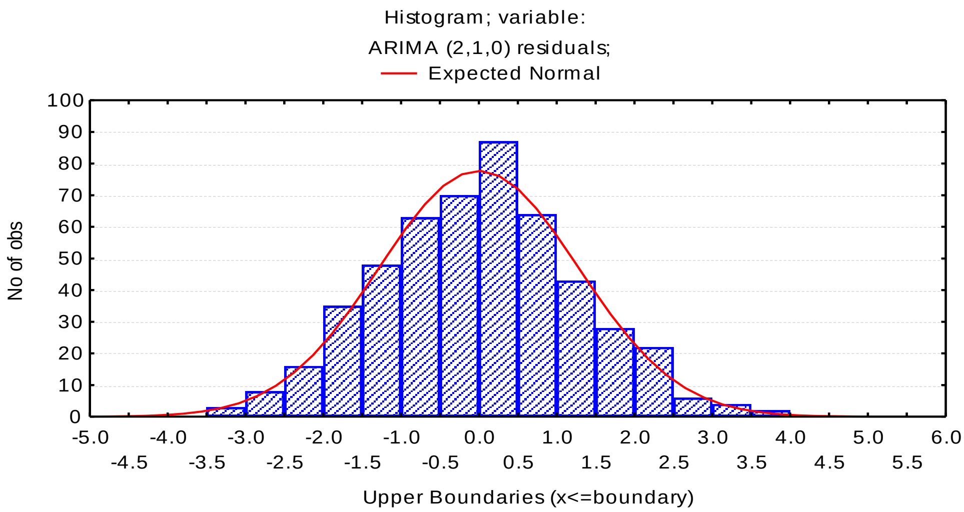

Figure 2.17: The histogram of AR (2) model residuals with one order of difference

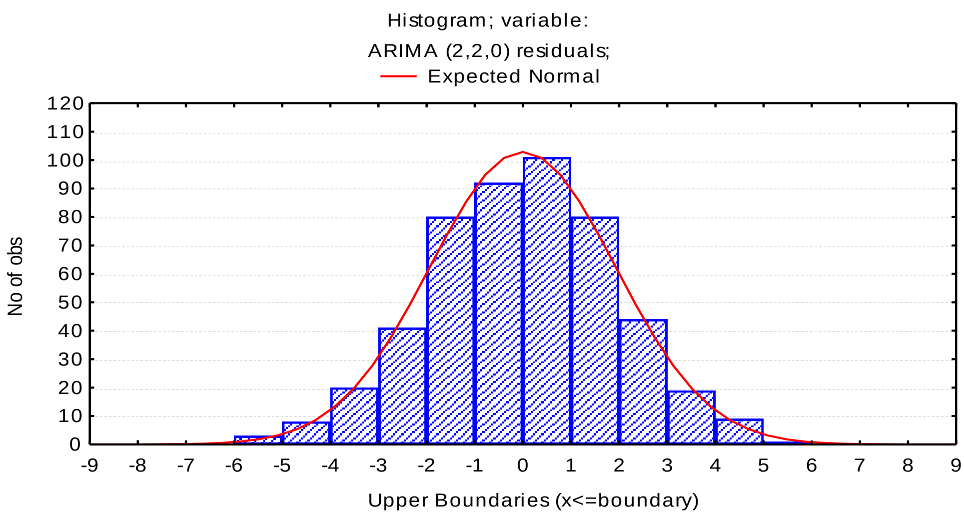

Figure 2.18: The histogram of AR (2) model residuals with two orders of difference

Similarly, Figure 2.1 through Figure 2.18 show the time series plots of the differenced series, ACF, PACF as well as the ACF, the box plots, and the histograms for the residuals calculated from the fitted models AR(2) respect to over-differencing. Above Figures reveal the following two important points 1) the plots of the differenced series still stationary, and 2) the histograms of the residuals still symmetrical and the corresponding box plots have extreme symmetry around its median, suggesting that the residuals computed from the fitted models to inappropriately differenced data are probably an approximate white noise which means that over-differencing may not be completely detected by the residuals analysis.

Finally, the result of the effect of Over-differencing on MA(2) model that are reported in Table 3.

Table 3: The result of the effect of Over-differencing on MA(2) model in which the model with no order, one order, and two orders of a nonseasonal difference

	Order of Difference		
Estimates	**D = 0**	**D = 1**	**D = 2**
$\hat{b}_1$	1.05	-0.06	- 0.32
$\hat{b}_2$	0.14	- 0.91	- 0.67
$SE\ \hat{b}_1$	0.044	0.018	0.951
$SE\ \hat{b}_2$	0.044	0.018	0.657
AIC	1414.67	1426.54	1676.37
Residual variance	0.98	1.00	1.67

Table 3 shows that the estimated coefficient $\hat{b}_1$ is decreasing as the order differencing increases, and the magnitude of $\hat{b}_1$ is largely affected when D = 0 as compared to higher degree of differencing D =1 and D = 2 (Positive effect 1.05 to negative effects -0.06 and -0.32). Additionally, Table 3 shows that the estimated coefficient $\hat{b}_2$ has inverse association with order differencing (0.14, -0.91 and -0.67). Furthermore, the AIC and residual variance measures are jumped from 1414.67 and 0.98 to 1676.37 and 1.67, respectively. Taken all together, the results of over differencing for MA(2) indicate that over differencing lead to completely different models.

The time series plots, ACF, PACF, residuals, box plots, and histogram for the fitted models of MA(2) calculated respect to the differenced series are shown in the following Figures.

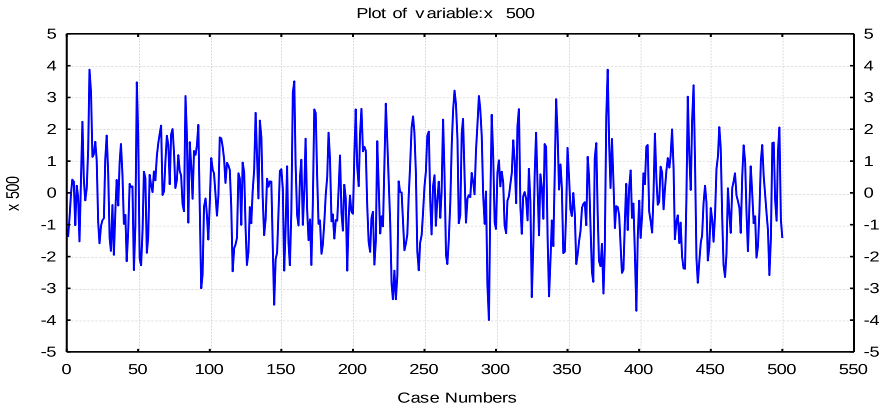

Figure 3.1: Plot of MA (2) model with no order of difference

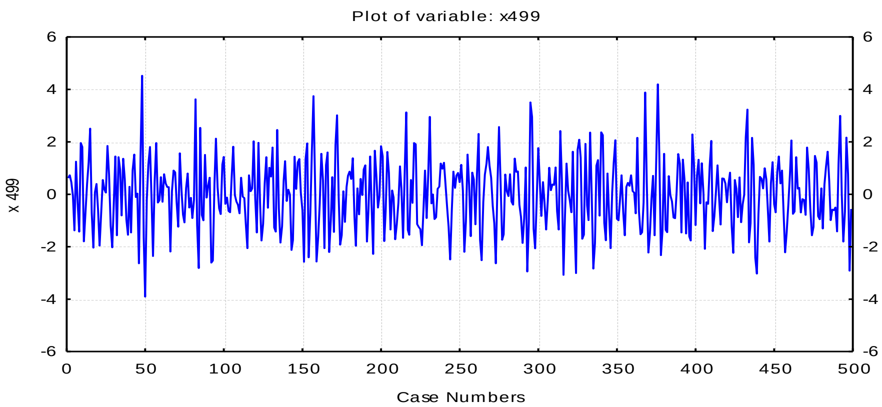

Figure 3.2: Plot of MA (2) model with one order of difference

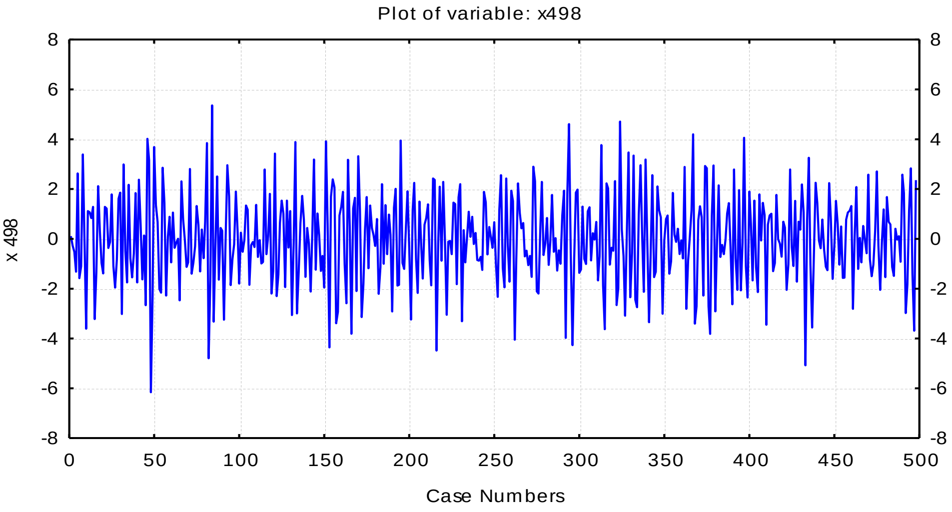

Figure 3.3: Plot of MA (2) model with two orders of difference

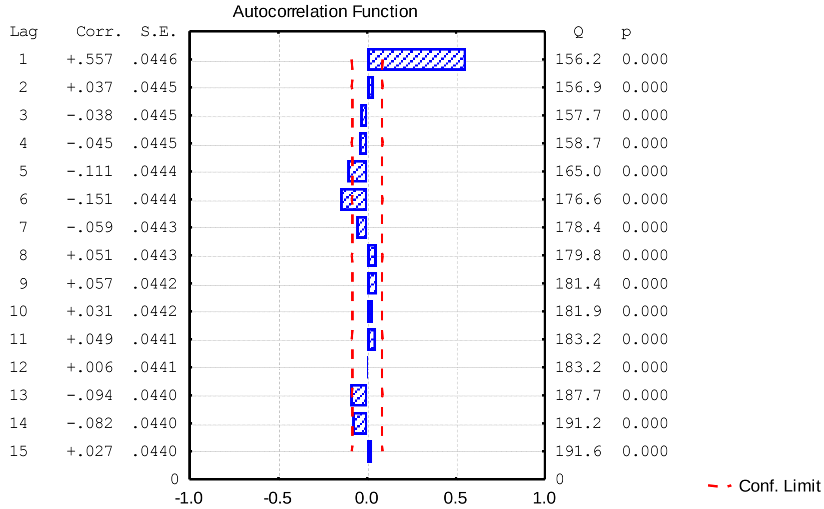

Figure 3.4: The ACF for MA (2) model with no order of difference

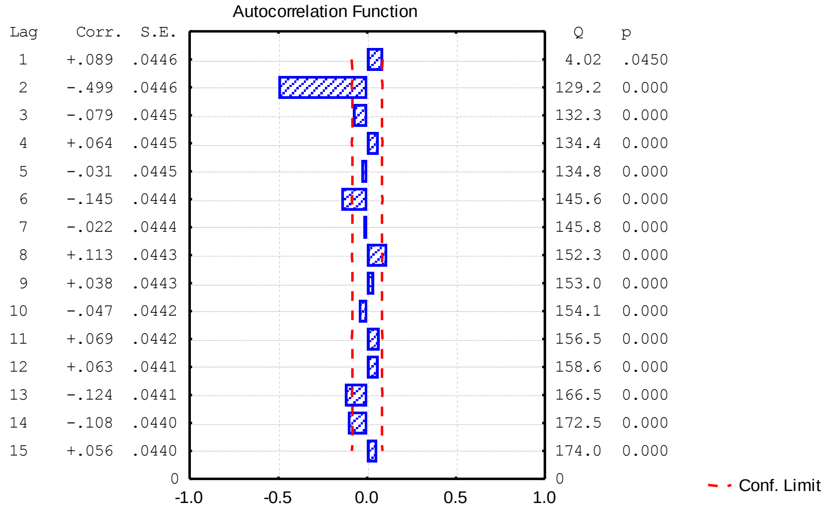

Figure 3.5: The ACF for MA (2) model with one order of difference

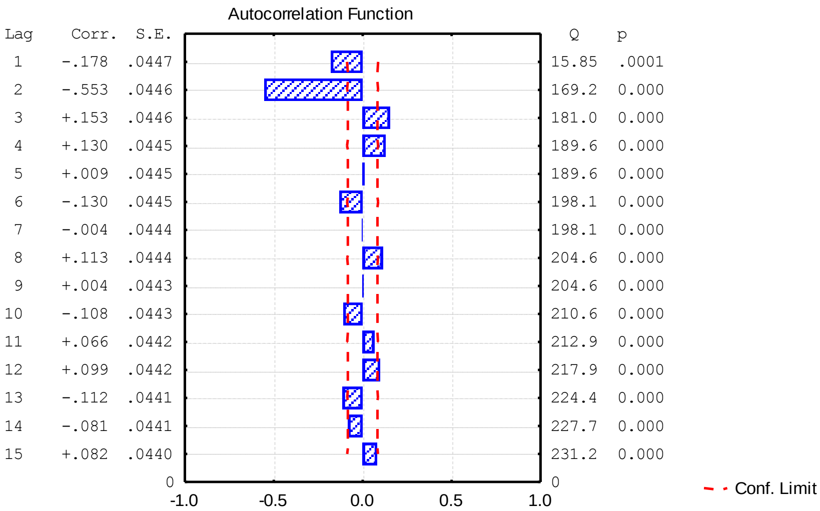

Figure 3.6: The ACF for MA (2) model with two orders of difference

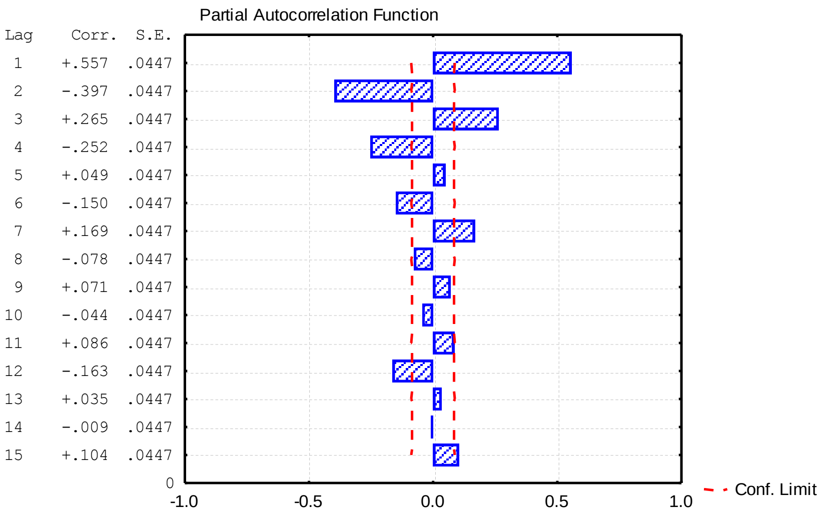

Figure 3.7: The PACF for MA (2) model with no order of difference

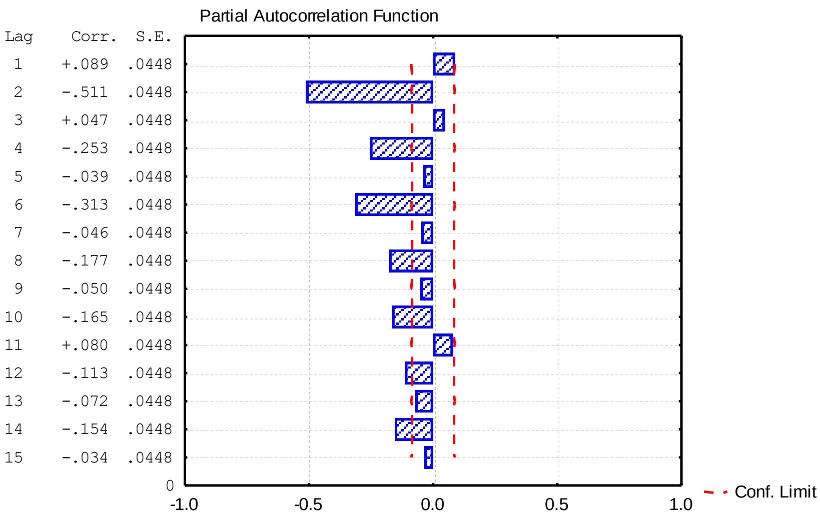

Figure 3.8: The PACF for MA (2) model with one order of difference

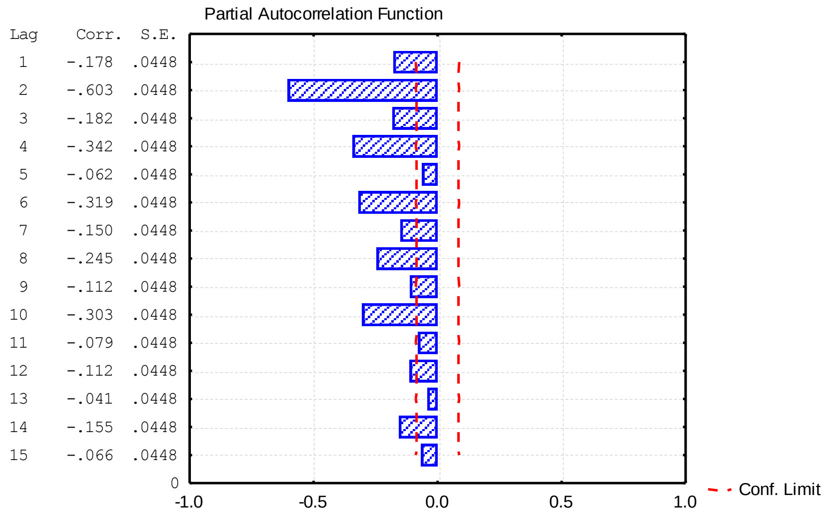

Figure 3.9: The PACF for MA (2) model with two orders of difference

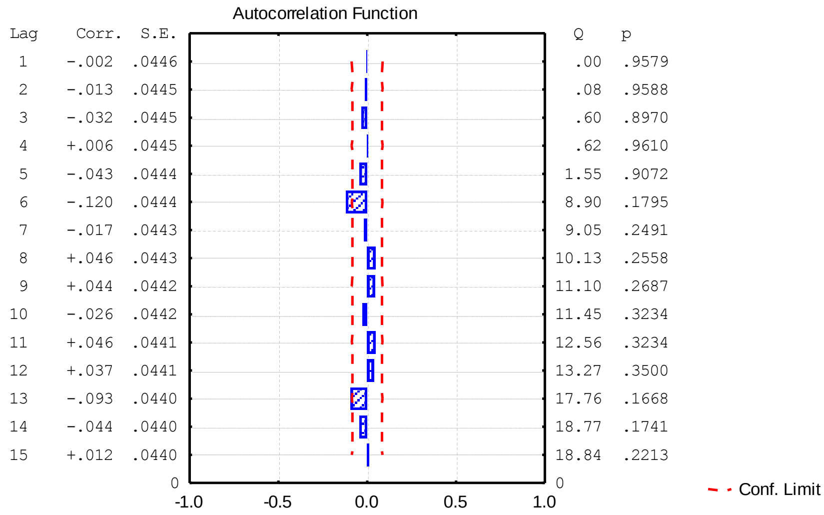

Figure 3.10: The ACF for MA (2) model residuals with no order of difference

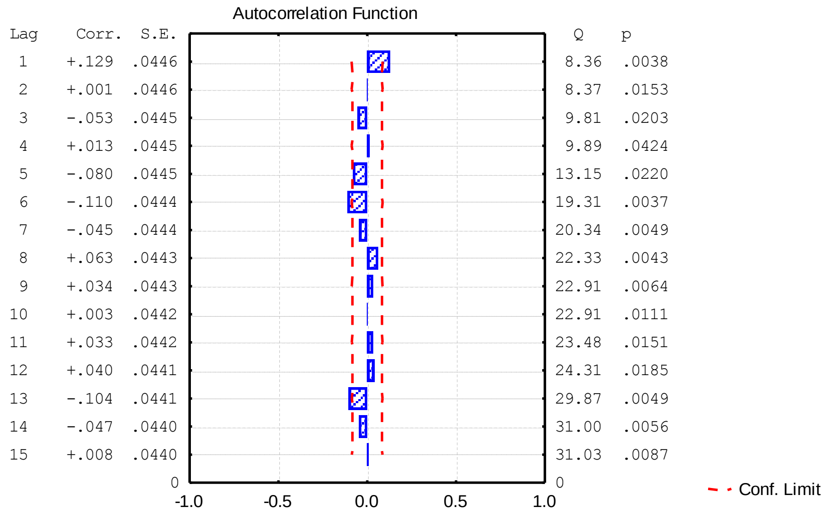

Figure 3.11: The ACF for MA (2) model residuals with one order of difference

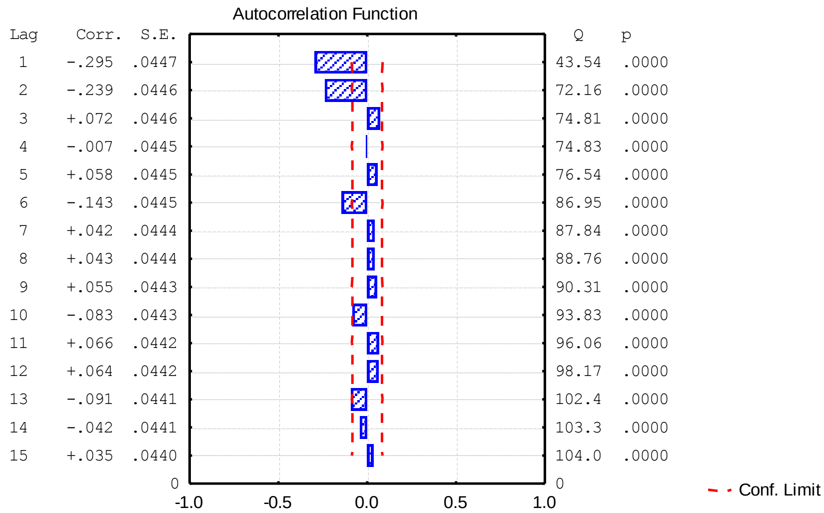

Figure 3.12: The ACF for MA (2) model residuals with two orders of difference

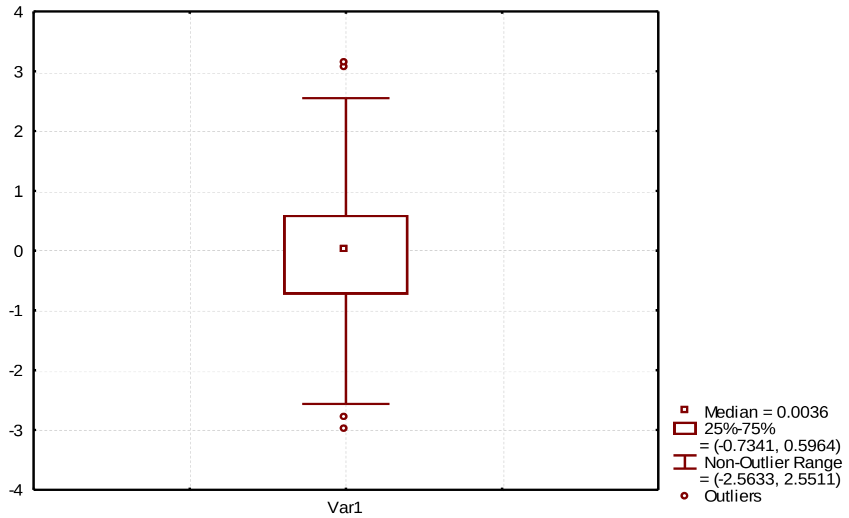

Figure 3.13: The box plot for MA (2) model residuals with no order of difference

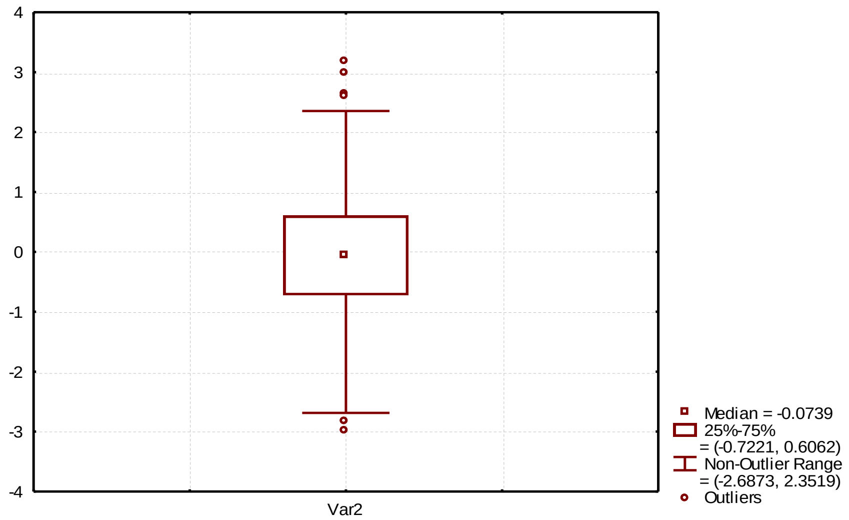

Figure 3.14: The box plot for MA (2) model residuals with one order of difference

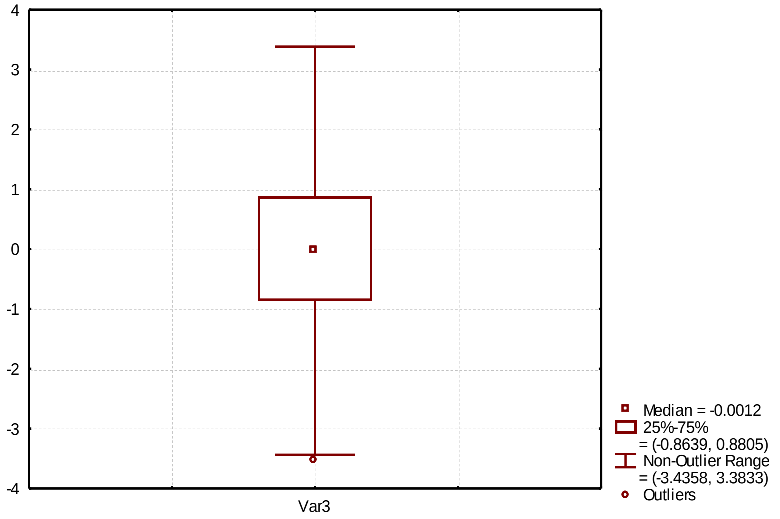

Figure 3.15: The box plot for MA (2) model residuals with two orders of difference

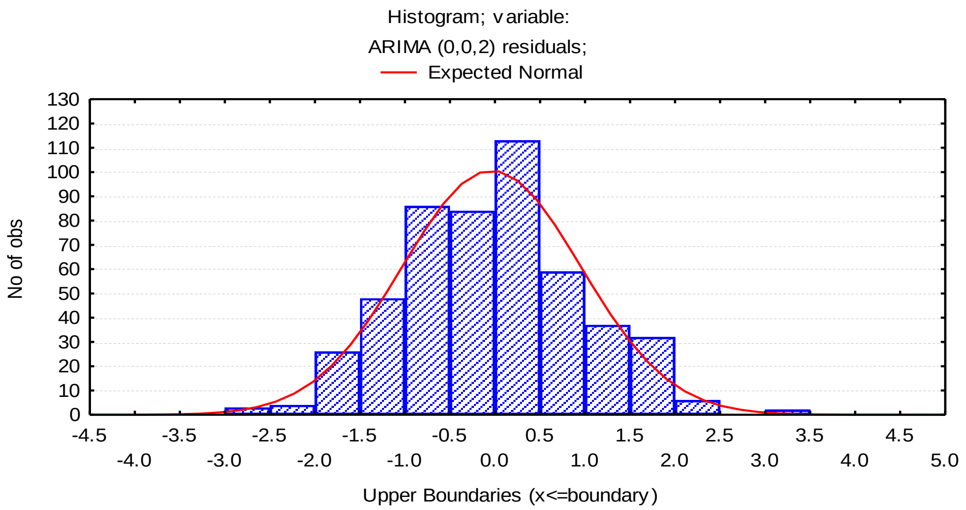

Figure 3.16: The histogram of MA (2) model residuals with no order of difference

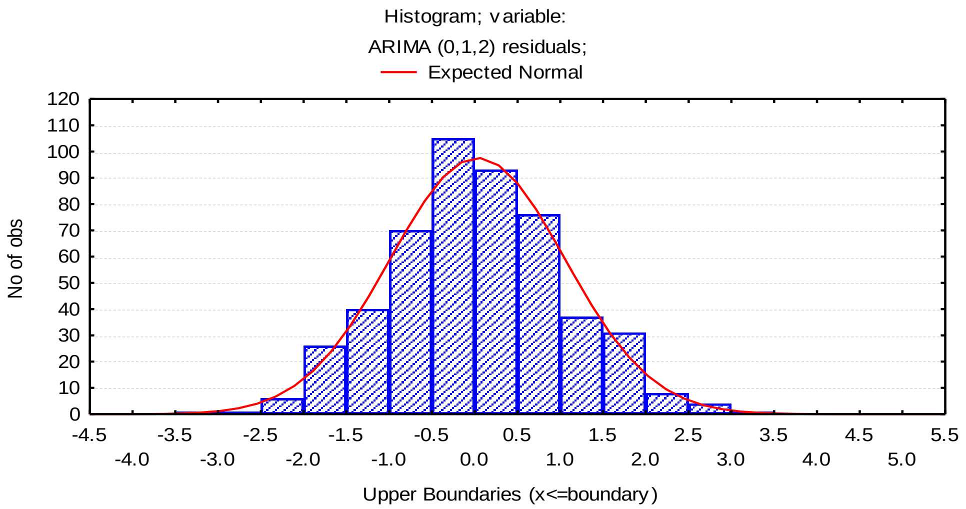

Figure 3.17: The histogram of MA (2) model residuals with one order of difference

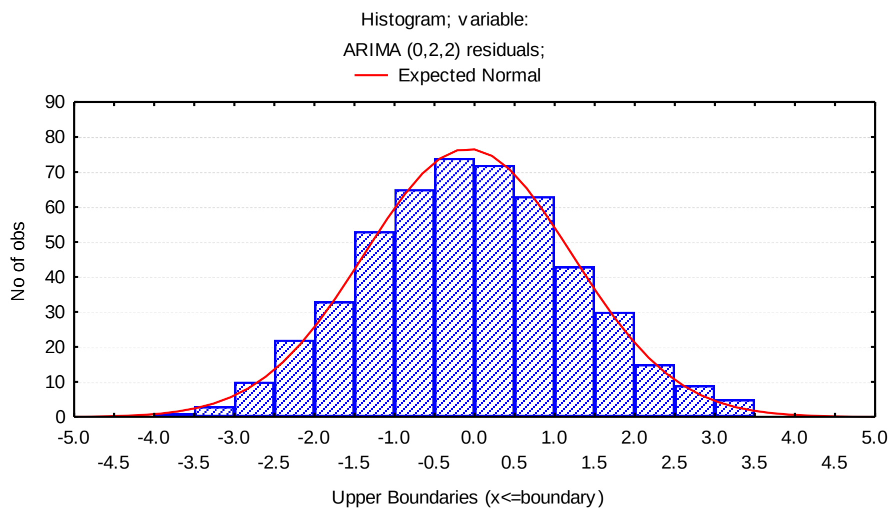

Figure 3.18: The histogram of MA (2) model residuals with two orders of difference

Also, Figure 3.1 through Figure 3.18 show the time series plots of the differenced series, ACF, PACF as well as the ACF, the box plots, and the histograms for the residuals calculated from the fitted models MA(2) respect to over-differencing. These Figures lead to similar conclusion that the plots of the differenced series still stationary, and the histograms of the residuals are symmetrically distributed and the corresponding box plots have extreme symmetry of the central portion where the median is equidistant from the lower and upper limits, indicating that the residuals computed from the fitted models to inadequately differenced data are probably due to white noise which means that over-differencing may not be entirely detected by just conducting the residual analysis.

7. SUMMARY AND CONCLUSIONS

This paper examines the sensitivity of unnecessary differencing on the appropriateness of the proposed model via simulation. Our findings indicted that over-differencing is a serious issue that warrants attention and validation because time series analysts unintentionally misunderstand data that has been improperly differenced. The results also show that over-differencing would result in entirely different model than the true model. Our findings further confirm that the true model is sensitive to over-differencing, however because we used simulated data, we are able to determine how sensitive the true model to the over-differencing, it becomes more challenging to detect the sensitivity of the true model when dealing with real data. It could be interesting to consider frequency domain approach rather than time domain approach. In other words, studying the changes which could happen to the spectrum of the true model due to unnecessary differencing.

REFERENCES

- Abonyi, J., Feil, B., Nemeth, S., & Arva, P. (2005). Modified Gath–Geva clustering for fuzzing segmentation of multivariate time-series. Fuzzy Sets and Systems, Data Mining Special Issue, 149, 39-56.
- Abraham, B., & Ledolter, J. (1983). Statistical methods for forecasting, John Wiley and Sons, Inc.
- Agrawal, R., Faloutsos, C., & Swami, A. (1993a). Efficient similarity search in sequence databases. In: Proceedings of the Fourth International Conference on Foundations of Data Organization and Algorithms, pp. 69–84.
- Anderson, T. W. (1971). The statistical analysis of Time Series; John Wiley and Sons, Inc.
- Berndt, D. J., & Clifford, J. (1996). Finding patterns in time series: a dynamic programming approach. Advances in Knowledge Discovery and Data Mining, 229–248.
- Beusekom D. P (2003). Forecasting and detruding of time series Models, Amsterdam University press.
- Box, G. E. P., & Jenkins, G. M. (1970). Time series Analysis, for forecasting and control. Holden-Day, Inc, San Francisco.
- Chan, K. P., & Fu, A. C. (1999). Efficient time series matching by wavelets. In: Proceedings of the 15th IEEE International Conference on Data Engineering, pp. 126–133.
- Chu, S., Keogh, E., Hart, D., & Pazzani, M. (2002). Iterative deepening dynamic time warping for time series. In: Proceedings of the Second SIAM International Conference on Data Mining.
- Cochrane, J. H. (2012). A brief parable of over-differencing." Manuscript. Chicago: University of Chicago, Booth School of Business.
- De Gooijer, J. G., & Hyndman, R. J. (2006). 25 years of time series forecasting. *International journal of forecasting*, *22*(3), 443-473.
- Faloutsos, C., Ranganathan, M., & Manolopoulos, Y. (1994). Fast subsequence matching in time-series databases. In: Proceedings of the 1994 ACM SIGMOD International Conference on Management of Data, pp. 419–429.

- Fuller, W. A. (1976); Introduction to statistical time series, John Wiley and Sons, Inc.
- Keogh, E. (1997a). A fast and robust method for pattern matching in time series databases. In: Proceedings of the Ninth IEEE International Conference on Tools with Artificial Intelligence, pp. 578–584.
- Keogh, E., & Pazzani, M. (2000a). A Simple dimensionality reduction technique for fast similarity search in large time series databases. In: Proceedings of the 4th Pacific-Asia Conference on Knowledge Discovery and Data Mining, pp. 122–133.
- Montgomery D. C., & Johnson, L. A. (1976). Forecasting and time series analysis, Mcgraw Hill, New York.
- Pristely M. B. (1981). Spectral analysis and time series and time series, Academic press, London.
- Shukla, A. K., Janmaijaya, M., Abraham, A., & Muhuri, P. K. (2019). Engineering applications of artificial intelligence: A bibliometric analysis of 30 years (1988–2018). *Engineering Applications of Artificial Intelligence*, *85*, 517-532.