PROPOSED HYBRID MODEL AR-HOLT (P+5) FOR TIME SERIES FORECASTING BY EMPLOYING NEW ROBUST METHODOLOGY

Firas Ahmmed Mohammed

Department of Statistics, College of Economics& Administration, University of Baghdad, Baghdad, Iraq
firasmohana@coadec.uobaghdad.edu.iq

Abstract

The optimal prediction or forecasting of time series values from the observations required many things such as checking the identification accuracy, model diagnosis, and data free from violations (outliers, for instance). Therefore, the researchers are always wondering if the used model or the supported method is sufficient to represent the data or there are more information that can be provided and probable increasing of precision as a consequence in the forecasting. This paper is an attempt to propose a new hybrid model building that can be denoted by AR-Holt (p+5). Also, suggest a new algorithm to estimate the parameters of this new hybrid model with its forecasting for inside and outside the series. Furthermore, the comparison has been done between this new hybrid model with AR(p) model which was identified as well as its parameters were estimated by many traditional methods which are Yule-Walker, Burg, robust RA, LS, Mcov and LMS methods for contaminated time series data. Simulation experiments have been conducted with different levels of contamination (p=0, 0.05, 0.15) to evaluate the superior of the performance of this new model according to different sample sizes (n=30, 70, 150). A real data application of the barley crops in Iraq is taken into consideration.

Keywords: Yule-Walker method, Burg method, RA method, least squares, modified covariance, LMS method, autoregressive time series model, Holt.

I. Introduction

The ambition to achieve higher accuracy in the construction of representative models of data faces a lot of complexity due to many reasons such as sample size determining and linear or non-linear data behavior. Therefore, statistics have played a major role in putting, developing, carrying out, comparing and forecasting plans and relying on the most accurate forecasting in wide areas of real life. Time series analysis is considered an effective tool in statistics to do so.

Initially, many methods were used in the forecasting approach. The smoothing methods had an important role in forecasting the values of the phenomenon. One of the most important of these methods is the double exponential smoothing method, which is called Holt method that was introduced by the researcher Holt in 1957 [IV]. This method deals directly regardless of their stationarity or not through two
smoothing parameters. Then it had been introduced a new approach based on data stationarity through the Box-Jenks models introduced by both Box and Jenkins in 1970 which have become the most popular in time series modeling and forecasting.

Several methods have been used to estimate the parameters of the model, including Yule-Walker and Burg in 1974 [IX], least square and modified covariance methods which are used to estimate the parameters of the Box Jenkins model in this paper. Sometimes the data may contain outliers, so research and development continued at all stages of construction. Robust methods were interactive in dealing with the contamination associated with the data, including Least Median of Square Regression (LMS) which introduced by Rousseeuw in 1984 [X] and Residual Auto-covariance (RA) which was introduced by Yohai and Bustos in 1989 [III].

However, many times we may face difficulties such as lack of data, inaccuracy in modeling, or the existence of non-linear behavior which is not taken into account, which loses some information about the phenomenon we are working on. Therefore, this paper is a continuation of the efforts that seek to develop models and the method used to estimate through the proposed AR-HOLT model. Moreover, an algorithm for estimating this model and comparing the results through comparison criterion with the traditional and robust methods showed above.

II. Methods and Materials

Yule-Walker method

Assume that $y_1, y_2, \ldots, y_n$ is a stationary time series formed from $n$ observations followed autoregressive model AR(p), and $\theta_i$ represent the parameters of the model which can be found through the following simultaneous equations [XII]:

\[
\begin{pmatrix}
1 & r_1 & \cdots & r_{p-2} & r_{p-1} \\
1 & r_2 & \cdots & r_{p-3} & r_{p-2} \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
r_{p-2} & r_{p-3} & \cdots & 1 & r_1 \\
r_{p-1} & r_{p-2} & \cdots & r_1 & 1 \\
\end{pmatrix}
\begin{pmatrix}
\hat{\theta}_1 \\
\vdots \\
\hat{\theta}_{p-1} \\
\hat{\theta}_p \\
\end{pmatrix}
=
\begin{pmatrix}
1 \\
1 \\
\vdots \\
r_{p-1} \\
r_p \\
\end{pmatrix}
\]

or

\[R\hat{\theta} = r\]

where is:

$\hat{\theta}_i$: parameter estimation, $(i=1, 2, \ldots, p)$
$r_i$: autocorrelation estimation at the shift $(i)$
$R$: autocorrelation matrix of order $p \times p$
$r$: column of the right hand side

then, \[\hat{\theta} = R^{-1}r\]
Burg Method

This method was developed by Burg, 1960 to estimate the spectral function which is called "maximum entropy" method where the model parameters estimated in a successive manner through minimizing forward $e^+_i$ and backward $e^-_i$ predictive error [IX]:

$$\varepsilon_i = \sum_{j=1}^{n}|e^+_i(j)|^2 + \sum_{j=1}^{n}|e^-_i(j)|^2$$

Forward predictive error: $e^+_i(j)$

Backward predictive error: $e^-_i(j)$

Using successive Levinson we can estimate AR(p) parameters successively, where:

$$\phi_p(j) = \phi_{p-1}(j) - \delta_p \phi_{p-1}(p-j) \quad 1 \leq j \leq p - 1$$

$$\delta_p = \frac{r(p) - \sum_{j=1}^{p-1} \phi_{p-1}^j(r(p-j))}{\varepsilon_{p-1}}$$

$$\varepsilon_i = \varepsilon_{i-1}(1 - \delta_i^2)$$

$$\varepsilon_0 = \frac{1}{n} \sum_{t=1}^{n} x_t^2$$

The Modified Covariance Method

Impose that $y_1, y_2, \ldots, y_n$ is a time series of size n observations followed autoregressive model AR(p). In order to estimate the parameters of this model, one can use the modified covariance method by minimizing forward predictive error and backward predictive error average [VII].

$$\rho_p = \sum_{k=p+1}^{n} (|e^+_p(k)|^2 + |e^-_p(k)|^2)$$

where is:

Forward predictive error: $e^+_p(k) = y(k) + \sum_{t=1}^{p} \phi_t y(k - t)$

Backward predictive error: $e^-_p(k) = y(k - p) + \sum_{t=1}^{p} \phi^*_t y(k - p + t)$

$\phi_i, \ i = 1, 2, 3, \ldots, p$ model parameters

To find the unknown parameters which minimize the above mean squared error, must taking the derivatives with respect the parameters to getting the following simultaneous equations system:
\[
\begin{bmatrix}
\tau_p(0,0) & \tau_p(0,1) & \ldots & \tau_p(0,p) \\
\tau_p(1,0) & \tau_p(1,1) & \ldots & \tau_p(1,p) \\
\vdots & \vdots & \ddots & \vdots \\
\tau_p(p,0) & \tau_p(p,1) & \ldots & \tau_p(p,p)
\end{bmatrix}
\begin{bmatrix}
1 \\
\phi_1 \\
\phi_2 \\
\phi_p
\end{bmatrix}
= \begin{bmatrix}
\rho_p \\
0 \\
0 \\
\vdots \\
0
\end{bmatrix}
\]

where:
\[
\tau_p(i,j) = \sum_{k=p+1}^{n} [y(k-i)y(k-j) + y(k-p+i)y(k-p+j)]
\]

then obtain later:
\[
\begin{bmatrix}
\phi_1 \\
\phi_2 \\
\phi_p
\end{bmatrix}
= \begin{bmatrix}
\tau_p(1,1) & \tau_p(1,p) & \tau_p(1,0) \\
\tau_p(2,1) & \tau_p(2,p) & \tau_p(2,0) \\
\vdots & \vdots & \vdots \\
\tau_p(p,1) & \tau_p(p,p) & \tau_p(p,0)
\end{bmatrix}^{-1}
\]

\[
\rho_p = \tau_p(0,0) + \sum_{k=1}^{p} \tau_p(0,k)\phi_k
\]

**Residual Auto-covariance: RA method**

As it is known, least squares methodology is based on the sum of squares minimization directly through AR(p) model estimation. While the estimation procedure of the parameters in RA method is based on a certain weighted error minimization through a robust variance to obtain robust estimators, as follows [III]:

\[
\hat{\phi} = \min(\sum e_i^*) , \text{ of order } (p \times 1) \text{ vector}
\]

where
\[
e_i^* = \Psi(\frac{e_i}{\sigma})
\]

\[
e_i(\hat{\phi}) = y_i - \hat{\phi}_1y_{i-1} - \hat{\phi}_2y_{i-2} - \ldots - \hat{\phi}_py_{i-p}
\]

\[
\Psi(v,k) = \begin{cases} 
  v(1 - \frac{v^2}{k^2})^2 & \text{if} \quad 0 \leq |v| \leq k \\
  0 & \text{o.w}
\end{cases}
\]

\[
\sigma = \frac{\text{median}(|e_{p+1}|, \ldots, |e_n|)}{0.6745}
\]

k=4.685

**Least Median Squares: LMS Method**

LMS is regarded as a robust method, especially with outliers' existence. The main idea is based on minimizing of the median of sum of squares error [X], i.e.:

\[
\text{Copyright reserved © J. Mech. Cont. & Math. Sci.}
\]

Firas Ahmmed Mohammed
\[ \bar{\sigma} = \min(\text{median}(e^{2})) \]

where \(e_i = y_i - \bar{\sigma}_1y_{i-1} - \bar{\sigma}_2y_{i-2} - \ldots - \bar{\sigma}_py_{i-p} \)

**Proposed Hybrid AR-Holt (p+5) Model**

In many cases, the identified models may not represent the data completely, i.e., there is some information still exists potentially in the data but they are not extracted because of one of the following reasons:

1. Small sample size relatively.
2. The mixture of the linear and non-linear behavior of these data, but the linear model is relied only upon because of its simplicity and ignoring the non-linear approach.
3. There are unidentified external effects.
4. Outliers existence.

Consequently, the rest information is embedded in the error term and cannot be useful in estimation precision increment.

In general, the main limitation is the difficulty of determining the model formulae of the remaining errors. Holt smoothing approach will be used in error term prediction inside the series and forecasting outside the series and then combining these estimators with AR(p) model in a linear regression to get a new hybrid model containing (p+5) of parameters, to be denoted finally by AR-Holt(p+5).

Another limitation is that the common robust methods were used to illustrate weights were relied upon to rescale the data by some modification of the nature of the original data [I]. While, by this new mathematical representation of AR-Holt(p+5) model the estimation procedure will not deal with the traditional robust methods, i.e., will deal with the original data directly.

Assuming \((y_1, y_2, \ldots, y_n)\) is a series of observations follows AR(p) model with \((e_1, e_2, \ldots, e_n)\), which are the resulting errors of this model. Then AR-Holt(p+5) model can be written as the following:

\[
y_t = B_0 + B_1(\hat{y}_{AR(p)}) + B_2(\hat{e}_{Holt(\alpha, \gamma)})
\]

where

\[
\hat{y}_{AR(P)} = \bar{\sigma}_1y_{t-1} + \bar{\sigma}_2y_{t-2} + \ldots + \bar{\sigma}_py_{t-p}
\]

are the estimated values of the data series according to AR(p).

\[
\hat{e}_{Holt(\alpha, \gamma)}
\]

are the estimated error values according to Holt approach.

\(B_0, B_1, B_2, \alpha, \gamma, \bar{\sigma}_1, \ldots, \bar{\sigma}_p\) are the model parameters.

**The new algorithm**

To build this model, initially, AR(p) model will be identified and estimated.
2. After that, the prediction inside the series \((\hat{y}_{p+1}, \hat{y}_{p+2}, \ldots, \hat{y}_n)\) will be obtained with their corresponding errors.

3. Obtain the smoothing Holt's parameters \((\alpha, \gamma)\) with regression coefficients \((B_0, B_1, B_2)\) at once by use the combining of the five resulting parameters \((\alpha, \gamma, B_0, B_1, B_2)\) to minimize the sum of squares error iteratively as follows:

\[
\lambda = \min_{B_0, B_1, B_2, \alpha, \gamma} \left( \sum_{t=p+2}^{n} (Y_t - B_0 - B_1 \hat{y}_{AR(P)} - B_2 \hat{\epsilon}_{Holt(\alpha, \gamma)})^2 \right)
\]

where

\[
\lambda = \begin{bmatrix}
\alpha \\
\gamma \\
B_0 \\
B_1 \\
B_2
\end{bmatrix}
\]

\[
\hat{\epsilon}_{Holt(\alpha, \gamma)}(t) = a(t-1) + b(t-1) t=2,3,\ldots,n
\]

\[
a_t^* = ae_t + (1-a)(a_{t-1}^* + b_{t-1}^*)
\]

\[
b_t^* = \gamma (a_t^* - a_{t-1}^*) + (1- \gamma)b_{t-1}^*
\]

where

\(a_t^*\) and \(b_t^*\) are Holt's coefficient which are based on smoothing Holt's parameters \((\alpha, \gamma)\).

While the initial parameters values will be as:

\[
a_1^* = e_1
\]

\[
b_t^* = \frac{(e_2-e_1)+(e_3-e_2)}{2}
\]

4. The prediction of the new proposed hybrid model will be as the following form:

\[
\hat{y}_{AR-Holt}(t) = B + B_1 \hat{y}_{AR(P)}(t) + B_2 (a_{t-1}^* + b_{t-1}^*) t=p+2,\ldots,n
\]

5. Forecasting outside the series range will be according to the following of \(L\) step:

\[
\hat{y}_{AR-Holt}(n+L) = B_0 + B_1 \hat{y}_{AR(P)}(n+L) + B_2 (a^* + (b^*L))
\]

### III. Simulation

Empirical study has been conducted by nine simulation experiments with: sample sizes \((n) = 30, 70, 150\), and contamination proportions \((p)\) = 0.00, 0.05, 0.15, for each of eight models:

\[
\begin{align*}
\text{AR(1)} & : \square_1 = 0.4, \quad \text{AR(1)} & : \square_1 = 0.1 \\
\text{AR(1)} & : \square_1 = -0.9, \quad \text{AR(1)} & : \square_1 = -0.4 \\
\text{AR(2)} & : \square_1 = 1.8, \square_2 = -0.9, \quad \text{AR(2)} & : \square_1 = 0.2, \square_2 = 0.6 \\
\text{AR(2)} & : \square_1 = -0.5, \square_2 = 0.4, \quad \text{AR(2)} & : \square_1 = -1.3, \square_2 = -0.6
\end{align*}
\]
The imposed parameters were selected which satisfied model stationary. The comparisons were based on mean squares error (MSE), and the number of occurring repetitions of the best method through \((r=500)\) replicates which is denoted by (Freq). The numerical results summarized in the following tables below.

Table (1): Simulation results of AR(1): \( \theta = 0.4 \)

| n   | p | Y-W MSE | Burg MSE | LS MSE | Mcov MSE | LMS MSE | RA MSE | Prop MSE | Freq |
|-----|---|---------|---------|--------|----------|--------|--------|----------|------|
| 30  | 0.00 | 0.0968 | 0.0967 | 0.0967 | 0.0967 | 0.0988 | 0.0980 | 0.0775 | 499  |
| 30  | 0.05 | 0.7019 | 0.6699 | 0.6187 | 0.6699 | 0.7007 | 0.6377 | 0.4624 | 500  |
| 30  | 0.15 | 1.6215 | 1.6012 | 1.5845 | 1.6012 | 1.6607 | 1.5877 | 1.0526 | 500  |
| 70  | 0.00 | 0.0991 | 0.099 | 0.099 | 0.099 | 0.0998 | 0.0996 | 0.0911 | 495  |
| 70  | 0.05 | 0.5988 | 0.5909 | 0.5845 | 0.5909 | 0.6042 | 0.5869 | 0.4622 | 500  |
| 70  | 0.15 | 1.5519 | 1.548 | 1.546 | 1.548 | 1.5902 | 1.5465 | 1.2298 | 498  |
| 150 | 0.00 | 0.1000 | 0.100 | 0.100 | 0.100 | 0.1004 | 0.1003 | 0.0962 | 497  |
| 150 | 0.05 | 0.6062 | 0.6033 | 0.6013 | 0.6033 | 0.6138 | 0.6018 | 0.4997 | 500  |
| 150 | 0.15 | 1.5943 | 1.593 | 1.5925 | 1.593 | 1.6092 | 1.5926 | 1.4009 | 500  |

In general, the number of preference occurring of the proposed model as the best method is not less than (495) of all iterations, which represents 99% for all samples and all contamination proportions. This number increases as the sample size and contamination increase to 100% of the 500 replicates. This is confirmed by MSE values. When the sample size is 30 and the contamination proportion is 0, i.e. (non-contamination), the value of MSE = 0.0775 is smaller than the remaining values of other methods. This difference is increased when the contamination rate increases at the same sample size. This work is presented at the same frequency at the sample size \((n=70)\) and \((n=150)\). This indicates the effect of the contamination proportion on increasing the effectiveness of the proposed model.

Table (2): Simulation results of AR(1): \( \theta = 0.1 \)

| n   | p | Y-W MSE | Burg MSE | LS MSE | Mcov MSE | LMS MSE | RA MSE | Prop MSE | Freq |
|-----|---|---------|---------|--------|----------|--------|--------|----------|------|
| 30  | 0.00 | 0.0967 | 0.0967 | 0.0967 | 0.0967 | 0.0983 | 0.0982 | 0.0784 | 497  |
| 30  | 0.05 | 0.6735 | 0.6512 | 0.6125 | 0.6512 | 0.6904 | 0.6326 | 0.4537 | 498  |
| 30  | 0.15 | 1.602 | 1.5923 | 1.5834 | 1.5923 | 1.6485 | 1.5865 | 1.0485 | 499  |
| 70  | 0.00 | 0.099 | 0.099 | 0.099 | 0.099 | 0.0996 | 0.0996 | 0.0911 | 496  |
| 70  | 0.05 | 0.5907 | 0.5866 | 0.5824 | 0.5866 | 0.6015 | 0.5845 | 0.4577 | 500  |
| 70  | 0.15 | 1.5515 | 1.55 | 1.5491 | 1.55 | 1.5786 | 1.5496 | 1.2221 | 500  |
| 150 | 0.00 | 0.1000 | 0.100 | 0.100 | 0.100 | 0.1002 | 0.1003 | 0.0961 | 495  |
| 150 | 0.05 | 0.6013 | 0.6002 | 0.5995 | 0.6002 | 0.6061 | 0.6000 | 0.4921 | 500  |
| 150 | 0.15 | 1.5928 | 1.5925 | 1.5924 | 1.5925 | 1.6022 | 1.5925 | 1.4003 | 500  |

It can be noticed that the minimum number of repetitions of the proposed model is 497 when the sample size \(n = 30\) and the proportion of contamination 0, i.e. (non-contamination) it excelled by 99.4% out of 500 replicates and increase this excellence to 100%. These results are confirmed by MSE values where the proposed model
obtained the lowest values for this criterion compared to the other methods. The lowest value of MSE was 0.0784 when the sample size was 30 and the contamination proportion was 0.

Table (3): Simulation results of AR(1): $\theta_1 = -0.9$

| n  | p  | Y-W MSE | Burg MSE | LS MSE | Mcov MSE | LMS MSE | RA MSE | Prop MSE | Freq |
|----|----|---------|----------|--------|----------|--------|--------|---------|------|
| 30 | 0.00 | 0.0977 | 0.0967 | 0.0964 | 0.0967 | 0.1000 | 0.0968 | 0.0784 | 498  |
| 30 | 0.05 | 0.8946 | 0.8014 | 0.8147 | 0.7552 | 0.4644 | 0.1000 | 0.0968 | 499  |
| 30 | 0.15 | 1.8902 | 1.7087 | 1.5815 | 1.7087 | 1.7989 | 1.6057 | 1.0395 | 500  |
| 70 | 0.00 | 0.0994 | 0.0992 | 0.0992 | 0.0992 | 0.1013 | 0.0993 | 0.0907 | 498  |
| 70 | 0.05 | 0.7027 | 0.6474 | 0.6056 | 0.6474 | 0.6602 | 0.6207 | 0.4552 | 500  |
| 70 | 0.15 | 1.6575 | 1.5968 | 1.5686 | 1.5968 | 1.6462 | 1.5702 | 1.217  | 500  |
| 150| 0.00 | 0.1001 | 0.1000 | 0.1000 | 0.1000 | 0.1017 | 0.1001 | 0.0959 | 498  |
| 150| 0.05 | 0.6602 | 0.6287 | 0.6124 | 0.6287 | 0.6393 | 0.6135 | 0.4958 | 500  |
| 150| 0.15 | 1.629  | 1.6107 | 1.6032 | 1.6107 | 1.6507 | 1.6040 | 1.3848 | 500  |

Note that the value of MSE increases with increasing contamination proportion in general, but the proposed model was carrying the lowest values for other methods were MSE = 1.0395 at the sample size 30 and the proportion of contamination 0.15 and this same behavior in the rest of the contamination proportions. In the above models, the results for the methods used in the value of MSE other than the proposed model were approximate to 0.0959, while the rest methods the corresponding MSE's were 0.1 at zero contamination (non-contamination) proportion. Furthermore, at (p=0.05) and (n=150) MSE of the proposed method was 1.3848, but for the other methods it was approximately 1.6 in general, and MSE for the proposed method equal 0.4958, while for the other methods it was approximately 0.6. This indicates that all methods have failed to improve the predictions of observations despite the increase in the sample sizes and the increase of contamination proportions, which refers to the goodness of this model and the robustness of estimations.

Table (4): Simulation results of AR(1): $\theta_1 = -0.4$

| n  | p  | Y-W MSE | Burg MSE | LS MSE | Mcov MSE | LMS MSE | RA MSE | Prop MSE | Freq |
|----|----|---------|----------|--------|----------|--------|--------|---------|------|
| 30 | 0.00 | 0.0966 | 0.0965 | 0.0965 | 0.0965 | 0.0986 | 0.0980 | 0.0789 | 498  |
| 30 | 0.05 | 0.7088 | 0.6746 | 0.6209 | 0.6746 | 0.7028 | 0.6508 | 0.4566 | 500  |
| 30 | 0.15 | 1.6172 | 1.5961 | 1.5795 | 1.5961 | 1.6811 | 1.5849 | 1.0282 | 499  |
| 70 | 0.00 | 0.0999 | 0.0999 | 0.0999 | 0.0999 | 0.0997 | 0.0996 | 0.0909 | 498  |
| 70 | 0.05 | 0.5999 | 0.5913 | 0.5819 | 0.5913 | 0.6015 | 0.586 | 0.4423 | 500  |
| 70 | 0.15 | 1.5618 | 1.5579 | 1.5554 | 1.5579 | 1.5818 | 1.556 | 1.2206 | 500  |
| 150| 0.00 | 0.1000 | 0.1000 | 0.1000 | 0.1000 | 0.1004 | 0.1003 | 0.096 | 498  |
| 150| 0.05 | 0.6032 | 0.6005 | 0.5989 | 0.6005 | 0.6084 | 0.5995 | 0.4872 | 500  |
| 150| 0.15 | 1.5969 | 1.5958 | 1.5954 | 1.5958 | 1.6066 | 1.5955 | 1.3924 | 500  |
Note that the minimum number of times exceeded the proposed model is 498, i.e. 99.6%, and this percentage increases with increasing of sample sizes and of contamination proportions to reach 100% of the 500 repetitions. The proposed model continued with the same characteristics as in the previous models, i.e. MSE criterion continued with the same excellence compared with the other methods where its MSE values were (0.096, 0.4872, and 1.3924) to the corresponding contamination proportions (0, 0.05, and 0.15) respectively when fixing the sample size at (n=150). At the opposite side, MSE values for the other methods were around (0.1, 0.6, and 1.6) with the corresponding contamination proportions (0, 0.05, and 0.15) respectively, which confirms the effectiveness of the proposed model in improving the predictions for the methods used and the robustness of those predictions for the contamination.

Table (5): Simulation results of AR(2): $\Box_1 = 1.8$ & $\Box_2 = -0.9$

| n  | p   | Y-W MSE | Burg MSE | LS MSE | Mcov MSE | LMS MSE | RA MSE | Prop MSE | Freq |
|----|-----|---------|----------|--------|----------|--------|--------|---------|------|
| 30 | 0.00| 0.1460  | 0.0947   | 0.0929 | 0.094    | 0.1043 | 0.1347 | 0.0731  | 497  |
| 30 | 0.05| 1.6349  | 1.1003   | 0.8046 | 0.8914   | 1.0541 | 1.525  | 0.5579  | 500  |
| 30 | 0.15| 5.3347  | 2.2412   | 1.7429 | 1.9221   | 2.2275 | 12.6662| 1.2642  | 499  |
| 70 | 0.00| 0.1155  | 0.098    | 0.0977 | 0.0979   | 0.1053 | 0.1369 | 0.0895  | 500  |
| 70 | 0.05| 1.8234  | 0.8496   | 0.7082 | 0.8296   | 2.0083 | 0.5723 | 498     |
| 70 | 0.15| 3.2534  | 1.7261   | 1.5908 | 1.6726   | 1.8689 | 8.2717 | 1.3296  | 498  |
| 150 | 0.00| 0.1040  | 0.0994   | 0.0994 | 0.0994   | 0.1038 | 0.1032 | 0.0953  | 500  |
| 150 | 0.05| 1.2763  | 0.6897   | 0.6355 | 0.6636   | 0.7081 | 1.2684 | 0.5456  | 500  |
| 150 | 0.15| 2.5722  | 1.6466   | 1.6009 | 1.6299   | 1.7296 | 5.5311 | 1.4182  | 499  |

The smallest result of MSE was marked by AR-Holt model for all sample sizes (n=30, 70, and 150) as well as all contamination proportions (0.00, 0.05, and 0.15). In addition to the superior outcome of (Freq) column with high repetitions (or frequency), i.e. AR-Holt is better with at least 99.4% when compared with the remaining methods.

Table (6): Simulation results of AR(2): $\Box_1 = 0.2$ & $\Box_2 = 0.6$

| n  | p   | Y-W MSE | Burg MSE | LS MSE | Mcov MSE | LMS MSE | RA MSE | Prop MSE | Freq |
|----|-----|---------|----------|--------|----------|--------|--------|---------|------|
| 30 | 0.00| 0.0941  | 0.0934   | 0.093  | 0.0934   | 0.0982 | 0.0955 | 0.0742  | 498  |
| 30 | 0.05| 0.7338  | 0.6975   | 0.6262 | 0.703    | 0.7742 | 0.7028 | 0.4567  | 499  |
| 30 | 0.15| 1.6318  | 1.5361   | 1.4188 | 1.5432   | 1.7099 | 1.4839 | 1.053   | 498  |
| 70 | 0.00| 0.0979  | 0.0977   | 0.0977 | 0.0978   | 0.0995 | 0.0984 | 0.0891  | 500  |
| 70 | 0.05| 0.6082  | 0.5747   | 0.5309 | 0.5788   | 0.6109 | 0.5606 | 0.4484  | 499  |
| 70 | 0.15| 1.5384  | 1.4947   | 1.4636 | 1.4985   | 1.5665 | 1.4719 | 1.2011  | 499  |
| 150 | 0.00| 0.0994  | 0.0994   | 0.0994 | 0.0994   | 0.1004 | 0.0997 | 0.0951  | 499  |
| 150 | 0.05| 0.6012  | 0.5835   | 0.5685 | 0.5854   | 0.5968 | 0.5719 | 0.4947  | 500  |
| 150 | 0.15| 1.5654  | 1.5547   | 1.549  | 1.5559   | 1.5867 | 1.5498 | 1.3653  | 500  |

Copyright reserved © J. Mech. Cont.& Math. Sci.
Firas Ahmed Mohammed
The smallest result of MSE was marked by AR-Holt model for all sample sizes (n=30, 70, and 150) as well as all contamination proportions (0.00, 0.05, and 0.15). In addition to the superior outcome of (Freq) column with high repetitions (or frequency), i.e. AR-Holt is better with at least 99.6% when compared with the remaining methods.

Table (7): Simulation results of AR(2): $\theta_1 = -0.5$ & $\theta_2 = 0.4$

| n  | p  | Y-W MSE | Burg MSE | LS MSE | Mcov MSE | LMS MSE | RA MSE | Prop MSE | Freq |
|----|----|---------|---------|-------|----------|--------|--------|---------|------|
| 30 | 0.00 | 0.0949  | 0.0937  | 0.0931 | 0.0937   | 0.0973 | 0.0965 | 0.0777  | 495  |
| 30 | 0.05 | 0.7626  | 0.6994  | 0.6139 | 0.6899   | 0.7553 | 0.7387 | 0.4674  | 500  |
| 30 | 0.15 | 1.6718  | 1.5458  | 1.4195 | 1.5468   | 1.7054 | 1.6601 | 1.0843  | 500  |
| 70 | 0.00 | 0.098   | 0.0978  | 0.0977 | 0.0978   | 0.0996 | 0.099  | 0.0907  | 498  |
| 70 | 0.05 | 0.6202  | 0.5815  | 0.5438 | 0.5784   | 0.6059 | 0.5961 | 0.4622  | 500  |
| 70 | 0.15 | 1.5652  | 1.508   | 1.4714 | 1.5101   | 1.5823 | 1.4832 | 1.2469  | 500  |
| 150| 0.00 | 0.0995  | 0.0994  | 0.0994 | 0.0994   | 0.1003 | 0.1001 | 0.0956  | 499  |
| 150| 0.05 | 0.6177  | 0.5926  | 0.5753 | 0.5935   | 0.6028 | 0.5784 | 0.5111  | 500  |
| 150| 0.15 | 1.5816  | 1.561   | 1.5511 | 1.5621   | 1.594  | 1.5533 | 1.3962  | 500  |

The smallest result of MSE was marked by AR-Holt model for all sample sizes (n=30, 70, and 150) as well as all contamination proportions (0.00, 0.05, and 0.15). In addition to the superior outcome of (Freq) column with high repetitions (or frequency), i.e. AR-Holt is better with at least 99% when compared with the remaining methods.

Table (8): Simulation results of AR(2): $\theta_1 = -1.3$ & $\theta_2 = -0.6$

| n  | p  | Y-W MSE | Burg MSE | LS MSE | Mcov MSE | LMS MSE | RA MSE | Prop MSE | Freq |
|----|----|---------|---------|-------|----------|--------|--------|---------|------|
| 30 | 0.00 | 0.0957  | 0.0932  | 0.0929 | 0.0931   | 0.0996 | 0.0955 | 0.0761  | 500  |
| 30 | 0.05 | 1.1008  | 0.8331  | 0.6306 | 0.7092   | 0.8328 | 1.1003 | 0.4694  | 499  |
| 30 | 0.15 | 1.9887  | 1.5837  | 1.3956 | 1.5417   | 1.7366 | 3.8349 | 1.0355  | 500  |
| 70 | 0.00 | 0.0981  | 0.0977  | 0.0976 | 0.0977   | 0.1023 | 0.0984 | 0.0897  | 497  |
| 70 | 0.05 | 0.7386  | 0.6054  | 0.5412 | 0.5828   | 0.6353 | 0.7023 | 0.4536  | 500  |
| 70 | 0.15 | 1.6315  | 1.5138  | 1.4837 | 1.505    | 1.629  | 1.6924 | 1.1936  | 499  |
| 150| 0.00 | 0.0995  | 0.0993  | 0.0993 | 0.0993   | 0.102  | 0.0996 | 0.0952  | 498  |
| 150| 0.05 | 0.6599  | 0.5838  | 0.5647 | 0.578    | 0.6202 | 0.6008 | 0.4781  | 500  |
| 150| 0.15 | 1.6091  | 1.559   | 1.5521 | 1.5568   | 1.641  | 1.6028 | 1.3678  | 500  |
IV. Application

Real data have been collected from Agricultural Statistics Directorate, Central Statistical Organization in Iraq about barley crops (in million tons) in Iraq for the period (1989-2018) [III]. The data have been identified to follow the autoregressive model of the first order, i.e., AR(1) which was considered to obtain the estimators of the parameters by applying the estimation methods. The final estimated model formula was:

\[ y_t = 0.81828y_{t-1} \]

So, the error term \( e_t = y_t - y_{t-1} \) can be computed for the above model AR(1), and then AR-Holt parameters are estimated as the following:

\[ y_t = 1.714526 + 0.07163 y_{t-1} + 1.045098 (\Delta_{t-1} + \Delta_{t-2}) \]
\[ \Delta_{t-1} = 0.177775 e_{t-1} + 0.822225 (\Delta_{t-2} + \Delta_{t-3}) \]
\[ \Delta_{t-2} = 0.65272 (\Delta_{t-1} - \Delta_{t-2}) + 0.34728 \Delta_{t-2} \]

The comparison's results among the different estimation method is summarized below according to MSE criterion.

Table (9): Application results of AR(1):

| Method | YW | Burg | LS | Mcov | LMS | RA | Prop. |
|--------|----|------|----|------|-----|----|-------|
| MSE    | 2.648428 | 2.648432 | 2.648428 | 2.648432 | 2.657819 | 2.648462 | 2.057143 |

It is noticed from the results listed in Table (9) that the proposed method of AR-Holt model yielded the best result regarding MSE outcome with (2.057143). Moreover, the forecasting at the upcoming year (2019) for the barley crop the results were found as the following:

Table (10): Forecasting results of barley crop (million tons)

| Data | Real data 2019 | Forecasting by AR(1) | Forecasting by AR-HOLT(6) |
|------|----------------|----------------------|-----------------------------|
| Barley Crop Product (million tons) | 1.518471 | 0.156002 | 1.708452 |
| Absolute Forecasting Error | - - - | 1.362468 | 0.189981 |

Through the above results in Table (10) it is obvious that the proposed model was better than the traditional model regarding absolute forecasting error. Figure (1) showed real data with their forecasting with concerning to both models.
V. Conclusion

Based on the previous numerical results which are summarized via tables (1-8), we can conclude the following essential points in both simulation experiments below.

1. The main merit of the proposed AR-Holt model is its performance to deal with the original data directly without using any weighting function as it was usually did with robust methods in statistical inference, i.e. the robustness is embedded into AR-Holt itself as a new alternative methodology.

2. At a large sample size (n=150) and a non-contamination proportion (p=0.00) all of the methods have the same MSE value, i.e. they have closeness except for the proposed AR-Holt method because of its excellence.

3. Generally, as contamination proportions increasing then the proposed method of AR-Holt will yield better outcomes.

4. For AR(1) and AR(2) simulations, the numerical results supported the proposed AR-Holt model outcomes concerning to all sample sizes and all contamination proportions as well with at least 99% preference amount.

While the conclusions related to the application can be listed below.

1. The data application coincided with the simulation conclusion in the superiority of AR-Holt performance through MSE outcomes.

2. AR(1) model has been identified and applied to achieve the forecasting of barley crop product amount which included forecasting error greater than AR-Holt which refers to the proposed model superior.
References

I. Abd El-Sallam, Moawad El-Fallah. "Methods of Estimation for Autoregressive Models with Outliers". Asian Journal of Mathematics and Statistics, Vol. 6, No. 2, 2013.

II. Agricultural Statistics Directorate. "Wheat and Barley Products (1989-2018)". Technical Report, Central Statistical Organization (CSO), Iraq, 2019.

III. Bustos, Oscar H. and. Yohai Victor J. "Robust Estimates for ARMA Models". Journal of The American Statistical Association, Vol. 81, No. 393, 1986.

IV. Hamood, Munaf Y.. Jumaa, Ahlam A. and Mohana, Firas A. "Time Series Analysis Part II", Al-Dhad Publishing and Press, Baghdad, Iraq, 2019.

V. Kay, S. M. "Modern Spectral Estimation: Theory and Application". Englewood Cliffs, NJ: Prentice-Hall, 1988.

VI. Marple, S. L., Jr. "Digital Spectral Analysis with Applications". Englewood Cliffs, NJ: Prentice-Hall, 1987.

VII. Marple, S. L., Jr. "A Fast Computational Algorithm for the Modified Covariance Method of linear Prediction". Digital Signal Processing, Vol. 1, 1991.

VIII. Monson, H. "Statistical Digital Signal Processing and Modeling". John Wiley & Sons, 1996.

IX. Pukkila, Tarmo M. "An Improved Estimation Method for Univariate Autoregressive Models". Journal of Multivariate Analysis 27, 422-433, 1988.

X. Rousseeuw, Peter J. "Least Median of Squares Regression". Journal of The American Statistical Association, Vol. 79, No. 388, 1984.

XI. Xiao, Han and Wu, Wei Biao. "Covariance Matrix Estimation for Stationary Time Series". The Annals of Statistics, Vol. 40, No. 1, 466–493, 2012.

XII. Zhang, Hui-Min, and Duhamel, Pierre. "On The Methods for Solving Yule-Walker Equations". IEEE Transactions on Signal Processing, Vol. 40, No. 12, 1992.