A Stepwise Approach to Automate the Search for Optimal Parameters in Seasonal ARIMA Models

Manisha Mukherjee, Diptarka Saha

WalmartLabs, Bengaluru

Abstract- Reliable forecasts of univariate time series data are often necessary in several contexts. ARIMA models are quite popular among practitioners in this regard. Hence, choosing correct parameter values for ARIMA is an imperative yet task. Thus, a stepwise algorithm is introduced to provide robust estimates for parameters $(p, d, q)(P, D, Q)$ used in seasonal ARIMA models. This process is focused on improvising the overall quality of the estimates and it alleviates the problems induced due to the unidimensional nature of the methods that are currently used such as 'auto.arima'. The fast and automated search of parameter space also ensures reliable estimates of the parameters that possess several desirable qualities and consequently resulting in higher test accuracy especially in the cases of noisy data. After vigorous testing on real as well as simulated data the authors have found sufficient evidence to conclude that the algorithm performs better than current state-of-the-art methods, all the while completely obviating the need for human intervention due to its automated nature.

Index Terms- Time Series, auto.arima, ARIMA parameters, Forecast, R

I. INTRODUCTION

Time Series analysis and forecasting plays an essential role in various fields such as business, finance, economics, science, and engineering. Due to its importance in solving practical problems, several methods have been proposed in the literature to model a time series so that past observations are carefully handled, and future can be predicted accurately with confidence. Time series forecasting is thus nothing but an endeavour to predict the future by an astute scrutiny of the past.

One of the most popular and frequently used time series models is ARIMA (Autoregressive Integrated Moving Average) suggested by George Box and Gwilym Jenkins in their seminal text-book *Time Series Analysis: Forecasting and Control* [1]. The modelling approach is well celebrated in the academic community due to its robust theoretical underpinnings. In fact, under certain assumptions, it has been shown that ARIMA models may yield the *optimal forecasts*, outperforming competing methods such as *Exponential Smoothing* [2]. Variants of ARIMA, such as seasonal ARIMA, has been in use as well, with additional sets of parameters to capture the seasonality present in the series.

When applied on real-world data however, ARIMA originally didn't receive such vogue from industry practitioners. This is partly because business data may not always conform to the necessary assumptions, but mostly due to the difficult, iterative time-consuming, and highly subjective procedure described by Box and Jenkins to identify the proper form of the model for a given data set. There have since been many attempts at automating the search for the optimal ARIMA model. However, as will be discussed in the later sections – all of them show several limitations.

This paper presents a stepwise algorithm which automates the iterative nature of Box and Jenkin's approach to find best seasonal ARIMA model for a particular time series and performs better than current state of the art algorithms in terms of various criterion on which a time series model is judged and thus may fail to eliminate the need of human intervention.

Section II summarizes the existing works which has been done to find solution to automatic ARIMA modelling. Section III describes the ARIMA model and the proposed algorithm. Section IV discusses the advantages of proposed algorithm over existing processes. Section V sums up the performance of the proposed algorithm on various time series data sets and section VI provides a salient conclusion.

II. LITERATURE REVIEW

Several attempts have been made in direction to automate the procedure to find ARIMA model in the past years, especially, in the eighties and nineties. The most recent method is given by Hyndman et. al [3] in which he proposed finding an optimum ARIMA model by minimizing AIC [4] by considering different combinations of model parameters. Presently, this procedure is most used commercially as a back-end algorithm in statistical software R's forecast package's *auto.arima* and in Scikit-Learn's *pmdarima* in Python.

Hannan [5] proposed a method to identify the order of an ARMA model for a stationary series by fitting the innovations as an autoregressive model to the data followed by computation of likelihood of potential models using a series of standard regressions. The asymptotic properties of the procedure under very general conditions were then derived.

Gömez [6] extended the Hannan-Rissanen identification method to include multiplicative seasonal ARIMA model identification. They implemented the automatic identification procedure in the softwares TRAMO and SEATS in which the algorithm fetched the model with minimum BIC.

Mélard and Pasteels [7] proposed an algorithm for univariate ARIMA models which also allows intervention analysis and has been implemented in the software package "*Time Series Expert*" (TSE-AX). Liu [8] also suggested a method for identification of seasonal ARIMA models using a filtering method and certain heuristic rules which is used in SCA-Expert software. Forecast Pro [9], which is partially based on [10] and is famous for its automatic ARIMA algorithm which was used in the M3-forecasting competition [11]. Another proprietary algorithm is implemented in Autobox [12].

Hwang [13] developed an automated time series cost forecasting system (ATMF) including both auto-selected procedures for determining a best-fitting model and an auto-extracting module for forecasting values using the Box-Jenkins approach. Amin et. al [14] proposed an automated forecasting approach based on ARIMA to capture linear and non-linear patterns to predict future values of Quality of Service (QoS) attributes that can assist in controlling in software intensive systems.

However, it needs to be highlighted that not much attempts have been proposed post Hyndman et. al [3] to automate the procedure of modelling in ARIMA. The lack of literature in recent times along with some serious limitations of the existing processes make it imperative to develop an automated forecast method based on the original Box and Jenkins approach.

III. MODEL

A. Framework

A seasonal ARIMA $(p,d,q)(P,D,Q)_f$ process is given by

$$\Phi(B^f)\phi(B)(1-B^m)^D(1-B)^d y_t \ = c + \Theta(B^f)\theta(B)\epsilon_t$$

Where, ϵ_t is a white noise process with mean 0 and variance $\sigma^2, \Phi(z)$ and $\Theta(z)$ are polynomials of order P and Q respectively, each containing no roots inside a unit circle, B is a backshift operator and f is the seasonal frequency of the series.

For a seasonal time series, $D = 0 \ or \ D = 1$ is decided on the basis of Canova-Hansen test [15]. This test tests the Null Hypothesis that no seasonal unit root is present. Unit-root tests to test for presence of stochastic trend such as ADF test [16] and KPSS test [17] are performed to choose the optimum value of the parameter d. ADF test tests the Null Hypothesis that unit root is present whereas KPSS test tests the Null Hypothesis that no unit root is present.

After optimum d and D is chosen, optimum $p,\ q,\ P,$ and Q are selected on the basis of the ACF and PACF of the series so that best model with minimum AIC and best performance on test data is obtained.

A time series Y_t has a mean,

$$\mu = E[Y_t]$$

and autocovariance function,

$$\gamma_Y(t+h,t) = Cov(Y_{\{t+h\}}, Y\_t) = E\big[\big(Y_{t+h} - \mu_{(t+h)}\big)(Y_t - \mu_t)\big]$$

It is stationary when both mean and autocovariance function are independent of t.

The Autocorrelation Function (ACF) is

$$\rho_y(h) = \frac{\gamma_Y(h)}{\gamma_Y(0)} \ = Corr(Y_{t+h}\,, Y_t)$$

When $d,\ D,\ p,\ q, P,$ and Q are known, a model can be evaluated via an information criterion such as AIC:

$$AIC \ = \ -2log(L) \ + \ 2(p+q+P+Q+k),$$

where, $k = 1$ if $c \neq 0$ and $k \ = \ 0$ otherwise and L is the maximized likelihood of the model fitted to the differenced data $\left(1 - B^f\right)^D (1-B)^d y_t$

The performance of a time series model is evaluated on the basis of *Mean Absolute Percentage Error* (MAPE). In a time series, if y_t is the actual value and $\widehat{y_t}$ is the predicted value (forecast) and the series conists of n time points then MAPE is defined as follows:

$$MAPE \ = \frac{100}{n} \sum_{t=1}^{n} \frac{|y_t - \widehat{y_t}|}{y_t}$$

An optimised seasonal ARIMA model should satisfy the following criteria:

1. Optimum difference order, d, should be on the basis of both KPSS and ADF test.
2. Optimum seasonal difference order, D, should be on the basis of Canova-Hensen test.
3. Optimum p, q, P and Q should be on the basis of ACF and PACF.
4. It should have minimum AIC.
5. It should have minimum test MAPE.
6. Model residuals should not have serial autocorrelation

B. Algorithm

Finding the best ARIMA Model has always been subjective and difficult. There is no hard and fast rule suggested in literature to find a best ARIMA model which makes developing an automated solution to it even more cumbersome. Robert Hyndman [3] has suggested finding optimum ARIMA models on the basis of AIC, which is currently being used in popular statistical softwares such as auto.arima in R and pmdarima in Scikit-learn in python. However, we note, this approach is unidimensional in nature as it only looks to find the model with minimum AIC.

We suggest an algorithm which has a multidimensional view and it has been broken down into three separate algorithms in this paper for the understanding of the reader. The Algorithm 1 finds optimum D and d of the ARIMA model. It consumes input data which is a time series data and tests for presence of seasonality in the data. It carries out Canova Hansen test to check for seasonal unit root and passes out the seasonally differenced data to next step. The algorithm automatically skips this test if the data is annual in nature. In next step, Augmented Dickey Fuller test and KPSS test are performed to find optimum d. In case, results of ADF test and KPSS test do not match, it throws a warning to user.

Algorithm 2 gives optimum $p, q, P,$ and Q of the model. The differenced data from algorithm 1 is taken as an input and ACF and PACF of the series are computed. Based on the values of ACF and PACF, optimum values of p, q, P and Q are found. There has been an upper bound put on the values of these parameters so that the resulting model is parsimonious.

Algorithm 3 fits the ARIMA model and performs diagnostic tests. It stores AIC of the resulting model and compute residuals. It then performs Ljung-Box test on residuals to check for presence of serial autocorrelation in the residuals. It throws the optimum parameters and AIC of the model along with the results of Ljung-Box test.

Let us, then, formally introduce the algorithms.

Algorithm 1: Finding Optimum D and d of ARIMA Model

Input: Time Series y_t, explanatory variable x_t, frequency f, cut off value V

Output: Optimum D, d

Initialize:

$$z_t = 0, d = 0, D = 0, w_t = 0, d_1 = 0, d_2 = 0$$

Seasonal Unit Root:

if $f > 1$:

Perform Canova Hansen Test on y_t

if H_0 *is rejected* then $z_t = y_t - y_{t-f}$; $D = 1$

else $z_t = y_t$; $D = 0$

Augmented Dickey Fuller Test:

while $i \in \{0,1,2\}$ do

$u_t = \Delta^i(z_t)$ Perform ADF test on u_t

if H_0 *is rejected* then $w_t = u_t$; $d_1 = i$

else $i = i + 1$

Kwiatkowski Phillips Schmidt Shin Test:

while $i \in \{0,1,2\}$ do

$v_t = \Delta^i(z_t)$ Perform KPSS test on v_t

if H_0 *is rejected* then $w_t = v_t$; $d_1 = i$

else $i = i + 1$

if $d_1 = d_2$ then $d = d_1 = d_2$

else $d = d_1$; output warning message

Algorithm 2: Finding Optimum p, q, P, and Q of ARIMA Model

Input: Time Series y_t, explanatory variable x_t, frequency f, cut off value V

Output: Optimum p, q, P, Q

Initialize:

$$a_t = 0, b_t = 0, p = 0, q = 0, P = 0, Q = 0$$

for $1 \leq h \leq f + 5$

$$a_t[h] = \rho(h)$$

for $1 \leq h \leq f + 5$

$$b_t[h] = \pi(h)$$

Optimum q:

```
for i ∈ {1,2,3,4,5} do
    if a_t[i] < V then
        q = i − 1; j = i + 1

        for j ∈ {1,2,3,4,5} do
            if a_t[j] < V then
                    q = j − 1
            else
                    q = 0
```

Optimum p:

```
for i ∈ {1,2,3,4,5} do
    if b_t[i] < V then
        p = i − 1; j = i + 1

        for j ∈ {1,2,3,4,5} do
            if b_t[j] < V then
                    p = j − 1
            else
                    p = 0
```

Optimum P:

```
for i ∈ {f + 1, f + 2, …, f + 5} do
    if b_t[i] < V then
        P = i − 1; j = i + 1

        for j ∈ {1,2,3,4,5} do
            if b_t[j] < V then
                    P = j − 1
            else
                    P = 0
```

Optimum P:

```
for i ∈ {f + 1, f + 2, …, f + 5} do
    if a_t[i] < V then
        Q = i − 1; j = i + 1

        for j ∈ {1,2,3,4,5} do
            if a_t[j] < V then
                    Q = j − 1
            else
                    Q = 0
```

Algorithm 3: Fitting optimal ARIMA model and performing diagnostic tests

---

Input: Time Series y_t, explanatory variable x_t, frequency f, cut off value V

Output: Optimum p, d, q, P, D, Q, R, AIC

Initialize:

$$AIC = 0, R = 0$$

Fit ARIMA $(p, d, q)(P, D, Q)[f]$ on w_t and explanatory variable x_t

Compute AIC of the model, store as AIC

Compute residuals $e_t = y_t - \widehat{y_t}$

Ljung Box test on Model Residuals:

if H_0 *is Rejected*

$R = 1$; Serial Autocorrelation Absent

else

$R = 0$; Serial Autocorrelation Present

IV. DISCUSSION

Literature suggests that a best time series model should satisfy all the criterion discussed in section III.A. Existing automated solutions for finding suitable ARIMA model used in popular software packages only take care of one aspect – namely, minimizing the AIC of the model and only conducts KPSS test to find optimum d. They consider different combinations of p, q, P, and Q and choose the one which provides minimum AIC.

The authors have sufficient evidence to believe that due to such neglect in consideration of ACF and PACF current methods fail to produce optimum model, especially in those cases where the data is noisy. It has also been observed that in some cases they give positive values of P and Q even though it is clearly evident from the data (and from ACF and PACF) that there is no seasonality present. In addition to this, they do not conduct Ljung-Box test on model residuals on their own. In literature, it is highly recommended in order to notify the user if the serial correlation is present in the residuals or not.

The algorithm discussed in previous section overcomes all these limitations and consequently shows better performance than existing solutions. First, it takes into account the results of both KPSS test and ADF test to find out optimum difference order and in case of contradictory results, it throws a warning to the user. Second, it iterates over ACF and PACF of the series to find out optimum p, q, P, and Q which makes this algorithm highly effective in case of noisy data to capture all the patterns and nuances present in the series. It tests the model on test data on its own to provide the user test MAPE. Third, this algorithm performs Ljung-Box test on model residuals to notify user if there is a serial correlation present in residuals generated by the model. Fourth, it gives the user liberty to choose the cut off value i.e. the lag(V) till which ACF and PACF are considered significant on his own, unlike the existing black box solutions. Fifth, in cases where ACF and PACF behaves abnormally, for e.g., ACF at lag 1 and lag 2 is insignificant but ACF at lag 3 is significant, the existing solutions would provide MA order as 1 or 2 but this algorithm goes to lag 3 and beyond to find optimum MA order while making sure that model remains parsimonious.

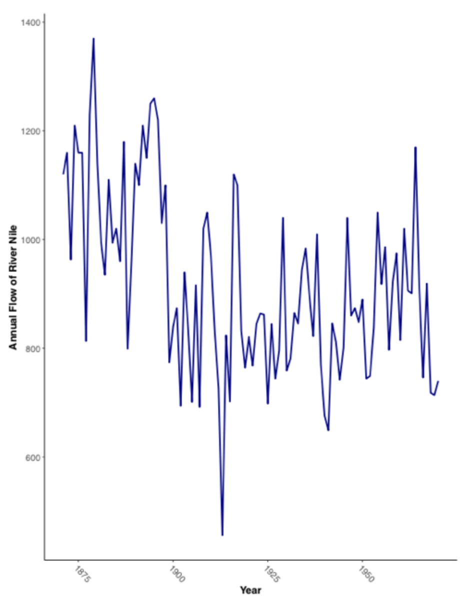
As a result, the model generated by this algorithm tend to perform better than the models generated by existing solution in terms of both AIC and test MAPE.

V. EXPERIMENTS

We show here three implementations of the proposed algorithm and demonstrate the superiority of our proposed solutions over currently popular solution such as auto.arima. Three cases have been chosen to highlight different aspect of our algorithm as will be discussed later.

A. Dataset 1: Nile Data

It is a standard time series data present in various statistical softwares summarizing annual flow of Nile River from year 1871 to 1970. It is illustrated in figure 1(a).



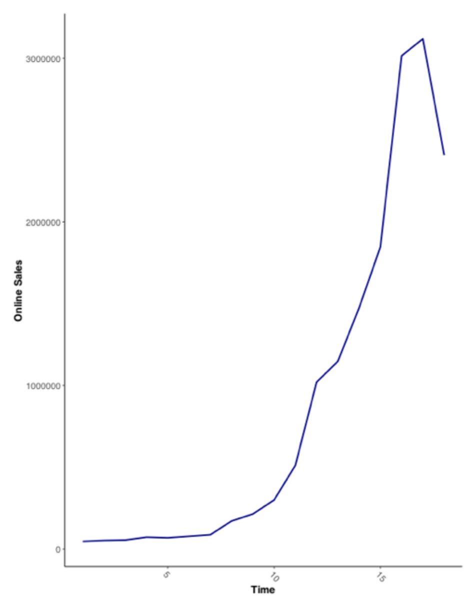


Figure 1(a): Annual Flow of Nile River

Figure 1(b): Quarterly Sales in Walmart

The data is split into train and test data, in which, train data is from year 1871 to 1941 and test data from year 1940 to 1970. To align with notations used in the algorithm in section II, we have, frequency $f = 0$, Train data as y_t and cut off value $V = 0.33$ as input. Canova-Hansen Test is not performed as time series is yearly in nature. Later, ADF test is performed followed by KPSS test, both of which, suggested optimum $d = 1$. ACF and PACF were then computed on the differenced series, w_t, to find optimum values of $p, q, P,$ and Q. The correlograms are depicted in figure 2(a).

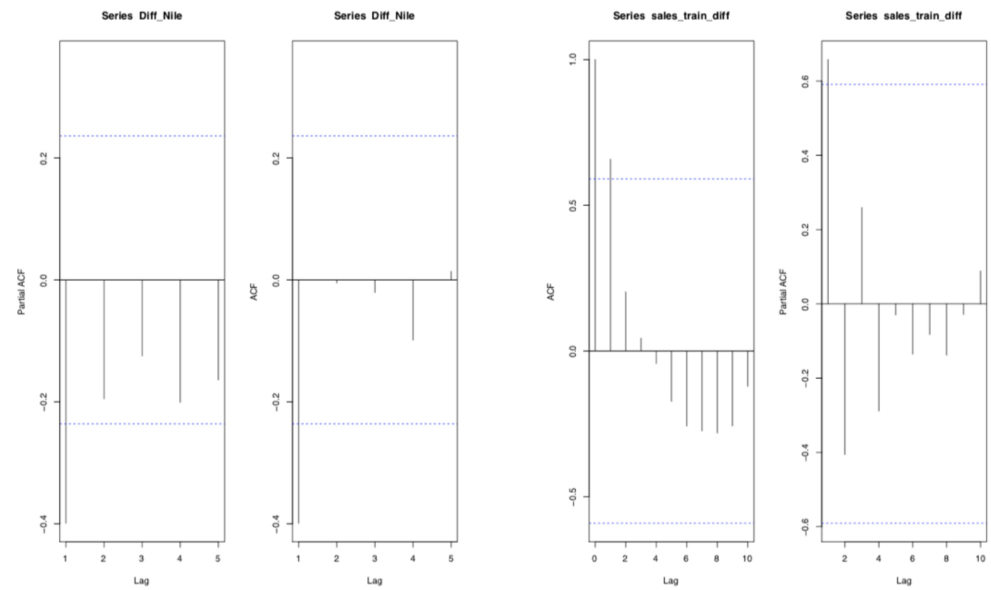


Figure 2(a): ACF, PACF in Dataset 1

Figure 2(b): ACF, PACF in Dataset 2

The algorithm derives ACF and PACF of the differenced series and gives optimum values of p, q, P, and Q as (1,1,0,0) respectively. This finding is consistent with the figure 2(a). There are significant spikes in correlograms of both ACF and PACF at lag 1 which diminish from lag 2 onwards. This step is followed by fitting an ARIMA model and performing Ljung-Box test on the model residuals which suggested no serial autocorrelation in model residuals with $p - value = 0.71$. The final output is $p = 1, d = 1, q = 1, P = 0, Q = 0, D = 0, R = 1, AIC = 894.13$. The test MAPE is 10.56 %.

To compare with existing automated solutions, auto.arima is implemented on the same train and test data. It suggests $p = 1, d = 0, q = 0, P = 0, D = 0, Q = 0, AIC = 910$. The optimum values of q is not consistent with values of ACF of the series. It does not perform Ljung-Box test on its own. When performed manually on model residuals generated by this model, the test suggests no significant serial autocorrelation at p-value 0.87. The test MAPE is 12.68 %. By looking at figure 2(a) and at the results of both KPSS test and ADF test, it is evident that the performance of the search algorithms used in auto.arima is suboptimal. Due to mis-specification of the model, the AIC obtained is not minimum and it performs relatively poor on the test data.

B. *Dataset 2: Real-life Retail Data*

We proceed by showcasing how the existing automated solutions tend to fail with real life seasonal data. Figure 1(b) presents Quarterly Sales data of a particular department of food section of Walmart U.S. The data has a non-linearity present which is evident from the slightly U-shaped form.

We execute the proposed algorithm on this data by specifying y_t as the Online Sales, $f = 4$, $V = 0.33$ We have following two explanatory variables as input all the hypothesis testing were carried out at 5% significance level.

1. x_{1t}: Real GDP of USA

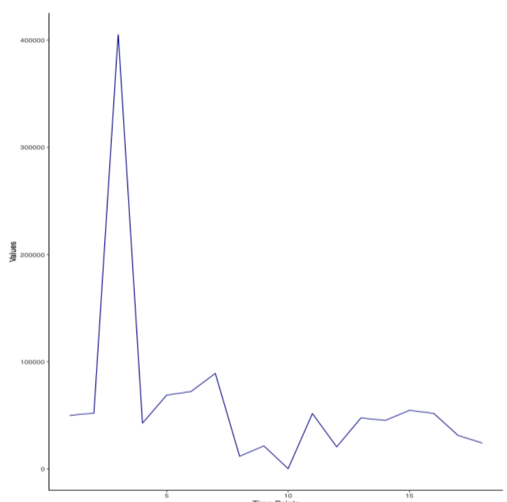
2. x_{2t}: Unemployment rate of USA

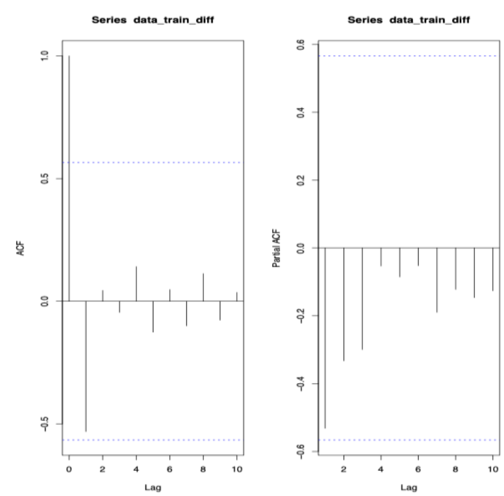
Canova-Hansen test is applied which suggested optimum $D = 0$ with $p - value = 0.17$. Afterwards, both KPSS test and ADF are test applied and both suggested optimum $d = 2$. ACF and PACF were then computed on the differenced series, w_t.The correlograms are presented in figure 2(b).

Optimum p, q, P, and Q are generated followed by fitting the ARIMA model and performing the Ljung-Box test on the model residuals. The output is $p = 2, q = 1, d = 2, P = 0, D = 0, Q = 0, R = 1, AIC = 302.67$. The optimum values of p, q, P, and Q match with the occurrence of spikes in the ACF and PACF of the series presented in fig. 2(b). The test MAPE is 18.67%.

For comparison with the existing process, we implement auto.arima on same train and test data which gives $p = 0, q = 0, d = 0, P = 0, D = 0, Q = 0, AIC = 381.43$. We notice once again, auto.arima fails to produce model with correct parameter values. Model residuals have insignificant autocorrelation according to Ljung-Box test. The test MAPE is 35.39%.

We note here that since auto.arima doesn't perform extensive tests for p, q and hence misses the mark while detecting the non-linearity mentioned previously. This is alleviated our algorithm and consequently,





better results are obtained.

C. Dataset 3: Simulated Data with Noise

Figure 3(a): Plot of Dataset 3

Figure 3(b): ACF, PACF in Dataset 3

A time series is simulated to showcase how the proposed algorithm is performing better when the series is noisy. The simulated data is quarterly in nature.

The data is presented in figure 3(a). An unusually high peak at time point 2 and small trough at time point 10 indicate that the series is noisy and has significant outliers present. We executed the proposed algorithm on this series with no exogenous variables at $V = 0.33\ and\ f = 4$. All the tests are done at 5% significance level.

First, Canova-Hansen test is executed which gave optimum $D = 0$ with $p - value = 0.33$. Following this, KPSS test and ADF test are executed. Here, the ADF test suggested optimum $d = 1$ and KPSS test suggested optimum $d = 0$, it threw warning to the user that '*ADF and KPSS results are different*' and went on to consider optimum $d = 1$. ACF and PACF are then computed on the differenced series, w_t. The correlograms are shown in figure 3(b).

Optimum $p, q, P,$ and Q are found out based on the values of ACF and PACF of the differenced series and ARIMA model is then fitted. This is followed by performing Ljung-Box test on the residuals generated. The output is $p = 2, d = 1, q = 1, P = 0, D = 0, Q = 0, R = 1, AIC = 320.67$. The optimum values of the parameters of the model match with the plots of ACF and PACF of the differenced series presented in fig. 3(b). The test MAPE is 29.52%.

For comparison purpose, we also implemented auto.arima on same train and test data which produced the output $p = 0, q = 0, d = 0, O = 0, Q = 0, D = 0, AIC = 340.03$. The test MAPE is 91.52% which is magnanimous and indicates that auto.arima fails miserably to read the irregularities present in the data.

To diagnose the reason behind this bump in performance, we must note that the proposed algorithm takes into account the behaviour of ACF and PACF of the series which makes it aware of the outliers present in the data. The resultant optimum value of $p = 2$ dictates that the predicted value of series at any time point depends on the values of past two points and hence the effect of the outlier will be subdued. Adding yo this, the coefficient attached to two AR components are 0.57 and 0.45, both of them less than unity in absolute terms. This implies that, although the occurrence of outlier at time point 10 goes on to affect the predicted values from time point 11 onwards, the effect of the outlier declines significantly as we move far from time point 10. Effectively, this means that when we consider test data from time point 14 onwards, the effect of outlier had already started getting dampened to the point that the predicted values do not get affected much. This results in high accuracy of the forecasts produced.

On the other hand, auto.arima, in order to reduce AIC without considering ACF and PACF of the series produces a model with AR and MA components equal to zero and a non-zero mean. It fails to match the crests and troughs present in a data. The estimated value of mean is obviously affected by the presence of outlier and is equal to 71766.15. It ends up predicting same value of the time series in both train and test data at all time points which are equal to 71766.15. As a result, it fares poorly in terms of MAPE on test data as well in terms of AIC.

The examples described above proves how the existing automated solutions get model specifications wrong at times due to not taking in to account the results of both ADF and KPSS test and the behaviour of ACF and PACF of the time series. The proposed algorithm outperforms these solutions in terms of both model AIC and test MAPE while taking same amount of computational time. It handles the noise as well as non-linearity in the data much better and produce best ARIMA models.

VI. CONCLUSION

Authors would like to reiterate here that the proposed methods will only be successful where ARIMA is able to capture the variability in the data. ARIMA, despite being statistically coherent, suffers from certain limitations. Petrica et al [18] had discussed in detail how ARIMA fares poorly in financial data. This is due to its inability to capture heteroscedasticity. In case of small data, ARIMA fails to capture a seasonal pattern which makes it data hungry and its dependency on assumption of normality of data makes it impractical.

That being said, time series is an indispensable part of research in several domains and seasonal ARIMA is one of the most common models used in all such works. However, as a result of the generality that SARIMA models offer parameter tuning becomes tiresome. As we discussed here, all previous efforts have been focused on optimising these parameters based on one or two criteria. Hence, in practice almost always the practitioner resorts to iterative optimisation through parameter space. We attempt to remedy that by introducing a series of tests in a need statement. As can be understood heuristically at each step we urge to fulfil one of the criteria of a good ARIMA fit as described in section III.A. The superiority of the process over the existing processes is also evident in the examples we have provided. Based on this and several other experiments the authors have conducted, we can conclude not only our model provides optimal parameter tuning resulting in better performance when executed on test data, but also the output parameters are more realistic and meaningful.

REFERENCES

[1] G. E. P. Box and Gwilym M. Jenkins, Time series analysis: forecasting and control, Holden-Day, 1970.

[2] C. C. Holt, "Forecasting seasonals and trends by exponentially weighted moving averages," *International Journal of Forecasting,* vol. 20, no. 1, pp. 5-10, 2004.

[3] R. Hyndman and Y. Khandakar, "Automatic Time Series Forecasting: The forecast Package for R," *Journal of Statistical Software,* vol. 26, 2007.

[4] H. Akaike, "A new look at the statistical model identification," *IEEE Transactions on Automatic Control* , vol. 19, no. 6, pp. 716 - 723, 1974.

[5] E. Hannan and J. Rissanen, "Recursive estimation of mixed autoregressive-moving average order," *Biometrika,* vol. 69, no. 1, pp. 81-94, 1982.

[6] V. Gomez and Maraval, "Statistical and Econometrics Software: TRAMO and SEATS," *Statistical and Econometrics Software,* 1996.

[7] G. Melard and J. Pasteels, "Automatic ARIMA modeling including interventions, using time series expert software," *International Journal of Forecasting,* vol. 16, no. 4, pp. 497-508, 2000.

[8] L.-M. Liu, "Identification of seasonal ARIMA models using a filtering method," *Communications in Statistics - Theory and Methods,* vol. 18, no. 6, pp. 2279--2288, 1989.

[9] R. L. Goodrich, "The Forecast Pro methodology," *International Journal of Forecasting,* vol. 16, pp. 533-535, 2000.

[10] R. L. Goodrich, "FOREX: A Time Series Forecast-ing Expert System," *Proceedings of Fourth International Symposium on Forecasting,* 1984.

[11] S. Makridakis and M. Hibon, "The M3-Competition: results, conclusions and implications," *International Journal of Forecasting,* vol. 16, no. 4, pp. 451-476, 2000.

[12] D. Reilly, "The AUTOBOX system," *International Journal of Forecasting*, vol. 16, no. 4, pp. 531-533, 2000.

[13] S. Hwang, M. Park, H.-S. Lee and H. Kim, "Automated Time-Series Cost Forecasting System for Construction Materials," *Journal of Construction Engineering and Management*, vol. 138, no. 11, pp. 1259-1269, 2012.

[14] A. Amin, L. Grunske and A. Colman, "An automated approach to forecasting QoS attributes based on linear and non-linear time series modeling," in *27th IEEE/ACM International Conference on Automated Software Engineering*, 2012.

[15] F. Canova and B. Hansen, "Are Seasonal Patterns Constant over Time? A Test for Seasonal Stability," *Journal of Business & Economic Statistics*, vol. 13, no. 3, pp. 237-52, 1995.

[16] D. Dickey and W. A. Fuller, "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root," *Econometrica*, vol. 49, no. 4, pp. 1057-72, 1981.

[17] D. Kwiatkowski, P. C. Phillips and P. Schmidt, "Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root: How Sure Are We That Economic Time Series Have a Unit Root?," Cowles Foundation for Research in Economics, Yale University, 1991.

[18] A.-C. Petrică, S. Stancu and A. Tindeche, "Limitation of ARIMA models in financial and monetary economics," *Theoretical and Applied Economics*, vol. 23, no. 4, pp. 19-42, 2016.