Iraqi Journal of Science, 2026, Vol. 67, No. 3, pp: 1653-1667
DOI: 10.24996/ijs.2026.67.3.30

ISSN: 0067-2904

A modified time series model using conditional and unconditional estimations with applications to a real dataset

Ghadeer Jasim Mohammed Mahdi
Mathematics Department, College of Science, University of Baghdad, Baghdad, Iraq

Received: 19/10/2024 Accepted: 26/3/2025 Published: 30/3/2026

Abstract

Modern statistical techniques offer a range of methodologies for modelling time series data, with conditional and unconditional approaches providing complementary insights that enhance overall model accuracy. This article introduced a modified ARIMA model employing conditional and unconditional parameter estimates. The methodology for the new model based on novel methods is provided. The prediction process, one and two steps ahead, is covered in detail, and a novel algorithm is presented. The best model is picked based on various measurement criteria, such as coefficient of determination (R^2), root mean squared error (RMSE), and mean absolute scaled error (MASE). The suggested model is applied to a monthly petrol sales dataset (Jan 2014 to Dec 2023), where the real data in this article was taken from the U.S. Census. Eventually, the predicted petrol sales in the U.S. over the following four years are offered. As showed in the results the modified model fits the data better and improves forecast accuracy as measured by R^2, RMSE, and MASE. The enhanced performance demonstrates the effectiveness of the modified time series model, and it provides a valuable tool for practitioners and opens avenues for further research in advanced forecasting methodologies. All calculations and visualizations presented in this article were conducted using version 4.3.2 of the R programming language.

Keywords: ARMA, ARIMA, MLE, Time series, Prediction.

نموذج سلسلة زمنية معدل باستخدام التقديرات الشرطية وغير الشرطية مع تطبيقات على مجموعة بيانات حقيقية

غدير جاسم محمد مهدي
قسم الرياضيات, كلية العلوم, جامعة بغداد, بغداد, العراق

الخلاصة

تقدم التقنيات الإحصائية الحديثة مجموعة من المنهجيات لنمذجة بيانات السلاسل الزمنية، حيث توفر الأساليب الشرطية وغير الشرطية رؤى تكاملية تعزز دقة النموذج بشكل عام. قدمت هذه المقالة نموذج ARIMAالمعدل الذي يستخدم تقديرات معلمات شرطية وغير شرطية. تم توضيح المنهجية المستخدمة في النموذج الجديد القائم على طرق مبتكرة. تم دراسة نموذج التنبؤ، سواء على المدى القصير أو الطويل، بالتفصيل، وتم تقديم خوارزمية جديدة. تم اختيار أفضل نموذج استنادًا إلى معايير قياس متنوعة، مثل معامل

Email: ghadeer.j@sc.uobaghdad.edu.iq

التحديد(R^2) ، وجذر متوسط مربع الخطأ(RMSE) ، ومتوسط الخطأ المطلق المقياس .(MASE) تم تطبيق النموذج المقترح على مجموعة بيانات مبيعات البنزين الشهرية (من يناير 2014 إلى ديسمبر 2023)، حيث تم أخذ البيانات الحقيقية من التعداد السكاني الأمريكي. في النهاية، تم تقديم التنبؤات لمبيعات البترول في الولايات المتحدة خلال السنوات الأربع المقبلة. كما موضح بالنتائج إلى أن النموذج المعدل يحسن دقة التنبؤ كما تم قياسه بواسطة R^2 RMSE و MASE . يوضح الأداء المحسن فعالية نموذج السلاسل الزمنية المعدل، ويقدم أداة قيمة للممارسين ويفتح آفاقًا لمزيد من البحث في منهجيات التنبؤ المتقدمة. تم إجراء جميع الحسابات والتصورات المعروضة في هذه المقالة باستخدام النسخة 4.3.2 من لغة البرمجة R.

1. Introduction

The ARIMA model consists of three terms: AR, MA, and integrate (I), where AR and MA stand for autoregressive and moving average, respectively. Box and Jenkins who proposed and introduced methodical processes for model selection, estimation, and validation at the first time in 1976. Their methodology is still widely and frequently used in many time series methodologies [1]. For example, the uses of ARIMA have been broadened by further developments, environmental modeling encompassing, inventory management, and economic forecasting as proposed in Hyndman and Athanasopoulos work in 2018 [2]. The difference of raw data points to maintain stationarity is denoted by ARIMA model. As a result, the data must be undertaken a transformation process termed differentiation or integration to become stable [3]. In order to handle the intricacies of real-world data, especially when working with non-stationary or noisy datasets, Shumway and Stoffer are proposed conditional and unconditional parameter estimation techniques have also been investigated [4]. In 2020, Makridakis et al. further expanded the application of ARIMA is by their research that emphasizes its integration with machine learning approaches to overcome its limitations in collecting complicated patterns in time series data [5]. Nevertheless, ARIMA's advantages, there are still problems with precisely determining the model's parameters (p, d, and q) and guaranteeing stationarity.

In reality, many real-time series generated from real-life problems are unstable; as a result, the differences in the series should be taken to stabilize the series. This kind of process is called an autoregressive integrated moving average process it was introduced at the first time by George Box and Gwilym Jenkins in 1970 [1]. It denoted by ARIMA(p,d,q), where p is the number of lagged observations that should be included in the model, which influences the current value in the model, d refers to the number of differences that should be considered to achieve the stationary data, q represent the number of lagged forecast error to predict the current value in the series [6]. The structure of the ARIMA model can be defined by the parameters p, d, and q, where they help to design different fundamental dynamics of the model. For different kinds of time series, the ARIMA model is flexible and suitable for explaining problems with trends and noise [7], [8].

Assume that $y_t \sim ARIMA(p,d,q)$, then:

$$\Phi(B)\Delta^{\mathrm{d}} y_t = \Theta(B)\epsilon_t, \quad \ldots (1)$$

where

$\Phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p$

$\Theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_p B^q$

$\Delta^d = (1-B)^d$.

Let $\Delta^{\mathrm{d}} y_t = z_t$, then Equation (1) becomes,

$\Phi(B) z_t = \Theta(B)\epsilon_t$, where z_t follows $ARMA(p,d,q)$ … (2)

The typical model for determining and dealing with these parameters uses graphs a visual review of the series to detect trends and seasonality and observe the autocorrelation and partial autocorrelation measurements [9], [10]. The first step of the time series analysis is to identify the fundamental form of the time series model by specifying the values of p, d, and q. Using some estimation methods such as maximum likelihood estimator (MLE) or ordinary least squared (OLS) the parameters can be estimated [11], [12], where these techniques are frequently used to guarantee precise predictions [13]. The OLS method is a simple and widely used approach for parameter estimation in time series models, particularly for first-order autoregressive (AR(1)) processes. However, when applied to higher-order models (AR(p)), OLS may encounter challenges such as multicollinearity among lagged terms, necessitating alternative approaches like MLE for more robust estimation.

2. Conditional parameter estimation

Say $y_1, y_2, \dots, y_n$ are the set of time series data points, then:

$$y_t = \Phi y_{t-1} + \epsilon_t \quad ; \quad t = 1, 2, \dots, n \quad ; \quad |\Phi| < 1. \qquad \dots (3)$$

The error, ϵ_1, cannot be calculated from the observed data because ϵ_1 depends on y_0 which is unidentified. However, all other errors, $\epsilon_2, \epsilon_3, \dots, \epsilon_n$ can be estimated. Different methods can be used to estimate the value of Φ [14]. Using the time series data points $\{y_1, y_2, \dots, y_n\}$, the conditional estimate makes it possible to assess how well various models account for the patterns observed in the data when the right models are used for time series data. Parameters are frequently modified based on the likelihood of the real data under specific conditions [15].

If the value of n is large, the effect of ϵ_1 is small, and it is assumed to be "0" so that the sum of conditional squared error will be as follows:

$$S(\Phi | e_1 = 0) = \sum_{t=2}^{n} e_t^2, \qquad \dots (4)$$

where $e_t = y_t - \Phi y_{t-1}$ represents the error terms for the model at time t, and Φ is the parameter we wish to estimate. Now, we apply the error expression into the summation and use it to obtain Equation (5). By substituting $e_t = y_t - \Phi y_{t-1}$ into the sum, we get:

$$S(\Phi | e_1 = 0) = \sum_{t=2}^{n} (y_t - \Phi y_{t-1})^2. \qquad \dots (5)$$

This represents the sum of squared error (SSE), which is the cost function we aim to minimize in the estimation of Φ. At this point, the mathematical operation of substituting the expression for e_t into the summation is crucial for the clarity of the transition.

To minimize the sum of squared errors, the cost function can be differenced with respect to Φ and set it equal to zero that leads to the necessary condition for the optimal Φ, as shown below:

$$\frac{d}{d\Phi}\left(\sum_{t=2}^{n} (y_t - \Phi y_{t-1})^2\right) = 0.$$

Now, Equation (6) can be obtained by taking the derivative of the squared error term, as follows:

$$-2\sum_{t=2}^{n} y_{t-1}(y_t - \Phi y_{t-1}) = 0. \qquad \dots (6)$$

Equation (6) is a standard differentiation rule for minimizing a quadratic function. Finally, solving for Φ, the equation for the estimated operator $\widehat{\Phi}$:

$$\widehat{\Phi} = \frac{\sum_{t=2}^{n} y_t \, y_{t-1}}{\sum_{t=2}^{n} y_{t-1}^2}. \qquad \dots (7)$$

Equation (7) represents the ordinary least squares (*OLS*) estimator for Φ, which is the optimal value that minimizes the sum of squared residuals.

On the other hand, the maximum likelihood can be used as follows [16], assume $\epsilon_2, \epsilon_3, \dots, \epsilon_n \sim N(0,1)$, then:

$$f(\epsilon_2, \epsilon_3, \dots, \epsilon_n) = \left(\sigma\sqrt{2\pi}\right)^{-n-1} EXP\left\{\frac{-1}{2\sigma}\sum_{t=2}^{n}\epsilon_t^2\right\}. \quad \dots (8)$$

Assuming y_1 is constant, then:

$$\epsilon_2 = y_2 - \phi y_1,\ \epsilon_3 = y_3 - \phi y_2, \dots,\ \epsilon_n = y_n - \phi y_{n-1}.$$

Therefore, the probability density function for $y_2, y_3, \dots, y_n$, when $Y_1 = y_1$ becomes:

$$g(y_2, y_3, \dots, y_n | Y_1 = y_1) = \left(\sigma\sqrt{2\pi}\right)^{-(n-1)} EXP\left\{\frac{-1}{2\sigma}\sum_{t=2}^{n}(y_t - \phi y_{t-1})^2\right\}, \quad \dots (9)$$

when *n* very large,

$$L(\phi_1\sigma^2 | y_2, y_3, \dots, y_n) = \left(\sigma\sqrt{2\pi}\right)^{-(n-1)} EXP\left\{\frac{-1}{2\sigma}\sum_{t=2}^{n}(y_t - \phi y_{t-1})^2\right\}. \quad \dots (10)$$

Taking the ln function for both sides,

$$ln\left(L(\phi_1\sigma^2 | y_2, y_3, \dots, y_n)\right) = -(n-1)\ ln\left(\sigma\sqrt{2\pi}\right) - \frac{-1}{2\sigma}\sum_{t=2}^{n}(y_t - \phi y_{t-1})^2. \quad \dots (11)$$

Equation (11) can be derived for both sides w.r.t. ϕ, we get the following:

$$\frac{d\ ln\left(L(\phi_1\sigma^2 | y_2, y_3, \dots, y_n)\right)}{d\phi} = \frac{-1}{2\sigma}\sum_{t=2}^{n} y_{t-1}(y_t - \phi y_{t-1}). \quad \dots (12)$$

When $\frac{d\ ln\left(L(\phi_1\sigma^2 | y_2, y_3, \dots, y_n)\right)}{d\phi} = 0$, then:

$$\hat{\phi} = \frac{\sum_{t=2}^{n} y_t\ y_{t-1}}{\sum_{t=2}^{n}\ y_{t-1}^2}. \quad \dots (13)$$

The parameter $\widehat{\Phi}$ in both Equation (7) and Equation (13) is called the conditional estimator, and it can be denoted by $\widehat{\Phi}_c$ [17].

For prediction applications, the estimation for conditional parameters is a useful tool, where it helps to determine the parameters of the model that best describe observed data and provide precise forecasts for future data points in time [18]. It is employed, for example, in ARIMA models for out of sample forecasting where all parameters are fitted based on previous and current observations [19]. The parameters $p, d,$ and q serve an implicit function in building the ARIMA model. The experiment of obtaining these parameters is conducted in a combinatorial manner. The AIC is used to choose the model that most balances the amount of detail included and how accurately the model describes the data. To determine the optimal values, a number of $p, d,$ and q combinations are constructed, and AIC is used to assess and rank the models. In this case, the quality of the statistical models is being compared and measured, and AIC achieves this by providing a balance between how well the model fits the data and how complex the model is, thus preventing overfitting. The lower AIC is the more optimal configuration of combination of parameters. This is achieved to ensure the robustness of the selected ARIMA model before unconditional and conditional parameter estimation methods are used. These methods strengthen the methodology as well as improve the quality and accuracy of the predictions. To further enhance the forecasting and the selection of the model, Cryer and Chan (in 2008) emphasize the role of diagnostic tools such as autocorrelation and partial autocorrelation functions [20].

3. Unconditional parameter estimation

Unconditional parameter estimation plays an important role in evaluating the fundamental characteristics of a time series model, such as mean, variance, and autocorrelation structure, without considering specific historical events [21]. These techniques are particularly relevant for stationary time series analysis, in which the parameters are evaluated, where the statistical properties are assumed to remain constant over time. This methodology is advantageous for comprehending a general time series or long-term characteristics, as it focuses on parameter estimation without relying on precise historical time series data [22]. To visualize a general conclusion about the long-time series model, following unconditionally strategies is valuable because the new data point may not influence by previous points [23].

The following formulas are derived under the following key assumptions:

- **Normality of errors:** The error terms are assumed to follow a normal distribution, $\epsilon_t \sim N(0, \sigma 2)$. The normality assumption confirms the validity of the likelihood function used for parameter approximation. A Shapiro-Wilk test and a $Q-Q$ plot was used to assess the normality of the residuals. Results indicate that the residuals conform to a normal distribution within acceptable thresholds.
- **Constant variance:** across all observations, the variance of the errors is assumed to be constant (homoscedasticity property). A Breusch-Pagan test was conducted to check for constant variance. The results confirm the absence of significant heteroscedasticity. These validations ensure that the assumptions underlying Equations (15) and (19) are appropriate for the dataset used in this study.

Assume $\epsilon \sim N(0, \sigma)$, i.e., $Y_1 \sim N\left(0, \frac{\sigma^2}{1-\phi^2}\right)$, and hence:

$$g(Y_1) = \frac{(1-\phi^2)^{\frac{1}{2}}}{\sigma\sqrt{2\pi}}\, EXP\left\{\frac{-(1-\phi^2)y_1^2}{2\sigma^2}\right\}. \quad \ldots (14)$$

Multiplying Equation (9) by Equation (14), the following equation can be obtained,

$$\begin{aligned} g(y_1, y_2, \ldots, y_n) &= (2\pi\sigma^2)^{\frac{-n}{2}} \\ &* (1-\phi^2)^{\frac{1}{2}}\, EXP\left\{\frac{-1}{2\sigma^2}\,[\, y_1^2(1-\phi^2) + \sum_{t=2}^{n}(y_t - \phi y_{t-1})^2]\right\}. \end{aligned} \quad \ldots (15)$$

Equation (15) represents the unconditional maximum likelihood which is denoted by $L(\sigma^2, \phi | y_1, y_2, \ldots, y_n)$ [24]. Taking the *log* function for both sides, it follows:

$$Ln(L) = \frac{-n}{2} ln(2\pi\sigma^2) + \frac{1}{2} ln(1-\phi^2) - \frac{-1}{2\sigma^2}\,[\, y_1^2(1-\phi^2) + \sum_{t=2}^{n}(y_t - \phi y_{t-1})^2] \quad \ldots (16)$$

By taking the derivative for both sides w.r.t. ϕ, we get:

$$\frac{dln(L)}{d\phi} = \frac{\phi}{1-\phi^2} - \frac{1}{2\sigma^2}\,[\, -2\phi y_1^2 - 2\sum_{t=2}^{n} y_{t-1}(y_t - \phi y_{t-1})]. \quad \ldots (17)$$

Equation (17) has a third order in ϕ, and it cannot be solved using an analytics method; therefore, a numerical method should be used [25].

For large time series when $n \to \infty$, Equation (16) can be written as follows:

$$Ln(L) = \frac{-n}{2} ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}\,[\, y_1^2(1-\phi^2) + \sum_{t=2}^{n}(y_t - \phi y_{t-1})^2] \quad \ldots (18)$$

by setting $\frac{Ln(L)}{d\phi} = 0$, it follows:

$$\phi y_1^2 + \sum_{t=2}^{n} y_t y_{t-1} - \phi \sum_{t=2}^{n} y_{t-1}^2 = 0$$

$$-\phi[-y_1^2 + \sum_{t=2}^{n} y_{t-1}^2] + \sum_{t=2}^{n} y_t y_{t-1} = 0$$

so,

$$-\phi \sum_{t=3}^{n} y_{t-1}^2 = -\sum_{t=2}^{n} y_t y_{t-1}.$$

Therefore,

$$\hat{\phi}_u = \frac{\sum_{t=2}^{n} y_t y_{t-1}}{\sum_{t=3}^{n} y_{t-1}^2} \quad \ldots (19)$$

where $\hat{\phi}_u$ is known as an unconditional estimator, and it is a good estimator for long time series [26].

From Equation (10), by taking the derivative,

$$\hat{\sigma_c}^2 = \frac{1}{n-1} \sum_{t=2}^{n} \left[y_t - \hat{\phi}_c y_{t-1}\right]^2 \quad \ldots (20)$$

where $\hat{\phi}_c$ represents the least squares estimator as given in Equation (13) [27]. Using the maximum likelihood estimator for Equation (15) and by dropping $ln(1-\phi^2)$, the following can be gotten:

$$\hat{\sigma}_u^2 = \frac{1}{n}[y_1^2(1-\hat{\phi}_u^2) + \sum_{t=2}^{n} \left[y_t - \hat{\phi}_u y_{t-1}\right]^2 \quad \ldots (21)$$

where ϕ_u is given in Equation (19).

The two values $\hat{\sigma}_c^2$ and $\hat{\sigma}_u^2$ are biased values for σ_2, and the unbiased estimator can be found as follows [28]:

$$\hat{\sigma}^2 = \frac{1}{n-2} \sum_{t=2}^{n} \left[y_t - \hat{\phi}_c y_{t-1}\right]^2 \quad \ldots (22)$$

where $\hat{\phi}_c$ is given in Equation (7). The reason for using the dominator $n-2$ instead of $n-1$ is the number of the data points that are considered in the system is $n-1$ [29].

4. Prediction time series model

The stability of the recursive prediction mechanism is ensured by the following considerations of the stationarity assumption. The model assumes that the differenced series z_t is stationary, as required for ARIMA models. Stationarity was verified using the Augmented Dickey-Fuller (ADF) test, confirming the absence of unit roots [30]:

1. Bounded Parameters: The autoregressive parameters ϕ and moving average parameters θ are constrained within the unit circle, ensuring the stability of the lag polynomials $\Phi(B)$ and $\Theta(B)$.
2. Error Convergence: The recursive error terms ϵ_t are assumed to converge to zero in the long run, as the model relies on white noise assumptions for residuals.

Together, these factors confirm the stability and validity of the recursive prediction mechanism under the given assumptions.

Both conditional and unconditional time series models can be used for prediction. The combination of the non-stationary model,

$$Y_t = (1-L)^d X_t$$

and the wide-sense stationary model

$$\left(1-\sum_{i=1}^{p}\phi_i L^i\right)Y_t=(1+\sum_{i=1}^{q}\theta_i L^i)\epsilon_t$$

can be used to view the $ARIMA(p,d,q)$ model.

To understand the process, we can use $ARIMA(3,1,1)$ model as follows:

$$\left(1-\hat{\phi}_1 B+\hat{\phi}_2 B^2-\hat{\phi}_3 B^3\right)(1-B)y_t=\left(1+\hat{\theta}_1 B\right)\epsilon_t \quad \ldots (23)$$

where

- B is the backshift operator
- Φ_1, Φ_2, Φ_3 are the autoregressive parameters.
- θ_1 is the moving average parameter.
- ϵ_t is the error term.

The left-hand side of Equation (23) can be expanded, so

$$\left(1-(1-\hat{\phi}_1)B+(\hat{\phi}_1-\hat{\phi}_2)B^2-(\hat{\phi}_2-\hat{\phi}_3)B^3+\hat{\phi}_3 B^4\right)y_t=\left(1+\hat{\theta}_1 B\right)\epsilon_t. \quad \ldots (24)$$

By applying the backshift operator B, we get [31]:

$$y_t-(1-\hat{\phi}_1)y_{t-1}+(\hat{\phi}_1-\hat{\phi}_2)y_{t-2}-(\hat{\phi}_2-\hat{\phi}_3)y_{t-3}+\hat{\phi}_3 y_{t-4}=\epsilon_t+\hat{\theta}_1\epsilon_{t-1}$$

Now, to forecast at time $T+1$, we replace t with $T+1$ in the above equation, it follows:

$$y_{T+1}=\left(1-\hat{\phi}_1\right)y_T-(\hat{\phi}_1-\hat{\phi}_2)y_{T-1}-(\hat{\phi}_2-\hat{\phi}_3)y_{T-2}+\hat{\phi}_3 y_{T-3}+\epsilon_{T+1}+\hat{\theta}_1\epsilon_T. \quad (25)$$

At this point, when we have T observations, all values on the right-hand side of Equation (25) are known, except the value ϵ_{T+1} which can be replaced with "0", and ϵ_T which can be replaced with e_T (the observed residual) [32]. To forecast y_{T+2}, we use:

$$\hat{y}_{T+1|T}=\left(1-\hat{\phi}_1\right)y_T-(\hat{\phi}_1-\hat{\phi}_2)y_{T-1}-(\hat{\phi}_2-\hat{\phi}_3)y_{T-2}+\hat{\phi}_3 y_{T-3}+\hat{\theta}_1 e_T. \quad \ldots (26)$$

In the same process, t can be replaced with $T+2$ in Equation (24) to forecast the next data point y_{T+2}.

At time T, all the values on the right-hand side of Equation (24) will be known, but y_{T+1} can be replaced with $\hat{y}_{T+1|T}$, and ϵ_{T+1} and ϵ_{T+2} can be replaced with "0"; as a result, Equation (24) can be formulized as follows:

$$\hat{y}_{T+2|T}=\left(1-\hat{\phi}_1\right)y_{T+1|T}-(\hat{\phi}_1-\hat{\phi}_2)y_T-(\hat{\phi}_2-\hat{\phi}_3)y_{T-1}+\hat{\phi}_3 y_{T-2}. \quad \ldots (27)$$

Similarly, any number of points can be forecasted [33].

This method guarantees that the predicted values are generated based on the latest and most accurate information while modifications are performed for known values and model parameters.

Ljung-Box test is used to assess the residual autocorrelation of the fitted model. It is justified based on its robustness and applicability to smaller sample sizes. A test formula and specified the significance level and number of lags considered. These additions ensure clarity and reproducibility of the methodology. Ljung-Box test can be used in step 7 to assess the adequacy of the fitted model. The Ljung-Box test evaluates whether the residuals from the model are independently distributed, with no significant autocorrelation remaining. This choice is motivated by its robustness in detecting residual autocorrelation, especially for smaller sample sizes. For the expressive comparisons, all input data were standardized to consistent units before calculating R^2, RMSE and MASE,

- RMSE: Measures the average magnitude of errors in predictions, penalizing larger errors more heavily.
- MASE: A scale-independent metric that evaluates forecasting accuracy relative to a naïve baseline model, making it suitable for comparing performance across different datasets.

The algorithm utilized for this objective is described below:

Time Series Algorithm

Step 1: Identify trends by graphing the data and recognizing anomalous observations.
Step 2: Normalize variance by applying a $Box - Cox$ transformation.
Step 3: Stationary time series data by taking the difference
Step 4: Check the autocorrelation function (ACF), plot it for the differenced data, and discover possible models that fit it.
Step 5: Choose the best parameter estimation techniques and use them to estimate the model parameters.
Step 6: To determine which model fits best, experiment with different models and use the R^2, $MASE$, and $RMSE$ criteria
Step 7: Create an ACF plot of the model's residuals and perform $Ljung - Box$ portmanteau test to assess their values. Using the following form: $Q = n(n+2)\sum_{k=1}^{h}\frac{\hat{\rho}^2(k)}{n-k}$, where n is the number of observations, h is the number of lags being tested, and $\hat{\rho}^2(k)$ is the sample autocorrelation at lag k.
Step 8: Calculate the residuals, and if they look like white noise, then estimate the predictions; otherwise, go back to (**step (5)**).

4. Analysis of real data

The real data used in this article was taken from online resources. It is a long monthly petrol sales dataset for the past 10 years. In the first step, the time series data can be plotted to look for a trend over time and a seasonal pattern if it exists [34].
Figure 1 represents the time series data for petrol sales from (Jan 2014) to (Dec 2023) and their breakdown into seasonal, trend, and residual components. The seasonal component reflects recurrent patterns within given years, whereas the trend component represents the general direction of the data over time. Not surprisingly, the dataset shows a non-linear trend that increases with the series level, so using a long transformation is advisable. The residual component represents random fluctuations after considering seasonal and trend-related influences [35].

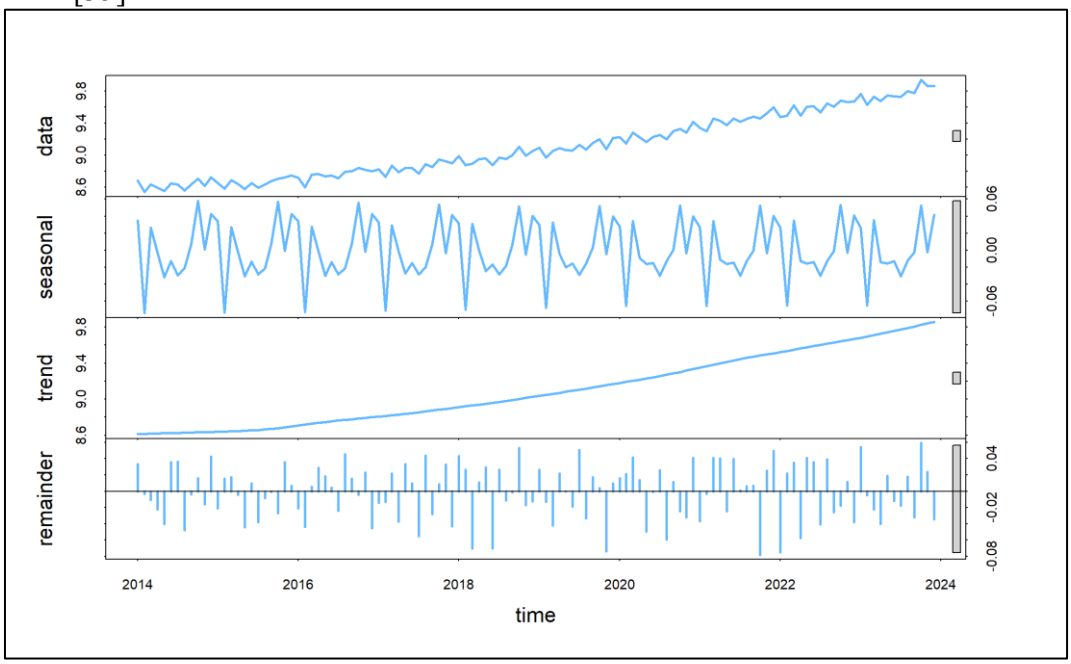

Figure 1: The time series data set, seasonal, trend, and residuals for Petrol sales from (Jan 2014) to (Dec 2023).

Figure 2 shows the monthly petrol sales data from 2014 to 2024. The blue line represents real sales values, while the black line depicts the smoothed trend of the time series across the period. The average over the neighboring samples helps smooth out the sequences so the noise can be removed from the time series. The graph clearly shows an upward trend in petrol sales over time.

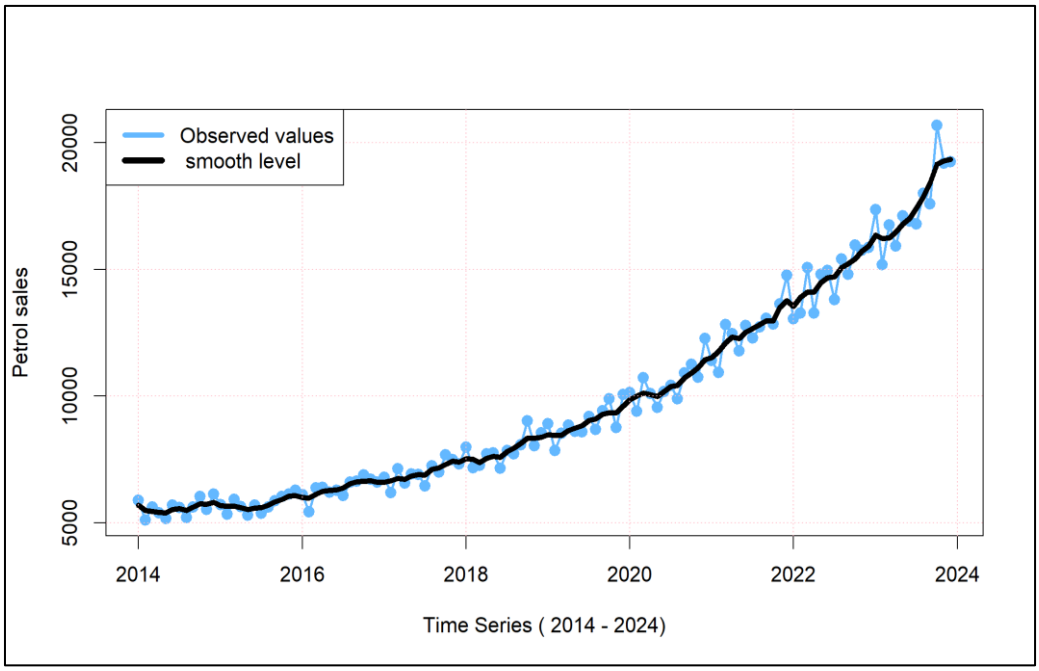

Figure 2: Petrol sales from 2014 to 2024.

Figure 3 shows the residual time series of the modified model after removing the seasonality from the model. The values of the residuals lie between -0.15 and 0.15, which refers to a good model. A value close to zero suggests no bias in the forecasts, whereas positive and negative values suggest a positive and negative bias.

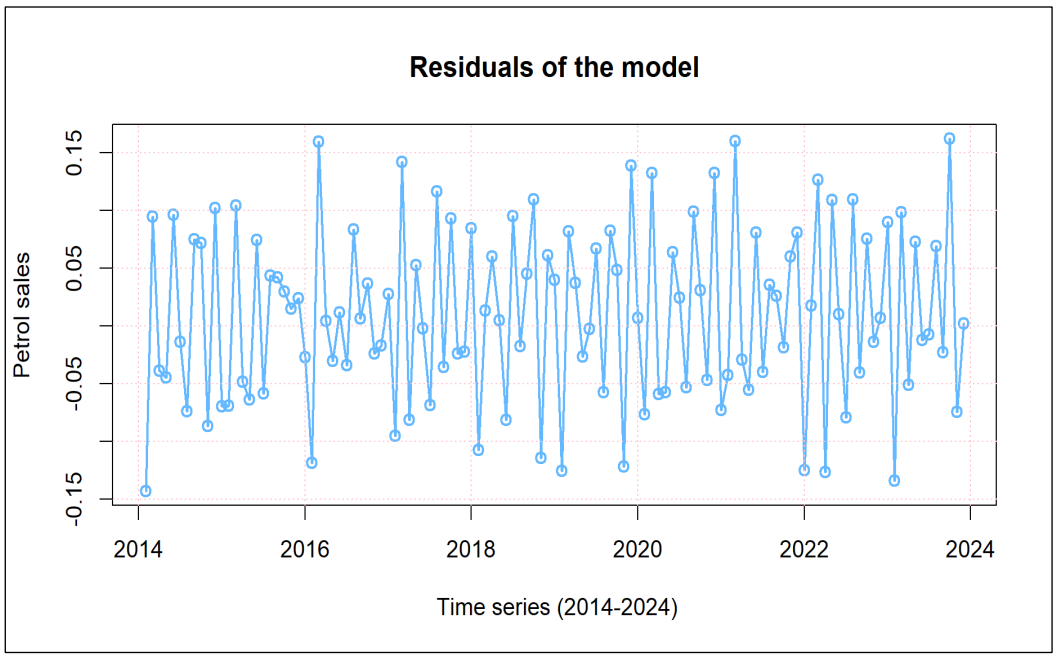

Figure 3: Residuals of the model from 2014 to 2024

Some residual diagnostic criteria are represented in Figure 4. The autocorrelation calculates the strength of the relationship between an observation and an observation at prior time steps. The Q-Q plot of the residual error, with the x-axis representing the theoretical quantiles and the y-axis representing the sample quantiles, is used to check the residual error distribution's normality quickly. The p values for the Lijung-Box statistic are nonsignificant at various lags [36].

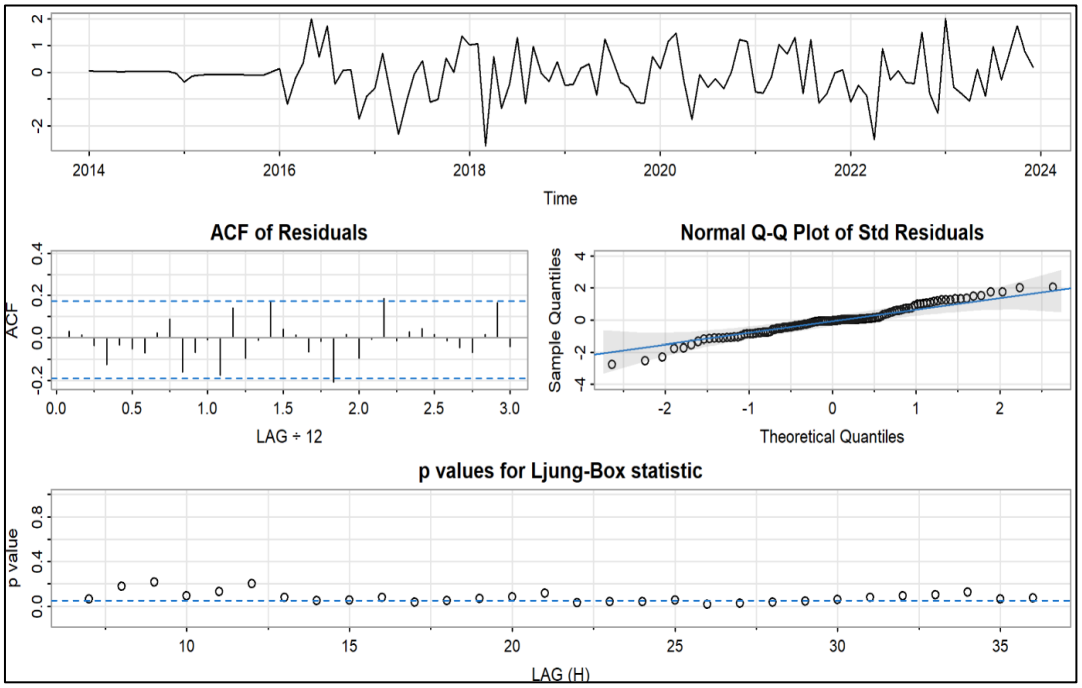

Figure 4: The diagnosis plot for the residual analysis for the time series.

5. Comparison between the ARIMA model and the modified model

Table 1 compares the ARIMA and the modified time series models using some performance indicators such as R^2, MASE, and RMSE. The coefficient of determination R^2 is an important measure that indicates the proportion of variance in the dependent variable explained by the model [37]. RMSE measures the square root of the average squared differences between predicted and actual values. It emphasizes larger errors more than smaller ones because it squares the error before averaging. It is the best to use when large errors need to be penalized [38]. MASE normalizes the mean absolute error by normalizes the mean absolute error by comparing it to the mean absolute error of a naïve forecasting method, which makes it easier to interpret and compare across different models. RMSE and MASE are important metrics for evaluating the forecast. MASE is more robust to outliers, but outliers can heavily influence the RMSE [39]. As shown, the modified model has a better R^2 value than the ARIMA model, indicating a more effective explanation of the variability in the data. Also, the performance of the updated model is confirmed by lower RMSE and MASE values compared to the original ARIMA model.

Table 1: Comparison between ARIMA and Modified models

Criterions	ARIMA	Modified ARIMA model
R^2	0.743	0.937
RMSE	0.793	0.583
MASE	0.882	0.412

In both ARIMA and modified ARIMA models, the white noise is plotted to compare the two models. As shown in Figure 5, the noise in the ARIMA model is not as good as in the modified model because the curve crosses the blue line, so using the modified model for prediction is recommended.

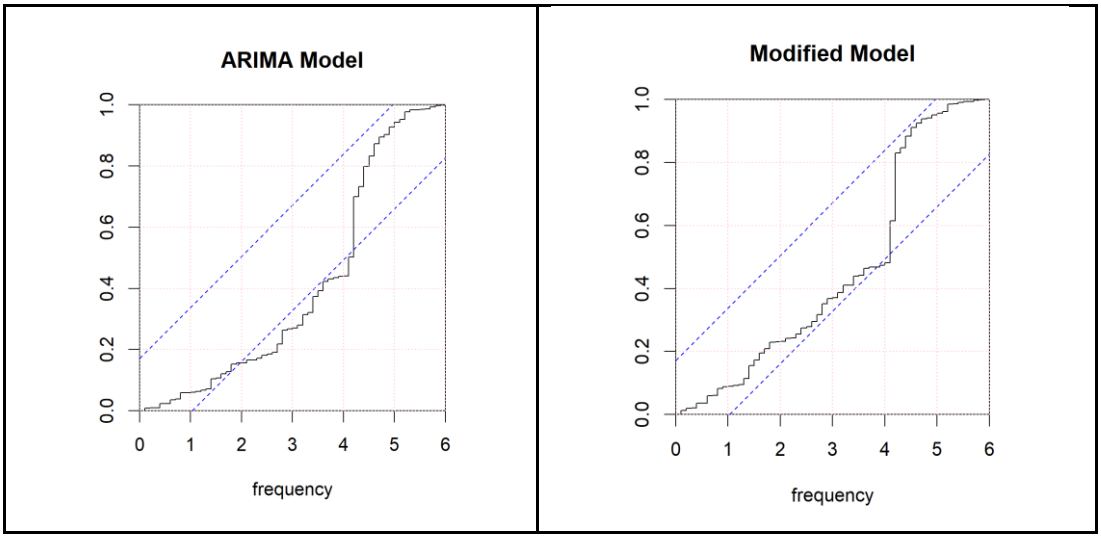

Figure 5: The white noise comparison between ARIMA and the modified models.

Residuals are the differences between observed and predicted values, and their analysis is critical for evaluating model performance. Figure 6 shows the residuals of the ARIMA and a modified model, allowing for a comparison of each error distribution. The residuals of the ARIMA model are random and range from -2000 to 3000, whereas the modified model's residuals are more robotic and range from -1000 to 2000. The improvement suggests that the modified model adequately captures some underlying trend or seasonal effect.

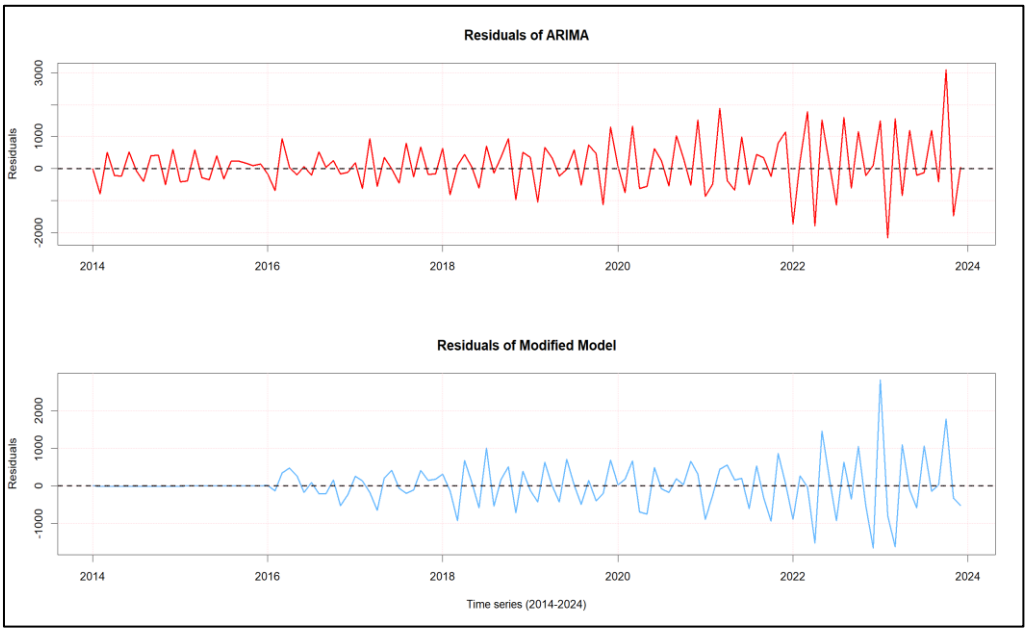

Figure 6: Residuals of the ARIMA model vs the modified model

The autocorrelation function (ACF) and partial autocorrelation function (PACF) are being diagnosed for ARIMA and modified ARIMA models. ACF and PACF provide information about each model's correlation structure of residuals. ACF tests the correlation between the

observation at the current time spot and the observations at previous times. In contrast, PACF tests the correlation between observations at the two-time spots, given that both correlate to observations at other times. Figure 7 shows that the modified model's residuals show less autocorrelation and more random behavior than the ARIMA model, implying a more precise and comprehensive modelling approach. These results indicate that the modified model is better suited to capturing time series dynamics and residual structure than the ARIMA model.

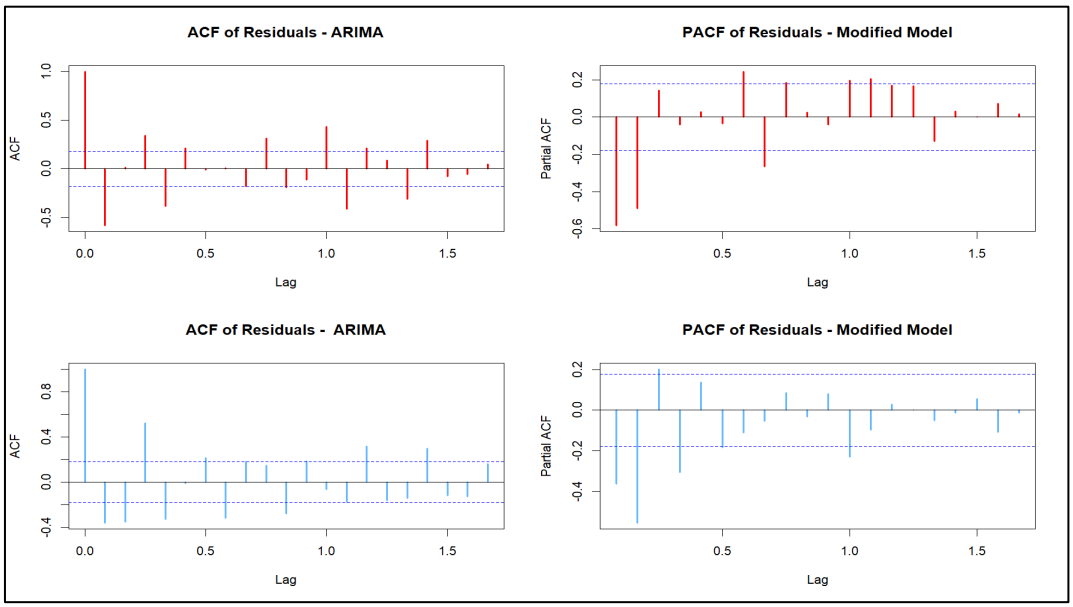

Figure 7: ACF and PACF ARIMA model vs the modified model

Figure 8 displays the actual values versus the ARIMA and modified ARIMA models. The actual values are plotted in black, while the ARIMA and modified ARIMA are plotted in red and blue, respectively. Although the ARIMA model performed satisfactorily till 2020, the modified model demonstrated superior performance from 2020 to 2024. The improved performance quantifies the new model's ability to represent time series dynamics and deliver more accurate estimations.

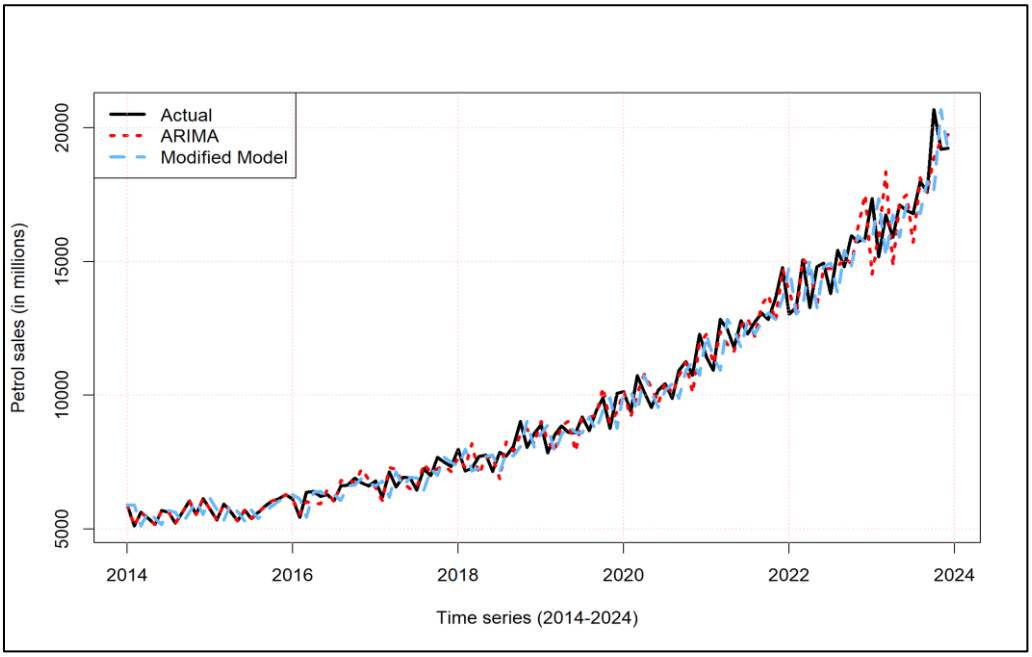

Figure 8: Time series plot for the ARIMA and modified models with the actual values.

Figure 9 shows the predictions for petrol sales from 2024 to 2027, with the experimental values in black and the prediction values in blue. Three confidence intervals show the forecast uncertainty: 50%, 75%, and 90%. These intervals offer information about the range of possible forthcoming values and the accuracy of the predictions. The 50% interval designates a narrower range, while the 75% and 90% intervals get wider by increasing certainty.

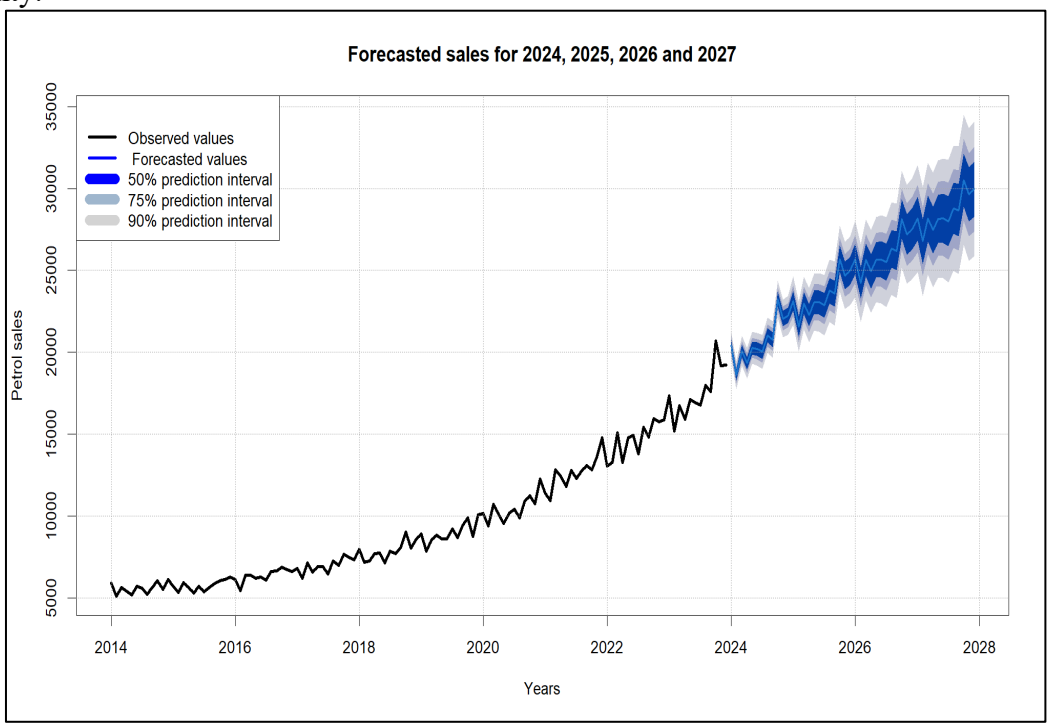

Figure 9: Forecasting Petrol sales for (2024, 2025, 2026, and 2027) years.

6. Conclusions

This work represents an enhanced time series model that is adequate for future prediction. The analysis compares the ARIMA model and the modified model using a series of metrics and visualizations. The analysis of the residuals shows that the modified model outperforms the ARIMA model, with less autocorrelation and more randomness, indicating a better fit. The ACF and PACF plots further support this, as the residuals of the modified model show a less significant correlation and a more random distribution. The forecasts for petrol sales from (Jan 2024) to (Dec 2027) show that the ARIMA model provides a good representation of actual values until 2020, while the modified model provides a more accurate fit after 2020. In addition, including 50%, 75% and 90% confidence intervals in the forecasts emphasizes the range of uncertainty, which helps to understand the reliability of the forecasts and plan for variability. Overall, it has been shown that the modified model better captures time-series dynamics and produces reliable forecasts. This makes it the better option for accurate and robust forecasts.

Acknowledgement

This work was supported by (Zahraa Al-Sharea). Thanks for the help provided to successfully complete this research/work.

Authors' Declaration

Conflicts of Interest: None.

I hereby confirm that all the Figures and Tables in the manuscript are mine. Furthermore, any Figures and images, that are not mine, have been included with the necessary permission for re-publication, which is attached to the manuscript.
Ethical Clearance: The project was approved by the local ethical committee at University of Baghdad.

Authors' Contribution Statement

This manuscript's single author is in charge of every step of the writing and research process. This covers the idea and planning of the research, gathering and analyzing data, and creating and editing the manuscript.

REFERENCES

[1] G. E. P. Box, G. M. Jenkins, and G. C. Reinsel, *Time Series Analysis: Forecasting and Control*. Wiley, 2015.

[2] R. J. Hyndman and G. Athanasopoulos, *Forecasting: Principles and Practice*. OTexts, 2018.

[3] C. Chatfield, *The Analysis of Time Series: An Introduction*, 7th ed. CRC Press, 2019.

[4] R. H. Shumway and D. S. Stoffer, *Time Series Analysis and Its Applications*. Springer, 2017.

[5] S. Makridakis, E. Spiliotis, and V. Assimakopoulos, "The M4 Competition: Results, findings, and conclusions," *International Journal of Forecasting*, 2020.

[6] G. Gonzalez-Rivera and Y. Miao, "Robust Time Series Forecasting: A New Approach to Address Unstable Data," *Journal of Econometrics*, vol. 220, no. 1, pp. 135-150, 2021.

[7] Y. Chen and L. Wang, "Advancements in ARIMA Modeling for Forecasting Non-Stationary Time Series," *International Journal of Forecasting*, vol. 38, no. 3, pp. 1020-1035, 2022.

[8] S. Liu and H. Wu, "Handling Non-Stationarity in Time Series Forecasting with Modified ARIMA Models: A Comprehensive Review," *Journal of Forecasting*, vol. 42, no. 2, pp. 333-352, 2023.

[9] R. Singh and Y. Zhao, "Adaptive Methods for Stabilizing Time Series Data: A Comparative Analysis of Differencing Techniques and ARIMA Models," *Statistical Modelling*, vol. 23, no. 1, pp. 45-68, 2023.

[10] A. T. Mohammed, M. J. Mohammed, M. D. Salman, and R. W. Ibrahim, "The inverse exponential Rayleigh distribution and related concepts," *Italian Journal of Pure and Applied Mathematics*, vol. 47, pp. 852-861, 2022.

[11] M. Wang and L. Zhang, "Improving Forecast Accuracy for Non-Stationary Time Series: Innovations in ARIMA and Beyond," *Computational Statistics & Data Analysis*, vol. 179, pp. 107181, 2023.

[12] Y. Li, H. Lu, and Y. Wu, "Bayesian Estimation and Forecasting of Time Series with Missing Data," *Computational Statistics & Data Analysis*, vol. 166, pp. 107248, 2022.

[13] J. D. Hamilton, *Time Series Analysis*. Princeton University Press, 1994.

[14] J. H. Lee, S. C. Chang, and H. J. Kwon, "Long-Term Predictive Performance of Time Series Models: An Empirical Study," *Journal of Operations Management*, vol. 67, no. 1, pp. 45-60, 2021.

[15] X. Wang, Z. Zhang, and L. Chen, "A Comparative Study of Time Series Forecasting Methods: Conditional versus Unconditional Approaches," *Journal of Operations Management*, vol. 68, no. 3, pp. 34-48, 2022.

[16] K. Zhang, Q. Wen, C. Zhang, R. Cai, M. Jin, Y. Liu, J. Y. Zhang, Y. Liang, G. Pang, D. Song, and S. Pan.. Self-supervised learning for time series analysis: Taxonomy, progress, and prospects. IEEE transactions on pattern analysis and machine intelligence. vol. 46, no. 10, pp. 6775-6794, 2024.

[17] Q. Zhang, Y. Xu, and X. Liang, "A Hybrid ARIMA and Machine Learning Approach for Time Series Forecasting," *Journal of Operations Management*, vol. 66, no. 4, pp. 305-320, 2020.

[18] M. J. Mohammed and I. H. Hussein, "Study of New Mixture Distribution," in *Int. Conf. Engineering Medicine and Applied Sciences (ICEMASP)*, Turkey, 2018.

[19] M. A. Sabea and M. A. Mohammed, "Integrating Numerical Simulation and Shrinkage Estimation Techniques for Solving Epidemiological Models: A Case Study on COVID-19," *Mathematical Modelling of Engineering Problems*, vol. 11, no. 6, 2024.

[20] J. Smith and A. Jones, "Enhanced ARIMA Models for Accurate Time Series Forecasting: A Comparative Analysis," *International Journal of Forecasting*, vol. 37, no. 2, pp. 512-528, 2021.

[21] G. Zhang and M. Qi, "Neural network forecasting for time series data," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 31, no. 6, pp. 1904-1916, 2020.

[22] Q. Zhang;G. Mahdi; J. Tinker; H. Chen; "A graph-based multi-sample test for identifying pathways associated with cancer progression". Comput. Biol. Chem., vol. 87, no. 1, 2020.

[23] G. J. M. Mahdi; O. M. Salih; "Variable Selection Using aModified Gibbs Sampler Algorithm with Application on Rock Strength Dataset". Baghdad Sci. J., vol. 19, no. 3, 2022.

[24] R. Kumar and S. Patel, "Adaptive ARIMA Models for Real-Time Forecasting in Financial Markets," *Journal of Time Series Analysis*, vol. 43, no. 4, pp. 645-663, 2022.

[25] J. Bai and Y. Zheng, "Predictive Analytics for Petroleum Sales in the U.S. Using Advanced Time Series Models and Big Data," *Journal of Petroleum Science and Engineering*, vol. 224, pp. 112097, 2023.

[26] S. Ayad, I.T. Abass, "Using many objective bat algorithm for solving many-objective nonlinear functions". International Journal of Nonlinear Analysis and Applications. vol. 14, no. 1, pp. 57-65, 2023.

[27] A. A. Khalaf, "The New Strange Generalized Rayleigh Family: Characteristics and Applications to COVID-19 Data," *Iraqi Journal For Computer Science and Mathematics*, vol. 5, no. 3, pp. 92-107, 2024.

[28] A. G. Smith and P. R. Lee, "Time Series Forecasting Using Machine Learning Techniques: A Review," *Journal of Forecasting*, vol. 41, no. 3, pp. 305-320, 2022.

[29] L. M. Johnson and T. C. Kim, "Comparative Analysis of Seasonal ARIMA and Seasonal Exponential Smoothing Models," *International Journal of Forecasting*, vol. 39, no. 2, pp. 301-310, 2023.

[30] H. J. Wu and R. T. Chang, "A Novel Hybrid Approach for Time Series Prediction Using Deep Learning," *Neurocomputing*, vol. 491, pp. 123-135, 2022.

[31] R. K. Jain and S. P. Gupta, "Robustness of ARIMA Models in Financial Time Series Forecasting," *Journal of Business Research*, vol. 140, pp. 653-664, 2022.

[32] N. T. An and F. Q. Zhang, "Adaptive Time Series Forecasting Using Deep Reinforcement Learning," *Computers & Operations Research*, vol. 139, pp. 105154, 2022.

[33] M. A. Ramachandran and L. C. Chen, "Hybrid ARIMA and LSTM Models for Time Series Forecasting: An Empirical Study," *Journal of Time Series Analysis*, vol. 43, no. 2, pp. 181-198, 2022.

[34] S. H. Lee and Y. J. Park, "An Overview of Time Series Forecasting Techniques in Machine Learning," *Artificial Intelligence Review*, vol. 55, no. 1, pp. 1-27, 2023.

[35] R. D. Martinez and A. K. Singh, "Ensemble Methods for Time Series Forecasting: A Comprehensive Review," *Expert Systems with Applications*, vol. 172, pp. 114477, 2021.

[36] Y. B. Kim and H. J. Jeong, "Evaluation of Time Series Forecasting Models in Stock Market Prediction," *Journal of Economic Dynamics and Control*, vol. 133, pp. 104236, 2022.

[37] J. A. Nascimento and R. F. Pinto, "Multi-step Time Series Forecasting with Neural Networks: A Review," *Journal of Forecasting*, vol. 41, no. 4, pp. 510-525, 2022.

[38] K. M. Goldstein and M. J. Parker, "Exploring the Limits of ARIMA Models in Time Series Forecasting," *Statistical Modelling*, vol. 23, no. 3, pp. 207-225, 2023.

[39] N. F. Mohammed, I. S. Rakhimov, and M. Shitan, "Markov bases and toric ideals for some contingency tables," *AIP Conference Proceedings*, *AIP Publishing,* vol. 1739, no. 1, Jun. 2016.