Charting the Unseen: Transformative Approaches in Time Series Forecasting in Real World Domains

Sanjeev Kumar [1*], Karandeep Saluja[1], Shashwat Agrawal [1], Dilkeshwar Pandey[2] and Saurav Chandra[2]
[1]School of Computer Sciences, UPES, Dehradun, Uttarakhand, 248007, India.
[2]CSE Dept., KIET Group of Institutions, Ghaziabad, India.
Contact email: sanjeev.kumar@ddn.upes.ac.in*, karandeepsaluja73@gmail.com, shashwat.agrawal0906@gmail.com, dilkeshwar.pandey@kiet.edu ,saurav.chandra@kiet.edu,
Corresponding author – Sanjeev Kumar, *e-mail : sanjeev.kumar@ddn.upes.ac.in

Abstract

This paper addresses the growing complexity of forecasting in an era where data volumes are expected to reach 180 zettabytes by 2025. It aims to bridge the gap between the theoretical aspects of time series analysis and their practical applications in fields like finance, healthcare, and environmental studies. The research covers foundational concepts such as Autocorrelation and White Noise and spans various methodologies from traditional models like ARIMA to advanced techniques involving machine learning. Special attention is given to the challenges of applying these theories to real-world, often irregular or incomplete data. The paper also explores the integration of technologies like AI in forecasting, emphasizing the need for robust and interpretable models. Concluding with a call for greater academia-industry collaboration, it suggests new research directions for innovative, practical forecasting solutions in a data-intensive world.

Keywords: Time series decomposition, Exponential smoothing, ARIMA models, Dynamic regression models

I. Introduction

Nowadays, there's so much data growing fast in all different forms. It's super important to get useful insights and guesses from time series data. This kind of data is all about predicting the future from past data. It's really useful in many areas like finance, business, energy, health, and environment. It helps people make smart choices, figure out where to use resources, and manage risks. This leads to better efficiency, new ideas, and progress. But, it's tough to use the big ideas from time series forecasting in real-life situations where there's a ton of data. This paper tries to really dig into this problem by looking at important topics. By 2025, there's going to be a huge amount of data, like 180 zettabytes (that's a really big number!). This is both exciting and challenging for time series forecasting. Now, we have new kinds of data from places like social media, sensors, and IoT (Internet of Things) gadgets. This data is special, but often it's mixed up, missing pieces, or keeps changing, so it needs good cleaning and organizing. Old but gold methods like ARIMA and smoothing stuff are still liked because they're easy to understand and use. Newer methods like deep learning are really accurate but sometimes hard to get. Putting different models together can make things more accurate and less random. There are also new cool things like Bayesian stuff for dealing with uncertainty and machine learning that's easier to understand. When you have lots and lots of data, you need big models and tools. Things like Spark, which is a computer thing, are getting popular. Automated tools like Auto-TS and TPOT help pick the right model. Knowing a lot about the area you're studying is key to picking out important info. Playing with model settings and knowing when to stop can make them work better. There are real examples where this stuff has worked, like in predicting what people will buy, energy use, and stock prices. It's really important to have models that are strong and make sense. It's also key to understand complicated patterns in the data. Adding in causes and effects can make predictions better. Mixing what you know about the area with machine learning can lead to better and clearer forecasts. More teamwork between people in universities and businesses can help make these big ideas work in real life. Data is everywhere today. It is very important. Forecasting, or predicting the future, is now key in many areas. The paper "Time Series Forecasting: Bridging Theory and Practice in a Data-Intensive World" looks at this important topic. It studies time series data. This type of data is a series of data points in time order. It's very useful for making predictions. This paper has two main goals. The first goal is to explain the theory behind time series analysis. It looks closely at ideas like Autocorrelation and White Noise. These are basic parts of time series data. The paper also looks at different tools and methods used in this field. It covers simple techniques like Judgmental Forecasts and more complex ones like Time Series Regression Models. The second goal is about using these theories in real life. It's one thing to understand these models. It's another to use them well in real situations. The paper talks about how to apply methods like Time Series Decomposition, Exponential Smoothing, ARIMA models, and Dynamic Regression Models. It pays extra attention to forecasting when you have time series data that are linked or in groups. This is becoming more important as data gets more

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

complex. The paper also looks at new forecasting methods. These new methods can help solve harder forecasting problems.

The paper also talks about real-life forecasting problems. These include issues like irregular data, big datasets, and changing data patterns. The introduction prepares us for a full talk on time series forecasting. It will cover both theory and real-world uses. It shows why this topic is important today. At the end, the paper looks at new trends and problems in forecasting. We need models that are strong and easy to understand. We should also think about cause and effect. The paper suggests new research areas. This includes creating new methods and using AI better. It says that academics and industries should work together. This will help make practical forecasting solutions faster. This whole approach is meant to help people who use forecasting. It gives them new ideas. This helps in many areas, like business and science.

II. Literature Survey

Autocorrelation is a tool used in time series data. It checks how similar data points are over time [2]. This helps in finding patterns and links in the data. There are methods like the Durbin-Watson statistics and the Ljung-Box test to measure this [2]. Granger introduced a concept called Granger causality [3]. It uses past data to predict future data. This shows how things are linked over time [3]. Autocorrelation is key for understanding and making models of time series data. It's used a lot in forecasting. White noise in time series data is like a bunch of random points that don't link up. It stays the same over time [4]. It's used to test other data against. Gardner talked a lot about smoothing techniques in time series data [4]. These techniques work well even with white noise. The Ljung-Box test, by Ljung and Box, checks if data has white noise or not [5]. White noise helps to see if a model's predictions are useful. There are many tools and techniques for forecasting. Hyndman and Athanasopoulos gave a lot of info on how to choose and use these tools [1]. They say it's important to look at different factors and pick the right one for your data. Makridakis and Hibon used the M3 test [7]. This test checks how well different methods predict at various times and with different data. Their work shows it's important to match the method to the data.

Judgmental forecasts use expert opinions and data together [8]. Armstrong wrote about how combining these two can improve forecasts [8]. He gives tips on how to gather and use expert advice. Goodwin and Fildes talked about the role of human judgment in forecasting [9]. They discuss its problems and how to make it better. Time series regression models mix time series data with other factors to get more accurate predictions [10]. Montgomery et al. explained how to make and test these models [10]. Wei gave a complete guide on time-varying and multiple factor analysis, including regression models [11]. Decomposing time series data means breaking it into main parts [12]. This includes trend, seasonal, and error parts. Cleveland et al. used the LOESS (STL) method to separate these parts [12]. They showed how it works for complex data. Shumway and Stoffer gave a detailed review of these decomposition methods [13]. They said it makes it easier to understand and model the data [13].

There are different smoothing methods in time series, like simple, Holt, and Winters' methods. Gardner [14] examined them and discussed their advantages and disadvantages. He talked about how user-friendly they are and how they can be utilized with various kinds of data. Hyndman and Koehler [15] examined the precision of their forecasts. To verify this, they employed metrics such as mean square error (MSE) and mean error (MAE). Time series benefit from the use of ARIMA models. They handle seasonal variations and trends. A thorough description of these models was provided by Kutu et al. [16]. They described the prediction process of ARIMA models. Additionally, Brockwell and Davis [17] discussed ARIMA models. They described ways to recognize, anticipate, and evaluate these models. Dynamic regression models are a bit different. They change over time. This makes them good for situations where things keep changing. Montgomery et al. [18] explained how these models work. They talked about how these models can adjust over time. Hyndman and Athanasopoulos [1] also studied these models. They said these models are flexible for complex problems.

Hierarchical forecasting deals with time series that are linked or in a hierarchy. Syntetos and Boylan [20] looked at how this works in inventory management. They talked about challenges in collecting data and making forecasts. Kourentzes, Barrow, and Petropoulos [21] focused on estimating sales for different products. They discussed how to make these estimates more accurate. Advanced forecasting uses new ideas beyond traditional methods. Chen and Yang [22] looked at using neural networks for forecasting. They talked about how these can spot complex patterns. Makridakis, Spiliotis, and Assimakopoulos [23] discussed how forecasting has changed. They gave advice on picking the right methods based on data and problems. Dealing with missing data and uncertainty is a big challenge. De Gooijer and Hyndman [24] reviewed these issues over 25 years. They talked about the importance of managing missing data. Willemain, Smart, and Schwarz [25] looked at how competition affects forecasts. They also discussed picking the best models. So, time series forecasting covers many ideas and challenges. There's a lot of research and advice out there on this topic. This paper gives a good starting point for anyone interested in learning more about forecasting.

III. Time Series Analysis

3.1 Autocorrelation

Drawing a parallel to correlation, which quantifies the linear association between two variables, autocorrelation assesses the linear dependence between the lagged values of a time series [1].

***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

$$r_k = \frac{\Sigma_{t=k+1}^{T}[(y_t - \bar{y})(y_{t-k} - \bar{y})]}{\Sigma_{t=1}^{T}(y_t - \bar{y})^2}$$

r_1: indicates the autocorrelation between the current observation and just previous observation or lag 1.
r_2: indicates the autocorrelation between the current observation and two period apart previous observation or lag 2. Where r_k: is a autocorrelation between y_t and y_{t-1}.

In essence, autocorrelation reveals the correlation between the present observation and preceding observations at varying lags. Autocorrelation is influenced by both trend and seasonality. In instances where the data exhibits a trend, autocorrelation tends to be positively substantial for shorter lags, gradually diminishing as the lags increase. Regarding seasonality, autocorrelation tends to be more pronounced for seasonal lags compared to non-seasonal lags.

3.2 White Noise

Time series exhibiting no correlation are known as White Noise. This concept is similar to white light containing all colors, White Noise includes all frequencies within its spectrum. It is statistically defined by a series of random values that are independent and identically distributed (i.i.d.). For white noise data, the sampling distribution of \(r_k \) (referenced in equation 1) tends to follow a normal distribution with a mean (\(\mu \)) of 0 and a variance (\(\sigma^2 \)) of \(1/T \), where \(T \) represents the time series length. This indicates that as the time series length increases, the sampling distribution increasingly resembles a normal distribution with a mean of 0 a variance inversely proportional to the time series length.

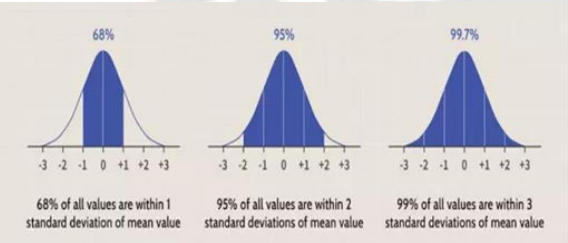

*Figure-1 (*Generalized Normal Distribution)

Figure-1 illustrates the Generalized Normal Distribution. Given that white noise conforms to a normal distribution with $\sigma^2 = 1/T$, approximately 95% of all r_k values for white noise should fall within the range of $-1.96/\sqrt{T}$ to $+1.96/\sqrt{T}$. Deviations beyond this range suggest the presence seasonality or trend in data, indicating that it may not strictly adhere to the characteristics of white noise.

IV. Decomposing Time Series

Time series frequently show a range of patterns, contributing to considerable variability that complicates analysis and prediction. Decomposing a time series into its unique components can greatly aid in understanding and improving forecast precision. Each component of this decomposition represents a specific type of pattern inherent in the data[1].

4.1 Transformations and adjustments

Changing the data helps us understand it better. This is often the first thing to do when breaking down time series data. It's helpful to change the data if it goes up and down a lot. For example, data with big changes at the end, like in power law distributions, is often changed using logs. This makes it easier to understand. When we use logs, changes in the data show up as percentage changes. If the original observations are labelled as $y_1 \dots y_T$, and the transformed ones as $w_1 \dots w_T$, then the transformation is represented as $w_t = \log(y_t)$.

Computations, such as the square or cube root, may be utilised in specific circumstances. "Power transformations" and they may be broadly defined as $w_t = y_p$[1].

Box-Cox transformations are a special kind of change you can make to data [2].

$$w_t = \begin{cases} \log(y_t) & if\ \lambda = 0 \\ \frac{sign(y_t)|y_t|^{\lambda} - 1}{\lambda} & otherwise \end{cases}$$

4.1.1 Classical Decomposition

Three-tiered segmentation of seasonal, trend, and residual data, enables for a more sophisticated and complete data analysis. It enables the identification of trend changing moments as well as the recurrence of seasonal patterns. This type of research is incredibly beneficial in generating more accurate forecasting models.
There are two types of "classical decomposition": "additive decomposition" and "multiplicative decomposition" [1].

For "Additive Decomposition": y_t-S_t
For "Multiplicative Decomposition": y_t / S_t

Adjustments made in time series analysis are customized based on the seasonal period, represented by \(m \) (for example, \(m=4 \) for quarterly data, \(m=12 \) for monthly data, or \(m=7 \) for daily data exhibiting weekly patterns). Traditional decomposition operates under the premise that the seasonal component stays consistent annually. In scenarios involving multiplicative seasonality, the values within the \(m \)-period forming the seasonal component are often referred to as seasonal indices..

4.1.2 Methods used by official statistics agencies

Official statistics agencies bear the responsibility of analysing extensive economic time series data. To effectively analyse these time series, these agencies have crafted their own decomposition techniques tailored for seasonal adjustment. It's noteworthy that these techniques are specifically designed

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

to handle quarterly and monthly data; however, they are not suitable for the analysis of daily, hourly, or weekly data.

X-11 Method

Even though this method is rooted in classical decomposition, it successfully addresses the limitations associated with classical decomposition. [1]

Comparison between the X-11 Method and Classical Decomposition:

Fixed Seasonal Component in Classical Decomposition:

- Drawback: "Classical decomposition" assumes a constant seasonal component over time. This poses challenges when dealing with time series that can exhibit irregular patterns in its seasonal components, such as in the case of "Electricity Consumption Demand."
- X-11 Solution: By including and regulating irregular patterns in the seasonality component, the "X-11 Model" provides a solution. It estimates the seasonal component using the moving average approach. This allows for response to changing seasonal trends.

Lack of Handling Outliers:

- The occurrence and treatment of outliers in time series data are not addressed by "classical decomposition." As a result, when outliers are present, the predicting results might suffer.
- The "X-11 Model" incorporates the issue of the outlier and gives a method for dealing with outliers. It detects and compensates for outlier data points to deliver more accurate decomposition results.

Inadequate Handling of Trading-Day Effects:

- The disadvantage is that "classical decomposition" may fail to account for trading-day effects, which can be significant in economic and financial time series.
- The "X-11 Method" provides procedures for correcting and accounting for trading-day impacts in time series data. This is particularly useful in economic and financial forecasting.

4.1.3 STL Method

Loess Seasonal and Trend Decomposition [1]. It can only result in an additive decomposition; for a multiplicative decomposition, the log of the additive decomposition must be calculated.

Comparisons between the X-11 Method and STL Method: Non-Linear Relationships

- The "X-11 Method" has the disadvantage of being unable of estimating non-linear connections.
- Solution: By estimating non-linear correlations with Loess, the "STL Approach" solves this issue. It is a versatile and dependable method for time series decomposition.

Seasonality:

- The "X-11 Method" is confined to dealing with seasonality in monthly or quarterly data series.
- Solution: In contrast, the "STL Method" is adaptable and capable of dealing with seasonality on a daily, weekly, monthly, or quarterly basis.

More Robustness to the Outlier:

- The "X-11 Method" has the disadvantage of being sensitive to outliers and extreme numbers, which may impact the accuracy of the decomposition.
- STL Solution: "STL Decomposition" estimates components using locally weighted regression (Loess). It reduces the importance of extreme values. It is resilient and stable decomposition in the presence of outliers.

The following images shows the comparison between the above discussed decomposed techniques, the following images are obtained by using tribble (us employment) available under the library fpp3.

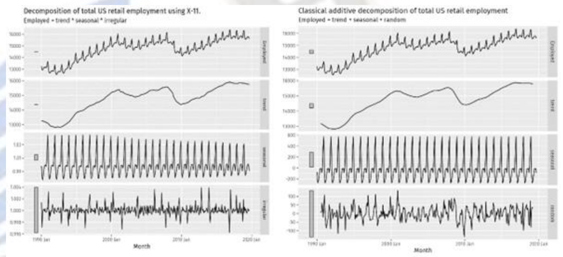

Figure-2: Classical, X-11, STL Decomposition applied to total Retail USA Employment

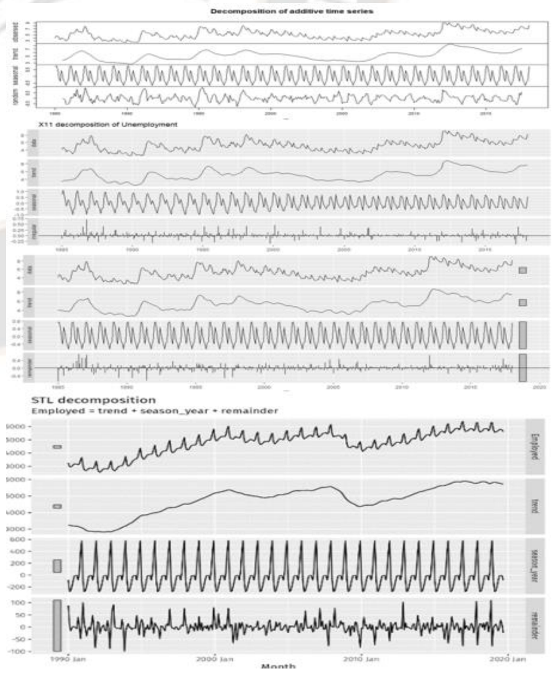

Figure-3: Classical, X-11, STL Decomposition applied to USA Unemployment

Article Received: *25 July 2023* ***Revised:*** *12 September 2023* ***Accepted:*** *30 September 2023*

V. Forecasting Workflow

- ***Data Preparation:***

The initial phase in forecasting entails data preparation, encompassing tasks such as loading the data, identifying missing values, and other pre-processing activities. Various forecasting models come with distinct requirements—some necessitate complete data without missing values, while others demonstrate robustness in the presence of missing values. Therefore, checking and addressing missing values is a crucial step for optimal forecasting.

- ***Plot the Data(visualise):***

It is strongly recommended to display the data to understand the intrinsic aspects, such as trend or seasonality. It helps in detection of common patterns which is essential for selection of a suitable model for study.

- ***Define a model(specify):***

Different forecasting models have their own pros and cons. It's important to choose a model that fits the data well for accurate forecasts.

- ***Model Training (Estimation):***

After choosing the model, the next step is training it with data. The model learns patterns and connections to make future forecasts.

- ***Check the model performance(evaluate):***

To evaluate a time series forecasting model, first split the data into training and test sets. Select metrics like MAE, MSE, RMSE, MAPE, and R^2 to assess performance. Fit the model to the training data, then make and evaluate predictions on the test set. Lower metric values and higher R^2 indicate better performance. Compare results with a basic forecast for context. Examine the forecasts against the actual data and look for unpredictability in the residuals. Perform time series cross-validation and change the model parameters as appropriate. The best model balances statistical performance with the application's specific needs, such as simplicity and computational efficiency.

- ***Generating Forecasts:*** After training, and checking the model, forecasting techniques begin computing. There is need to include all crucial supplementary data in the forecast-focused dataset to create reliable forecasts.

VI. Some Simple Forecasting Methods

Basic forecasting approaches are frequently used as benchmark models for comparison. These benchmarks enable us to assess the performance of more complex models, indicating whether our chosen model outperforms these simpler alternatives. This comparative analysis helps in gauging the effectiveness of our forecasting approach.

- ***Mean Method:***

According to this strategy, the "average" (or "mean") of the provided historical data will be the forecast value for all future values. The following expression can be used to indicate the future forecast values if we represent the provided historical data as $y_1, y_2 ..., y_T$:

$$y_{\widehat{\{T+h|T\}}} = \overline{\{y\}} = {y_1 + \cdots . + y_T}/{T} \text{ [1]}$$

The estimate of y_{T+h} based on the data $y_1, y_2 \ldots, y_T$ can be denoted as $y_{\widehat{\{T+h|T\}}}$

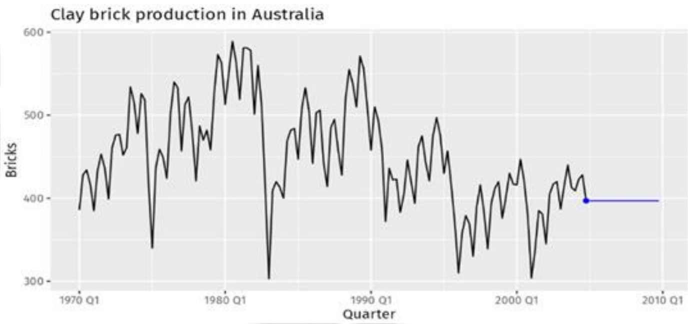

Figure4: Average forecasts are employed for the projection of clay brick production in Australia.

- ***Naïve Method:***

For the "Naïve Method", all forecast values are set to value of the last observation. In other words,

$$y_{\widehat{\{T+h|T\}}} = y_T \text{ [1]}$$

This method works very well with the economical and finical data:

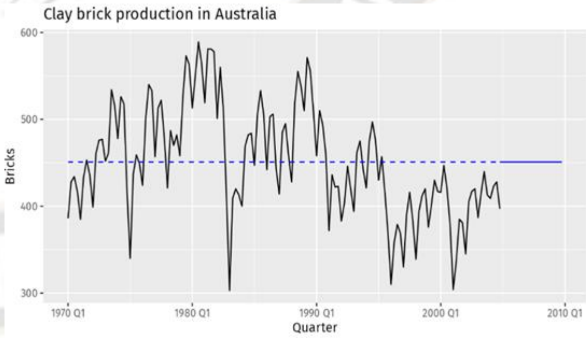

Figure5: Naïve forecasts are utilized for projecting clay-brick production of Australia.

- ***Seasonal Naïve Method:***

Similar to the "Naïve Method," this approach is employed when the data demonstrates a strong seasonal pattern.
In this scenario, the approach involves establishing each forecast to match the last observed value from the corresponding season (e.g., the identical month of the preceding year).

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

$$y_{\{\widehat{T+h|T}\}} = y_{T+h-m(k+1)} \; [1]$$

The forecast for all future values in the formula, where "k" is the integer component of "h-1/m" and "m" is the seasonal period, equals the last observed value for each season. When working with monthly data, for instance, the forecast is set to equal the latest observed value for January for all future January values.

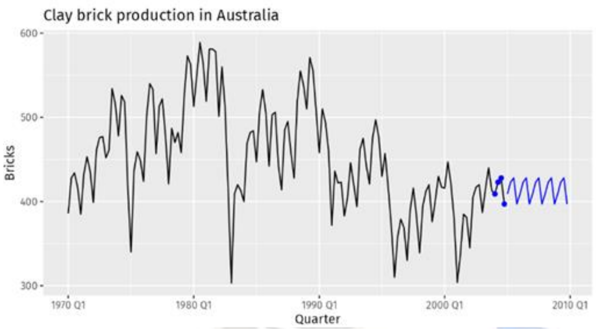

Figure6: For the estimation of clay brick production in Australia, seasonal naïve forecasts are implemented.

- ***Drift Method:***

Permitting prediction values to show a growing or declining trend over time is one way to apply a variation of the "Naïve Method." The average change shown in previous data is used in this method to determine the amount of change over time, or "Drift."

$$y_{\{\widehat{T+h|T}\}} = y_T + \frac{h}{T-1} \sum_{t=2}^{T}(y_t - y_{t-1}) = y_T + h\left(\frac{y_T - y_1}{T-1}\right) \; [1]$$

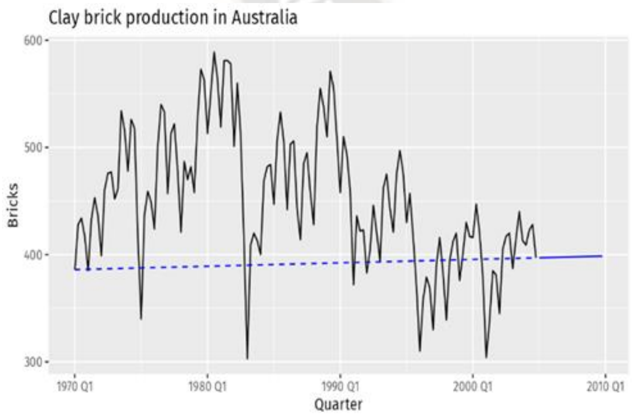

Figure7: Drift forecasts are employed for projecting clay brick production in Australia.

Google's daily Closing Stock Price

In Figure 8, Non-seasonal benchmark models (Mean, Naïve, and Drift) are implemented on Google's Daily Closing Stock Price Data for the year 2015. These models are utilized to forecast one month ahead. It's worth noting that the Seasonal Naïve Model was not applied in this context, as there is no discernible seasonality in the Stock Price Data.

Figure8: Forecasts are generated based on Google's daily closing stock price in the year 2015.

Every now and then, one of these straightforward models can end up being the best option for forecasting. But these approaches are frequently used as standards rather than as the best option.If your selected model outperforms these benchmark models, it is considered valuable and worth using. On the contrary, if your model does not surpass the performance of these benchmark methods, it may not be considered suitable for the forecasting task at hand.

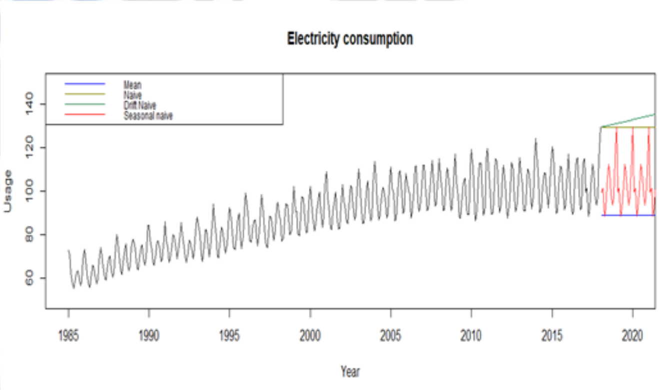

Figure9: Forecasts are generated based on Electricity Consumption in USA

VII. Fitted Values Explanation

In time series analysis, each data point can be predicted using its preceding historical observations. These predictions are known as fitted values, represented as $y_{\{\widehat{t|t-1}\}}$ or simply $\hat{y}_t$. This indicates that the forecast for $\hat{y}_t$ relies on the observations $y_1, \ldots, y_{t-1}$. It's important to understand that fitted values are not actual forecasts because they are computed using parameters estimated from the entire time series, including future data points.

***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

- ***Mean Method - Fitted Values:***

For the Mean Method, fitted values are calculated as $\hat{y}_t = \hat{c}$, where $\hat{c}$ is the average calculated using all observations, even those beyond the time t.

- ***Fitted Values of Drift:***

Fitted Values of "Drift Method" are given by $\hat{y}_t = y_{t-1} + \hat{c}$ where $\hat{c} = \frac{y_T - y_1}{T-1}$. [1]

- ***Fitted Values of Naïve and Seasonal Naïve Method:***

Both methods, "Seasonal Naïve Method" and "Naïve Method", do not involve any parameters in their forecasts. As a result, their fitted values are considered actual or true forecasts since no parameter estimation is employed, making the predictions solely based on observed data.

VIII. Residuals

The "Residuals" in a Time Series are what is left over after fitting the Model. The Residuals are the difference between the observations and corresponding fitted values.

$e_t = y_t - \hat{y}_t$ [1]

If the transformations have been used in the Model, then we should look residuals at the transformed scale, these are called "innovation residuals". Suppose we modelled the logarithms of the data, $w_t = \log(y_t)$, then the innovation residuals are given by $w_t - \hat{w}_t$ whereas regular residuals are given by $y_t - \hat{y}_t$. Residuals are used to check whether the data has captured the information of the data properly or not. For this purpose, we use the innovation residuals. If the patterns are observable in the innovation residuals, the model can be improved.

IX. Residuals Diagnostics

A robust "Forecasting Method" should produce "Innovation Residuals" with the following properties:

1. Uncorrelated "Innovation Residuals" are essential. If there is correlation among them, valuable information remains untapped, urging the need for improvement in forecasting computations.
2. "Innovation residuals" should have a zero mean to avoid bias in forecasts.

Any forecasting method lacking these properties can be enhanced. While these ensure optimal use of information, they are not ideal for selecting a forecasting method. If any property is not met, adjustments can be made. Addressing bias involves adding the mean *m* to all forecasts. Resolving correlation issues will be discussed later.
Additionally, although not necessary, it is beneficial for the residuals to exhibit the following properties:

3. It is beneficial for the "Innovation Residuals" to maintain a constant variance, known as "Homoscedasticity".
4. While not mandatory, it is advantageous for the innovation residuals to approximate a "Normal Distribution".

Forecasting Google Daily Closing Stock Price

Continuing with the example of Google's Daily Closing Stock Price mentioned earlier, for the stock market prices and indexes, the "Naïve Method" is considered the best forecasting approach/Method as compare other simple forecasting Methods. In this method, each projected value value is set equal to $\hat{y}_t = y_{t-1}$ which is the most recent observed value.

Hence the "Residuals" are $e_t = y_t - \hat{y}_t = y_t - y_{t-1}$

Figure10: Predictions are derived from the daily closing stock price of Google in the year 2015.

The Figure 10 illustrates the closing daily stock price of Google in 2015. The notable large jump on July 17, 2015, corresponds to a 16% increase attributed to unexpectedly robust second-quarter results.

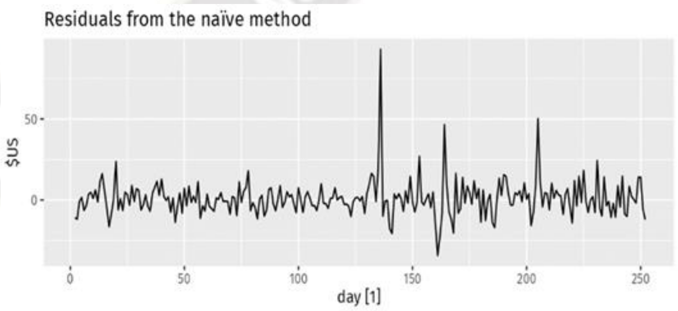

Figure11: Residuals arise from estimating the price of Google's stock using the "Naïve Method".

Figure 11, the "Residuals" obtained from this series using the "Naïve Method". The substantial positive residual is a consequence of the unexpected jump in July.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

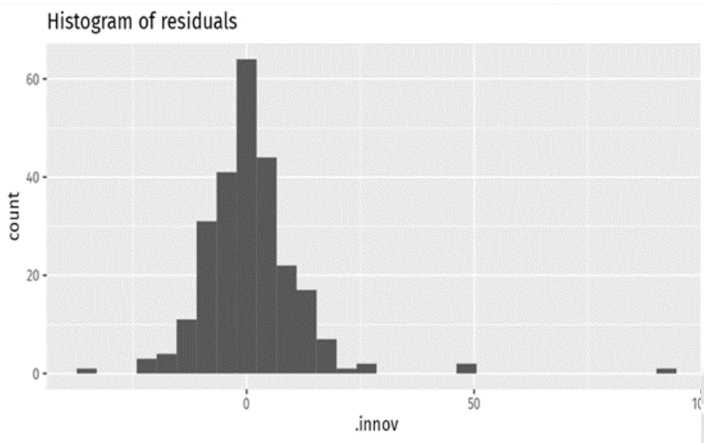

Figure12: A histogram depicting the residuals from applying the "Naïve Method" to forecast the Google stock price reveals that the right tail appears somewhat elongated for a normal distribution.

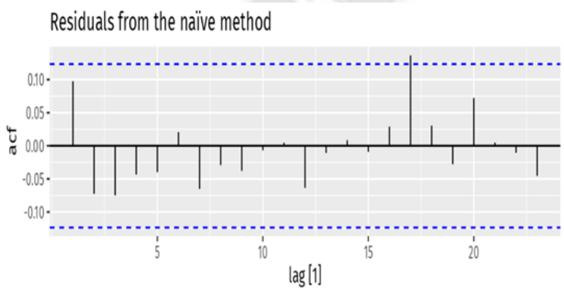

Figure13: The "Autocorrelation Function (ACF)" of the residuals resulting from applying the "Naïve Method" to the Google stock price illustrates a lack of correlation, indicating that the forecasts are sound.

The forecasts produced by the "Naïve Method," which makes use of all the data in the provided time series, are shown in the above graphs. There are no discernible correlation exits in the residual series, and the mean of the residuals for the specified time series is very near to zero.

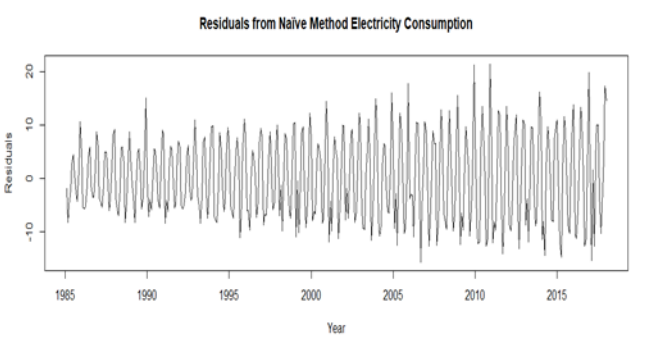

Figure14: Residuals arise from forecasting the Electricity Consumption in USA using the "Naïve Method".

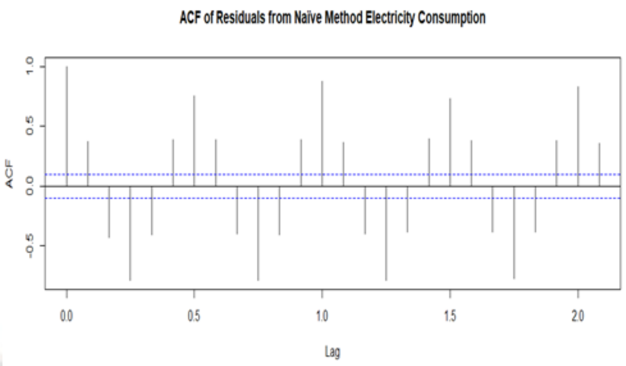

Figure15: The "Autocorrelation Function (ACF)" of the residuals resulting from applying the "Naïve Method" to the Electricity Consumption in USA illustrates a high amount of correlation, indicating that the forecasts are not sound.

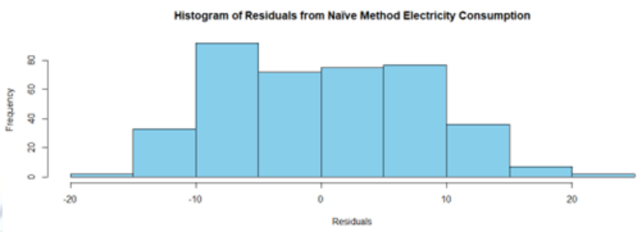

Figure16: A histogram depicting the residuals from applying the "Naïve Method" to forecast the Electricity Consumption reveals that residual does not follow a normal distribution.

The forecasts produced by the "Naïve Method" are shown in the figure 16, which do not fully utilise the information contained in the provided time series. There is a substantial correlation in the residual series, and the mean of the residuals of the provided time series is not close to zero.

X. Prediction Intervals and Distributional Forecasts

- ***Forecast Distributions***

A probability distribution, which characterises the chance of witnessing future values of the provided time series using the fitted model, can be used to quantify forecast uncertainty. The mean of this distribution is represented by the Point Forecast. Notably, the majority of time series fitted models produce projections that are regularly distributed [1].

- ***Intervals of Prediction***

A prediction interval denotes the range that we expect or forecast y_t to fall with a certain probability. It can be expressed as $\widehat{y_{T+h\,|\,T}} \pm c\widehat{\sigma_h}$, where the coverage probability affects c. Prediction intervals are important because they can communicate forecast uncertainty. Point forecasts alone don't reveal the accuracy of forecasts, but by providing prediction

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

intervals, we can communicate the associated uncertainty with each forecast.

- ***Prediction Intervals in One Step***

When predicting one step ahead, the standard deviation of the forecasting error can be estimated using the standard deviation of the residuals, as indicated by the formula.

$$\hat{\sigma} = \sqrt{\frac{1}{T-K-M}\sum_{t=1}^{T}\widehat{e_t^2}}$$

Here, T is the time series' length, K denotes the number of parameters that were calculated using forecasting techniques, and M denotes the quantity of missing values. Since the first observation cannot be predicted, M equals 1 for the forecast in the "Naïve Method."

The prognosis for the next value of the price in the instance of a "Naïve Forecast for the Google Daily Closing Stock Price Data," where the last observed value is 758.88, is also 758.88. According to the equation above, the residuals' standard deviation is 11.19.

Hence the 95 % prediction interval for the next value is

$$758.88 \pm 1.96(11.19) = [736.9, 780.8]$$

Similarly, an 80% prediction interval for the next value is

$$758.88 \pm 1.28(11.19) = [745.5, 773.2]$$

- ***Multiple-Step Forecasting Periods***

Prediction intervals are characterised by their common property of lengthening with an increase in the forecast horizon (h). The prediction intervals go broader the further out we project since there is more uncertainty involved. In other words, h grows with *σ-h*. An estimate of *σ-h* is required in order to generate the prediction intervals, and this can be obtained using the formula in the "One-Step Prediction Intervals".

- ***Benchmarking Techniques***

Assuming uncorrelated residuals, the predicted standard deviation for the four benchmark approaches is as follows:

If $\widehat{\sigma_h}$ denotes standard deviation of the h-step forecast distribution, and $\hat{\sigma}$ is residual standard deviation.
If, $\hat{\sigma}$ is the residual standard deviation and, $\widehat{\sigma_h}$ is the standard deviation of the h-step forecast distribution.

Note h = 1 and T is very large.

Benchmark Methods	h-step forecast Standard Deviation
Mean	$\widehat{\sigma_h} = \hat{\sigma}\sqrt{1 + 1/T}$
Naïve	$\widehat{\sigma_h} = \hat{\sigma}\sqrt{h}$
Seasonal Naïve	$\widehat{\sigma_h} = \hat{\sigma}\sqrt{k+1}$
Drift	$\widehat{\sigma_h} = \hat{\sigma}\sqrt{h(1+h/(T-1))}$

Figure17: Prediction intervals using a naïve technique for the closing stock price of Google, are established at 80% and 95% confidence levels.

When visualized, intervals of prediction are depicted as dim regions, and the colour intensity within these regions corresponds to the associated probability. This graphical representation serves to convey a comprehensive understanding of the uncertainty linked to the forecasts.

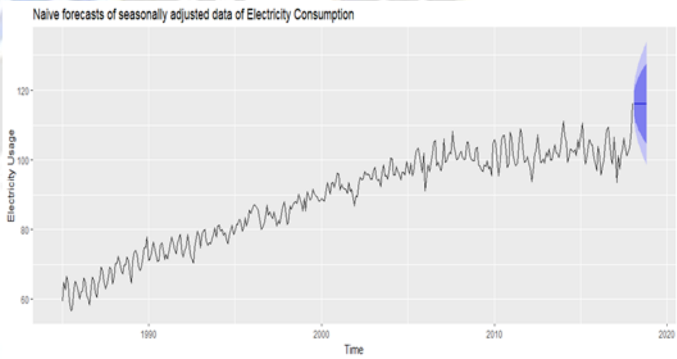

Figure18: Prediction intervals for the Electricity Consumption in USA, based on a naïve method, are established at 80% and 95% confidence levels.

XI. Forecasting using Transformations

Some common transformations which can be used when modelling was discussed early. When forecasting from a model with transformations, we first produce forecasts of the transformed data. Then we need to reverse the transformations to obtain forecasts on the original scale.

For the Box-Cox Transformation given by (), the reverse transformation is given by

$$y_t = \begin{cases} exp(w_t) & if\ \lambda = 0 \\ sign(\lambda_{wt}+1)|\lambda_{wt} + 1|^{1/\lambda} & otherwise \end{cases}$$

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

- ***Intervals of prediction with transformations***

In case a transformation has been applied, the prediction interval is initially calculated on the transformed scale. Subsequently, the endpoints are reverted to the original scale through back transformation, ensuring that the intervals maintain their probability coverage. However, it's important to note that after back transformation, the intervals may no longer retain symmetry around the point forecast.

- ***Bias Adjustments***

One drawback of using the back-transformed point forecast is known as the Box-Cox transformation may no longer represent the projected distribution's mean. Assuming that the forecast distribution is symmetrical on the transformed scale, it is often the median of the forecast distribution.

$$\hat{y}_{T+h\,|\,T} = \begin{cases} exp(\widehat{w}_{T+h\,|\,T})\left[1 + \dfrac{{\sigma_h}^2}{2}\right] & if\ \lambda = 0; \\ (\lambda\widehat{w}_{T+h\,|\,T} + 1)^{1/\lambda}[1 + \dfrac{{\sigma_h}^2(1-\lambda)}{2(\lambda\widehat{w}_{T+h\,|\,T} + 1)^2}] & else \end{cases}$$

The magnitude of the forecast variance directly influences the disparity between the mean and median.

This disparity between the basic back-transformed prediction and (as per equation 5.2) also the mean (as per equation 5.3) is termed bias. When the mean is used instead of the median, it implies that the point forecast has undergone bias-adjustment.

To assess the impact of this bias-adjustment, take into consideration the following scenario: we use the drift method with a log transformation to anticipate the average annual price of eggs (λ= 0). The log transform is employed in this instance, to guarantee that forecasting as well as forecast intervals remain positive.

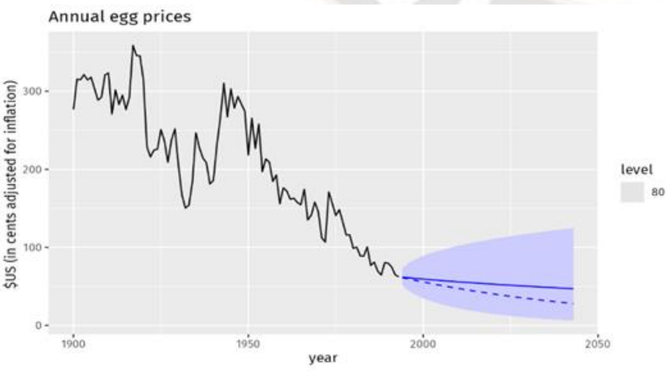

Figure 19: Egg price forecasts, utilizing the "Drift method applied to the logged data", depict bias-adjusted mean forecasts represented by a solid line, and median forecasts indicated by dashed lines.

In Figure 19, the dashed line represents the median forecasts, while the solid line represents the mean forecasts. The upward shift of the dashed line to the solid line is attributed to the bias adjustment. This adjustment is made to align the point forecasts with the mean, providing a clearer representation of the forecast distribution.

XII. Forecasting with Decomposition

The earlier discussion on Time Series Decomposition reveals its significant role in the forecasting process. When employing a Decomposition in Addition, we express the temporal sequence as:

$$y_t = \hat{S}_t + \hat{A}_t$$

Here, $\hat{S}_t$ represents the Temporal Aspect of the Time Series, and $\hat{A}_t = \hat{T}_t + \hat{R}_t$ denotes the Seasonally Adjusted a part of the Time Series, combining the Component of Trend ($\hat{T}_t$) and the Residual Component ($\hat{R}_t$).

In cases of Multiplicative Decomposition, the formulation changes to:

$$y_t = \hat{S}_t \times \hat{A}_t$$

and here $\hat{A}_t = \hat{T}_t \times \hat{R}_t$.

To forecast using these decomposed series, the Seasonal Component $\hat{S}_t$ and Seasonally Adjusted Component $\hat{A}_t$ are projected independently. The assumption is that the seasonal component either remains constant or changes very gradually. Consequently, the Seasonal Naïve Method is often the preferred approach for forecasting the Seasonal Component. For the Seasonally Adjusted Component, various non-seasonal forecasting methods can be effectively utilized.

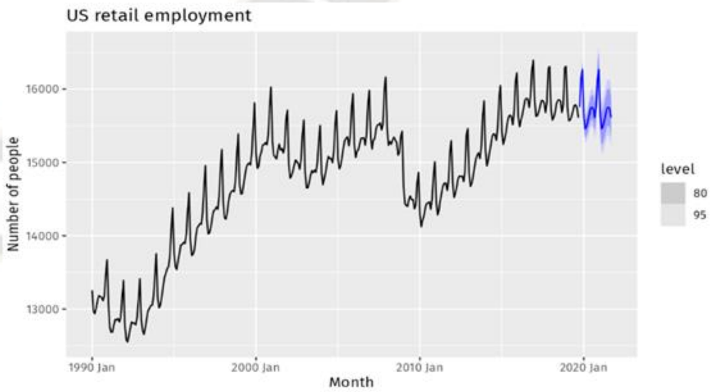

Figure20: Forecasts for the total US retail employment data are generated by employing a naïve forecast for the "Seasonally adjusted component and a seasonal Naïve forecast for the seasonal component", following an STL decomposition of the data.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

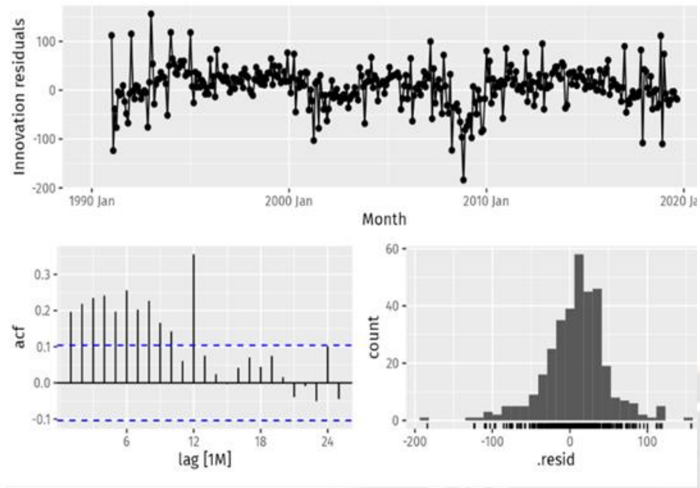

Figure21: Checking the residuals.

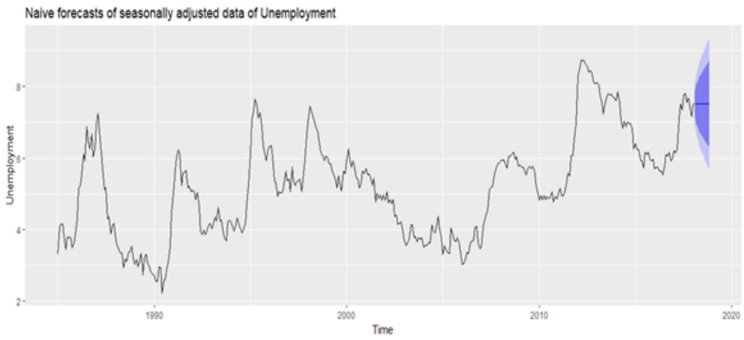

Figure22: Forecasts for the unemployment data of USA are generated by employing a naïve forecast for the seasonally adjusted component and a seasonal naïve forecast for the seasonal component, following an STL decomposition of the data.

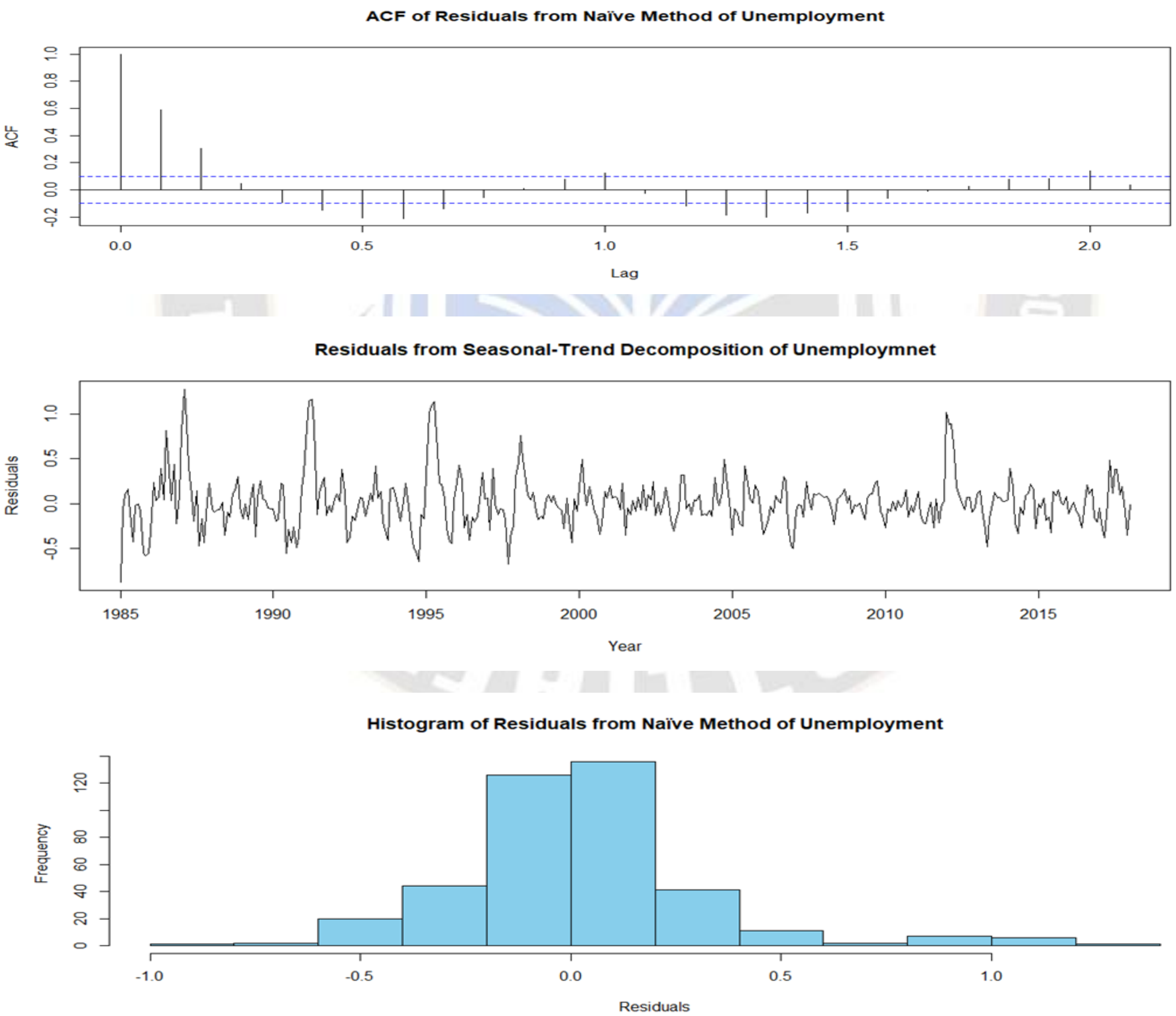

Figure23: Checking the residuals.

Conclusion and Future Scope

This paper has delved into the intricate world of time series forecasting, applying transformation and decomposition techniques across a diverse array of data sets, including US trade employment and unemployment figures, Google stock trends, electricity consumption patterns, and egg price fluctuations in the market. We have proved the adaptability and usefulness of these approaches from time series data through these many applications. In analyzing US trade and employment, we look at unemployment rates. We use transformation and decomposition in this analysis. These methods reveal complex economic dynamics. They provide valuable insights for policy-making. They also help in economic planning. The study focuses on Google stock trends. These trends are analysed using specific methods. This analysis gives deeper insights into market behaviors. It helps investors. It also aids analysts. They can make more informed decisions. Electricity consumption patterns are important. They matter for planning and sustainability efforts. These patterns have been forecasted more accurately. This shows the potential of certain techniques. These methods are utilised in the energy industry. The application of these strategies in egg pricing marketing demonstrates their utility in agriculture. Applying transformation and decomposition improve forecast accuracy and reveal underlying trends and patterns.

Future Scope

More study in this area is promising. It is possible to combine powerful machine learning and AI with classical approaches. This may result in more advanced models. These models might perform better with large datasets. It would boost prediction accuracy and dependability. It is critical to investigate real-time forecasting in volatile markets. Stock and energy markets are critical. This would allow for informed, fast judgments. Decisions would be made using current information. It is also advantageous to use these strategies to varied, worldwide datasets. It would also help with environmental conservation. transformation and deconstruction approaches show promise in predicting. They have enormous potential for use in time series analysis. It is critical to keep innovating in this field. Ongoing research is also necessary.

References:

1. George Athanasopoulos and Rob J. Hyndman Forecasting: Principles and Practice.
2. Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (2015). Time series analysis: forecasting and control. John Wiley & Sons.
3. Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. Econometrica, 37(3), 424-438.
4. Gardner Jr, E. S. (1985). Exponential smoothing: the state of the art. Journal of Forecasting, 4(1), 1-28.
5. Ljung, G. M., & Box, G. E. P. (1978). On a measure of lack of fit in time series models. Biometrika, 65(2), 297-303.
6. Makridakis, S., & Hibon, M. (2000). The M3-competition: Results, conclusions and implications. International Journal of Forecasting, 16(4), 451-476.
7. Armstrong, J. S. (2001). Principles of forecasting: A handbook for researchers and practitioners. Springer Science & Business Media.
8. Goodwin, P., & Fildes, R. (1999). Judgmental forecasting. International Journal of Forecasting, 15(3), 187-195.
9. Montgomery, D. C., Jennings, C. L., & Kulahci, M. (2015). Introduction to time series analysis and forecasting. John Wiley & Sons.
10. Wei, W. W. S. (2006). Time series analysis: univariate and multivariate methods. Pearson Education.
11. Cleveland, R. B., Cleveland, W. S., McRae, J. E., & Terpenning, I. (1990). STL: A seasonal-trend decomposition procedure based on Loess. Journal of Official Statistics, 6(1), 3-73.
12. Shumway, R. H., & Stoffer, D. S. (2017). Time series analysis and its applications: With R examples. Springer.
13. Gardner Jr, E. S. (1985). Exponential smoothing: the state of the art. Journal of Forecasting, 4(1), 1-28.
14. Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. International Journal of Forecasting, 22(4), 679-688.
15. Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (2015). Time series analysis: forecasting and control. John Wiley & Sons.
16. Brockwell, P. J., & Davis, R. A. (2016). Introduction to time series and forecasting. Springer.
17. Montgomery, D. C., Jennings, C. L., & Kulahci, M. (2015). Introduction to time series analysis and forecasting. John Wiley & Sons.
18. Syntetos, A. A., & Boylan, J. E. (2005). The accuracy of intermittent demand estimates. International Journal of Forecasting, 21(2), 303-314.
19. Kourentzes, N., Barrow, D. K., & Petropoulos, F. (2016). Forecasting with multivariate temporal aggregation: The case of promotional modeling. International Journal of Production Economics, 181, 200-208.
20. Chen, C. H., & Yang, C. H. (2019). Time series forecasting with neural network ensembles: An application for exchange rate prediction. Expert Systems with Applications, 129, 201-215.
21. Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2018). Statistical and Machine Learning forecasting methods: Concerns and ways forward. PloS One, 13(3), e0194889.

22. De Gooijer, J. G., & Hyndman, R. J. (2006). 25 years of time series forecasting. International Journal of Forecasting, 22(3), 443-473.
23. Willemain, T. R., Smart, C. L., & Schwarz, H. F. (2004). A new approach to forecasting intermittent demand. International Journal of Forecasting, 20(3), 375-387.
24. Please note that these references provide a comprehensive foundation for understanding various aspects of time series forecasting.