iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 Issue: IV Month of publication: April 2024

DOI: https://doi.org/10.22214/ijraset.2024.60658

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue IV Apr 2024- Available at www.ijraset.com

Forecasting US Inflation Trends: Insights from Time Series Analysis

Raheel Hassan[1], Kanishk Bhandari[2], Rajkumar Singh[3], Yokshit Sankhyan[4], Deepanshu Kumar[5], Yash Gupta[6]
Department of Computer Science and Engineering, Chandigarh University, Gharuan, Punjab, INDIA,

Abstract: In our research, we dive into the world of forecasting US inflation using the ARIMA (Auto-Regressive Integrated Moving Average) model. We crafted a predictive framework by meticulously analyzing data, conducting tests like the Augmented Dickey-Fuller, and crunching metrics such as RMSE and MSE. Our meticulous model selection procedures involve carefully examining many ARIMA configurations to identify the ideal parameters, ensuring robustness and accuracy in forecasting inflation trends over the study period. Our findings show that the ARIMA (0,1,2) model outperforms others, offering reliable forecasts. These results not only showcase the model's effectiveness but also provide valuable insights for both policymakers and market players. By understanding potential inflationary trends, they can confidently navigate risks and make informed decisions, ultimately fostering stability and growth in the economy.
Keywords: Time series analysis, ARIMA, Forecast, US Inflation, Statistical model

I. INTRODUCTION

Inflation is a fundamental economic phenomenon that profoundly influences various aspects of the economy, ranging from consumption patterns to investment decisions and monetary policy formulation. Understanding the dynamics of inflation and its underlying drivers is crucial for policymakers, economists, and market participants alike. This paper presents a comprehensive time series analysis of inflation in the United States spanning the period from 1947 to 2022.

Over the past decades, the US economy has experienced significant fluctuations in inflation rates, ranging from periods of high inflationary pressures to periods of subdued inflation or even deflationary concerns. These fluctuations have been influenced by a multitude of factors, including changes in monetary policy stance, fiscal policy measures, oil price shocks, and broader macroeconomic trends.

The study of inflation dynamics is inherently complex due to its multifaceted nature and the interplay of various economic forces. Time series analysis offers a powerful framework for examining the behavior of inflation over time, allowing researchers to identify long-term trends, cyclical patterns, and potential drivers.

The primary objectives of this study are twofold: first, to provide a comprehensive understanding of inflation dynamics in the US over the specified historical period, and second, to assess the forecasting performance of the developed models in predicting future inflationary trends. By achieving these objectives, the main aim of this research is to contribute valuable insights to the current literature on inflation analysis and provide practical guidance for policymakers and market participants in managing inflation risks and making informed decisions.

Financial time series can be forecasted using a variety of techniques. Univariate forecasting is one technique that solely estimates time. The autoregressive integrated moving average, or ARIMA model is a unique kind of modeling where the moving average component (the moving average component) of the time series is determined by taking the historical values of the autoregressive component and adding the current and lag values of the white noise error term. The ARIMA model and how we may forecast or anticipate future values based on past values are the main topics of the paper. In the subsequent sections of this paper, we will delve into the methodology employed for data collection and analysis, present the findings of our time series analysis, discuss the implications of our results, and conclude with recommendations for future research and policy considerations.

II. LITERATURE SURVEY

The study of inflation dynamics has been the subject of extensive research in economics literature, with scholars employing various methodologies to understand the complexities of inflation behavior. Numerous studies have investigated inflation trends, drivers, and forecasting techniques, offering valuable insights into the dynamics of inflation in different economic contexts.

One seminal work in this field is the study conducted by Fisher (1972), which emphasized the role of monetary factors in driving inflationary pressures.

Fisher's analysis highlighted the significance of money supply growth as a key determinant of inflation, laying the groundwork for subsequent research on the monetary theory of inflation. However, while Fisher's study provided valuable insights into the relationship between money supply and inflation, it has been criticized for its narrow focus on monetary factors and its neglect of other important determinants of inflation, such as fiscal policy measures and supply-side shocks (Taylor, 1979). Taylor argued that a more comprehensive approach to inflation analysis is needed to capture the full range of factors influencing inflation dynamics.

In response to these criticisms, subsequent research has expanded the scope of inflation analysis to include a broader set of variables and factors. For example, Blinder (1997) conducted a comprehensive study of inflation dynamics in the United States, examining the impact of various macroeconomic variables, including monetary policy, fiscal policy, oil price shocks, and exchange rate movements.

Blinder's study provided valuable insights into the multifaceted nature of inflation dynamics and underscored the importance of considering multiple factors in inflation analysis. However, one of the drawbacks of Blinder's study was its reliance on traditional econometric models, which may have limitations in capturing the nonlinear and dynamic nature of inflation processes (Stock & Watson, 2007).

Recent developments in time series analytic methods have created new avenues for a more thorough examination of inflation dynamics. Researchers have increasingly turned to sophisticated econometric models, such as ARIMA, to analyze inflation time series data and uncover hidden patterns and trends.

Overall, the literature on inflation analysis provides a rich body of knowledge that continues to evolve with advancements in economic theory and empirical techniques. While past studies have contributed valuable insights into inflation dynamics, there remains a need for further research to refine existing models, incorporate new data sources, and improve forecasting accuracy.

III. MATERIALS AND METHOD

George Box and Gwilym Jenkins created the ARIMA model in the 1970s and used mathematics to describe changes in time series data. The terms ARIMA and Box-Jenkins are used interchangeably in certain instances. A statistical analysis technique called Auto-regressive Integrated Moving Average, or ARIMA, makes use of time series data to forecast future trends or to get a deeper understanding of the data set. If a statistical model forecasts future values by using historical data, it is said to be autoregressive. An ARIMA model could, for instance, attempt to estimate a company's profitability based on historical periods or predict a stock's future pricing based on its historical performance. Lagged moving averages are used by ARIMA to smooth time series data.

Auto-Regressive (AR): The abbreviation for autoregressive models is AR. The time-series' lagged values determine the future values. The AR model estimates future data values by looking at historical data values. The order "p" indicates that current values are predicted using previous data from period "p." The following is a pth-order AR process:

$$y_t = \alpha + \emptyset_1 y_{t-1} + \emptyset_2 y_{t-2} + \ldots + \emptyset_p y_{t-p} + \varepsilon_t \quad \text{.............(1)}$$

In this case, $\alpha = \mu(1 - \emptyset_1 - \emptyset_2 - \ldots - \emptyset_p)$, with $\emptyset_1, \emptyset_2, \ldots, \emptyset_p$ is the parameter estimation component and ε_t representing the error term. Where y_t is the stationary response variable at time t, $y_{t-1}, y_{t-2}, \ldots, y_{t-p}$ is the dependent variable at different time lags.

Integrated (I): The I component refers to differencing the time series data to make it stationary, i.e., removing trends and seasonality. Stationarity is essential for many time series models, including ARIMA. The order of differencing, denoted by (d), represents how many times differencing is applied to achieve stationarity.

Moving Average (MA): The next step in the ARIMA modeling process is the MA term. Moving average is shortened to MA. It is the data lag from a random process (white noise process or term) induced by another random process.

The MA (q) is expressed as follows:

$$y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \ldots + \theta_q \varepsilon_{t-q} \quad \text{.............(2)}$$

where the coefficients for parameter estimation are $\theta_1, \theta_2, \ldots, \theta_q$, and q is the lag count of the moving average.

The US inflation is the subject of the investigation. US inflation data spanning 75 years, from 1947 to 2022, was gathered online. Here, the Auto Correlation Function (ACF), Partial Auto Correlation Function (PACF), and Augmented Dickey-Fuller [Dickey and Fuller (1979)] tests are used to determine whether the data are stationary.

The ADF Test is employed to ascertain the stationarity of the time series. The description of a random walk with trend and drift is as follows:

$$\Delta y_t = \alpha + \beta_t + \gamma y_{t-1} + \delta_1 \Delta y_{t-1} + \cdots + \delta_p \Delta y_{t-p+1} + \varepsilon_t \quad \text{.............(3)}$$

The amount of included p, which represent the AR process's lag order, should be large enough to render the residuals serially uncorrelated [9]. The time series will be not stationary as the null hypothesis claims if the ADF test result's p-value is more than 0.05, and this should be differentiated.

This study also uses the autocorrelation function (ACF) and partial autocorrelation function (PACF) in addition to the ADF test to evaluate the stationarity of the time series. The results of these tests are plotted on the correlograms for the time series analysis. In mathematics, the degree of continuity over the relevant variable lags is represented by the ACF expression:

$$r_k = \frac{\sum_{t=1}^{T-k}(y_t-\bar{y})(y_{t+k}-\bar{y})}{\sum_{t=1}^{T}(y_t-\bar{y})^2} \quad \ldots\ldots\ldots\ldots(4)$$

The degree of link between two variables can be determined mathematically using the PACF expression:

$$\begin{cases} r_1 & \text{if } k = 1 \\ \dfrac{r_k-\sum_{j=1}^{k-1}(P_{k-1,j} \cdot r_{k-j})}{1-\sum_{j=1}^{k-1}(P_{k-1,j} \cdot r_j)} & \text{if } k = 2, 3, \ldots \end{cases} \quad \ldots\ldots\ldots\ldots(5)$$

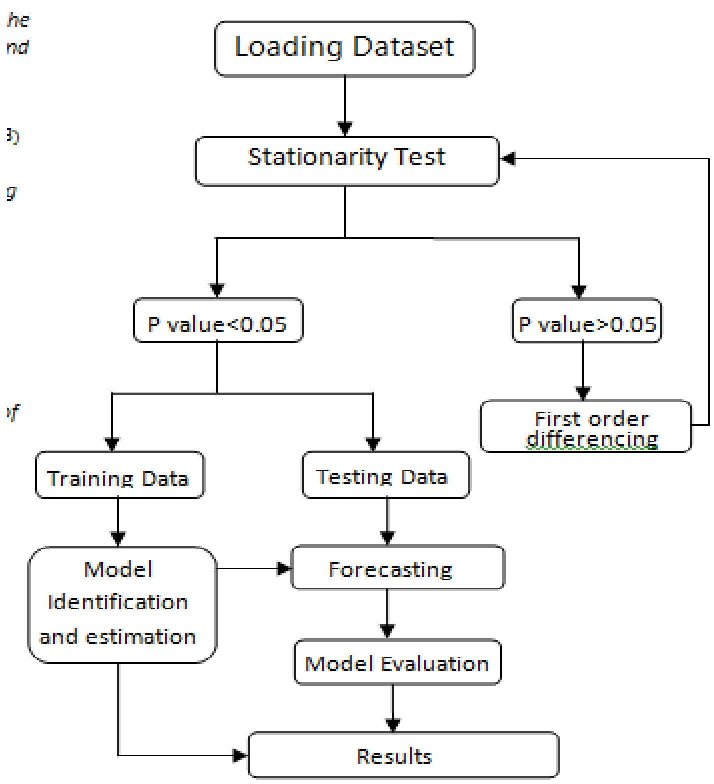
Where $P_{k,j} = P_{k-1,j} - P_{k,k}P_{k-1,k-j}$ for $j = 1,2,\ldots,k-1$. The time series can be considered stationary if the ACF and PACF plots indicate that the majority of the coefficients are within critical levels.

Finally, the forecasting of inflation is done here using the Box and Jenkins (1970) ARIMA methodology. ARIMA (p, d, q) is the standard notation for an ARIMA model. The ARIMA model (p,1,q) is given by

$y_t = \beta_0 + \emptyset_1 y_{t-1} + \emptyset_2 y_{t-2} + \ldots + \emptyset_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \ldots + \theta_q \varepsilon_{t-q}$ ....(6)

where the coefficients for parameter estimation are $\emptyset$, β, and θ, and y_t is the first order differenced series.

The best-fitted model has been chosen here with the help of some model choice criteria like AIC, and BIC. For model accuracy, the values of Error measures like Root Mean Square Error (RMSE), Mean Squared Error (MSE), etc. are calculated here. The normality of the residuals of the ARIMA Model is tested here and a line and density plot of residual is evaluated.



International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue IV Apr 2024- Available at www.ijraset.com

The above flowchart depicts the purposed approach for inflation prediction. It requires following procedures in order to predict future values:

1) *Data Preparation*: Gathering and pre-process the time series data, using methods like differencing to make sure it is stationary. Dividing the data into sets for testing and training. The testing set is used for evaluating the model's performance, and the training set is used to fit the model.
2) *Model Identification and estimation*: Finding the differencing order (d) that is required to make the series stationary. Plots of the ACF and PACF can be used to figure out this. Determining the values of p, d, and q and using them to fit the ARIMA model to the training set. Find the optimal model by using measures such as AIC and BIC.
3) *Forecasting and Model Evaluation:* Using the model to produce forecasts for next time periods. To measure the forecasts' accuracy, we can use measures like Mean Squared Error (MSE) and Root Mean Squared Error (RMSE).

IV. DATA ANALYSIS

The past inflation rates for the US from January 1947 until 2022 are included in this dataset. An important economic indicator that can reveal information about the state of the economy and the buying power of a currency is the inflation rate. The CPI is a gauge of how prices for a market basket of goods and services have changed on average over time for urban consumers. CPI is frequently used to monitor the purchasing power of money over time or to account for inflation in other economic stats. An examination of this data can provide information on the nation's monetary policies and overall state of economy.

There are two columns in the dataset:

Date: The end of that month (spelled out in YYYY-MM).

Value: The Consumer Price Index, or CPI, at the conclusion of *the corresponding month.*

Statistics is provided by the Federal Reserve Bank of St. Louis' Federal Reserve Economic Data (FRED) and are taken from Kaggle website.

V. RESULTS

Here, we consider the inflation data for 75 years in the US from 1947-2022. The basic descriptive statistics of the data are shown in Table 1 and the data is shown in Fig. 1.

Table 1: Basis Descriptive Statistics of the Inflation data

Measures	Observations
Total observations	912
Minimum observation	21.48 (January 1947)
Maximum observation	298.99 (December 2022)
Mean	116.5807
Median	105.6
SE Mean	2.737044
Standard Deviation	82.65693

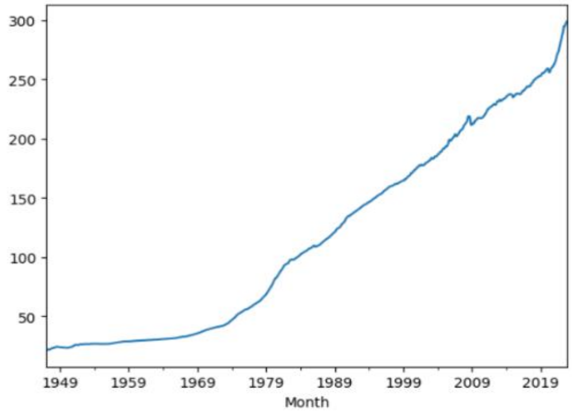


Fig. 1: Time series plot of US inflation data

To check stationarity of the data the Augmented Dickey-Fuller test is performed and the results are shown in Table 2.

Table 2: Results of the ADF test

ADF Statistic	2.660602
p-value	0.999084

Since the p-value is greater than 0.05 we will try to make it stationary. To make the data stationary we have to remove the stochastic trend from the data. To do this we have taken the simple difference of the data by first-order differencing *i.e.*

New values = $Y_t - Y_{t-1}$............ (7)

After taking the differences the new results obtained are shown in Table 3.

ADF Statistic	-2.953629
p-value	0.039547

Table 3: Results of ADF test after first order differencing

Since the p-value is less than 0.05, it suggests that we can reject the null hypothesis, indicating that the time series data does not have a unit root and the data is stationary. To test the stationarity of the data graphically we have plotted the Auto Correlation Function (ACF) and Partial Auto Correlation Function (PACF) which are shown in Figs. 2 and 3.

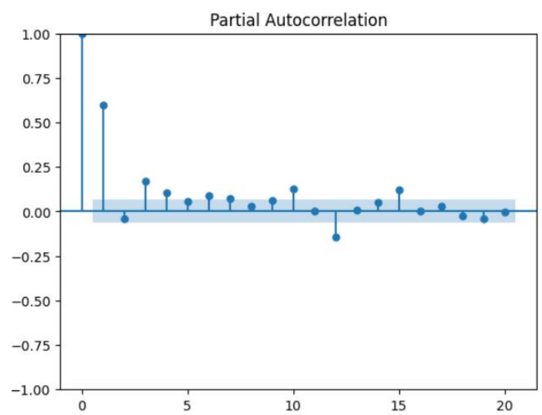


Fig. 2: Partial Auto Correlation Function (PACF)

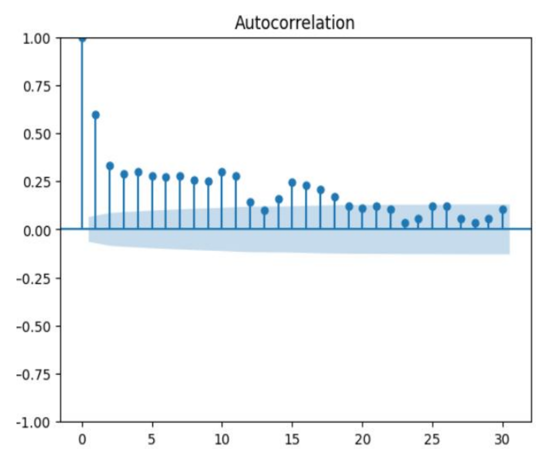


Fig. 3: Auto Correlation Function (ACF)

So in the ARIMA (p,d,q) model, we will try possible combinations of p,d,q and based on Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) we can get the best model that can fit into the prediction. The models proposed are shown in Table 4 with their AIC and BIC values.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue IV Apr 2024- Available at www.ijraset.com

Order	AIC	BIC
(1, 1, 2)	430.716	444.1044
(2, 1, 1)	430.9498	444.3382
(0, 1, 2)	431.3289	441.3702
(3, 1, 1)	432.3259	449.0614
(2, 1, 2)	432.5481	449.2837
(1, 0, 2)	432.5486	449.3079
(3, 0, 1)	432.8894	453.0005
(2, 0, 2)	433.2761	453.3872
(3, 0, 0)	433.3843	450.1436
(1, 0, 1)	433.4417	446.8491

Table 4: Different (p,d,q) models and their AIC and BIC

From Table 4, it is observed that the AIC and BIC values for Model 3: ARIMA(0,1,2) are slightly lower and the test rmse and mse values for this model are lowest. So, here Model 3 is used to forecast the data. To check the residuals of the ARIMA(0,1,2) model "line plot and density plot of residual" is constructed here and is depicted in Figs. 4 and 5 and the summary statistics of residuals is shown in table 5.

Count	911.0000
Mean	0.010480
Std	0.360583
Min	-3.143813
25%	-0.096918
50%	0.007996
75%	0.119352
Max	2.013422

Table 5: Summary statistics of residuals

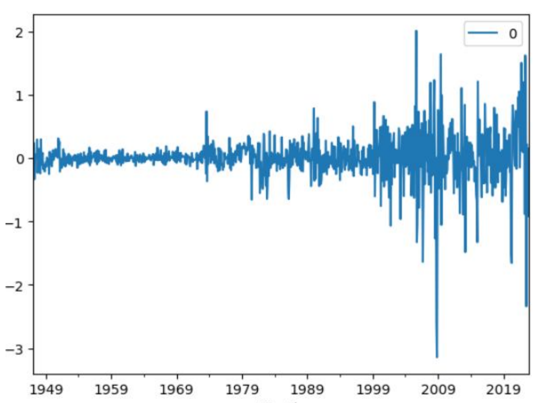


Fig. 4: Line plot of residual

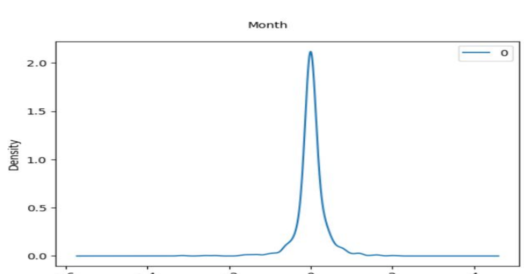


Fig. 5: Density plot of residual

The prediction is done using the ARIMA(0,1,2) model and the rmse and mse calculated are 0.703 and 0.495 respectively. Fig. 6 illustrates the comparison between the actual and predicted values, providing insights into the model's performance, while Fig. 7 visually represents the predicted results, offering a clear depiction of the forecasted trajectory.

```
predicted=0.595424, expected=0.521000
predicted=0.321728, expected=1.054000
predicted=0.705407, expected=0.670000
predicted=0.346883, expected=0.048000
predicted=0.220694, expected=0.650000
predicted=0.565578, expected=0.565000
predicted=0.364258, expected=0.226000
predicted=0.298456, expected=0.196000
predicted=0.314244, expected=0.449000
predicted=0.426634, expected=0.519000
predicted=0.407627, expected=0.590000
predicted=0.451492, expected=-0.178000
predicted=0.065990, expected=0.173000
```

Fig. 6: Predicted results by using ARIMA(0,1,2)

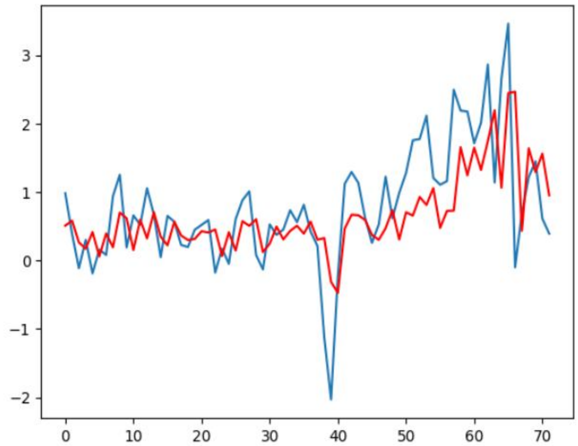


Fig. 7: Graph of predicted results

VI. SUMMARY AND CONCLUSION

In this paper, we analyzed US inflation dynamics from 1947 to 2022, using advanced time series techniques. We tested data stationarity with the Augmented Dickey-Fuller (ADF) test, then applied differencing for stationarity. Next, we employed ARIMA for forecasting, testing various models to find the best fit based on AIC, BIC, RMSE, and MSE. Having selected the most suitable ARIMA model, we proceeded to generate inflation forecasts for the specified period. Visualizations, including line plots and residual plots, were utilized to present the forecasted results and assess the model's performance.

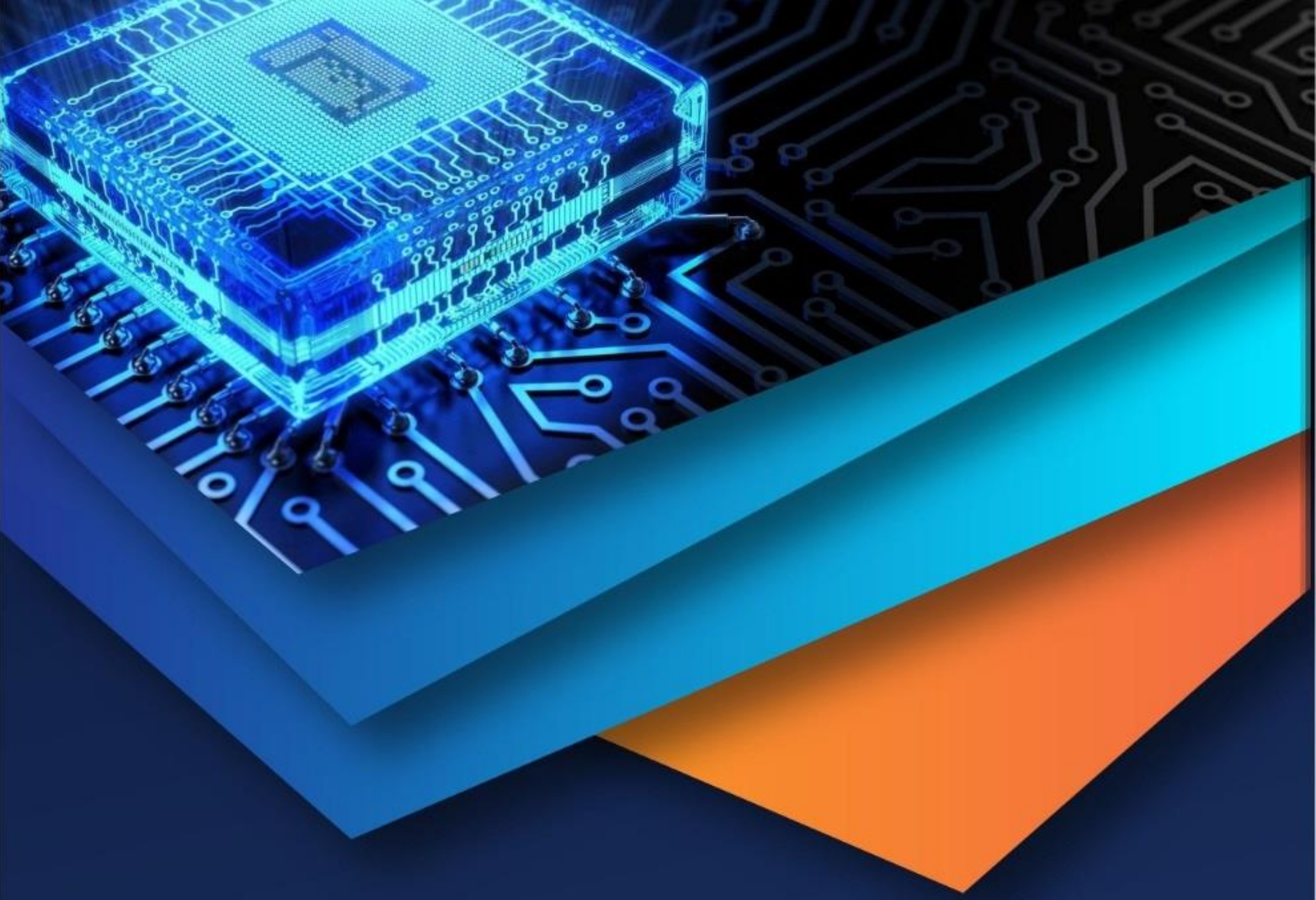
In conclusion, our research offers a strong forecasting framework for inflation trends, benefiting policymakers, economists, and market participants. However, it is essential to acknowledge the limitations of our study, including the assumptions underlying the ARIMA model and potential uncertainties in forecasting future inflation trends. Future research may explore alternative modeling approaches, incorporate additional variables, or employ more sophisticated techniques to enhance forecasting accuracy.

VII. ACKNOWLEDGEMENTS

We express sincere gratitude to all contributors to this research on "FORECASTING US INFLATION TRENDS: INSIGHTS FROM TIME SERIES ANALYSIS". Special thanks to Ms. Raheel Hassan for invaluable guidance and mentorship. We thank Chandigarh University for resources and this opportunity to work on this. We would like to acknowledge the contributions of the researchers whose seminal works and studies lays the foundation for our research. Their insights and findings have enriched our understanding of inflation dynamics and guided our analytical approach. We appreciate Kaggle.com for data access. Lastly, thanks to friends, colleagues, and family for unwavering support.

REFERENCES

[1] Fisher, I. (1972). The Theory of Interest. Dover Publications.
[2] Taylor, J. B. (1979). Staggered wage setting in a macro model. The American Economic Review, 69(2), 108-113.
[3] Blinder, A. S. (1997). Monetary policy and the inflation process: Some lessons from the 1980s. Economic Review, Federal Reserve Bank of Kansas City, 82(4), 5-22.
[4] Stock, J. H., & Watson, M. W. (2007). Introduction to Econometrics (3rd ed.). Addison Wesley.
[5] Box, George E. P. Time series analysis: forecasting and control. -- Fifth edition / George E.P. Box, Gwilym M. Jenkins, Gregory C. Reinsel, Greta M. Ljung.
[6] Pincheira P, Medel CA (2012b): Forecasting Inflation with a Simple and Accurate Benchmark: A Cross-Country Analysis. Central Bank of Chile, Working Paper, no. 677.
[7] Proietti T (2011): Direct and Iterated Multistep AR Methods for Difference Stationary Processes. International Journal of Forecasting, 27(2):266-280.
[8] Stock JIH, Watson MW (1996): Evidence on Structural Stability in Macroeconomic Time Series Relations. Journal of Business and Economic Statistics, 14:11-30.
[9] Wagenmakers E-J, Farrell S (2004): AIC Model Selection Using Akaike Weights. Psychonomic Bulletin and Review, 11:192-196.
[10] West K (1996): Asymptotic Inference about Predictive Ability. Econometrica, 64:1067-1084.
[11] Cosgrove, J. (2020). The Russian invasion of the Crimean peninsula, 2014-2015: A post-cold war nuclear crisis case study. Applied Physics Lab, Johns Hopkins University Laurel Md Laurel.
[12] Ara Stepanyan, Agustin Roitman, Gohar Minasyan, Dragana Ostojic, and Natan Epstein.(2015). Spillover. THE SPILLOVER EFFECTS of Russia’s Economic Slowdown on Neighboring Countries.
[13] Russia inflation rate in 2014. Russian Federation Inflation Rate in 2014. (n.d.).
[14] Ozili, P. K. (2022). Global economic consequence of Russian invasion of Ukraine.SSRN, 4064770.
[15] Yurina, E. A., Tishchenko, E. S., Fedotova, E. V., Markov, B. E., & Kuzmenko, N. I. (2022). New Approaches to the Economic Security of Russian Regions in Terms of the COVID-19 Pandemic. In Geo-Economy of the Future: Sustainable Agriculture and Alternative Energy. Cham: Springer International Publishing, 699-715.
[16] Kasyan, E. A., & Nemirovich-Danchenko, M. M. (2022). Selection of a forecasting model for Russia consumer price index based on the application of Analytic Hierarchy Process. In Journal of Physics: Conference Series, 2291, 1, 012006.
[17] Benkraiem, R., Garfatta, R., Lakhal, F., & Zorgati, I. (2022). Financial contagion intensity during the COVID-19 outbreak: A copula approach. International Review of Financial Analysis, 102136.
[18] GAUTAM, R. S., & KANOUJIYA, J. (2022). Inflation Targeting: An Application of ARIMA Modelling Using Forecasting of CPI and WPI. Iconic Research and Engineering Journals, 5(11), 195-198.
[19] Jalil, A., & Rao, N. H. (2019). Time series analysis (stationarity, cointegration, and causality). In Environmental Kuznets Curve (EKC).Academic Press, 85-99.



10.22214/IJRASET

INDEX COPERNICUS
45.98

IMPACT FACTOR:
7.129

TOGETHER WE REACH THE GOAL
IMPACT FACTOR:
7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY