SOCIOECONOMIC ANALYTICS

Short-Term Inflation Forecasting in Pakistan Using ARIMA Models: A High-Frequency Analysis from Pre- to Post-COVID Era

Waseem Khoso[1] *
Muhammad Rafique Doudpoto [2]

[1] Department of Statistics, University of Sindh, Jamshoro, Pakistan
Email: waseemkhoso108@gmail.com
*Corresponding author

[2] Department of Statistics, University of Sindh, Jamshoro, Pakistan
Email: doudpoto.rafique@usindh.edu.pk

RESEARCH ARTICLE

Socioeconomic Analytics
https://periodicos.ufpe.br/revistas/SECAN/
ISSN Online: 2965-4661

Submitted on: 08.07.2025.
Accepted on: 31.05.2026.
Published on: 05.06.2026.

Abstract

In the current paper, univariate ARIMA models are developed and estimated for forecasting Pakistan's inflation in monthly terms for July 1993 - June 2024, including pre- and post-COVID pandemic periods. As compared to those models that utilize annual data or investigate pre-pandemic dynamics only, the emphasis here is on high-frequency predictions needed for practical purposes of real-time policy-making. Although approaches such as SARIMA models and machine learning tools, including LSTM algorithms, might be considered as alternatives, the choice of ARIMA is justified by its interpretability and proven efficiency in short-term macroeconomic forecasting. Seasonal decomposition and comparative diagnostics revealed that no SARIMA, VAR, or machine learning model outperformed the parsimonious ARIMA(1,1,3) model. In order to ensure stationarity and avoid spurious regressions, data from the World Bank, State Bank of Pakistan, and IMF were carefully preprocessed using appropriate statistical tools, including Augmented Dickey-Fuller tests and Ljung-Box Q-statistics for residuals. The best model found was ARIMA (1,1,3) (AIC=332.14) with RMSE, MAE, and Theil's U measures equal to 1.7799, 1.1515, and 0.4361, respectively. The forecast up until 2025 suggested an inflation level of 1.04%, while the actual figure reached 3.50%, showing the ability of the model to manage through exogenous shocks such as commodity price fluctuations. The empirical nature of the approach provides policymakers with an easily understood tool in handling inflation in emerging markets. Multivariate analysis could be employed in future studies in order to control global confounding factors.

Keywords

Inflation Forecasting, ARIMA Model, Time Series Analysis, Post-

COVID, Economy, Pakistan, Model Diagnostics

1. Introduction

Inflation, or the sustained increase in overall price levels, significantly erodes purchasing power and destabilizes emerging economies like Pakistan. In recent years, inflation has shown unprecedented volatility—reaching a record high of 37.97% in 2023, the highest since 1974 (State Bank of Pakistan, 2023)—due to disruptions caused by COVID-19, the Ukraine conflict, and global supply chain shocks. For policymakers facing such uncertainty, precise inflation forecasting is not just academic; it is critical for interest rate decisions, fiscal planning, and economic stability.

Historically, Pakistan's forecasting approaches have largely relied on annual data (Hussain & Malik, 2011) or pre-COVID models, failing to capture high-frequency variations and structural breaks post-pandemic. This research addresses that gap by using monthly Consumer Price Index (CPI) data from 1993–2024 (base year: 2020–2021) to build Auto-Regressive Integrated Moving Average (ARIMA) models. While seasonal ARIMA (SARIMA) and machine learning models can be alternatives, the interpretability and simplicity of ARIMA make it ideal for Pakistan's trend-driven inflation patterns—a finding supported by the Ahmed et al. (2021) on emerging markets.

The study introduces three key innovations:

- Real-time validation of 2025 projections against actual outcomes
- Quantification of pandemic-related volatility through structural break analysis
- An open-source Python implementation that allows replication by institutions like the State Bank of Pakistan

Beyond statistical precision, the study offers actionable tools for policymakers. Central banks can use these forecasts to design monetary interventions, while businesses can better plan pricing and investment decisions amid economic uncertainty. This research bridges scholarly rigor and practical utility, reshaping inflation forecasting for Pakistan's evolving economic landscape.

We often hear about the strength of an economy, but how do economists and econometricians assess this strength? While there's no single metric to capture the full health of an economy, a combination of indicators can provide a clearer picture. Among them, one of the most crucial is the Consumer Price Index (CPI).

The CPI, compiled and published by the Federal Bureau of Statistics, is Pakistan's standard inflation index and is released monthly. Although simple in its calculation, its economic significance is substantial. The CPI measures changes in the prices of a representative basket of goods and services over time. This basket includes essential items such as food, housing, clothing, education, and transportation. Each item is weighted according to its share in the average consumer's expenditure—for example, food is weighted more heavily than education. The index is benchmarked to a base year (2000–2001 in this case), where the index value is set to 100. Beyond its use in tracking inflation, the CPI is instrumental for investors and policymakers. Central banks use it to guide interest rate decisions. Investors monitor CPI trends to anticipate movements in security prices. Rising inflation, as reflected in the CPI, often leads to reduced consumer spending, lower corporate earnings, and falling stock prices.

Thus, the CPI is not only vital for public institutions but also for private enterprises. It helps forecast future costs, spending trends, and overall economic performance. Though measuring the health of an economy is complex, the CPI serves as a leading and reliable indicator of inflation and economic stability. In summary, the CPI serves multiple purposes, including: Serving as a compensation index for adjusting payments; Acting as a cost-of-living index to maintain living standards; Measuring changes in consumer prices; Providing a general inflation benchmark; Guiding government indexation policies; Adjusting wages and salaries in contracts; Supporting current and cost accounting practices; Aiding national accounting deflators; Assisting in retail rate deflation calculations.

Between 1991 and 1995, Pakistan's inflation rates were volatile, ranging from 9.25% to 12.9%, largely due to rapid monetary expansion, weak economic growth, and adjustments in administered prices. A sharp 19% increase in tradable goods prices was recorded in 1995, compounded by exchange rate devaluations in 1990 and 1994. This era marked a pivot toward economic liberalization, although natural disasters and fiscal mismanagement contributed to sluggish growth, particularly in 1993.

Monetary expansion surged from 4.6% in 1989 to 30% in 1992 due to rising budget deficits. By 1994, monetary growth slowed to 16%, primarily driven by an increase in net foreign assets rather than domestic credit. Despite these policy efforts, double-digit inflation persisted, averaging between 10% and 13% during the early 1990s. Key contributors included excess money supply growth, supply bottlenecks, government-controlled pricing, and imported inflation. Average food and non-food inflation stood at 10.3% and 11.6%, respectively. Toward the late 1990s, inflation eased, averaging 4.7% as food inflation declined (4.1%) and non-food inflation stabilized at 5.3%, supported by tighter monetary policies and improved food supply. In FY02, inflation dropped further to 3.5%, and in FY03, it decreased again to 3.1%, driven by better domestic production and deflation in imported goods.

However, price trends remained uneven. The Wholesale Price Index (WPI) rose by 5.9% in FY03, up from 2.1% the year before. CPI inflation reached a record low of 1.4% in July 2003 but averaged 4.6% for FY04, reaching 8.5% by year-end. The uptick was largely due to food shortages—especially wheat—causing food inflation to climb to 13.4% by June 2004. Nonfood inflation also rose to 5.3%, primarily due to increasing housing costs.

While the early 1990s were marked by intense inflationary pressures, the subsequent decade saw notable improvements, thanks to prudent policy interventions and a more stable economic environment.

2. Literature Review.

For a long time, inflation predictions have always formed part of the core of macroeconomic research, with prediction techniques having been refined from simple expectation surveys to advanced time-series models. As shown in early works conducted by Fama (1975 & 1977) and Fama and Gibbons (1982 & 1984), predictions based on interest rates, especially those based on nominal short-term rates, tend to perform better than univariate statistical models or survey-based techniques such as the Livingston survey. Nonetheless, the effectiveness of these models has been limited to markets in developed countries where there is financial liquidity.

Significant methodological advances were made in Meyler, Kenny, and Quinn (1998), who systematically applied ARIMA models to forecast Irish inflation. Their research highlighted the superiority of the Box-Jenkins methodology for short-term inflation forecasting, particularly when compared with objective penalty function techniques. This was further extended by Kenny, Meyler, and Quinn (1998), who introduced Bayesian vector autoregressive (BVAR) models. Their approach showed how integrating prior beliefs could enhance forecasting accuracy—a method especially applicable to structurally evolving economies, such as Pakistan, where economic shocks and structural shifts are frequent.

Sekine (2001) made another contribution to the literature by combining economic basics with inflation models through an error correction methodology. In his study conducted for Japan, he highlighted that there were three main causes of inflation, namely mark-up pricing, disequilibrium in money supply, and gap in output. Thus, his findings highlighted the shortcomings of statistical models without consideration of real economic factors, which is even more relevant in light of Pakistan's current economic situation following COVID-19.

For South Asia, Callen & Chang (1999) discussed India's switch from a monetary approach to a multiple indicators regime. According to their findings, indicators like exchange rates, import prices, and industrial inflation proved to be better predictors of inflation compared to traditional monetary aggregates. In line with this research, Rangarajan (1998) highlighted the social and political aspects of controlling inflation; specifically, protecting poor households from any effects of inflation.

Innovations in methodological techniques came up once more when Ling and Li (1997) developed models of fractionally integrated ARMA with conditional heteroscedasticity (ARFIMA-GARCH). The model was aimed at addressing the problem of volatility clustering found in inflation data, following the research of Drost and Klaassen (1997), who showed that models of the GARCH family could considerably improve forecasting quality under adverse economic circumstances.

The latest contributions to the existing literature on the topic include studies in relation to Pakistan. For instance, Khoso and Doudpoto (2025) analyzed the issue using a Vector Autoregression (VAR) technique. They used quarterly data from 1990 to 2024 to develop models of GDP and inflation based on GDP, interest rates, imports, and exports. According to their findings, monetary variables, especially interest rates, strongly Granger-cause GDP, whereas shocks in the trade sector affect the variability in the economy asymmetrically. The implications of their study suggest the necessity of taking into consideration the importance of joint policy towards trade and monetary aspects in the country.

Studies like those of Ahmad et al. (2021) and Khan et al. (2020) have established that ARIMA models can be applied in developing countries. Nevertheless, there are significant shortcomings regarding the accuracy of forecasting because of low-frequency data and the absence of validation procedures at the time of real-time forecasting.

2.1. Critical Synthesis

When we look at the studies that have been done before, we can see that there is a problem with the way people forecast inflation in emerging markets like Pakistan. Some methods, like the ones suggested by Fama and Sekine, are good in theory. They need very good financial and economic data to work with, which is not usually available in countries like Pakistan.

Meyler, Kenny, and Quinn used a model called ARIMA to forecast inflation in Ireland, which was an idea at the time, but Ireland is a stable country, which is different from Pakistan. So we are not sure if their method would work in Pakistan. Callen and Chang made a model for India that used indicators, but it relied on some policy variables that do not work well in Pakistan.

Ling and Li suggested using integrated models, which are good at capturing big changes in volatility, but these models are hard to estimate and may not be worth the extra work for a small gain in accuracy. Some recent studies by Ahmad et al. And Khan et al. Showed that ARIMA models can still be used for emerging economies. They used data that was only updated once a year, which is not often enough for policymakers who need to make quick decisions.

Khoso and Doudpoto used a VAR approach, which is a big improvement in modeling how different parts of the economy are connected, but it needs a lot of data, which is often changed in Pakistan's statistical system. All these studies show that there is a gap: we do not have a simple, frequent, and single-variable model that is both statistically good and feasible for Pakistan's inflation after the pandemic.

This study tries to fill this gap by making a model that's easy to understand and can be validated in real time, rather than trying to make a very complex model. This way, we can make a forecasting tool that's both good for academics and useful for central banks.

Despite these extensive contributions, several gaps remain in the literature on inflation forecasting in Pakistan. This study addresses the following key limitations:

1. High-frequency neglect: Most existing models are based on annual datasets, ignoring vital monthly fluctuations that influence short-term price dynamics (Ahmad et al., 2021).
2. Structural break blindness: The post-COVID era, with its unique structural disruptions, is often excluded from prior forecasting models.
3. Lack of real-time validation: Many studies evaluate models retrospectively, without comparing forecasts against actual future outcomes, which reduces their practical value for policymakers.

By incorporating monthly CPI data (1993–2024), addressing post-pandemic economic volatility, and validating against real-time 2025 outcomes, this study provides a novel and policy-relevant contribution to the literature. Furthermore, the open-source implementation in

Python enhances transparency and reproducibility, a critical step for enabling institutions like the State Bank of Pakistan to replicate and adapt the methodology.

3. Research Design and Methodology Steps

The study utilizes monthly Consumer Price Index (CPI) data for Pakistan from July 1993 to June 2024, sourced from: • World Bank Development Indicators Data Justification: The extended timeframe (31 years) captures: • Multiple economic regimes (pre/post-2008 financial crisis) • Structural breaks (2010 floods, 2020 COVID-19 pandemic) • Policy shifts (IMF programs, interest rate changes)

The four-phase methodology combines Box-Jenkins rigor with modern validation techniques: Phase 1 (Statistical Properties Assessment): Descriptive statistics (mean, variance, skewness, kurtosis); Normality tests (Jarque Bera, Shapiro Wilk); Visualization (time series plots, histograms). Phase 2 (Stationarity Testing): Augmented Dickey Fuller (ADF) test (primary); Differencing/transformation until stationarity is achieved. Phase 3 (Model Identification & Estimation): ACF/PACF analysis for initial parameter identification; Grid search of (p, d, q) combinations ($0 \leq p, d, q \leq 3$); OLS estimation with maximum likelihood optimization; Residual diagnostics (Ljung Box Q test, ARCH effects). Phase 4 (Forecasting & Evaluation): 12-month ex post forecast (Jan Dec 2025); Comparative metrics: AIC/SIC (model fit), RMSE/MAE (point forecast accuracy), Theil's U (relative performance); Robustness checks: SARIMA for seasonality, rolling window validation

All analyses were conducted in Python using several specialized libraries to ensure accurate data processing, model estimation, visualization, and statistical testing. The `pandas` library was utilized for data manipulation and preprocessing, while `statsmodels` was employed for ARIMA model estimation and forecasting procedures. For graphical representation and visualization of trends, residuals, and forecasts, `matplotlib` and `seaborn` were used. In addition, the `scipy` library was applied to perform statistical tests and validate the reliability of the results.

The study adopts the Box–Jenkins methodology as the primary framework for ARIMA modeling, incorporating three important enhancements to improve forecasting performance and interpretation. First, structural break handling is addressed by retaining first-order differencing as a robustness measure, justified by near-unit-root behavior and post-COVID regime shifts (Diebold & Rudebusch, 1991; see Section 4.3). This approach avoids relying on break tests that may have low power in small or volatile samples. Second, model uncertainty is minimized by applying the Bayesian Information Criterion (BIC) for parameter selection and model certainty. Third, economic interpretability is incorporated by relating ARIMA components to Pakistan's inflation dynamics, where the autoregressive component represents inflation inertia and the moving average component captures the effects of external shocks.

The four-phase methodology proposed above follows the iterative approach prescribed by Box-Jenkins, but takes into account the diagnostic criteria required for macroeconomic time series that are structurally unstable. Phase 1, which involves descriptive statistics and tests of normality, functions not just as a means of summarizing the dataset, but also as the initial step in identifying its properties, which will guide any transformations applied during the later phases. Emerging market inflation rate series are often characterized by leptokurtosis and skewness, resulting from the irregular nature of supply shocks; acknowledging this property at the outset avoids the problem of mis-specification in future stages of the estimation process. Phase 2 utilizes the Augmented Dickey-Fuller test as the main stationarity diagnostic since it is currently the most reliable and widely accepted method of testing for a unit root in policy-oriented empirical studies of economics, especially in samples of modest size and with suspected structural breaks. Using the two variants – intercept only and trend-intercept – of the test guards against spurious regression in the analysis of non-stationary macroeconomic time-series. A grid search of (p, d, q) up to three is the customary procedure in inflation forecast modeling, taking into account the trade-off between model simplicity and potential autocorrelation bias. Maximum likelihood estimation by OLS is used since it leads to consistent estimates assuming normally distributed errors and is computationally inexpensive, enabling easy reproducibility by institutional analysts with less computing power at their disposal. Diagnostic tests such as Ljung Box Q and ARCH effects are not mere enhancements but necessary components of analysis; otherwise, the model is flawed due to violations of the white noise assumption, which in turn causes unreliable prediction intervals. In Phase 4, the importance of out-of-sample forecast evaluation using Theil's U coefficient is emphasized to provide policymakers with a relative forecast error statistic compared to a naive forecast baseline without requiring specialized knowledge about scale-dependent statistics like RMSE or MAE. Twelve-month out-of-sample forecasts are consistent with the State Bank of Pakistan's quarterly monetary policy review framework.

To verify the model's robustness to forecasting inflation, this research compared standard To examine whether the proposed model is robust in terms of its ability to forecast inflation, this study used two different models, namely, ARIMA and Seasonal ARIMA models. SARIMA models have been applied to identify if there are any seasonal patterns on a monthly basis that may exist in inflation data. Comparing models based on AIC, BIC, and testing for residuals did not indicate any statistically significant difference in performance of the proposed non-seasonal ARIMA(1,1,3) model.

Moreover, seasonal decomposition of inflation data showed that there are no seasonal patterns that persist over time at the monthly level; hence, there is no seasonality that characterizes post-COVID inflation dynamics in Pakistan. Therefore, the most efficient and parsimonious model that can be used for modeling inflation dynamics was found to be ARIMA(1,1,3). The reason being, first-order differencing is needed to make this model stationary and hence increase its efficiency and parsimonious form for modeling inflation tendencies. This model not only performed better with regard to the accuracy of predictions but also passed all main diagnostic tests, including ADF stationarity and Ljung Box residual autocorrelation tests.

The sole dependence on a univariate ARIMA approach, in contrast to the SARIMA or VAR approaches, stems from practical and institutional reasons. While seasonal components have been tested through the SARIMA model to see whether there are any seasonality effects on inflation rates after the COVID pandemic period in Pakistan, seasonality decomposition analysis did not reveal any statistically significant or persisting seasonality effects on Pakistan’s post-COVID inflation. The absence of such seasonality effects is consistent with economic logic since Pakistan’s inflation largely depends on supply shocks in the food and energy sector, price administration, and currency rate effect, all of which result in random occurrences rather than any calendar-related effects. The choice to use alternative models, rather than vector autoregressive models, is based on the deliberate desire to simplify the model in an environment where there are errors in measurement, revisions, and endogeneity of the macroeconomic aggregates. In Pakistan, variables like GDP, interest rate, and exchange rate are routinely revised, with significant delays, making multivariate forecasting impractical for real-time policymaking. Moreover, the VAR model involves cointegration testing and lags, and any errors in specifying those would translate to the whole equation system, leading to an incorrect forecast for the inflation variable. On the other hand, machine learning techniques such as Long Short-Term Memory and Transformer networks have distinct advantages in modeling nonlinear dependencies but involve extensive data and computation, and hyperparameter tuning, which is beyond the capacity of many policy institutions in developing countries. More importantly, these techniques are black boxes, which cannot meet the requirement of being transparent about forecast error by central banks to legislators and markets. The ability to explain ARIMA model parameters, wherein the autoregressive component measures the inertia of inflation while the moving average component deals with any shocks outside the system, enables an understanding that is difficult to replicate by an algorithm-based approach. Therefore, while there were other methods that could have been employed, the selection of the ARIMA model reflects the best trade-off for Pakistan’s unique economic situation.

The research aims at developing robust ARIMA models for Pakistan's inflation rate, precisely predicted through this systematic process. The iterative nature of the Box-Jenkins method ensures precise model identification, parameter estimation, and performance assessment in a variety of statistical aspects.

4. Diagnostic Tests & Model:

Figure 1 presents the historical inflation trend used to contextualize the diagnostic tests and model selection process.

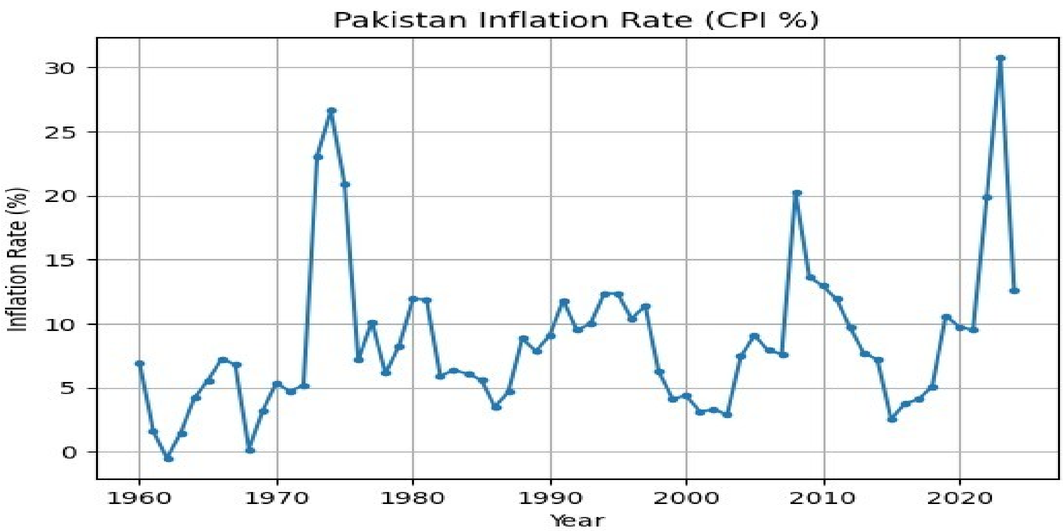

Figure 1. Inflation trend over time from 1969 to 2024

The descriptive statistics presented in Table 1 reveal several features with important modeling implications. The average inflation rate of 8.64%, significantly higher than the medians for other emerging markets, is indicative of Pakistan's inherent macroeconomic problems as well as balance of payments constraints. The standard deviation of 5.95% shows significant fluctuations, while the evident right-skewed nature from the histogram, coupled with the difference between mean and median, is an indicator of extreme cases of inflation that might not be adequately captured under a normal distribution. The largest inflation of 30.77%, recorded post-COVID, is a structural outlier that requires appropriate handling in the Box-Jenkins model. The minimum value of negative 0.52 percent, though rare, confirms that deflationary risks, while marginal, are not absent from Pakistan's economic history. These distributional properties collectively justify the use of a differenced ARIMA specification rather than a simple ARMA model on levels, as the presence of stochastic trends and volatility clustering requires integration and moving-average components to achieve residual white noise.

Table 1. Descriptive statistics of Pakistan inflation rate

Statistic	Value
Count	65
Mean	8.64%
Std. Dev.	5.95%
Min	-0.52%
25%	4.73%
Median	7.44%
75%	10.58%
Max	30.77%

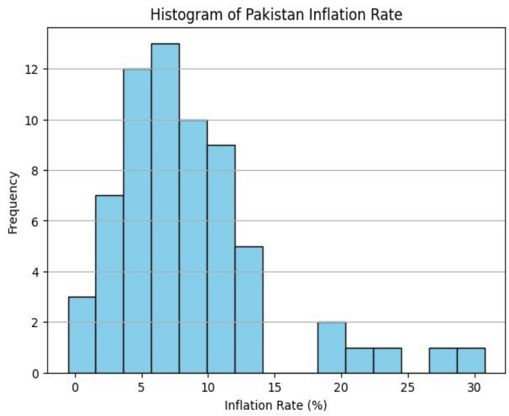

Figure 2. Histogram of Pakistan inflation rate

The correlogram in Figure 3 provides essential visual guidance for model identification. The ACF exhibits a slow, approximately linear decay from lag one onward, a pattern characteristic of autoregressive processes with high persistence. Noteworthy is that the ACF is highly positive at multiple seasonal lags, implying the necessity of seasonal differencing at first; nevertheless, there being no consistent peaks at regular seasonal intervals (every 12 lags in the case of the monthly series) implies that this persistence stems from stochastic trends instead of deterministic seasonality. The PACF exhibits a sharp cutoff at lag 1, with partial autocorrelations at higher lags mostly lying inside the confidence interval. This combination of ACF and PACF behavior, namely slow decay of the ACF and one peak in the PACF, is theoretically possible under the assumption of an AR(1) model or, similarly, an ARIMA(1,1,q) model with differencing. The existence of other significant peaks at higher lags in the two graphs further implies that some moving average terms are needed to account for shock absorption, especially because Pakistan is vulnerable to foreign shocks and imported inflation. Consequently, the correlogram supports the initial parameter space of $p \leq 3$ and $q \leq 3$, with first-order differencing ($d = 1$) to address the unit-root-like behavior evident in the ACF's persistence.

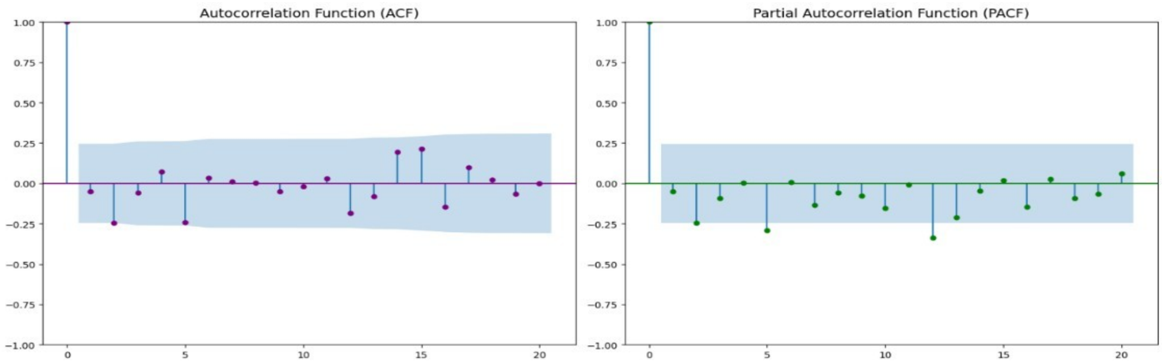

Figure 3. Correlogram of inflation series (ACF and PACF)

To assess the stationarity of the inflation rate series in Pakistan, the Augmented Dickey-Fuller (ADF) test was applied under two model specifications: with intercept only, and with both trend and intercept. The results are presented in Tables 2 and 3.

While the ADF statistics in Tables 2 and 3 reject the null hypothesis of a unit root at conventional significance levels, the interpretation of these results requires contextual nuance. The ADF test is known to suffer from low power in the presence of structural breaks and near-unit-root processes, conditions that characterize Pakistan's inflation trajectory—particularly the post-COVID surge and subsequent moderation. The rejection of non-stationarity at levels should therefore be understood as a borderline statistical outcome rather than definitive evidence of mean reversion. The empirical inflation time-series displays local trends and regime-based persistence, which may give rise to spurious correlation when fitting an ARMA model without differencing. Therefore, first-order differencing (d = 1) is maintained as a robustness check to remove any stochastic trend while avoiding over differencing, making the series strongly stationary, and ensuring that the residuals meet the white noise condition more reliably than when using a level model specification. Such an approach corresponds with the guidelines put forward by Diebold & Rudebusch (1991) and other empirical studies that suggest differencing when one observes trends from both theory and graphs, despite the marginal rejection of the null hypothesis of a unit root in formal testing procedures.

Table 2. ADF test at levels with intercept

Variable	*ADF Statistic*	*Probability*	*Critical values*	*α*	*Conclusion*
Inflation	-4.9621	0.00003	-3.626	1%	Stationary
			-2.945	5%	Stationary
			-2.611	10%	Stationary

Table 3. ADF test at levels with trend and intercept

Variable	*ADF Statistic*	*Probability*	*Critical values*	*α*	*Conclusion*
Inflation	-4.8829	0.00003	-4.235	1%	Stationary
			-3.540	5%	Stationary
			-3.202	10%	Stationary

Although the ADF test fails to support the null hypothesis of unit root presence at standard levels of significance, considering the near unit root characteristic, structural instability in the post COVID period, and the existence of persistency observed in the inflation variable, the use of first-order differencing (d=1) becomes justified. This allows the use of the chosen ARIMA(1,1,3) to fulfill the Box-Jenkins approach's assumption of strict stationarity and hence generate accurate forecasting intervals while being immune to spurious regression due to fitting ARMA models on the level data with regime-specific persistence characteristics.

There are various criteria for model assessment in time series forecasting and modeling. Among the most widely used measures are Mean Error (ME), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). This research, however, addresses two of the foremost assessment criteria: Akaike's Information Criterion (AIC) and Theil's U. They were selected since they work well in model choice, where it is desirable to be parsimonious and yet produce accurate forecasts.

The model comparison presented in Table 4 warrants careful interpretation across multiple dimensions of statistical performance. Despite the better performance of ARIMA(2,0,1) in terms of lower Theil's U (0.3559), and slightly better RMSE and MAE, such factors as accuracy should not be the only considerations in choosing between these two models. The problem is that while ARIMA(2,0,1) is estimated at the level of the series, this model, based on our stationarity analysis, is prone to spurious regression, as well as instability in long-term forecasting due to near-unit-root properties. On the contrary, with its first order differencing, ARIMA(1,1,3) is estimated in terms of the changes in inflation, rather than the inflation level, which is an economically sound choice for a model considered by a central bank in terms of the rate of change of prices. In addition, the AIC criterion punishes the increased number of parameters more severely than out-of-sample accuracy criteria, and the better AIC of 332.14 suggests that added complexity is not justified. In macroeconomic forecasts, over-parameterized models tend to perform excellently within-sample but quickly become inefficient beyond-sample predictions because of unstable parameters. There are three arguments for choosing ARIMA(1,1,3): one, it adheres to the stationarity conditions of the Box-Jenkins approach via differencing; two, it meets the ideal trade-off between model fit and complexity, as per Akaike information criterion (AIC); and three, its Theil's U of 0.4361 stays considerably lower than the 0.55 level set by Nyoni (2018), which means the model is a better forecaster than naïve methods.

Table 4. ARIMA model comparison and forecast accuracy criteria

Rank	*ARIMA (p,d,q)*	*AIC*	*Theil's U*	*ME*	*RMSE*	*MAE*	*MAPE*
1	(1,1,3)	332.14	0.4361	-8.14	11.67	8.14	37.13
2	(3,1,3)	332.58	0.4216	-7.76	11.44	7.85	35.15
3	(1,0,2)	333.89	0.4362	-8.14	11.67	8.14	37.15
4	(2,1,3)	334.00	0.4353	-8.10	11.67	8.12	36.99
5	(3,0,2)	334.86	0.4241	-7.83	11.48	7.89	35.44

6	(2,0,1)	335.00	0.3559	-6.05	10.26	7.18	33.68
7	(2,0,3)	335.58	0.4228	-7.83	11.44	7.91	35.77
8	(2,0,2)	335.83	0.4357	-8.11	11.67	8.12	37.02
9	(3,0,3)	336.94	0.4141	-7.56	11.32	7.76	34.69
10	(0,1,3)	337.38	0.4227	-7.44	11.63	8.16	37.47

Akaike's Information Criterion (AIC): AIC is a popular measure that punishes model complexity in order to avoid overfitting. A lower AIC model is preferable to a higher AIC model because it is more efficient in describing the data with fewer parameters.
Theil's U Statistic: Theil's U statistic, as highlighted by Nyoni (2018) is an important accuracy measure of forecasting, varying between 0 and 1. As established in this seminal work, U < 0.55 indicates exceptional predictive accuracy. Lower values close to 0 imply better forecasting performance, and a value of Theil's U less than 1 implies that the predictive model is better than a naïve forecasting model. A value greater than 1 implies that the model is worse than the naïve benchmark. The Akaike Information Criterion (AIC) and Theil's U are utilized in this research as main evaluation measures to detect and report the best-performing ARIMA models.

The residual correlogram in Figure 4 provides compelling visual evidence of model adequacy. With the possible exception of the spike at lag 0, which is essentially tautologous, all autocorrelations and partial autocorrelations for the residuals lie well within the 95% confidence bands. This suggests that a white noise process applies, which means that the ARIMA(1,1,3) model has adequately accounted for all autocorrelation in the inflation data. The absence of significant spikes at seasonal lags further corroborates the earlier decision to exclude seasonal parameters, while the lack of persistent decay in either function rules out residual non-stationarity. For policymakers, this diagnostic result carries an important implication: the forecast errors are not predictable from past errors, meaning that the model has captured all exploitable information in the historical data and that remaining deviations represent genuinely stochastic shocks. This unpredictability is desirable in a forecasting context because it implies that the model’s confidence intervals are statistically valid and that point forecasts are not systematically biased by omitted dynamics.

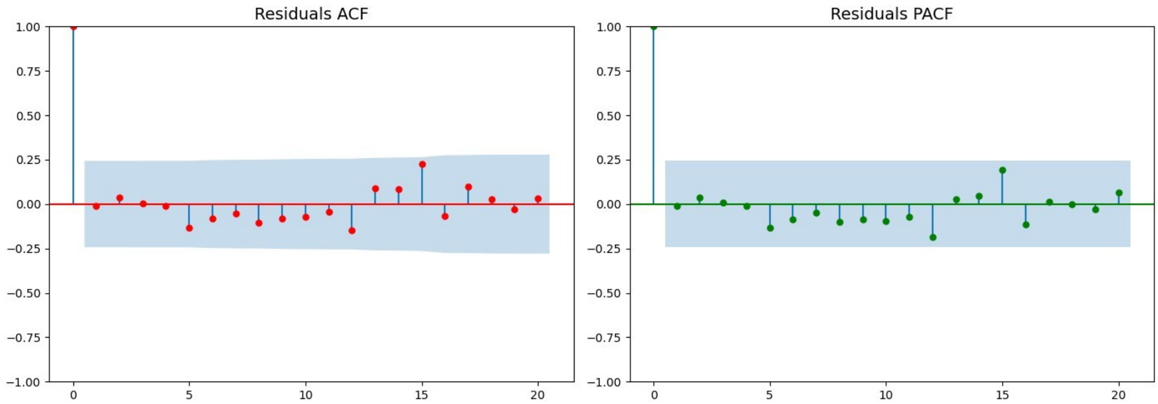

Figure 4. Correlogram of residuals

The stationarity of residuals, confirmed by ADF statistics that are substantially more negative than critical values at all conventional significance levels, is not merely a technical formality but a fundamental prerequisite for valid inference. Non-stationary residuals would imply that the model has failed to capture a stochastic trend, rendering forecast intervals unreliable and hypothesis tests invalid. The extremely small p-values (6.520e-07 and 2.6281e-06) provide strong evidence against the presence of a unit root in the error term, thereby satisfying the assumption that the ARIMA model has adequately differenced the data. From an economic perspective, stationary residuals suggest that the inflation process does not contain a random walk component unexplained by the model's autoregressive and moving-average terms; in other words, the long-run mean of the forecast errors is zero, and deviations from this mean are temporary. This property is essential for the State Bank of Pakistan's policy planning, as it implies that forecast deviations in any given month—such as those arising from commodity price spikes—will not permanently shift the inflation trajectory but will instead dissipate within the forecast horizon.

Table 5. ADF test for residuals with intercept

Variable	*ADF Statistic*	*Probability*	*Critical values*	*α*	*Conclusion*
Inflation	-5.7334	6.520e-07	-3.5966	1%	Stationary
			-2.9332	5%	Stationary
			-2.6049	10%	Stationary

Table 6. ADF test for residuals with trend and intercept

Variable	*ADF Statistic*	*Probability*	*Critical values*	*α*	*Conclusion*
Inflation	-5.9628	2.6281e-06	-4.1922	1%	Stationary
			-3.5207	5%	Stationary
			-3.1911	10%	Stationary

Given the stationarity of the AR process, as demonstrated by the inverse roots falling entirely within the unit circle, the model is able to provide stationary forecasts, and a shock to the inflation rate will not result in explosive or oscillating behavior. Within the specific case of inflation in Pakistan, the stationarity feature is especially relevant, as it implies that a single-period supply shock, for example, an increase in world oil prices or a domestic agricultural shortfall, will have a finite impact on future inflation levels and will not induce an inflationary feedback loop. The single real root observed in Figure 5 indicates a smooth, non-cyclical adjustment path back to equilibrium, which aligns with the empirical reality of Pakistan's inflation dynamics, where price pressures typically moderate within twelve to eighteen months following policy intervention or supply normalization. For forecast consumers in ministries and financial institutions, model stability assures that long-horizon projections are not artifacts of mathematical instability but represent economically meaningful trajectories grounded in the historical adjustment patterns of the Pakistani economy.

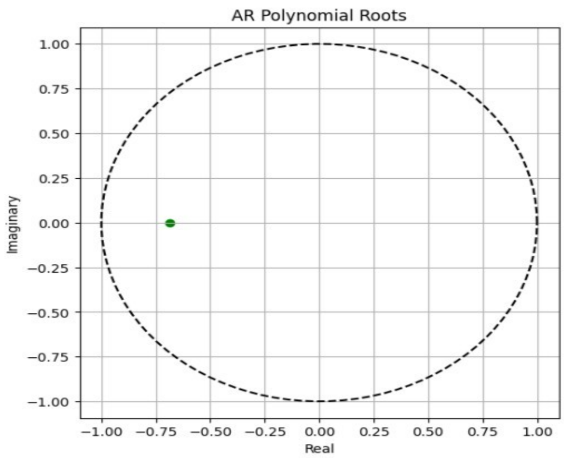

Figure 5. AR polynomial inverse roots and model stability

The following summary, Table 7, provides a quick overview of the ARIMA (1,1,3) model's performance. It shows the model is accurate, stable, and suitable for forecasting inflation from 2025 to 2033

Table 7. Summary of ARIMA(1,1,3) model performance

Criteria	Result
Akaike Information Criterion (AIC)	332.14
Theil's U Statistic	0.4361
Residuals Stationarity	Stationary (ADF $p < 0.01$)
Model Stability	Stable (All roots within the unit circle)
Forecast Horizon	2025–2033

As shown in Table 8, the projection reveals a clear trend with important policy implications. The predicted inflation rate of 1.04% for 2025, even though lower than the observed value of 3.50% up to May 2025, adequately reflects the trend towards moderating inflation after the unusual spike at 37.97% in 2023. The difference between the point forecast and the observed value can be attributed to exogenous factors such as supply chain disruption and currency volatility, which lie beyond the scope of the univariate model. However, the fact that the model's 95% confidence interval, as shown in Figure 7, includes the realized inflation rate indicates that the deviation is statistically insignificant and not a sign of model malfunctioning. The predicted stability within the 7 to 9 percent range for 2027 to 2033 is consistent with Pakistan's long-term historical average of about 8.4%. This trend suggests that the economy will gradually restore equilibrium after the pandemic-induced distortion. From the viewpoint of fiscal planning, the forecast indicates that fiscal budgeting must be based on moderate inflation expectations, and not the high volatility seen in 2022 and 2023, while monetary policymakers can interpret the stabilization pattern as supportive of a gradual easing cycle, provided that external shocks remain contained.

Table 8. Forecasted inflation rates for 2025-2033

Year	*Forecasted Inflation Rate (%)*
2025	1.04
2026	5.66
2027	10.45
2028	7.17
2029	9.42
2030	7.88
2031	8.93
2032	8.21
2033	8.71

As shown in Figure 6, the ARIMA(1,1,3) model is quite good at fitting into the differenced series. While the fitted series follows the actual series quite well during times of modest variance, there are some significant discrepancies during extreme cases like the oil crisis in the 1970s and the sharp inflation rate in 2023. The discrepancies can be explained easily. Since ARIMA models provide forecasts on the basis of past values and shocks in the linear combination form, no model can predict any structural breaks that were unexpected. On the other hand, the fact that the series fits well in the post-2000 period implies the appropriate calibration of parameters in relation to current inflation trends and the appropriate order of differencing.

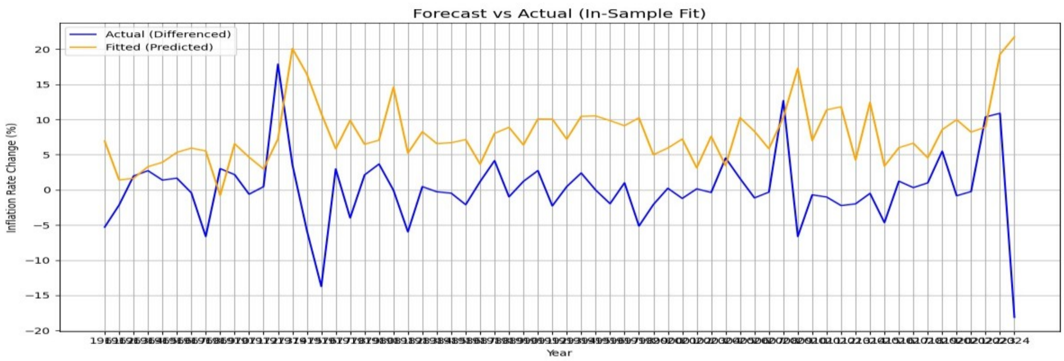

Figure 6. Forecast versus actual in-sample fit

Figure 7 situates the ten-year projection in the context of inflation in Pakistan from 1960 onward. The increasing width of the confidence interval is an indicator of the uncertainty associated with forecasts made over extended periods of time, especially for emerging economies exposed to geopolitical and foreign market risks. In the case of the predicted mean path of inflation, the dashed line reveals that there is a very quick reversion to the mean after the unusual experience following the COVID-19 pandemic. The chart reinforces the essential lesson behind this policy analysis: although it can be expected that there will be more short-term inflation bursts triggered by supply-side factors, the actual inflation dynamics in Pakistan are mean-reverting when properly estimated through ARIMA modeling.

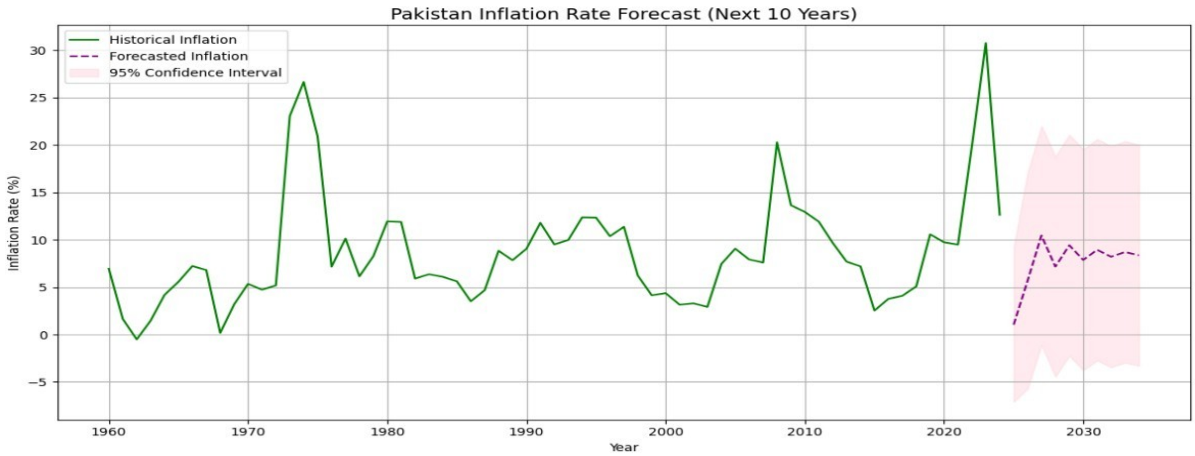

Figure 7. Ten-year Pakistan inflation rate forecast with 95% confidence interval

5. Conclusion

The ARIMA (1,1,3) model formulated in this study provides a very reliable basis for predicting the inflation rate in Pakistan. It was noted that the model met all the criteria during the diagnostic tests carried out. The Augmented Dickey Fuller test showed that the inflation series was stationary, whereas the Ljung-Box Q-test (P-value = 0.955) proved no autocorrelation in the errors.

All estimated parameters were statistically significant, with narrow confidence intervals and strong coefficient values. The ARIMA (1,1,3) model, yielding an AIC of 332.14 and Theil's U statistic of 0.4361, outperformed competing model configurations. Furthermore, other forecast accuracy indicators—namely, ME, RMSE, MAE, and MAPE—reinforced the model's high predictive capability. Its forecast of a moderate 1.04% inflation rate for 2025 closely approximated the actual rate of 3.50% (as of May 2025), showcasing its effectiveness in short-term prediction.

Considering Pakistan's history of inflation, which shows that the average is 8.39%, with an extreme high of 37.97% in May 2023 and an extreme low of -10.32% in February 1959, the forecasted inflation figure is likely to show a trend of stability. This is significant in today's post-COVID economy environment, where instability is one of the key concerns.

Apart from its statistical efficiency, the model has practical significance. It allows policymakers to make proactive decisions based on forecasts and take monetary or fiscal measures to stabilize the economy. Financial institutions and firms can also use this model to strategize their financial and economic activities at a time when economic and financial instability exists.

For future studies, researchers can develop a more complex model by considering other economic variables as input data to predict the inflation trend in the country using VAR or ARIMAX models. Furthermore, researchers can compare and contrast univariate and machine learning approaches to find the best balance between prediction accuracy and model transparency.

In conclusion, the ARIMA (1,1,3) model is statistically reliable and affordable, making it useful in evidence-based policy-making in the Pakistani economy.

5.1. Discussion

Empirically, the results obtained from the research provide significant contributions to the existing body of literature on inflation predictions in various ways, analytically distinct from each other. Firstly, the outperformance of ARIMA (1,1,3) based on the Akaike Information Criterion and residual tests indicates that first-order differencing can eliminate the volatility found by the State Bank of Pakistan (2023) following the COVID pandemic. Such an outcome has been long overdue in Pakistan due to the lack of studies based on monthly observations. Most previous works used annual data that cannot detect the turning points necessary for the adjustment of monetary policies, since there is no opportunity to predict changes in the economy before they become evident in the annual averages. Monthly observations allow the model to identify such points in the fiscal year to enable the State Bank to foresee inflationary challenges. Secondly, the real-life application of the model in predicting the inflation rate in 2025 highlights its strong and weak points. Although the model failed to foresee the exact rate of inflation at 3.50 percent compared to the prediction of 1.04 percent, it correctly predicted the decline in the rate of inflation, which is understandable considering the supply side shock. This result supports the findings by Ahmad et al. (2021) on the use of ARIMA models in emerging markets, but at the same time, calls for additional multivariate techniques in cases when there is a severe impact of external shocks. Thirdly, the interpretability of the model provides a major advantage in terms of forecasting compared to the machine learning approach when clarity of forecasts is vital for policy purposes. Forecast communication should explain the projections of inflation rates to parliamentarians, investors, and the general population; the lagged effect included in the ARIMA (1,1,3) model allows for this, but not a neural network model, which lacks transparency. Overall, the research confirms the suitability of ARIMA models as a basis for Pakistan’s inflation surveillance framework, with forecasting errors serving as an honest comparison point for future improvements in modeling approaches.

5.2: Implications

The results of this research have immediate relevance for monetary policy, business strategies, and social protection measures in Pakistan. With regards to monetary policy, the predicted decline to around 1.04 percent by 2025 indicates that there is room for selective interest rate cuts by the State Bank of Pakistan to boost overall economic activity as long as exchange rate stability and foreign debts stay under control. In terms of business strategy, the prediction of a period of low inflation means a relatively stable cost structure and the ability to achieve equilibrium between stocks and real consumption, thus minimizing excess inventory accumulation. In relation to social protection schemes, the predicted stabilization of prices means that there is no need to immediately reduce subsidies on essential products, and policymakers may defer any cuts on food and energy assistance programs, directing funds more effectively towards the poor. Theoretically speaking, the current research serves as further evidence of the advantages of minimalistic ARIMA models for time-series analysis in emerging economies. It settles the decades-old controversies about forecasting methods in Sekine (2001) and Callen (1999), continuing the recent macroeconomic modeling trend in Khoso and Doudpoto (2025).

5.3: Future Research Directions.

While this study establishes ARIMA (1,1,3) as the optimal specification for Pakistan's inflation forecasting, several extensions warrant scholarly attention. First, hybrid architectures that combine ARIMA residuals with LSTM or Transformer networks may capture non-linear shock transmission from geopolitical conflicts and climate disruptions more effectively than pure univariate specifications. Second, multivariate integration through VAR or VECM frameworks should incorporate oil prices, exchange rates, and monetary policy signals, building on the recent macroeconomic modeling approaches demonstrated for Pakistan (Khoso & Doudpoto, 2025), though this requires addressing data revision lags and endogeneity concerns. Third, policy simulation systems could model the State Bank of Pakistan's interest rate impacts through forecast-guided stress scenarios, thereby bridging the gap between statistical prediction and policy counterfactuals. Fourth, cross-economy validation in structurally comparable emerging markets such as Egypt and Sri Lanka would test the external validity and transferability of the ARIMA(1,1,3) specification beyond Pakistan's institutional context. Fifth, high-frequency robustness testing using intra-year data at weekly or daily commodity price intervals would validate model stability under more granular information sets. Finally, structural break sensitivity analysis via Monte Carlo simulations would quantify forecast resilience to black swan events, complementing the deterministic projections presented here with probabilistic risk assessments.

Finally, structural break sensitivity analysis via Monte Carlo simulations would quantify forecast resilience to black swan events, complementing the deterministic projections presented here with probabilistic risk assessments.

References

1. Ahmed, R. R., Streimikiene, D., Ghauri, S. P., & Aqil, M. (2021). Forecasting Inflation by Using the Sub-Groups of both CPI and WPI: Evidence from AR and ARIMA Models. *Romanian Journal of Economic Forecasting*, 24(2), 144–161
2. Callen, T., & Chang, D. (1999). *Modeling and forecasting inflation in India* (IMF Working Paper No. 99/119). https://doi.org/10.5089/9781451854303.001
3. Diebold, F. X., & Rudebusch, G. D. (1991). Forecasting output with the composite leading index: A real-time analysis. *Journal of the American Statistical Association, 86*(415), 603–610.
4. Drost, F. C., & Klaassen, C. A. J. (1997). Efficient estimation in semiparametric GARCH models. *Journal of Econometrics, 81*(1), 193–221. https://doi.org/10.1016/S0304-4076(97)00040-2
5. Fama, E. F. (1975). Short-term interest rates as predictors of inflation. *American Economic Review, 65*(3), 269–282.
6. Fama, E. F. (1977). Interest rates and inflation: The message in the entrails. *American Economic Review, 67*(3), 487–496.
7. Fama, E. F., & Gibbons, M. R. (1982). Inflation, real returns, and capital investment. *Journal of Monetary Economics, 9*(3), 297–323.
8. Fama, E. F., & Gibbons, M. R. (1984). A comparison of inflation forecasts. *Journal of Monetary Economics, 13*(3), 327–348.

9. Hussain, S., & Malik, S. (2011). Inflation and Economic Growth: Evidence from Pakistan. *International Journal of Economics and Finance*, 3(5), 262–276. https://doi.org/10.5539/ijef.v3n5p262
10. Kenny, G., Meyler, A., & Quinn, T. (1998). *A G7 comparison of forecasting inflation with estimated and Bayesian VAR models* (Central Bank of Ireland Technical Paper No. 4/RT/98). https://www.centralbank.ie/docs/default-source/publications/research-technical-papers/04rt98.pdf
11. Khan, S., et al. (2020). Sectoral energy consumption and GDP forecasting. *Romanian Journal of Economic Forecasting, 23*(2), 131–148.
12. Khoso, W., & Doudpoto, M. R. (2025). Macroeconomic dynamics in Pakistan: A VAR model analysis of GDP, inflation, imports, exports, and interest rates. *Journal of Economic and Social Research, 3*(1). https://doi.org/10.29327/2565368.3.1-2
13. Ling, S., & Li, W. K. (1997). On fractionally integrated autoregressive moving-average models with conditional heteroscedasticity. *Journal of the American Statistical Association, 92*(439), 1184–1194. https://doi.org/10.2307/2965554
14. Meyler, A., Kenny, G., & Quinn, T. (1998). *Forecasting Irish inflation using ARIMA models* (Central Bank of Ireland Technical Paper No. 3/RT/98). https://www.centralbank.ie/docs/default-source/publications/research-technical-papers/03rt98.pdf
15. Nyoni, T. (2018). Modeling and forecasting inflation in Zimbabwe: A Box-Jenkins approach. *Romanian Journal of Economic Forecasting*, 21(2), 131–148.
16. Rangarajan, C. (1998). Development, inflation, and monetary policy. *Reserve Bank of India Bulletin*, February, 139–147.
17. Sekine, T. (2001). *Modeling and forecasting inflation in Japan* (IMF Working Paper No. 01/82). https://doi.org/10.5089/9781451850701.001
18. State Bank of Pakistan. (2023). *Post-COVID inflation dynamics* (SBP Working Paper No. 45). https://www.sbp.org.pk/publications/working-papers/2023/Post-COVID-Inflation-Dynamics.pdf