

International Journal of Statistics and Applied Mathematics

ISSN Online: 2456-1452
NAAS Rating (2026): 4.49
IJSAM 2026; 11(10): 26-29
www.mathsjournal.com
Received: 14-08-2026
Accepted: 26-09-2026
Published: 07-10-2026

Nguyen Hong Nhat
Faculty of Mathematical
Economics, College of
Technology, National
Economics University, Ha Noi,
VietNam

Analysis and forecasting of the consumer price index for food and catering services in Vietnam using an ARIMA model

Nguyen Hong Nhat¹

DOI: <https://www.doi.org/10.22271/math.2026.v11.i10a.2713>

ABSTRACT

This study analyses and forecasts Vietnam's consumer price index (CPI) for food and catering services using 180 monthly observations from January 2010 to December 2024. Following the Box–Jenkins procedure, the series was examined for stationarity, candidate autoregressive integrated moving-average models were estimated, and the preferred specification was assessed through residual diagnostics and short-horizon forecast errors. The augmented Dickey–Fuller statistic of -7.577747 ($p < 0.001$) indicates that the reported index series is stationary in levels. Among the retained candidates, ARIMA(12,0,1) has the lowest Akaike, Schwarz, and Hannan–Quinn information criteria and the smallest regression standard error. Its forecasts for June–December 2024 differ from actual values by only 0.01–0.56%. Forecasts for 2025 remain within 100.12–100.74, implying limited month-to-month variation under the continuation of historical dynamics. Residuals are stationary and show no autoregressive conditional heteroskedasticity, but short-lag autocorrelation and non-normality remain; consequently, the forecasts should be treated as a baseline rather than as a complete representation of food-price shocks. The results provide concise quantitative evidence for short-term price monitoring, supply planning, and inflation surveillance in Vietnam.

Keywords: ARIMA, Consumer price index, Food and catering services, Inflation, Forecasting

Introduction

The consumer price index (CPI) is a central measure of changes in household living costs and inflation. In Vietnam, food and catering services form a major CPI group because they cover staple food, other food products, and meals consumed away from home. Their prices directly affect household purchasing power and can transmit supply, transport, fuel, weather, disease, and seasonal shocks to headline inflation. In 2024, prices in this group rose by 3.87% and contributed 1.30 percentage points to overall CPI growth.¹

Autoregressive integrated moving-average (ARIMA) models are widely used for short-term macroeconomic forecasting because they exploit the serial dependence in a variable's own history. International applications have forecast CPI in China, Jordan, and Zambia, as well as food-price inflation in Malaysia.^{2–5} Vietnamese studies have also applied ARIMA to headline CPI and inflation, obtaining different lag structures across sample periods.^{6–9} These studies confirm the usefulness of the Box–Jenkins framework, but most concern aggregate CPI or inflation rather than the food and catering services component.

This article addresses that narrower empirical gap using monthly data for 2010–2024. Its objectives are to describe the series, determine its order of integration, select a parsimonious ARIMA specification, evaluate residual behaviour and forecast accuracy, and generate a 12-month baseline for 2025. The contribution is practical: food-related CPI is both socially salient and exposed to shocks that may not be visible in an aggregate price index.

2. Materials and Methods

2.1. Data

The study uses 180 monthly observations on Vietnam's CPI for food and catering services from January 2010 through December 2024, obtained from the General Statistics Office of Vietnam (GSO). The CPI is compiled from a representative basket and fixed expenditure weights for a specified base period.^{10,11} The sample spans post-crisis recovery, the COVID-19 disruption, global food and energy-price volatility, and changes in domestic price administration. Estimation and diagnostic testing were performed in EViews.

Corresponding Author:
Nguyen Hong Nhat
Faculty of Mathematical
Economics, College of
Technology, National
Economics University, Ha
Noi, VietNam

2.2. ARIMA specification and model selection

For a time series y_t , an ARIMA(p, d, q) model applies d differences and represents the resulting stationary series using p autoregressive and q moving-average terms. In compact form,

$$\phi(B)(1-B)^d y_t = c + \theta(B)\varepsilon_t,$$

where B is the backshift operator, $\phi(B)$ and $\theta(B)$ are the autoregressive and moving-average polynomials, and ε_t is an innovation. Model construction followed the Box–Jenkins cycle of identification, estimation, diagnosis, and forecasting.^{12,13}

Stationarity was tested with the augmented Dickey–Fuller (ADF) test. The autocorrelation function (ACF) and partial autocorrelation function (PACF) then guided candidate lag selection. Specifications with statistically insignificant principal coefficients were removed, and the remaining models were compared using adjusted R^2 , regression standard error, Akaike information criterion (AIC), Schwarz criterion (SC/BIC), and Hannan–Quinn criterion (HQ). Lower information criteria and standard error were preferred, subject to diagnostic adequacy.

Residual assessment comprised: (i) correlogram and Ljung–Box Q-statistics for serial correlation; (ii) an ADF test for residual stationarity; (iii) an ARCH test for conditional heteroskedasticity; and (iv) the Jarque–Bera test for normality. Forecast performance was evaluated from June to December 2024 using the absolute percentage error,

$$APE_t = \left| \frac{y_t - \hat{y}_t}{y_t} \right| \times 100.$$

The reported errors are complemented by qualitative comparison of actual and fitted movements; this avoids presenting in-sample fit as

definitive evidence of performance under future structural shocks.

3. Results and Discussion

3.1. Stationarity and model identification

Visual inspection showed monthly fluctuations around a stable level rather than a persistent deterministic trend. The ADF result in Table 1 rejects the unit-root null at the 1% level. The series was therefore modelled in levels, so $d = 0$.

The PACF was prominent at lags 1 and 13, while the ACF was strongest at lag 1, decayed over lags 2–3, and displayed additional seasonal signals around lags 12 and 18. These patterns motivated combinations involving autoregressive lags 1, 12, and 13 and moving-average lags 1, 2, 3, 12, and 18. After excluding candidates with insignificant principal terms, five specifications remained.

3.2. Model estimation and selection

Table 2 shows that ARIMA(12,0,1) performs best on every reported comparison criterion: it has the highest adjusted R^2 , the lowest standard error, and the lowest AIC, SC, and HQ values.

Table 1. Augmented Dickey–Fuller test for the CPI series

Statistic	Value	p-value
ADF test statistic	-7.577747	0.0000
Critical value, 1%	-3.466994	
Critical value, 5%	-2.877544	
Critical value, 10%	-2.575381	

Note: Null hypothesis: the series has a unit root. Source: authors' calculations in EViews.

The improvement over ARIMA(1,0,12) is modest, but consistent. Higher-order alternatives fit materially worse.

Table 2. Comparison of retained ARIMA specifications

Model	Adjusted R^2	S.E.	AIC	SC	HQ
ARIMA(1,0,12)	0.290002	0.814794	2.454158	2.525113	2.482927
ARIMA(12,0,1)	0.295070	0.811881	2.449391	2.520345	2.478160
ARIMA(12,0,2)	0.066102	0.934478	2.726462	2.797416	2.755231
ARIMA(12,0,3)	0.049832	0.942582	2.743870	2.814825	2.772639
ARIMA(12,0,12)	0.160323	0.886085	2.708046	2.779001	2.736815

Note: S.E.: regression standard error. Source: authors' calculations in EViews.

All reported coefficients of the selected model are statistically significant (Table 3). Using the parameterization reported by EViews, the

fitted relationship is summarized as

$$X_t = 100.4201 + 0.251225X_{t-12} + 0.550072\varepsilon_{t-1} + \varepsilon_t,$$

where X_t is the food and catering services CPI at month t . The lag-12 term captures annual dependence, while the first-order moving-average

term absorbs short-run innovation persistence.

Table 3. Estimated ARIMA(12,0,1) coefficients

Parameter	Coefficient	Standard error	t-statistic	p-value
Constant	100.420100	0.145162	691.7800	0.0000
AR(12)	0.251225	0.055958	4.489520	0.0000
MA(1)	0.550072	0.073317	7.502637	0.0000
σ^2	0.644502	0.043408	14.84755	0.0000

Note: Source: authors' calculations in EViews.

(continued on next page)

3.3. Residual diagnostics

The diagnostics give mixed evidence (Table 4). The residual ADF statistic is -12.22900 ($p < 0.001$), confirming residual stationarity. The ARCH test does not reject constant conditional variance: the F -test probability is 0.3769 and the $\text{Obs.} \times R^2$ probability is 0.3687 . These results support the stability of the residual variance.

However, correlogram Q -statistic probabilities at lags 3, 4, and 5 are 0.005 , 0.014 , and 0.031 , respectively. Thus, short-lag dependence has not been eliminated completely, although probabilities exceed 0.05 from lag 6 onward. Residuals are also non-normal: skewness is 0.706656 , kurtosis is 6.394018 , and the Jarque–Bera statistic is 101.3761 ($p < 0.001$). ARIMA(12, 0, 1) is therefore useful as a compact forecasting baseline, but it is not a fully adequate stochastic description of extreme or asymmetric price movements.

Table 4. Summary of residual diagnostics

Diagnostic	Result	Interpretation
Residual ADF	-12.22900 ; $p = 0.0000$	Residuals are stationary
ARCH LM	$F = 1.084861$; $p = 0.3769$	No evidence of ARCH effects
$\text{Obs.} \times R^2$ ARCH	10.85828 ; $p = 0.3687$	Constant variance not rejected
Residual correlogram	$p < 0.05$ at lags 3–5	Some short-lag autocorrelation remains
Jarque–Bera	101.3761 ; $p = 0.0000$	Residual normality is rejected

Note: Source: authors' calculations in EViews.

3.4. Forecast evaluation and 2025 outlook

The holdout-style comparison for June–December 2024 is shown in Table 5. Absolute percentage errors range from 0.01% in October to 0.56% in November. The mean of the seven reported absolute per-

centage errors is approximately 0.30% . The fitted series follows normal month-to-month movements closely, but it tends to understate abrupt peaks and troughs, as expected from a univariate model that has no direct information on harvest conditions, fuel and transport costs, disease outbreaks, or policy interventions.

Table 5. Short-term forecast evaluation, June–December 2024

Month	Actual	Forecast	Error	Month	Actual	Forecast	Error
Jun.	100.75	100.27	0.48%	Oct.	100.55	100.56	0.01%
Jul.	100.26	100.74	0.48%	Nov.	99.78	100.34	0.56%
Aug.	100.27	100.25	0.02%	Dec.	99.87	100.04	0.17%
Sep.	100.92	100.51	0.41%				

Note: Error is the absolute percentage difference reported in the original estimation output. Source: authors' calculations in EViews.

The 2025 baseline forecasts remain near 100 throughout the year (Table 6). The maximum is 100.74 in February and the minimum is 100.12 in March, a range of 0.62 index points. This profile implies no

sustained acceleration under historical dynamics. It should not be read as a structural inflation projection because the model cannot anticipate exogenous supply disruptions.

Table 6. Forecast CPI for food and catering services in 2025

Month	CPI	Month	CPI	Month	CPI	Month	CPI
Jan.	100.28	Apr.	100.28	Jul.	100.38	Oct.	100.45
Feb.	100.74	May	100.41	Aug.	100.38	Nov.	100.26
Mar.	100.12	Jun.	100.50	Sep.	100.55	Dec.	100.28

Note: Source: authors' ARIMA(12, 0, 1) forecasts in EViews.

The results are broadly consistent with prior evidence that ARIMA can provide accurate short-horizon CPI baselines.^{2,3,8} Nevertheless, the remaining residual autocorrelation and heavy-tailed errors are consequential for food prices, which are especially exposed to shocks. Future work should compare seasonal ARIMA, intervention models, and multivariate specifications incorporating fuel prices, transport costs, weather, exchange rates, and food-supply indicators. Forecast evaluation should also use a longer rolling-origin test period.

4. Conclusion

Monthly food and catering services CPI in Vietnam is stationary in levels over 2010–2024 according to the reported ADF test. Of five retained candidates, ARIMA(12, 0, 1) provides the strongest combination of fit and parsimony, with AIC of 2.449391 , SC of 2.520345 , and HQ of 2.478160 . Forecast errors for June–December 2024 are small, and the 2025 baseline remains in a narrow 100.12 – 100.74 interval.

The model is informative but not definitive. Residual stationarity and the absence of ARCH effects support its use, whereas autocorrelation at lags 3–5 and pronounced non-normality limit statistical inference and the treatment of extreme shocks. Authorities may therefore use the forecasts as one input to monthly monitoring, accompanied by real-time information on crop supply, inventories, transport and fuel costs, weather, disease, and seasonal demand.

For policy and business planning, the central implication is preparedness rather than mechanical reliance on a point forecast. Public agencies can strengthen early-warning systems and coordinate essential-food supply around holidays and adverse weather. Food producers, distributors, and catering firms can use the baseline for procurement and inventory planning while maintaining contingency suppliers and financial buffers. These measures would reduce the pass-through of temporary supply disruptions to household living costs.

References

1. General Statistics Office of Vietnam. Press release on price developments in December, the fourth quarter and 2024. Hanoi: GSO; 2025.
2. Dongdong W. A consumer price index forecast based on ARIMA model. In: 2010 WASE International Conference on Information Engineering. 2010:307-310.
3. Ghazo A. Applying the ARIMA model to the process of forecasting GDP and CPI in the Jordanian economy. *Int J Financ Res*. 2021;12(3):70.
4. Jere S, Banda A, Chilyabanyama R, Moyo E. Modeling consumer price index in Zambia: a comparative study between multicointegration and ARIMA approach. *Open J Stat*. 2019;9(2):245-257.
5. Mohamad Rosni Q, Othman MR. Time-series modelling of food security inflation in Malaysia using an ARIMA model as a machine learning approach. *J Marit Logist*. 2022:62-78.
6. Du N. Application of the ARIMA model to forecast the consumer price index in Vietnam. 2017.
7. Duy VQ, Au HH. Forecasting Vietnam's inflation for August 2013–July 2014. *Can Tho Univ J Sci*. 2014;30:34-41.
8. Huyen NTT, Son NT. Forecasting the consumer price index in Vietnam using an ARIMA model. Hanoi: Academy of Finance; 2022.
9. Banking Review. Application of an ARIMA model to forecast Vietnam's inflation and policy recommendations. *Banking Review*. 2023.
10. General Statistics Office of Vietnam. Method for calculating the consumer price index: closely reflecting consumer-price movements in the market. Hanoi: GSO; 2021.
11. General Statistics Office of Vietnam. Consumer-price survey plan and CPI weights for the 2020–2025 period. Hanoi: Ministry of Planning and Investment; 2023.
12. Box GEP, Jenkins GM, Reinsel GC, Ljung GM. Time series analysis: forecasting and control. 5th ed. Hoboken: Wiley; 2015.
13. Hyndman RJ, Athanasopoulos G. Forecasting: principles and practice. 3rd ed. Melbourne: OTexts; 2021.

How to cite this article: Nguyen HN. Analysis and forecasting of the consumer price index for food and catering services in Vietnam using an ARIMA model. *International Journal of Statistics and Applied Mathematics*. 2026;11(10):26-29. doi: <https://www.doi.org/10.22271/math.2026.v11.i10a.2713>