Proceedings of the 8th IEOM Bangladesh International Conference on Industrial Engineering and Operations Management, Dhaka, Bangladesh, December 20-21, 2025

Proceedings of the International Conference on Industrial Engineering and Operations Management

Publisher: IEOM Society International, USA DOI: 10.46254/BA08.20250178
Published: December 20, 2025

Forecasting Onion Prices in Indian Cities: A Comparative Analysis of Time Series Models

Syed Tahmid Jamil
Department of Industrial Engineering
BGMEA University of Fashion and Technology
Dhaka, Bangladesh
tahmid.jamil@buft.edu.bd

Mashiat Iqbal
Department of Statistics and Data Science
Jahangirnagar University
Dhaka, Bangladesh
mashiatiqbal12@gmail.com

Abstract

Forecasting perishable commodity prices is inherently difficult due to strong volatility and intricate seasonal patterns. Therefore, forecasting models should effectively capture the underlying structure of the data. This study focuses on forecasting daily onion prices across three key Indian cities using open-source data from 2020 to 2024. The dataset includes daily observations of minimum, maximum, modal, and retail prices, along with market arrival quantities. The modal price, representing the most frequent transaction value, was selected as the primary forecasting variable. Three univariate time series models were applied to assess their forecasting performance. Univariate models were used to isolate forecasting performance and evaluate the effectiveness of decomposition in stabilizing highly volatile daily price series. These models include Seasonal Autoregressive Integrated Moving Average (SARIMA), Seasonal and Trend decomposition using Loess combined with ARIMA (STL-ARIMA), and STL combined with Exponential Smoothing State Space models (STL-ETS). Model performance was evaluated using standard accuracy metrics such as RMSE, MAE and MAPE. Results show that the STL-ETS model consistently achieved the lowest forecast errors across all three cities, outperforming both SARIMA and STL-ARIMA. The findings highlight the suitability of decomposition-based models for capturing high-frequency variations in agricultural price data. This research establishes an evidence-based comparison of univariate forecasting models and demonstrates that decomposition-based methods offer superior accuracy for highly volatile perishable price series.

Keywords
Onion price forecasting, Time series analysis, STL decomposition, Model prices.

1 Introduction
Precise forecasting of prices within the agricultural sector is vital for ensuring economic stability and food security. It is also crucial for governments and supply chain stakeholders to be able to make timely decisions. In developing economies, where a significant portion of household income is allocated to essential food items, even minor fluctuations in the prices of staples can exert substantial impacts on society and the economy. Among all these food products, onions hold significant importance in South Asian households due to their frequent use in daily cuisine. However, their prices remain highly volatile due to fluctuations in their product, particularly when weather conditions are unpredictable and adequate storage and transportation options are lacking. Additionally, they are

accessible only during specific periods of the year. India is among the largest producers, consumers, and exporters of onions. When prices rise abruptly, it often leads to undesirable consequences such as political instability, increased media scrutiny, and the implementation of emergency policy measures including export restrictions, buffer stock adjustments, or import relaxations. These factors demonstrate the difficulty and significance of understanding fluctuations in onion prices.

Individuals have employed ARIMA and SARIMA to analyze historical commodity prices, based on the assumption that patterns in prior data exhibit minimal variation. But daily price variations are more affected by short-term events than weekly or monthly data. Because of this they become harder to predict. However, it is now possible to look into more effective methods because high-frequency datasets are becoming more common. Researchers can use decomposition-based models to breakdown a time series into components that show trends, seasons and anomalies. This makes it easier to understand and could help make predictions more precise for things that change a lot.

It is difficult to predict the cost of onions for each day, even though they are essential. This is because the market is always changing and there are different types of markets where things can happen rapidly. Most of the study that has been done till now has focused on trends at the national level instead of differences between cities. This indicates that there are limited methods available for forecasting events in specific cities. This study aims to fill that gap by examining daily model price data from 2020 to 2024 for three significant Indian markets like- Hyderabad, Mumbai and Ahmedabad. The study uses the SARIMA, STL-ARIMA and STL-ETS models in R to check how accurate the forecasts are. It does this by using RMSE, MAE and MAPE to find the best methods in different market situations.

Policymakers who want to be able to predict short-term price spikes and plan timely interventions will find the results of this study useful. If traders, wholesalers, and supply chain planners have better predictions, they can also make better choices about how to handle their stock, when to buy new items and how to get them to customers. The study improves our understanding of methodologies by focusing how decomposition-based univariate time series models can work with irregular, high-frequency agricultural data and how city-specific factors can affect forecasting performance.

1.1 Objectives

1. Apply and compare SARIMA, STL-ARIMA, and STL-ETS models to forecast daily onion prices in selected Indian cities.
2. Evaluate model performance using standard forecasting metrics.
3. Identify the most accurate and consistent approach for practical price prediction.

2. Literature Review

Since agricultural markets are so unstable, accurate forecasting is essential. This is particularly applicable for onions in South Asia, which are politically important and very unstable. They are affected by the weather, supply chain problems and frequent price changes that affect households and legislative choices. But it is hard to predict agricultural prices because of the way they operate. Due to inconsistent supply and the fragility of onions, prices may fluctuate rapidly in reaction to shifts in demand or supply chain disruptions. Any forecasting method must take into consideration significant and repetitive seasonal components. Onions have certain seasonal production cycles, mostly during the harvest seasons. External shocks, such as problems with transportation, extreme weather and sudden changes in policy make it even harder to make predictions. These make the changes happen in a way that regular models can't easily show. Because of differences in supply chain efficiency, storage capacity and consumption habits, a model that works well in a particular environment might not work well in another. This also causes significant pricing variance between different areas. Furthermore to these structural problems, many developing countries still have trouble accessing data. Traditional statistical models are inadequate for dealing with the unreported fluctuations caused by an important number of unrecorded transactions, particularly in rural markets. Previous study has highlighted these limitations, for example, Zhang and Na (2018) emphasized the need for adaptable forecasting models that can accommodate uncertainty and nonlinear characteristics in unstable market conditions, while Saxena et al. (2020) illustrated the distribution of fluctuations and price variations across various regions. Singh et al. (2022) recommend the forecasting methods applied in this study by offering customized modeling solutions that accurately incorporate regional supply and demand fluctuations.

Traditional time series models have been the foundation for predicting agricultural prices for many years. ARIMA is a traditional model that uses autoregressive, differencing and moving average parts. It is still widely used because it is good at showing trends and autocorrelation. SARIMA, its periodic expansion, adds more seasonal conditions to make it work for commodities that have cycles that repeat every year or month. Chuan-xi (2009) and Oliveira and Teles (2022) have shown that ARIMA and SARIMA work well for predicting agricultural prices, especially when the data is mostly trend-based. But ARIMA-based methods may not work well in the real world because agricultural time series often show unpredictable behavior, unexpected shocks and changing seasonal patterns.

Decomposition-based models have become more popular as an approach to overcome these drawbacks. The ETS (Error, Trend, Seasonal) framework makes predictions by repeatedly modifying its parts. This helps the model adapt to changes in economic behavior or changes in policy. ETS is especially useful when data show seasonal patterns that change gradually, which is common in agricultural markets. STL (Seasonal–Trend Decomposition using Loess) is another method that uses a flexible, non-parametric way that is more effective in finding fluctuating seasonal patterns than traditional decomposition methods. STL is reliable for working with large-scale agricultural datasets where seasonal patterns can change in size or shape from day to day. Studies by Dokumentov and Hyndman (2020) and Ugebe et al. (2025) have proven that STL works effectively for modeling complex seasonality and enhancing the accuracy of predictions in hybrid models.

Hybrid methodologies are now recognized as a practical improvement in forecasting, since no particular model can completely represent all components of a complex time series. These models utilize decomposition techniques such as STL. To identify both structural and residual anomalies predictive models like ARIMA or ETS are used. Gumani et al. (2017) demonstrated significant improvements in sales forecasting through hybrid models that integrate decomposition and statistical learning. While several recent studies explore machine-learning-based hybrids for agricultural markets, much of the evidence emphasizes that ML models require large datasets, extensive feature engineering, and stable seasonal structure to outperform statistical methods—conditions not always met in daily agricultural price series. Research by Makridakis et al. (2018) showed that in many univariate forecasting settings, machine learning models fail to consistently outperform classical statistical approaches due to overfitting and challenges in learning long-term seasonality from limited data. Furthermore, Similarly, Hewamalage, Bergmeir, and Bandara (2021) found that neural network architectures such as LSTM struggle with long seasonal periods unless seasonal features are explicitly engineered, making decomposition-based statistical models a more robust choice for smaller, noisy datasets like daily onion prices. Therefore, classical decomposition driven approaches remain a strong baseline in agricultural forecasting, particularly when data availability is constrained. Their interpretability and minimal data requirements further reinforce their relevance in real-world market environments where comprehensive feature sets are rarely available.

Overall, the literature shows a clear shift from traditional ARIMA-based forecasting toward decomposition-based approaches that improve accuracy by separating the series into trend, seasonal, and irregular components. Onion prices are strongly influenced by seasonal patterns, external shocks, regional differences, and unpredictable market movements, making them difficult to model with a single technique. Decomposition-based forecasting frameworks such as STL-ARIMA and STL-ETS offer a suitable structure for handling both stable seasonal effects and more irregular or nonlinear behavior in the data. This shift in methodology highlights the need for flexible, robust, and city-specific time series models, which directly supports the objectives of the present study.

3. Methodology

The analysis was conducted in the R programming environment due to its mature ecosystem for time series modeling and its extensive collection of forecasting packages such as forecast, fable, and tsibble. All computations, diagnostics, and visualizations adhered to reproducible workflows. Before model development, the daily onion price datasets for Ahmedabad, Hyderabad, and Mumbai were examined for missing observations and outliers. The tsclean() function from the forecast package was used to treat anomalous values by applying robust interpolation techniques. This ensured that the input series maintained continuity while preserving the underlying structure essential for seasonal and trend-based models. A preliminary time plot Figure 1, Figure 2 and Figure 3 revealed strong annual seasonality and volatility across cities.

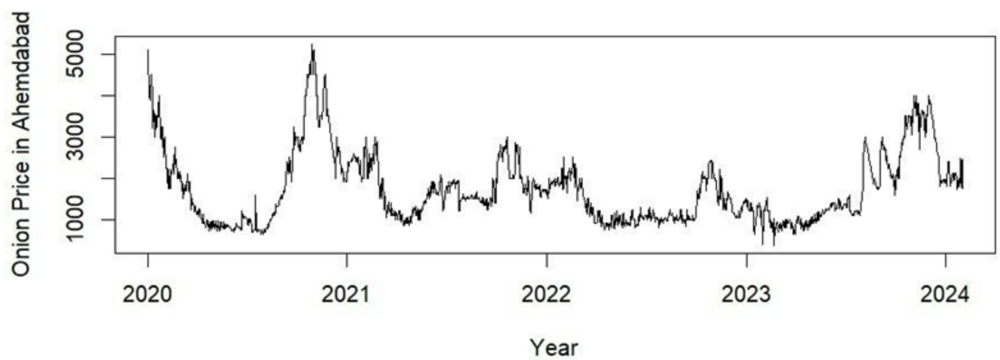

Figure 1. Time plot of onion Price of Ahmedabad.

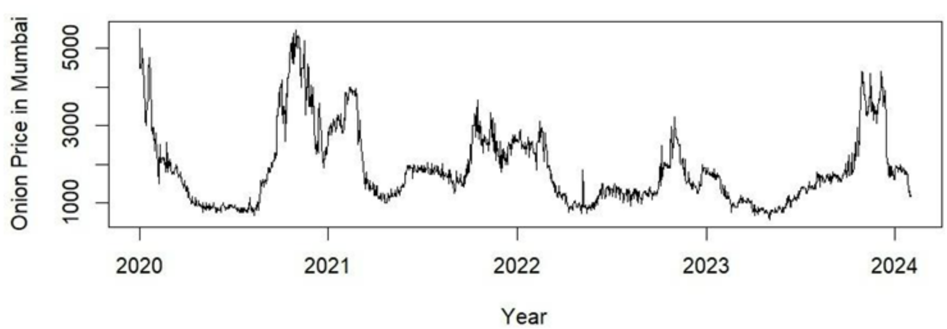

Figure 2. Time plot of onion Price of Hyderabad.

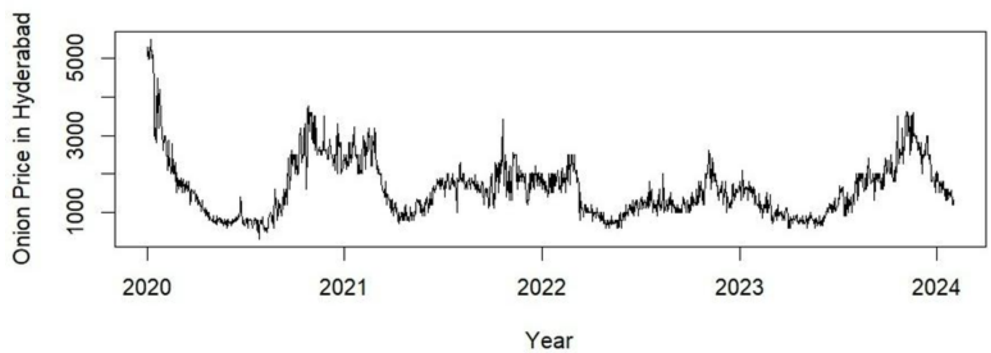

Figure 3. Time plot of onion Price of Mumbai.

3.1 Model development
To ensure a comprehensive understanding of onion price dynamics, multiple forecasting frameworks were developed and evaluated using both classical time-series models and decomposition-based hybrid approaches. The following subsections outline the modeling procedures applied to each method, including SARIMA, STL-ARIMA, and STL-ETS.

3.1.1 Seasonal ARIMA (SARIMA):
Seasonal Autoregressive Integrated Moving Average (SARIMA) models were developed to capture both the short-term dynamics and the annual cyclical behavior present in daily onion price series. Prior to model estimation, the stationarity properties of each city's dataset were examined through Augmented Dickey–Fuller (ADF) and KPSS tests. These diagnostics indicated the need for both non-seasonal and seasonal differencing, motivating the inclusion of a seasonal periodicity of 365 to reflect agricultural cycles, weather-driven supply fluctuations, and annual consumption patterns.

Model identification followed a systematic approach in which candidate autoregressive and moving-average orders were explored using a combination of autocorrelation structures, partial autocorrelation behavior, and automatic selection via AICc minimization. The auto.arima() function in R assisted in generating a set of competitive models.

Based on this selection process, SARIMA(3,1,0)(0,1,0) 365 was identified as the best-fitting model for Ahmedabad, while Hyderabad's price behavior aligned most closely with a SARIMA(1,1,1)(0,1,0) 365 structure. For Mumbai, the SARIMA(1,1,2)(0,1,0) 365 model provided the strongest balance of fit and residual quality. These three models collectively reflect the variability in city-specific price dynamics while consistently emphasizing the importance of annual seasonality in shaping onion market behavior across India (Table 1).

Table 1. Selected SARIMA model for each city

City under observation	Model (SARIMA)
Ahmedabad	ARIMA (3,1,0) (0,1,0) [365]
Hyderabad	ARIMA (1,1,1) (0,1,0) [365]
Mumbai	ARIMA (1,1,2) (0,1,0) [365]

1.1.1 Utilization of decomposition method
To better capture the complex dynamics embedded in daily onion prices, the Seasonal–Trend decomposition using Loess (STL) was applied to the datasets for Ahmedabad, Hyderabad, and Mumbai. Daily agricultural price series often exhibit nonlinear seasonal variations and irregular fluctuations driven by weather conditions, storage limitations, transport delays, and shifting supply–demand pressures. STL is well-suited for such data because it adaptively separates the series into three components, trend, seasonal, and remainder, using locally weighted smoothing. This decomposition offers greater flexibility than classical methods, allowing the seasonal pattern to evolve over time rather than remain fixed.

By decomposing the data beforehand, we gain a clearer understanding of the dominant long-term movements and recurring annual cycles present in each city's market. The resulting STL plots (Figures 4–6) highlight meaningful contrasts across cities, particularly in the strength and stability of seasonal effects and the variability of underlying trends. These decomposed components form the basis for the subsequent hybrid modeling approaches, where ARIMA and ETS models are applied to the seasonally adjusted series to enhance forecasting performance.

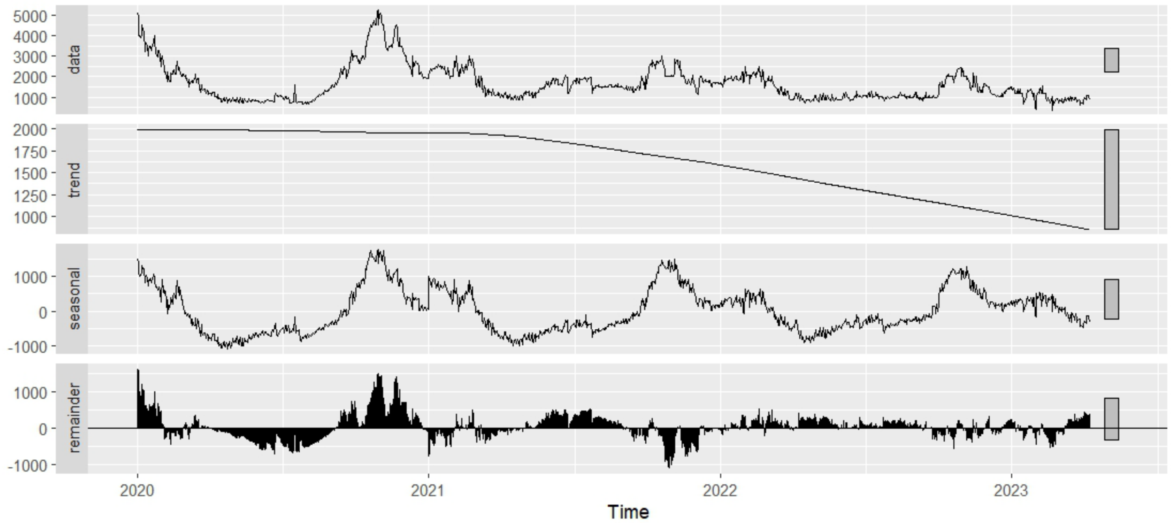

Figure 4. STL Decomposed data of onion price of Ahmedabad

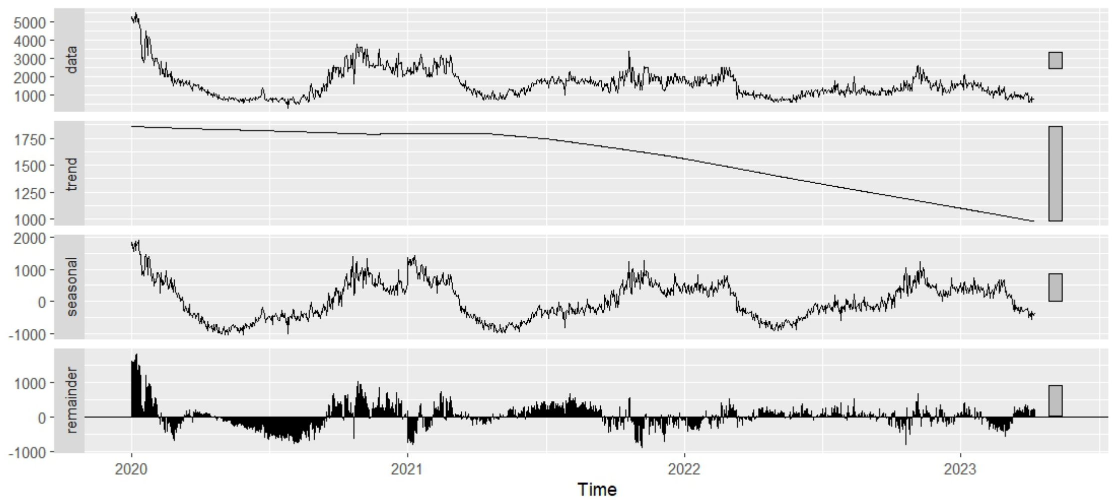

Figure 5. STL Decomposed data of onion price of Hyderabad

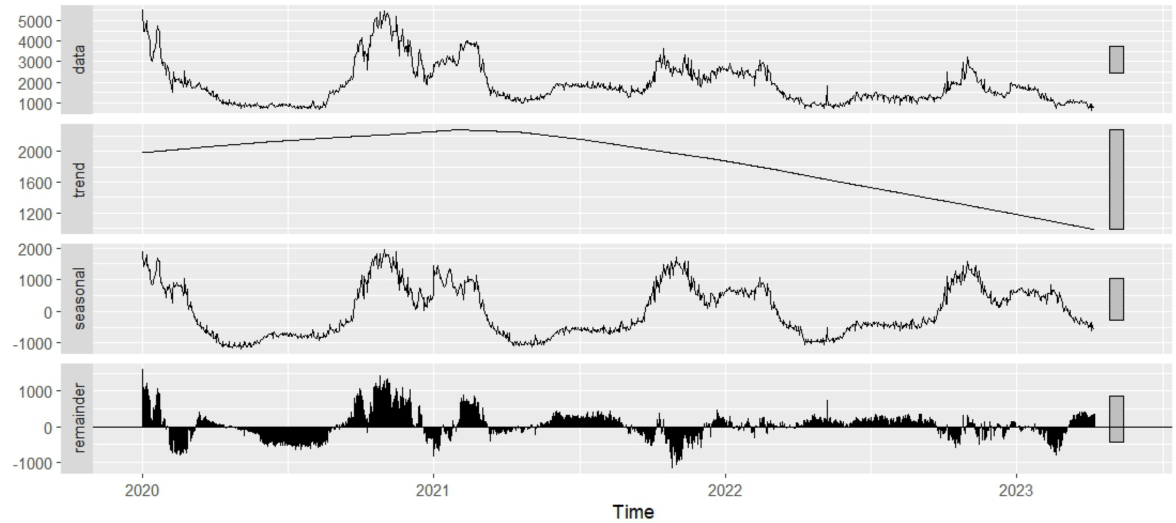

Figure 6. STL Decomposed data of onion price of Hyderabad

1.1.1.1 STL Decomposition with ARIMA

In the STL-ARIMA hybrid framework, the time series is first decomposed using STL to isolate the stable seasonal pattern from the underlying trend and irregular fluctuations. Once the seasonal component is extracted, the remaining seasonally adjusted series is modeled using an ARIMA structure, allowing the ARIMA model to focus exclusively on capturing short-term autocorrelation, shocks, and non-seasonal dynamics that are typically masked by strong seasonality in daily agricultural data. Model identification for the deseasonalized series followed an AICc-based selection process supported by diagnostic checks to ensure well-behaved residuals. Through this procedure, ARIMA(4,1,0), ARIMA(5,1,4), and ARIMA(2,1,1) were identified as the most suitable models for Ahmedabad, Hyderabad, and Mumbai, respectively. These models reflect the differing levels of volatility and autocorrelation across the markets and demonstrate how STL decomposition enhances ARIMA's ability to model the non-seasonal structure more effectively. The final forecasts are obtained by recombining the projected seasonal component with the ARIMA-based forecasts of the adjusted series, resulting in a more flexible and interpretable hybrid forecasting approach (Table 2).

Table 2. ARIMA model selection of decomposed data

City under observation	Model (ARIMA)
Ahmedabad	ARIMA (4,1,0)
Hyderabad	ARIMA (5,1,4)
Mumbai	ARIMA (2,1,1)

1.1.1.2 STL Decomposition with Exponential Smoothing State Space (ETS):

The STL-ETS hybrid model extends the decomposition framework by fitting an Exponential Smoothing State Space (ETS) model to the seasonally adjusted series obtained from STL. After the seasonal component is isolated through decomposition, the ETS model focuses on capturing variations in level, trend, and damping behavior—features that are often prominent in agricultural commodity prices due to changes in supply conditions, policy interventions, and episodic market disruptions. ETS models offer a flexible structure defined by additive or multiplicative components, making them well-suited for series where trend dynamics evolve gradually over time. Through an AICc-guided selection process and diagnostic evaluation, ETS(A,Ad,N) was chosen for Ahmedabad, while ETS(A,N,N) was selected for both Hyderabad and Mumbai, indicating that additive error structures with either damped or no trend provide the best representation of their depersonalized movements. Once forecasts are generated for the adjusted series, they are combined with the projected seasonal component from STL, resulting in a hybrid model that effectively integrates structural decomposition with adaptive state–space forecasting (Table 3).

Table 3. ETS model selection of decomposed data

City under observation	Model (ETS)
Ahmedabad	ETS (A, A_d, N)
Hyderabad	ETS (A,N,N)
Mumbai	ETS (A, N, N)

4. Results and Discussion

To assess the performance of each model, we employed standard evaluation metrics, including Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE). These metrics provide a robust measure of the accuracy and reliability of the forecasts. The results for Mumbai, presented in Table 4, indicate a clear performance advantage for STL-based hybrid models. Both STL-ARIMA and STL-ETS substantially reduce prediction errors when compared with the standalone SARIMA model. Among the three approaches, STL-ETS achieves the lowest RMSE (586.48) and MAPE (22.48%), marginally outperforming STL-ARIMA. In contrast, SARIMA records a considerably higher MAPE of 30.14%, reflecting weaker predictive capability. These improvements suggest that decomposing the series prior to model fitting allows better handling of Mumbai's complex seasonal and irregular price fluctuations.

Table 4. Comparison between SARIMA, STL-ARIMA and STL-ETS model (Mumbai)

Model	MSE	RMSE	MAE	MAPE
SARIMA	895557.5	946.339	654.5289	30.14103
STL-ARIMA	351343.9	592.7427	448.1124	23.0191
STL-ETS	343955.6	586.4772	440.2529	22.47657

A similar pattern is observed in the Hyderabad dataset (Table 5). STL-ARIMA and STL-ETS again outperform SARIMA by a significant margin across all error metrics. The STL-ETS model demonstrates the highest accuracy, achieving a MAPE of 26.98%, compared to 32.49% for SARIMA. The consistency of error reductions across multiple metrics reinforces the superiority of decomposition-based approaches, indicating that Hyderabad's onion price dynamics also benefit from the flexible seasonal extraction provided by STL.

Table 5. Comparison between SARIMA, STL-ARIMA and STL-ETS model (Hyderabad)

Model	MSE	RMSE	MAE	MAPE
SARIMA	671676.9	819.559	635.5767	32.48797
STL-ARIMA	346243.5	588.4246	499.8981	27.84189
STL-ETS	334157.2	578.0633	487.9248	26.98499

For Ahmedabad (Table 6), STL-based models continue to exhibit improved accuracy relative to the SARIMA model, though the magnitude of improvement is slightly smaller than in the other two cities. STL-ETS yields the best performance overall, with STL-ARIMA closely matching it. SARIMA, on the other hand, produces the highest error values, with a MAPE of 39.18%. The comparatively larger error levels observed in Ahmedabad may reflect greater market volatility or more irregular price behavior, which can constrain forecasting precision.

Table 6. Comparison between SARIMA, STL-ARIMA and STL-ETS model (Ahmedabad)

Model	MSE	RMSE	MAE	MAPE
SARIMA	1341285	1158.138	930.5712	39.1769
STL-ARIMA	814882.3	902.7083	792.903	36.81761
STL-ETS	812128.1	901.1815	791.1897	36.718

Across all three markets, a consistent hierarchy of model performance emerges: STL-ETS performs best, followed closely by STL-ARIMA, while SARIMA yields the weakest results. This pattern highlights the advantage of decomposing daily onion prices into trend, seasonal, and irregular components prior to model estimation. Since onion markets are heavily influenced by weather, harvesting cycles, and supply disruptions, their seasonal structure is neither static nor perfectly periodic. In such conditions STL decomposition is particularly effective (Figure 7).

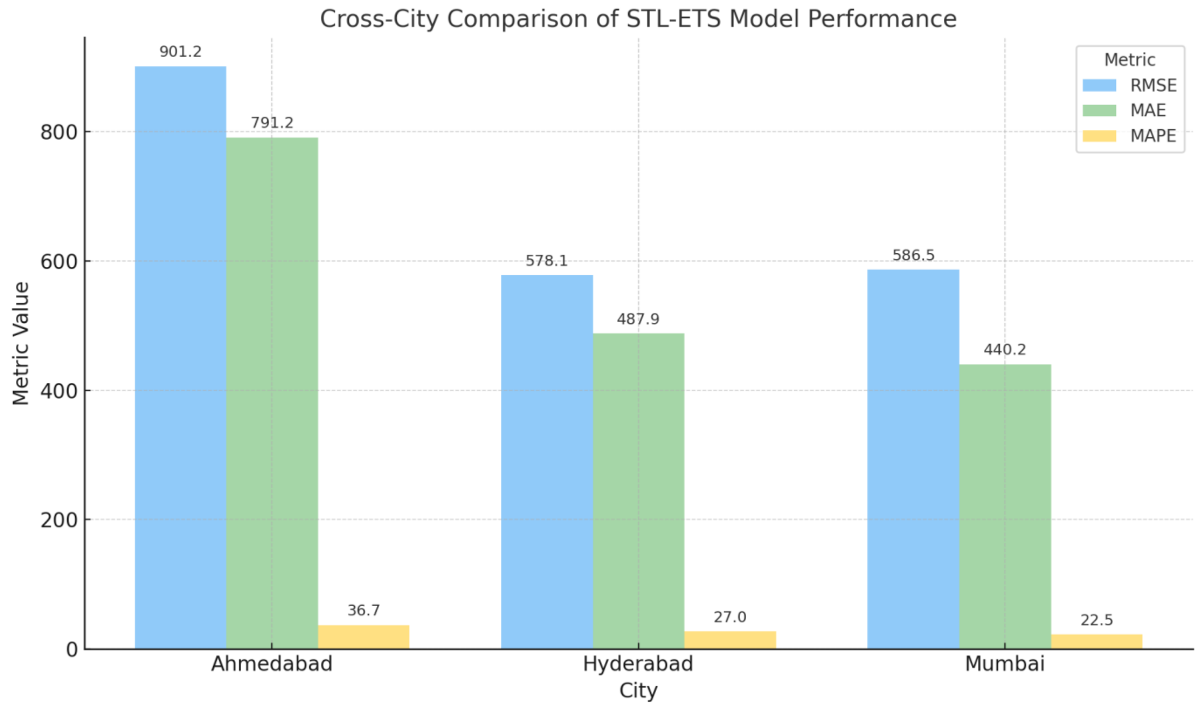

Figure 7. Cross-City Comparison of STL-ETS Forecasting Accuracy Across Ahmedabad, Hyderabad, and Mumbai

The reduction in forecast errors achieved by STL-based models has been often seen exceeding 20–30% relative to SARIMA which demonstrates the practical value of hybrid approaches for agricultural price forecasting. At the same time, the variation in error magnitudes across cities underscores the need to account for regional market characteristics when interpreting model outcomes.

5. Conclusion

This study evaluated the predictive accuracy of SARIMA, STL-ARIMA and STL-ETS models using daily onion price data from Mumbai, Hyderabad and Ahmedabad. The study dealt with critical issues associated with high-frequency and unpredictable agricultural pricing by combining decomposition methodologies with traditional time series models. The results showed that STL-ETS always measure the most accurate predictions, using STL's improved seasonal and trend gathering features. Mumbai had the fewest mistakes in its forecasts, while Ahmedabad experienced the most. This shows that the markets and the data were not behaving the same way for every city. The strong effectiveness of STL-based hybrid models shows that they can be used in daily life to predict agricultural goods which have changing seasonal patterns and prices vary significantly.

References

Areef, M., *Price behaviour and forecasting of onion prices in Kurnool Market, Andhra Pradesh State*, Economic Affairs, Vol. 65, No. 1, pp. 27–33, 2020, https://doi.org/10.30954/0424-2513.1.2020.6

Banaś, J., & Utnik-Banaś, K., *Evaluating a seasonal autoregressive moving average model with an exogenous variable for short-term timber price forecasting*, Forest Policy and Economics, Vol. 131, Article 102564, 2021, https://doi.org/10.1016/j.forpol.2021.102564

Chuan-xi, L., *Application of ARIMA model in forecasting agricultural product price*, Computer Engineering and Applications, Vol. 45, No. 12, pp. 235–238, 2009.

Dokumentov, A., & Hyndman, R. J., *STR: Seasonal-trend decomposition using regression*, INFORMS Journal on Data Science, Vol. 2, No. 2, pp. 1–19, 2020, https://doi.org/10.1287/ijds.2021.0004

Gumani, M., Korke, Y., Shah, P., Udmale, S., Sambhe, V., & Bhirud, S., *Forecasting of sales by using fusion of machine learning techniques*, in Proceedings of the International Conference on Data Management, Analytics and Innovation (ICDMAI), pp. 93–101, 2017, https://doi.org/10.1109/ICDMAI.2017.8073492

Hewamalage, H., Bergmeir, C., & Bandara, K., *Recurrent neural networks for time series forecasting: Current status and future directions*, International Journal of Forecasting, Vol. 37, No. 1, pp. 388–427, 2021.

Makridakis, S., Spiliotis, E., & Assimakopoulos, V., *Statistical and machine learning forecasting methods: Concerns and ways forward*, PLOS ONE, Vol. 13, No. 3, Article e0194889, 2018.

Palazzi, R., Maçaira, P., Meira, E., & Klotzle, M., *Forecasting commodity prices in Brazil through hybrid SSA-complex seasonality models*, Production, Vol. 33, Article e20220025, 2023, https://doi.org/10.1590/0103-6513.20220025

Saxena, R., Paul, R., & Kumar, R., *Transmission of price shocks and volatility spillovers across major onion markets in India*, Agricultural Economics Research Review, Vol. 33, pp. 45–56, 2020, https://doi.org/10.5958/0974-0279.2020.00005.1

Ugbehe, P., Diemuodeke, O., & Aikhuele, D., *Electricity demand forecasting methodologies and applications: A review*, Sustainable Energy Research, Vol. 10, Article 149, 2025, https://doi.org/10.1186/s40807-025-00149-z

Zhang, Y., & Na, S., *A novel agricultural commodity price forecasting model based on fuzzy information granulation and MEA-SVM model*, Mathematical Problems in Engineering, Vol. 2018, Article ID 2540681, 2018, https://doi.org/10.1155/2018/2540681

Biographies

Syed Tahmid Jamil is a Lecturer in the Department of Industrial Engineering at BGMEA University of Fashion & Technology (BUFT), Dhaka, Bangladesh. He holds a Bachelor of Science in Industrial and Production Engineering from the Military Institute of Science and Technology (MIST) and a Master's degree in Applied Statistics and Data Science from Jahangirnagar University. He teaches courses in supply chain management, simulation, and material handling. His research interests include supervised machine learning, time series analytics, ergonomics and human factors and machining process optimization. He also holds professional certifications in project management, psychology, SAP, and prompt engineering.

Mashiat Iqbal is currently a research assistant at MIST, working under the guidance of Professor Dr. A.K.M. Nurul Amin. She successfully earned her bachelor's degree in Industrial and Production Engineering from the Military Institute of Science and Technology (2018-2021). Her undergraduate thesis focused on parameter optimization in high-speed machining of titanium alloy (Ti6Al4V) using nanofluid-based MQL. Currently enrolled in a Master's program in Applied Statistics and Data Science at Jahangirnagar University. She has a strong passion for High-speed machining of advanced materials, nano-lubrication (Nano-MQL), parameter optimization, data-driven process modeling and machine-learning techniques. She is also interested in time-series forecasting, data mining and applied machine learning for engineering and industrial analytics.