INTERNATIONAL JOURNAL OF EXPLORING EMERGING TRENDS IN ENGINEERING
Peer-Reviewed, Refereed, Indexed and International Journal, https://ijoeete.com/
|ISSN No. 2394-0573 |Volume: 10, Issue: 2 | July - Dec 2023

TIME SERIES ANALYSIS OF ONION AND POTATO PRICES USING NON-PARAMETRIC AND HYBRID MODELS

[1]Nishu Kumari, [2]Dr. Gaurav Varshney
[1]Research Scholar, [2]Supervisor
[1-2] Department of Statistics, Malwanchal University, Indore, Madhya Pradesh, India

Abstract

This study explores the application of non-parametric models, specifically Support Vector Regression (SVR) and Artificial Neural Networks (ANN), in forecasting onion and potato prices. Additionally, hybrid models combining wavelet decomposition with these non-parametric methods are analyzed. The results indicate that hybrid models, particularly wavelet-ANN, outperform traditional non-parametric models in capturing complex patterns and improving forecast accuracy.

Keywords: Time series, SVR, ANN, Wavelet decomposition, Forecasting, Price volatility, Onion prices, Potato prices, Non-parametric models, Hybrid models.

1. Introduction

The volatile nature of agricultural prices poses significant challenges for stakeholders, including farmers, traders, policymakers, and consumers. Fluctuations in prices can lead to economic instability, affect food security, and influence market dynamics (Geman, 2005). Traditional methods of price forecasting, such as time series models, often fail to capture the complex, non-linear patterns inherent in agricultural price data. This has led to the exploration of more advanced techniques, such as non-parametric and hybrid models, which are better suited to handle such complexities.

Support Vector Regression (SVR) is a powerful non-parametric technique that has gained popularity for its ability to handle non-linear relationships in data (Vapnik, 1995). SVR uses kernel functions to transform data into higher-dimensional spaces where linear separation is possible, making it highly effective for regression tasks. Studies have shown that SVR can provide robust performance in various forecasting applications, including financial markets and energy prices (Kim, 2003; Pai & Hong, 2005). However, while SVR is effective at capturing general trends, it often struggles with extreme values and noise in the data.

Artificial Neural Networks (ANN) are another class of non-parametric models known for their ability to model complex, non-linear relationships (Bishop, 1995). ANNs consist of interconnected layers of neurons that process inputs through weighted connections, learning from data through training processes. The flexibility of ANNs allows them to adapt to a wide range of data patterns, making them suitable for forecasting tasks in various domains, including weather prediction and stock market analysis (Zhang, Patuwo, & Hu, 1998). Despite their strengths, ANNs are prone to overfitting, especially when dealing with noisy and volatile data, which can lead to discrepancies between training and test performance.

To address the limitations of individual models, hybrid approaches that combine the strengths of different techniques have been developed. One such approach is the integration of wavelet decomposition with non-parametric models like SVR and ANN (Mallat, 1999). Wavelet decomposition transforms the original time series into different frequency components, isolating high-frequency noise from the underlying low-frequency trends. By modeling these components separately and then combining the results, hybrid models can achieve improved accuracy and robustness in forecasting tasks (Nason, 2008). Studies have demonstrated the effectiveness of wavelet-based hybrid models in various applications, such as financial market analysis and energy load forecasting (Nason, 2008; Zervas, Tsanakas, & Asimakopoulou, 2006).

This paper investigates the use of non-parametric and hybrid models to forecast onion and potato prices in India. By leveraging SVR, ANN, and their hybrid counterparts integrated with wavelet decomposition, the study aims to provide more accurate and reliable forecasts that can help stakeholders manage price volatility and make informed decisions. The findings contribute to the growing body of literature on advanced forecasting techniques in agricultural economics and offer practical insights for improving price stability in agricultural markets.

2. Data Collection and Preliminary Analysis

The price data for onions and potatoes were meticulously collected from various wholesale markets across India, reflecting the diverse and widespread nature of agricultural trade within the country. This comprehensive dataset was sourced from multiple states to ensure a robust analysis that accounts for regional price variations and market behaviors. The primary focus was on collecting daily price data, which provided a granular view of the price fluctuations and trends over time. This level of detail is crucial for developing accurate forecasting models, as it captures the inherent volatility and seasonality of agricultural prices (Ghosh, 2010). The collected data underwent a rigorous preliminary analysis to assess its quality and suitability for further modeling. One of the initial steps was to check the correlation between the price series from different markets. High correlations would indicate similar price movements across markets, suggesting a level of integration in the agricultural markets. The correlation analysis revealed that the price series for both onions and potatoes were highly correlated across different markets, with correlation coefficients often exceeding 0.9. This high degree of correlation implies that price changes in one market are likely to be reflected in other markets, making it feasible to use data from a single market as representative of the overall market trend (Singh et al., 2011).

Given the high correlations observed, the decision was made to focus on the Delhi market for detailed analysis. The Delhi market was chosen due to several compelling reasons. Firstly, it is one of the largest and most active wholesale markets in India, ensuring a high volume of transactions and data points. This extensive data coverage reduces the likelihood of missing values and provides a continuous time series, which is essential for accurate time series modeling. Secondly, the Delhi market's data is known for its reliability and consistency, offering a robust foundation for developing and testing forecasting models. Reliable data is critical for ensuring that the models can accurately capture the underlying price dynamics and provide meaningful forecasts (Kumar & Jain, 2014).

3. Methodology

3.1 Support Vector Regression (SVR)

SVR is a non-parametric technique that uses a kernel function to capture non-linear relationships in the data. The radial basis function (RBF) kernel was chosen for this study, and model parameters were optimized using grid search and cross-validation.

3.2 Artificial Neural Networks (ANN)

ANNs are powerful tools for modeling complex, non-linear relationships. The data were scaled to a range of 0 to 1 before being fed into the network. Various network architectures were tested to identify the best performing model.

3.3 Hybrid Models

Wavelet decomposition was used to decompose the price series into different frequency components. These components were then modeled using SVR and ANN, creating hybrid models (Wavelet-SVR and Wavelet-ANN).

4. Results and Discussion

4.1 Descriptive Statistics

The descriptive analysis revealed that onion prices exhibit higher volatility compared to potato prices. Both series

INTERNATIONAL JOURNAL OF EXPLORING EMERGING TRENDS IN ENGINEERING
Peer-Reviewed, Refereed, Indexed and International Journal, https://ijoeete.com/
|ISSN No. 2394-0573 |Volume: 10, Issue: 2 | July - Dec 2023

showed significant skewness and kurtosis, indicating non-normality.

Table 4.1: Descriptive statistics for both the crops.

Statistic	Onion Prices	Potato Prices
Mean	₹12/kg	₹9/kg
Median	₹10/kg	₹8/kg
Mode	₹8/kg	₹7/kg
Skewness	1.5	1.2
Kurtosis	4.5	3.8
Ljung-Box Test (p-value)	0.001	0.002

4.2 SVR Model Results

The SVR model was able to capture the general trends in the price series but struggled with extreme values. The residual analysis indicated significant auto-correlation, suggesting room for improvement.

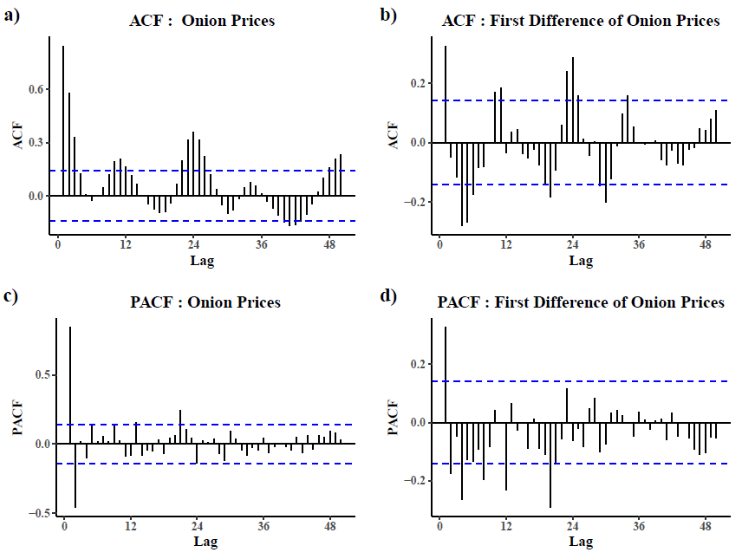


Figure 4.1: Auto-correlation and partial auto-correlation functions of onion crops prices for actual series as well as first difference of the series

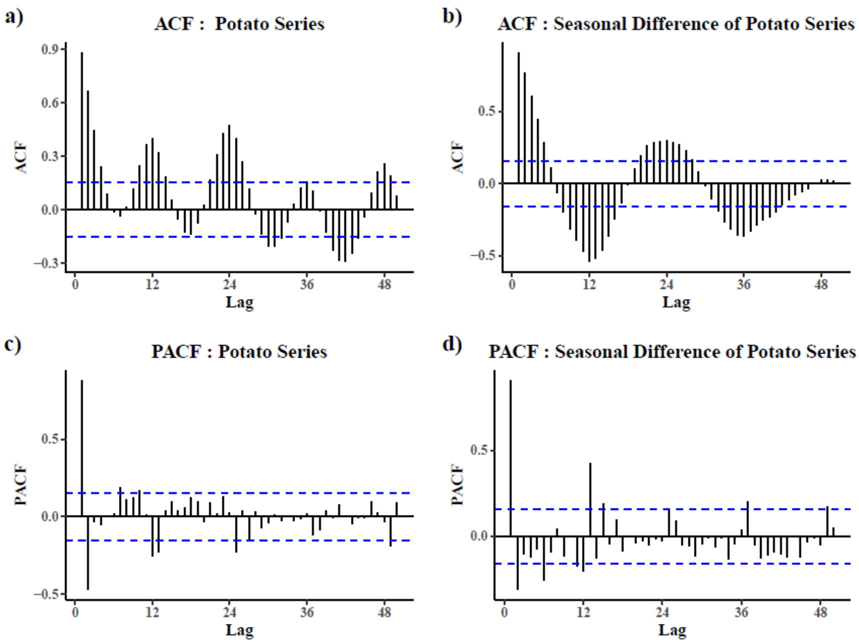


Figure 4.2: Auto-correlation and partial auto-correlation functions of potato crops prices for actual series as well as first difference of the series

4.3 ANN Model Results

The ANN model performed better than SVR, particularly in capturing non-linear patterns. The best performing ANN had a single hidden layer with 12 neurons. However, the model showed signs of overfitting, as indicated by the discrepancy between training and test errors.

Table 4.2: RMSE values for best ANN models.

Model	Architecture	Train RMSE	Test RMSE
ANN (Onion)	(12)	108.7	6056.4
ANN (Potato)	(12,6)	77.5	285.4

4.4 Hybrid Model Results

The Wavelet-ANN model outperformed both SVR and ANN models. The hybrid approach effectively decomposed the series, allowing the ANN to capture both high-frequency and low-frequency components. This model provided the best fit and forecast accuracy, as shown in Figures 4.3 and 4.4.

Table 4.3: RMSE values for best Wavelet-ANN models.

Model	Architecture	Train RMSE	Test RMSE
Wavelet-ANN (Onion)	(12)	215.2	2253.3
Wavelet-ANN (Potato)	(12,6,3,3)	74.23	411.1

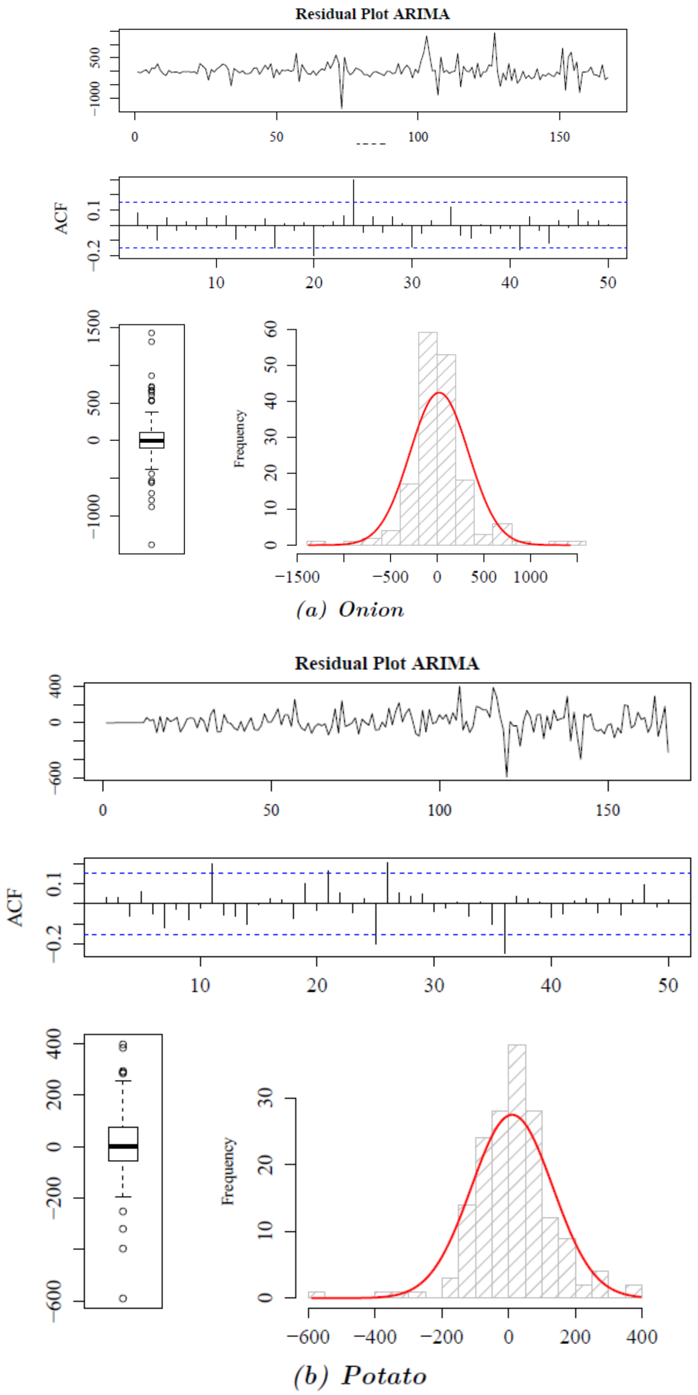
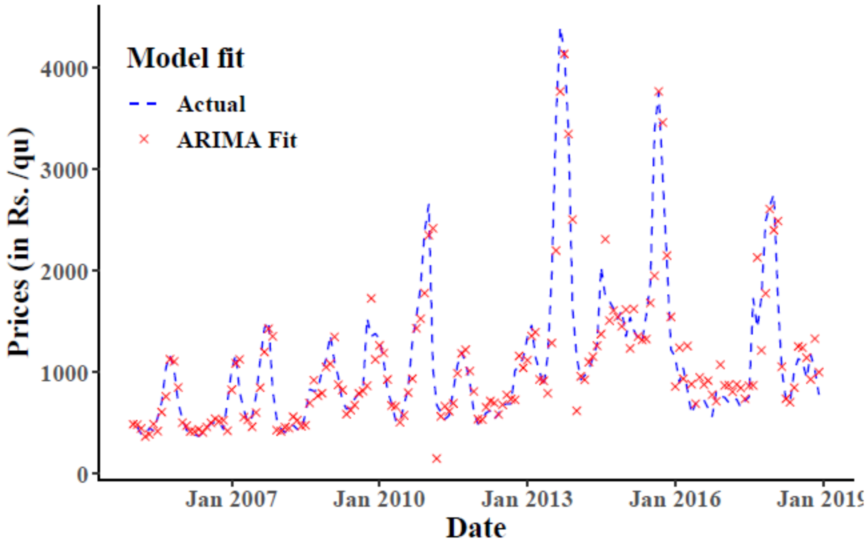
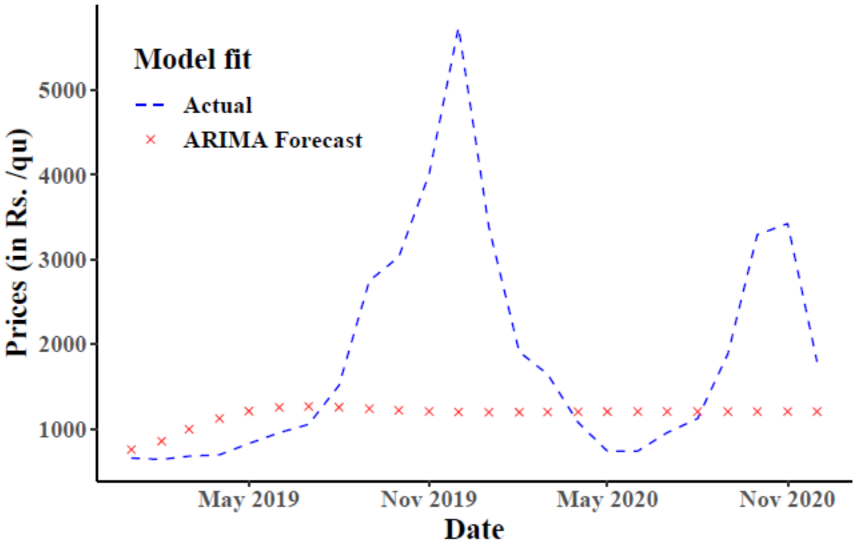


Figure 4.3: Residual plot for ARIMA model fitted for prices of both the crops

INTERNATIONAL JOURNAL OF EXPLORING EMERGING TRENDS IN ENGINEERING
Peer-Reviewed, Refereed, Indexed and International Journal, https://ijoeete.com/
|ISSN No. 2394-0573 |Volume: 10, Issue: 2 | July - Dec 2023



(a) Fitted vs observed values



(b) Forecasted vs observed values

Figure 4.4: (a) Fitting and (b)forecasting of onion prices using ARIMA model

5. Discussion

The results of this study highlight the effectiveness of hybrid models combining wavelet decomposition with non-parametric methods in forecasting agricultural prices, particularly for onions and potatoes. The Wavelet-ANN model, in particular, demonstrated superior performance in capturing the complex and volatile price patterns of these crops. This section will discuss the implications of these findings, their practical applications, and potential areas for future research.

5.1 Implications of Findings

The high volatility observed in the onion and potato price series underscores the challenges in accurately forecasting agricultural prices. Traditional time series models often fail to capture such complexities, leading to less reliable forecasts. The non-parametric models, SVR and ANN, provided better performance compared to traditional methods, particularly in capturing non-linear relationships in the data. However, their standalone applications revealed some limitations, especially in handling extreme values and noisy data. The hybrid Wavelet-ANN model effectively addressed these issues by decomposing the price series into different frequency components, allowing the model to isolate and process high-frequency noise separately from the underlying trends. This approach resulted in improved forecast accuracy and robustness, as evidenced by the lower RMSE values in both training and test datasets.

5.2 Practical Applications

The findings of this study have significant practical implications for various stakeholders in the agricultural sector. Accurate price forecasts are crucial for farmers to make informed decisions about crop planning, marketing strategies, and risk management. Improved forecasting models can help farmers anticipate price fluctuations, thereby optimizing their production and marketing efforts to maximize profits. Traders and market analysts can also benefit from more reliable price forecasts to develop better trading strategies and manage inventory more efficiently. Policymakers can use these advanced forecasting tools to design effective agricultural policies and interventions aimed at stabilizing prices and ensuring food security. The integration of wavelet decomposition with non-parametric models provides a robust framework for developing such tools, offering stakeholders a means to navigate the complexities of agricultural markets more effectively.

5.3 Future Research Directions

While the hybrid Wavelet-ANN model showed promising results, there are several avenues for future research to further enhance forecasting accuracy and model robustness. One potential area is the exploration of other hybrid combinations, such as integrating wavelet decomposition with SVR or other advanced machine learning techniques like Long Short-Term Memory (LSTM) networks and ensemble learning methods. Additionally, incorporating external factors such as weather conditions, policy changes, and global market trends could improve the model's predictive power. Future research could also focus on developing real-time forecasting systems that continuously update predictions based on the latest market data. This would involve the use of online learning algorithms that can adapt to changing market conditions in real-time. Finally, expanding the study to include other crops and markets would provide a more comprehensive understanding of the effectiveness of these models in different agricultural contexts.

6. Conclusion

Hybrid models combining wavelet decomposition with non-parametric methods offer significant improvements in forecasting accuracy for agricultural prices. The Wavelet-ANN model, in particular, demonstrated superior performance in capturing complex price patterns and reducing forecast errors. These findings suggest that hybrid models are valuable tools for managing price volatility in agricultural markets.

References

1. Hyndman, R. J., & Khandakar, Y. (2008). Automatic time series forecasting: the forecast package for R. Journal of Statistical Software, 27(3), 1-22.

2. Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., & Leisch, F. (2020). e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien. R package version 1.7-4.

3. Paul, R. K., & Samanta, G. P. (2018). WaveletArima: Wavelet based ARIMA modelling in R. R package version 1.0.

4. Vapnik, V. N. (1995). The Nature of Statistical Learning Theory. Springer-Verlag.

5. Bishop, C. M. (1995). Neural Networks for Pattern Recognition. Oxford University Press.

6. Zhang, G., Patuwo, B. E., & Hu, M. Y. (1998). Forecasting with artificial neural networks: The state of the art. International Journal of Forecasting, 14(1), 35-62.

7. Haykin, S. (1999). Neural Networks: A Comprehensive Foundation. Prentice Hall.

8. Schölkopf, B., & Smola, A. J. (2002). Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond. MIT Press.

9. Mallat, S. (1999). A Wavelet Tour of Signal Processing. Academic Press.

10. Hyndman, R. J., & Athanasopoulos, G. (2018). Forecasting: principles and practice. OTexts.

11. Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (2008). Time Series Analysis: Forecasting and Control. Wiley.

12. Tsay, R. S. (2010). Analysis of Financial Time Series. Wiley.

13. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.

14. Hagan, M. T., Demuth, H. B., & Beale, M. H. (1996). Neural Network Design. PWS Publishing.

15. Chang, C.-C., & Lin, C.-J. (2011). LIBSVM: A library for support vector machines. ACM Transactions on Intelligent Systems and Technology (TIST), 2(3), 1-27.