



Long Memory Modelling of Weekly Jute Prices in Cooch Behar Market, West Bengal

Chowa Ram Sahu¹, Satyananda Basak², Deb Sankar Gupta², Vinay H.T.²

10.18805/ag.D-6181

ABSTRACT

Background: The objective of this paper is modeling and forecasting the weekly jute prices in the Cooch Behar market of West Bengal in the presence of long memory process.

Methods: A fractionally integrated autoregressive moving average (ARFIMA) model is fitted to 672 weekly data (from January, 2009 to December, 2022). The wavelet method has been used to estimate the fractional difference parameter in the ARFIMA model. Furthermore, we compared the forecasting abilities of the ARIMA and ARFIMA models.

Result: The ARFIMA (1,0.327,1), ARFIMA (3,0.327,1) and ARIMA (3,1,0) models are selected on the basis of minimum AIC and BIC values using 538 training set data. Validation results indicate that the forecasting performance of the ARFIMA (3,0.327,1) model is strongly better than that of the ARIMA (3,1,0) model. Finally, the ARFIMA (3,0.327,1) model is found to be the best optimal model to forecast the jute prices for the Cooch Behar market in terms of RMSE (164.42), MAE (105.66) and MAPE (1.70) criteria using 134 testing set data.

Key words: ARFIMA, ARIMA, Forecasting, Jute prices, Long memory.

INTRODUCTION

Jute (*Corchorus capsularis* L.) known as the “golden fiber” is one of the most important commercial cash crops grown in India. It is one of the cheapest and strongest of all natural fibers and is considered the fiber of the future. After cotton, it is the most important fiber crop grown in India. India is the world's largest producer of raw jute, accounting for more than half of global jute production. Among the states, West Bengal ranks first in area and production of jute in the country with a total area of 0.52 million hectares (78.49 per cent) and a total production of 7.75 million bales (82.49 per cent) with a 2701 kg/hectare productivity during 2022-23. (Directorate of Economics and Statistics, DA and FW). Agricultural prices play an important role in the whole national economy of India. Commodities price forecasts are critical to market participants who make production and marketing decisions and to policymakers who administer commodity programs and assess the market impacts of domestic or international events (Das *et al.*, 2022; Sahu and Basak, 2024). Time series modelling is one such approach that collects and analyses past values to develop appropriate models that describe the inherent structure and characteristics of the series. Most of the research in time-series analysis assumes that the observations that are far apart in time are basically independent, but in many practical situations it is seen that many empirical economic series show that the distant observations are dependent; this is characterized as a long memory. This long memory component of the market cannot be adequately explained by systems that work with short memory parameters *i.e.*, AR (p), MA(q), ARMA(p,q) and ARIMA (p,d,q). The autoregressive fractional integral moving average (ARFIMA) model has been widely employed for long memory time series forecasting in a divergent domain

¹Sant Kabir College of Agriculture and Research Station Kawardha, Kawardha-491 995, Chhattisgarh, India.

²Uttar Banga Krishi Viswavidyalaya, Coochbehar-736 165, West Bengal, India.

Corresponding Author: Chowa Ram Sahu, Sant Kabir College of Agriculture and Research Station Kawardha, Kawardha-491 995, Chhattisgarh, India. Email: chowasahu100@gmail.com

How to cite this article: Sahu, C.R., Basak, S., Gupta, D.S. and Vinay, H.T. (2025). Long Memory Modelling of Weekly Jute Prices in Cooch Behar Market, West Bengal. *Agricultural Science Digest*. 1-7. doi: 10.18805/ag.D-6181.

Submitted: 03-09-2024 **Accepted:** 22-09-2025 **Online:** 04-11-2025

for several decades. Several studies have been conducted in different parts of India and abroad for modelling time series in the presence of long memory. Booth *et al.* (1982), Helms *et al.* (1984) and Sahu *et al.* (2023) applied the rescaled range (R/S) method to detect the existence of long memory in the prices data and provided evidence of long-memory behavior in financial markets. Granero *et al.* (2008) investigated the Hurst exponent for long memory processes in capital markets. Erfani and Samimi (2009) studied the long memory of the stock price index (TSIP) and established the ARFIMA and ARIMA models, concluding that the ARFIMA is a much better model in long memory. Paul (2017) worked in long memory on maximum and minimum temperature series and analysed them using the ARFIMA model and concluded that the ARFIMA model could be used successfully for modelling the temperature series. Similar works have been done by Liu *et al.* (2017), Safitri *et al.* (2019) and Sahu *et al.* (2024).

However, in this study, a wavelet domain based ARFIMA model is used for adequate modelling of the long memory

features of weekly jute prices in the Cooch Behar market of West Bengal. Furthermore, we compared the forecasting abilities of the ARIMA and ARFIMA models.

MATERIALS AND METHODS

Data description

In order to carry out our analysis, historical jute price data of Cooch Behar market has been taken from the Agricultural Marketing Information Network (<https://agmarknet.gov.in>) portal from January 2009 to December 2022 (672 weeks) with weekly data. In the present study, statistical analyses have been carried out using the powerful software “RStudio” (<https://www.rstudio.com>). The first 80% observations (*i.e.*, from the 1st week of 2009 to the 10th week of 2020) are used for the model building purpose and the rest 20% observations (*i.e.*, from the 11th week of 2020 to the 48th week of 2022) are used for model validation.

Test for stationarity

Testing data for stationarity is very important in research where the underlying variables are based on time. A time series is said to be stationary if its underlying generating process is based on a constant mean and constant variance, with its auto correlation function (ACF) essentially constant through time. If these conditions do not hold, the series is non-stationary. In order to check the data for stationarity, the Augmented Dickey-Fuller (ADF) test (Dickey and Fuller, 1979) and the Phillips-Perron (PP) unit root test (Phillips and Perron, 1988) have been used.

Long memory process

The concept of long memory, or long-range dependence, is an important one in time series analysis. A long memory feature occurs when the autocovariances for a stationary time series tend to zero like a power function but more slowly than an exponential decay. In this study, the Hurst exponent method (Hurst, 1951) has been used to test the long memory. The value of Hurst exponent (*H*) varies from 0 to 1, When $0.5 < H < 1$, it implies that persistent long memory of a time series is strong.

Autoregressive integrated moving average (ARIMA) model

The ARIMA model (Box and Jenkins, 1976) is a widely recognized statistical forecasting model that predicts future observations of a time series on the basis of some linear function of past values and white noise terms. The ARIMA (*p*, *d*, *q*) process is expressed by:

$$\phi_p(B)(1-B)^d y_t = \mu + \theta_q(B) \varepsilon_t$$

Where, μ is the mean of the series, y_t is the Observed time series values at time *t*; $\phi_p(B) = (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)$ and $\theta_q(B) = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q)$ are the autoregressive operator and moving average operator respectively in which ϕ_i (*i* = 1, 2, ... , *p*) and θ_j (*j* = 1, 2, ... , *q*) are the autoregressive and moving average coefficients

respectively; *p* and *q* are integers and often referred to as orders of autoregressive and moving average respectively; ε_t is the *t*th white noise and *B* is the backshift operator ($By_t = y_{t-1}$). If the series is not stationary, the first difference or higher-order differences will produce a stationary time series.

Autoregressive fractionally integrated moving average (ARFIMA) model

One of the most challenging issues in statistical modelling is time series analysis and forecasting. In general, models for time series data can have many forms and represent different stochastic processes. In all linear time series forecasting approaches, one of the most important models is ARIMA, which has been widely used to forecast social, economic, agricultural, engineering and financial problems (Box and Jenkins, 1976). The ARIMA model can only capture the short-range dependence (SRD) property, but many practical agricultural datasets, principally commodity prices data, show the typical feature of a long memory process. Therefore, there is a need for a model that has the property of long memory. The autoregressive fractional integral moving average (ARFIMA) model meets this property.

The ARFIMA (*p*, *d*, *q*) model (Granger and Joyeux, 1980; Hosking, 1981) is similar to that of the ARIMA (*p*, *d*, *q*) model, except for the differencing value (*d*), which is given as follows:

$$\phi_p(B)(1-B)^d y_t = \mu + \theta_q(B) \varepsilon_t \text{ for } 0 < d < 0.5$$

There are three steps in the procedure of establishing an ARFIMA model. First, testing for long term memory in the time series and determining the fractional differencing parameter. Second, imposing fractional differencing on the series and obtaining an ARMA process. Third, determining the other two parameters of the ARFIMA model, namely *p* and *q*. After fractional differencing is determined, we obtained the fractional differencing time series as $(1-B)^d y_t$ with binomial expansion. The partial auto correlation function (PACF) is most useful for identifying the order (*p*) of an autoregressive model and the autocorrelation function (ACF) is useful for identifying the order (*q*) of moving average model.

Wavelet for long memory parameter estimation

The maximum overlap discrete wavelet transform (MODWT) is a modified DWT where the subsampling process is avoided, leading to a higher level of information in the resulting wavelet and scaling coefficients when compared to the DWT (Ghaemi *et al.*, 2019). The MODWT determines scaling coefficient $V_\phi(m, n)$ and wavelet coefficient $W_\psi(m, n)$ (where scale parameter and translations parameter *n*) by applying low and high pass filters to the original dataset (Kumar *et al.*, 2020).

For estimating the long memory parameter of ARFIMA model, the algorithm based on wavelet (Jensen, 1999) is followed. Let y_t be a mean zero fractionally differenced process with $0 < d < 1/2$. Using the autocovariance

function of the $I(d)$ process, Jensen (1999) found that as $j \rightarrow 0$, the wavelet coefficients, $W_\psi(m, n)$, associated with a mean zero $I(d)$ process with $0 < d < 1/2$ are distributed $N(0, \sigma^2 2^{-2md})$, where σ^2 is a finite constant. The wavelet coefficients from an $I(d)$ process have a variance that is a function of the scaling parameter m , but is independent of the translation parameter, n . The correlation of the wavelet coefficients from an $I(d)$ process decay exponentially over time and scale. Hence, define $R(m)$ to be the wavelet coefficients' variance at scale m , i.e., $R(m) = \sigma^2 2^{-2md}$. Taking the logarithmic transformation of $R(j)$, we obtain the relationship $R(m) = \ln \sigma^2 - d \ln 2^m$. Where $\ln R(m)$ is linearly related to $\ln 2^{-2m}$ by the fractional differencing parameter, d . Hence, the unknown d of a fractionally integrated series can be estimated by the ordinary least squares estimator d .

Information criteria and accuracy measures

In this paper, we used two widely applied criterion Akaike information criterion (AIC) and the Bayesian information criterion (BIC) to select the best model among a set of candidate models. Furthermore, the RMSE, MAE and MAPE are used as an accuracy measure to evaluate the performance of the models.

RESULTS AND DISCUSSION

Primary statistical analysis

The weekly series of jute prices for the Cooch Behar market of West Bengal is shown in Fig 1. The Cooch Behar market series depicts an up-and-down pattern, with two sharp rises between 2013 and 2016 and another between 2020 and 2022. The descriptive statistics to summarize information from the weekly jute price data are listed in Table 1. As Table 1 shows, the series of jute prices in the Cooch Behar market has a mean of 4188, a standard deviation of 1518.44 and, 36.26% coefficient of variation, suggesting that it has been volatile. In addition, skewness

and kurtosis statistics show that the price series is not normally distributed.

To begin with the implementation of ARIMA and ARFIMA models, the data series are divided into two sets: The training set and the testing set. First, the model is fitted using the training data with 538 observations (i.e., from the 1st week of 2009 to the 10th week of 2020) and then it is predicted over the validation period using a testing set with the last 134 observations (i.e., from the 11th week of 2020 to the 48th week of 2022).

Test for stationarity

The first step in applying ARIMA and ARFIMA models is to check whether the time series is stationary or not. In order to test for stationarity, we first conducted Augmented Dickey-Fuller and Phillips-perron unit root tests on the training dataset of the series. According to the results, the p-values (Test statistic) of the ADF and PP tests are 0.494(-2.20) and 0.251(-15.39), respectively, indicating that the time series under consideration is clearly non-stationary. The study, therefore, proceeded to find the stationary series.

Test for long memory and estimation

The presence of long memory in a time series (training set) was confirmed by investigating the autocorrelation function (ACF) plot, which shows that the correlations decay very slowly towards zero up to 250 lags (Fig 2), indicating the presence of long memory processes. Accordingly, the presence of long memory is tested as discussed in methodology and it is found that the R/S Hurst value ($H = 0.873$) is higher than 0.5, which firmly concludes the existence of the long memory characteristic of the jute prices. The models that consider the long memory property are very sensitive to the estimation of the long-memory parameter (i.e., the fractional differencing parameter) and for this reason, in this study, it has been estimated by using the wavelet-based ordinary least squares estimator (d_{wavelet}) and is found to be 0.327.



Fig 1: Weekly jute price series for the Cooch Behar market, including all breaks and confidence intervals.

Table 1: Descriptive statistics.

No. of observations	Min	Max	Mean	Median	SD	CV (%)	Skewness	Kurtosis
672	1569	7448	4188	4648	1518.44	36.26	0.29	2.07

After determining the fractional differencing parameter we obtained the fractional and first-order differencing time series shown in Fig 3. For that, the stationary test results are shown in Table 2. The t -values of the ADF and PP tests are less than 5%, which reveal the series has become stationary, which is also confirmed by Fig 3.

Model Identification

To establish ARIMA and ARFIMA models, the values of d and H must be determined. In the above section, we have identified the value of H . Now in this section, we are going to find the optimal value of d and q which are order of autoregressive and moving average terms. We used the training set as in-sample data for the determination of the parameters and q of the ARIMA and ARFIMA models. First, we computed the values of autocorrelation and partial autocorrelation for fractionally differenced series and first-order differenced time series, as illustrated in Fig 4-5. On

computation of ACF and PACF for each estimated difference parameter, it is observed that the decay rate of ACF has improved as compared to the decay of ACF in the actual training set (Fig 4). The orders of non-seasonal parameters and (q) are obtained by looking for significant spikes in autocorrelation and partial autocorrelation functions.

In the identification stage, we estimated different ARIMA and ARFIMA specifications with different combinations of $(AR \text{ terms})$ and $(MA \text{ terms})$, which are listed in Table 3 and selected the appropriate model from each method as having the minimum values of AIC and BIC values. Thus, the models selected for the training period are ARFIMA $(1,0.327,1)$, ARFIMA $(3,0.327,1)$ and ARIMA $(3,1,0)$.

Validation and diagnostic checking

After appropriate ARFIMA and ARIMA models have been obtained, the next step is to see their ability to forecast the data. The model verification process is concerned with

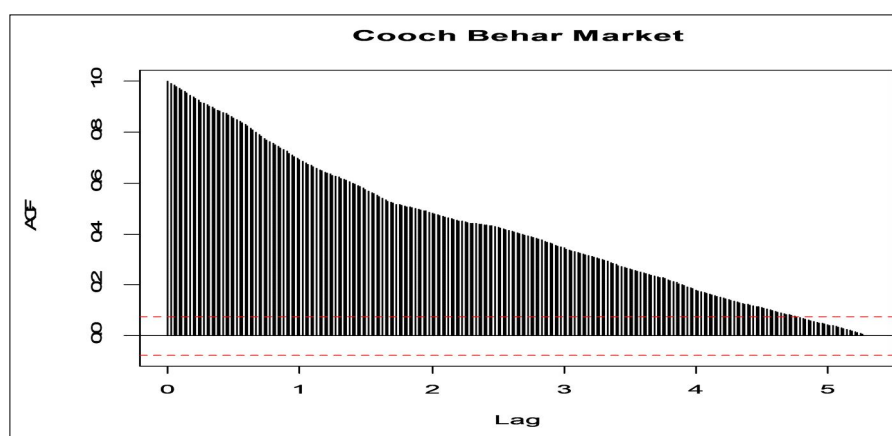


Fig 2: ACF plot of the weekly series of jute prices.

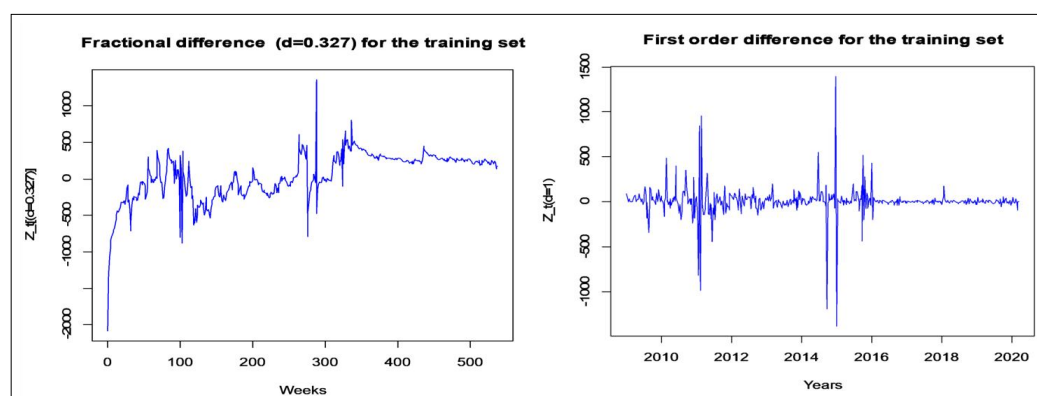


Fig 3: Fractional difference series and first order difference series.

Table 2: Stationary test for the fractional difference series and first order difference series.

Method	Test for Z_t and (Z_t^1) series	Test statistic	Lag order	p-value
$\hat{d}_{\text{wavelet}}$ (0.327)	ADF test	-4.377	8	0.01
	PP test	-190.39	6	0.01
d (1)	ADF test	-8.7626	8	0.01
	PP test	-673.99	6	0.01

examining residuals obtained from fitted models to see if they contain any systematic pattern that could still be removed to improve the chosen models. This has been done through the Ljung-box diagnostic test and it is found that the $-value$ of the Ljung-box test is more than 5% (Table 4), which means that the model residual meets the assumption of white noise residuals. The evaluation of forecasting performance has been done for the test set as an out of-sample period of 134 observations. Table 4 represents the results of the models based on the three different accuracy performance measures: RMSE, MAE and MAPE.

As shown in Table 4, comparing the validation results of all three models indicates that all are likely to perform well in the forecasting phase and, it is observed that the ARFIMA(3,0.327,1) model produces the lowest RMSE, MAE and MAPE, which are 164.42, 105.66 and 1.70, respectively. It can be concluded that the wavelet method based ARFIMA (3,0.327,1) model is the most accurate compared to other models, where predictions indicate that there are narrow variations between the actual and predicted values of jute prices (Fig 6). The strength of the ARIMA model in forecasting jute prices in the Cooch Behar market is considerable, but Table 4 shows that the ARIMA does not

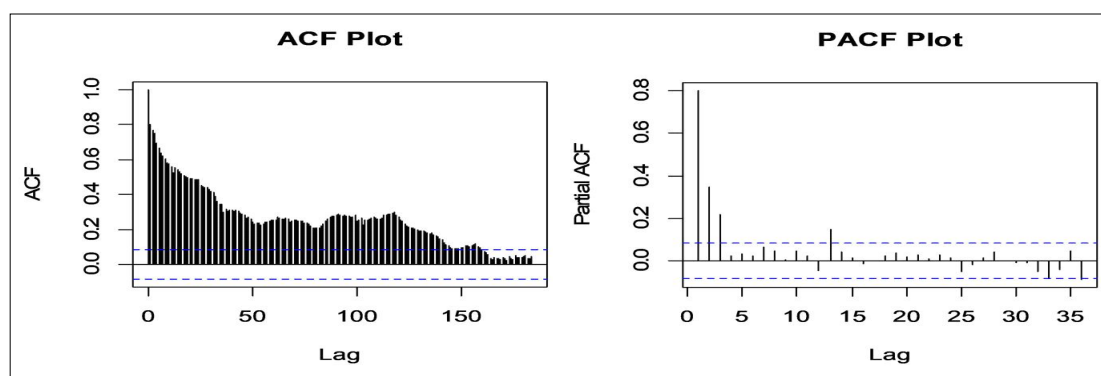


Fig 4: ACF and PACF plot of Wavelet method based fractional difference series.

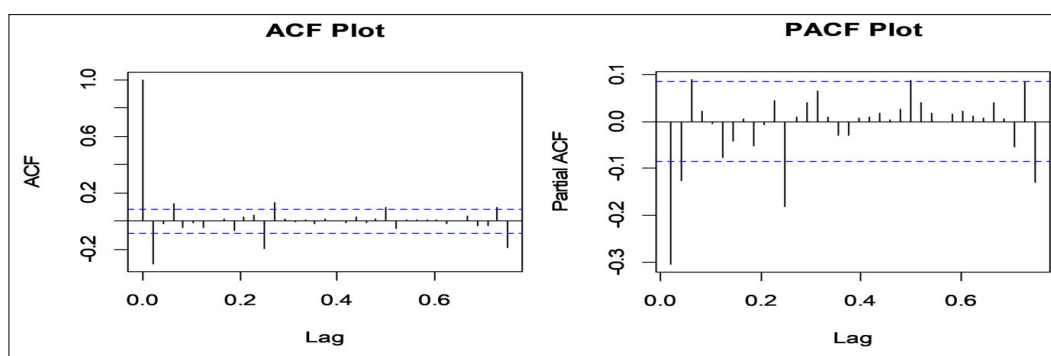


Fig 5: ACF and PACF plot of first order difference series.

Table 3: AIC and BIC values of the ARFIMA and ARIMA models.

$d_{\text{wavelet}} = 0.327$ ARFIMA (p,0.327,q)			ARIMA (p,1,q)		
Model	AIC	BIC	Model	AIC	BIC
(1,0.327,0)	7052.88	7070.03	(1,1,0)	6872.243	6880.815
(2,0.327,0)	6941.13	6962.57	(2,1,0)	6866.222	6879.08
(1,0.327,1)	6893.98	6915.42	(3,1,0)	6863.569	6880.713
(2,0.327,1)	6895.22	6920.94	(1,1,1)	6869.522	6882.38
(3,0.327,1)	6893.97	6923.99	(2,1,1)	6865.125	6882.269
(1,0.327,2)	6895.18	6920.90	(3,1,1)	6865.347	6886.777
(2,0.327,2)	6897.68	6927.69	(1,1,2)	6869.139	6886.283
(1,0.327,3)	6896.88	6926.89	(2,1,2)	6864.765	6886.195
(0,0.327,1)	7426.24	7443.39			
(0,0.327,2)	7319.81	7341.25			

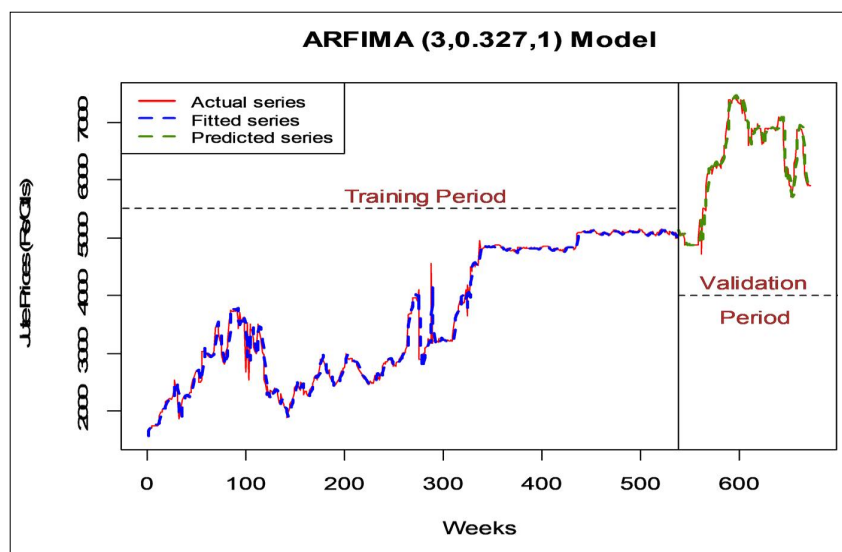


Fig 6: Plot of ARFIMA (3,0.327,1) with training and validation period.

Table 4: Validation of estimated models.

Method	RMSE	MAE	MAPE	Ljung-box test
$\hat{d}_{\text{wavelet}} = 0.327^{(0.131)**}$ ARFIMA(1,0.327,1)	168.11	107.66	1.73	0.205[0.651]
$\hat{d}_{\text{wavelet}} = 0.327^{(0.104)***}$ ARFIMA(1,0.327,1)	164.42	105.66	1.70	0.011[0.917]
ARMA(3,1,0)	166.72	106.61	1.71	0.020[0.888]

* Significant at 5%; ** Significant at 1%; *** Significant at 0.1%; the value in the parenthesis is the SE of coefficient; values of the Ljung and Box statistics are reported between square brackets.

perform well. That the most accurate model is conclude to forecast the weekly jute prices in the Cooch Behar market of West Bengal is the ARFIMA (3,0.327,1) model, which is given as:

$$(1 - 0.578B - 0.188B^2 - 0.210B^3)(1 - B)^{0.327} y_t = 1569.681 + (1 + 0.212B) \varepsilon_t$$

CONCLUSION

The aim of this paper is to introduce an appropriate model for modeling and forecasting the weekly jute prices in the Cooch Behar market of West Bengal in the presence of long memory feature. The presence of long memory behavior is confirmed in jute price series using an ACF plot and Hurst rescaled range (R/S) analysis indicates that it would be better to develop and employ ARFIMA models. We considered the wavelet methods for estimating the fractional differencing parameter. ARIMA and ARFIMA models are fitted to the jute price series and the models selected are ARFIMA (1,0.327,1), ARFIMA (3,0.327,1) and ARIMA (3,1,0) on the basis of the minimum AIC and BIC values. A comparative study has been made between the forecasting performances of the ARIMA and ARFIMA models and it is found that the ARFIMA (3,0.327,1) model out performs the best fitted ARIMA (3,1,0) model in terms of the RMSE, MAE and MAPE. Hence, it is evident that long memory plays an

important and dominant role in describing and modeling the jute prices. Finally, the ARFIMA (3,0.327,1) model is found to be the best optimal model to forecast the jute prices for the Cooch Behar market. The model demonstrated good performance in terms of explained variability and predicting power. The study has revealed that the ARFIMA model could be used successfully for modelling as well as forecasting, especially for data with the long memory (long-range dependency) property.

Conflict of interest

The authors declare that they have no conflict of interest.

REFERENCES

- Booth, G.G., Kaen, F.R. and Koveos, P.E. (1982). R/S analysis of foreign exchange rates under two international monetary regimes. *Journal of Monetary Economics*. **10**(3): 407-415.
- Box, G.E. and Jenkins, G.M. (1976). *Time Series analysis: Forecasting and Control*. Holden Day, San Francisco, New Jersey, USA.
- Das, P., Jha, G.K. and Lama, A. (2022). "EMD-SVR" hybrid machine learning model and its application in agricultural price forecasting. *Bhartiya Krishi Anusandhan Patrika*. **37**(1): 1-7. doi: 10.18805/BKAP385.
- Dickey, D. and Fuller, W. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*. **74**: 427-431.

- Erfani, A. and Samimi, A.J. (2009). Long memory forecasting of stock price index using a fractionally differenced arma model. *Journal of Applied Sciences Research*. **5(10)**: 1721-1731.
- Ghaemi, M.S. et al. (2019). Multiomics modeling of the immunome, transcriptome, microbiome, proteome and metabolome adaptations during human pregnancy. *Bioinformatics*. **35(1)**: 95-103.
- Granero, M.A.S., Segovia, J.E.T. and Pérez, J.G. (2008). Some comments on Hurst exponent and the long memory processes on capital markets. *Physica A: Statistical Mechanics and its Applications*. **387**: 5543-5551.
- Granger, C.W.J. and Joyeux, R. (1980). An introduction to long memory time-series models and fractional differencing. *Journal of Time Series Analysis*. **4**: 221-238.
- Helms, B.P., Kaen, F.R. and Rosenman, R.E. (1984). Memory in commodity futures contracts. *The Journal of Futures Markets*. **4(4)**: 559-567.
- Hosking, J.R.M. (1981). Fractional differencing. *Biometrika*. **68**: 165-176.
- Hurst, H.E. (1951). Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*. **116(1)**: 770-799.
- Jensen, M.J. (1999). Using wavelets to obtain a consistent ordinary least squares estimator of the long-memory parameter *Journal of Forecasting*. **18**: 17-32.
- Kumar, A., Babu, B.M., Satishkumar, U. and Reddy, G.V. (2020). Comparative study between wavelet artificial neural network (WANN) and artificial neural network (ANN) models for groundwater level forecasting. *Indian Journal of Agricultural Research*. **54(1)**: 27-34. doi: 10.18805/IJARE. A-5079.
- Liu, K., Chen, Y.Q. and Zhang, Xi. (2017). An Evaluation of ARFIMA (Autoregressive Fractional Integral Moving Average) Programs†. *Axioms*. **6**: 1-16.
- Paul, R. (2017). Modelling long memory in maximum and minimum temperature series in India. *Mausam*. **68**: 317-326.
- Phillips, P.C.B. and Perron, P. (1988). Testing for unit roots in time series regression. *Biometrika*. **75**: 335-346.
- Safitri, D., Mustafid, Ispriyanti, D. and Sugito, H. (2019). Gold price modeling in Indonesia using ARFIMA method. *Journal of Physics: Conference Series*. **1217**: 1-11.
- Sahu, C.R. and Basak, S. (2024). Performance comparison of time delay neural network and support vector regression for forecasting of jute prices in coochbehar District, West Bengal. *Bhartiya Krishi Anusandhan Patrika*. **39(3-4)**: 245-253. doi: 10.18805/BKAP737.
- Sahu, C.R., Basak, S. and Gupta, D.S. (2023). Modelling long memory in volatility for weekly jute prices in the Malda district, West Bengal. *International Journal of Statistics and Applied Mathematics*. **8(3)**: 118-124.
- Sahu, C.R., Basak, S. and Gupta, D.S. (2024). Long memory time-series model (ARFIMA) based modelling of jute prices in the samsi market of malda district, West Bengal. *Journal of Scientific Research and Reports*. **30(6)**: 600-614.