An Improved ARFIMA Model using Maximum Overlap Discrete Wavelet Transform (MODWT) and ANN for Forecasting Agricultural Commodity Price

Santosha Rathod
ICAR Indian Institute of Rice Research

Kamalesh Narain Singh
Indian Agricultural Statistics Research Institute

Ranjit Kumar Paul
Indian Agricultural Statistics Research Institute

Saroj K. Meher
Indian Statistical Institute

Available online at www.isas.org.in/jisas
JOURNAL OF THE INDIAN SOCIETY OF
AGRICULTURAL STATISTICS 71(2) 2017 103–111

An Improved ARFIMA Model using Maximum Overlap Discrete Wavelet Transform (MODWT) and ANN for Forecasting Agricultural Commodity Price

Santosha Rathod[1], K.N. Singh[1], Ranjit K. Paul[1], Saroj K. Meher[2], G.C. Mishra[3], Bishal Gurung[1], Mrinmoy Ray[1] and Kanchan Sinha[1]
[1]ICAR-Indian Agricultural Statistics Research Institute, New Delhi
[2]Indian Statistical Institute, Bengaluru
[3]Banaras Hindu University, Varanasi

Received 23 May 2016; Revised 07 January 2017; Accepted 10 January 2017

SUMMARY

Autoregressive fractionally integrated moving average (ARFIMA) is widely employed model for long memory time series forecasting in divergent domain from several decades. One of the main pitfall of this model is the presumption of linearity. As many long memory time series data in real world are not purely linear, therefore there is an opportunity to enhance the prediction ability of ARFIMA models by fusing with other nonlinear models. With this reasoning, the present article attempts to estimate the parameters of ARFIMA model by maximum overlap discrete wavelet transform (MODWT) and long memory time series prediction was made by combining ARFIMA-MODWT and ANN for forecasting spot prices of mustard. Experimental study justified the superiority of the proposed hybrid model over ARFIMA model in terms of several measurement indices.

Keywords: ARFIMA, Long memory time series, MODWT, ANN, Hybrid methodology.

1. INTRODUCTION

The sample autocorrelation function (ACF) of the time series is expected to disappear rapidly as the observations are distance apart in time, for example in ARMA model (Box and Jenkins 1970) the ACF exhibit short range dependence or decreases exponentially as the time lag increases and in some series the decay can occur at much slower hyperbolic rate and the correlations remain positive for long lags. Such series are said to have long memory and commonly prevail in stock market prices, economic growth rate, inflation rate, oil price and GDP figures etc. Classical time series models namely ARIMA models cannot describe such long memory phenomenon. Therefore, to overcome this difficulty set of models has been established, among which most popular is autoregressive fractionally integrated moving average (ARFIMA) model given by Granger and Joyeux (1980). Ramalingam (2010) made overall review of long-term memory independently. Long memory studies in housing prices were carried out by Gil-Alana *et al.* (2014), Lima and Xiao (2010) and Tzouras *et al.* (2015) in financial time series. Not much research work has been done in long memory time series pertaining to agriculture. Paul 2014 and Paul *et al.* 2015 carried out long memory studies in pigeonpea and mustard for price forecasting.

As far as parameter estimation of ARFIMA model is concerned, GPH method of Geweke and Porter-Hudak (1983) and the Gaussian semi parametric method developed by Robinson (1995) are widely used. Sowell (1992) gave maximum likelihood estimation for stationary univariate fractionally integrated time series model, Mohamed (2009) made comparison of non-parametric and semiparametric tests in detecting long memory. Hsua and Tasib (2009) gave semiparametric estimation methods for seasonal

Corresponding author: Santosha Rathod
E-mail address: santosha.rathod@icar.gov.in

long memory time series using generalized exponential models. In recent years, wavelet method has also been used for estimation of long memory parameter. Based on the discrete wavelet transform (DWT) coefficients; Jensen (1999) developed estimation technique for long memory parameters, Nielsen and Frederickson (2005) applied wavelet estimators in fractional integration model, Lu and Guegan (2011) estimated long memory parameters in time varying series using wavelet method and Paul *et al*. (2015), estimated parameters of ARFIMA model using Maximal overlap discrete wavelet transform (MODWT).

Sometimes the time series often contain both linear and nonlinear components, rarely they are pure linear or nonlinear under such condition neither ARFIMA nor artificial neural network (ANN) are adequate in modeling and forecasting of long memory time series (Gooijer and Kumar 1992). Since the ARFIMA model cannot deal with nonlinearity, while the ANNs are alone not able to capture both linear and nonlinear behavior equally. To overcome these difficulty, hybrid methods were evolved. Applications of hybrid methods in the literature (Khashei *et al*. 2003, Zhang 2003, Faruk 2010, Asadi *et al*. 2012, Khashei *et al*. 2012, Pektas and Cigizoglu 2013, Chaabane 2014, Jha and Sinha 2014, Shan *et al.* 2015, Ray *et al*. 2016) show that combining different methods can be an effective and efficient way to improve forecasts.

In this paper, attempt has been made to investigate the structure of long memory in daily wholesale price of mustard in Mumbai market, India during the period 1st January, 2009 to 31st December, 2012. The data is collected from Ministry of Consumer's Affairs, Government of India.

Estimation of long memory parameter of ARFIMA model is done by wavelet method using MODWT and daily wholesale price of mustard in Mumbai market are forecasted. In the next step; the residuals obtained from ARFIMA models are modelled and forecasted using ANN. Finally, forecasted values obtained from ARFIMA model and forecasts of residuals obtained from ANN are combined and forecasting accuracies are compared between ARFIMA and hybrid model.

In Section 2, a brief description of model has been given following the procedure of hybrid methodology in Section 3. Data description, Results and discussions are reported in Section 4. Finally, Section 5 includes the concluding remarks.

2. METHODOLOGY

Methodology section comprises the long memory process and its detection, ARFIMA model and its parameter estimation by MODWT method, testing of stationarity and fundamentals of ANN model.

2.1 Long Memory Process

Long memory in time-series can be defined as autocorrelation at long lags (Robinson 1995). Mathematically, time series X_t is said to be long memory series if the autocorrelation function ρ_t satisfies the condition:

$$\lim_{t\to\infty} \sum_{t=-n}^{n} |\rho_t| \to \infty \tag{1}$$

where, n is the sample size, for detection of long memory many statistical tests are available in literature *viz.,* the classical rescaled range series test commonly represented as R/S test, the modified R/S test, KPSS (Kwiatkowski, Phillips, Schmidt and Shin) method, logarithmic diagram method also known as GPH method, and Gauss semi-parametric estimation (GSP) method (Aarthi *et al.* 2012). Among them R/S analysis method is most popular one. A brief description of the test is given below.

Let us consider the time series X_t of the sample length T which is divided into k sub intervals of length n and the average of n series observed values is $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} x_i$. The range of each subinterval is defined as

$$R(n) = \max_{1\le k\le n} \sum_{j=1}^{k}(x_j - \overline{x_n}) - \min_{1\le k\le n} \sum_{j=1}^{k}(x_j - \overline{x_n}) \tag{2}$$

and the standard deviation is

$$S(n) = \left|\frac{1}{n} \max_{1\le k\le n} \sum_{j=1}^{k}(x_j - \overline{x_n})^2\right|^{1/2} \tag{3}$$

For a given n there exists a statistic

$$Q_n = \frac{R(n)}{S(n)} \tag{4}$$

which is equivalent to

$$\lim_{n\to\infty} n^{-H} Q_n = C \tag{5}$$

where C is a constant, and H is Hurst index, so we can get approximate estimate of H as follows:

$$H = \frac{\ln Q_n}{\ln n} \tag{6}$$

In general, the R/S analysis method holds following relationship

$$(R/S)_n = C.n^H \tag{7}$$

where, R is rescaled range, S is the standard deviation, H is Hurst index i.e. the parameter that relates mean R/S values for subsamples of equal length of the series to the number of observations within each equal length subsample, C is a constant, and n is sample observation number. The logarithmic form of equation (7) can be expressed as follows.

$$log(R/S)_n = \log(C) + H\log(n) \tag{8}$$

When $0.5<H<1$, the long memory structure exists. If $H \geq 1$, the process has infinite variance and is non-stationary. If $0<H<0.5$, anti-persistence structure exists. If $H=0.5$, the process is white noise thus the trend is gradually becoming random. The relationship between Hurst exponent and long memory parameter (d) is: $H=1–d$. Positive values of d indicate sort of long memory known as persistence may have infinite conditional variance. An application of long memory model in agriculture can be found in Paul *et al*. (2015).

2.2 The ARFIMA Model

ARFIMA model (Granger and Joyeux 1980) is given as follows

$$\varphi(B)(1-B)^d X_t = \theta(B)e_t,\ -0.5<d<0.5 \tag{9}$$

where, B is the back-shift operator such that $BX_t=X_{t-1}$ and e_t is a white noise process with $E(e_t)=0$ and variance is $\sigma_e^{\,2}$. The polynomials $\varphi(B) = (1-\varphi_1 B - \cdots - \varphi_p B^p)$ and $\theta(B) = (1-\theta_1 B - \cdots - \theta_q B^q)$ have orders p and q respectively with all their roots outside the unit circle. The process is stationary if d=0 and the effect of shock to $\boldsymbol{e}_t$ on $X_{(t+j)}$ decays geometrically as j increases. For d=1, the process is non-stationary.

2.3 Maximal Overlap Discrete Wavelet Transform (MODWT)

Wavelets are fundamental building block functions, analogous to the trigonometric sine and cosine functions. A good description of wavelets can be found in Percival and Walden (Percival and Walden 2010). Some applications of this method can be found in Ghosh *et al*. (2010), Paul *et al*. (2011, 2013). The MODWT is a linear filtering operation that transforms a series into coefficients related to variations over a set of scales. It is similar to the discrete wavelet transform (DWT) in that both are linear filtering operations producing a set of time-dependent wavelet and scaling coefficients. The MODWT is well defined for all sample sizes n, whereas for a complete decomposition of J levels the DWT requires N to be a multiple of 2^J. For a time-series X_t with arbitrary sample size n, the j^{th} level MODWT wavelet ($\tilde{W}_j$) and scaling ($\tilde{V}_j$) coefficients are defined as,

$$\tilde{W}_{j,t} \equiv \sum_{l=0}^{L_j-1} \tilde{h}_{j,l} X_{t-l \bmod N} \text{ and } \tilde{V}_{j,t} \equiv \sum_{l=0}^{L_j-1} \tilde{g}_{j,l} X_{t-l \bmod N} \tag{11}$$

where $\tilde{h}_{j,l} \equiv h_{j,l}/2^{j/2}$ the j^{th} level wavelet filters, and $\tilde{g}_{j,l} \equiv g_{j,l}/2^{j/2}$ are the j^{th} level scaling filters. L_j is the width of j^{th} level filter.

2.3.1 Estimating long memory by wavelets

For estimating the long memory parameter of ARFIMA model, the algorithm based on wavelet (Jensen 1999) is followed. Let X_t be a mean zero *I(d)* process with $0<d<1/2$. Using the autocovariance function of the *I(d)* process, Jensen (Jensen 1999) found that as $j \rightarrow 0$, the wavelet coefficients, W_{jk} associated with a mean zero *I(d)* process are distributed as $N(0, \sigma^2 2^{-2jd})$, 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, j, but is independent of the translation parameter, k. The correlation of the wavelet coefficients from an *I(d)* process decay exponentially over time and scale. Hence, define *R(j)* to be the wavelet coefficients variance at scale j, i.e. $R(j) = \sigma^2 2^{-2jd}$. Taking the logarithmic form of $\boldsymbol{R(j)}$, we obtain the relationship $\ln R(j) = \ln\sigma^2 - d\ln 2^{2j}$ where $\ln \boldsymbol{R(j)}$ is linearly related to $\ln\ 2\text{-}2j$ by the fractional differencing parameter d. Hence, the unknown parameter d of a fractionally integrated series can be estimated by the OLS estimator $\hat{d}$.

2.4 Testing of Stationarity

For testing of stationarity most popularly used methods *viz.,* Augmented Dickey Fuller (ADF) unit root test and Phillips-Perron unit root tests are used. The details of these tests are found in many literature (Dickey and Fuller 1979, Phillips and Perron 1988).

2.5 Fundamentals of ANN Model

Artificial neural networks (ANNs) are nonlinear model that are able to capture various nonlinear structures present in the data set. ANN model specification does not require prior assumption of the data generating process, instead it is largely depending on characteristics of the data. Single hidden layer feed forward network is the most popular for time series modeling and forecasting. The ANN model is characterized by a network of three layers of simple processing units, and thus termed as multilayer ANNs. The first layer is input layer and the last layer is output layer of dependent variable. The remaining layer in the model is called as hidden layer. The relationship between the output (X_t) and the inputs ($X_{t-1}, X_{t-2}, \ldots, X_{t-p}$) can be mathematically represented as follows:

$$X_t = \alpha_0 + \sum_{j=1}^{q} \alpha_j \, g\left(\beta_{0j} + \sum_{i=1}^{p} \beta_{ij} X_{t-i}\right) + \varepsilon_t \tag{12}$$

where, α_j ($j = 0,1,2,\ldots,q$) and β_{ij} ($i = 0,1,2,\ldots, p$, $j = 0,1,2, \ldots , q$) are the model parameters often called the connection weights, p is the number of input nodes and q is the number of hidden nodes. Activation function defines the relationship between inputs and outputs of a network in terms of degree of the non-linearity. Most commonly used activation function is logistic function which is often used as the hidden layer transfer function, i.e.

$$g(x) = \frac{1}{(1 + exp(-X_t))} \tag{13}$$

Thus ANN model performs a nonlinear functional mapping between the input and output which characterized by a network of three layers of simple processing units connected by acyclic links

$$X_t = f\left(X_{t-1} + X_{t-2}, \ldots, X_{t-p}, w\right) + \varepsilon_t \tag{14}$$

where, w is a vector of all parameters and f *is* a function of network structure and connection weights. Therefore, the neural network resembles a nonlinear autoregressive model. Expression (12) indicates one output node in the output layer which is commonly used as one-step-ahead forecasting in out of sample forecast (Zhang *et al.* 1998). Graphically, the ANN model can be expressed in Fig.1.

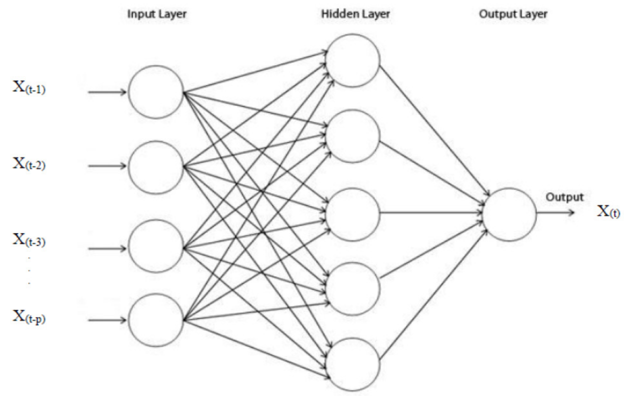

Fig.1. Neural network structure

The selection of appropriate number of hidden nodes as well as optimum number of lagged observation p for input vector is important in ANN modeling for determination of the autocorrelation structure present in a time series. Though there are no established theories available for the selection of p and q, various training algorithms have been used for the determination of the optimal values of p and q. The objective of training is to minimize the error function that measures the misfit between the predicted value and the actual value. The error function which is widely used is mean squared error which can be written as

$$E = \frac{1}{N}\sum_{t=1}^{N}(e_t)^2$$

$$= \frac{1}{N}\sum_{t=1}^{N}\left\{X_t - \left(w_0 + \left(\sum_{J=1}^{Q} w_J g(w_{0j} + \sum_{i=1}^{p} w_{ij} X_{t-i})\right)\right)\right\}^2 \tag{15}$$

where N is the total number of error terms. The parameters of the neural network w_{ij} are changed by an amount of changes in $_{ij}$ as

$$\Delta w_{ij} = -\eta \frac{\partial E}{\partial w_{ij}} \tag{16}$$

where, η is the learning rate. The error surface of multilayer feed forward neural network with non-linear activation function is complex in nature and believed to have many local and global minima.

3. PROPOSED HYBRID METHODOLOGY

ARFIMA models are well suited for modeling linear relation in a data generating process and may not be appropriate for nonlinear problems. On the other hand, ANN models may not be useful for modeling linear behavior present in a data. In real life it is difficult to completely understand the characteristics of data, for such problems hybrid methods can be an effective and efficient alternative. Schematic representation of proposed methodology is expressed in Fig. 2.

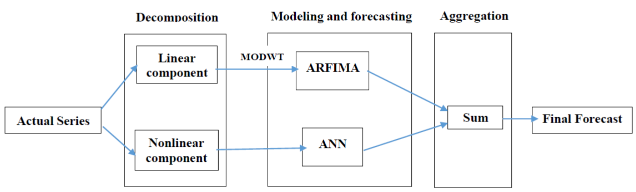

Fig. 2. Proposed methodology

Hybrid methodology consists of both linear and non-linear components which can be a good platform for practical purposes. The model can be written as follows

$$X_t = L_t + N_t \tag{17}$$

where X_t is the original time series, L_t denotes the linear part and N_t denotes the non-linear part. In this work the linear component (L_t) is estimated by ARFIMA model using MODWT and residuals obtained from the ARFIMA model are considered as non-linear part (N_t) and which are obtained as follows

$$e_t = y_t - \hat{L}_t \tag{18}$$

are examined for linearity assumption of the model. The linearity of the residuals is tested using BDS test (Brock *et al*. 1996). If the residuals are found to be non-linear then ANN model can be used for modeling and prediction of these residuals. The ANN model for the residuals can be written as

$$e_t = f(e_{t-1}, e_{t-2}, \ldots\ldots, e_{t-n}) + \varepsilon_t \tag{19}$$

where, f is a non-linear function obtained by ANN and ε_t is the random error. The estimation of equation (19) will result in prediction of non-linear component of time series. Finally, the residuals predicted from ANN are summed with forecasts obtained from ARFIMA model. Therefore, the combined forecast will become

$$\hat{X}_t = \hat{L}_t + \hat{N}_t \tag{20}$$

The technique of hybrid methodology has the strength of unique characteristics of ARFIMA and ANN model to capture the linear and non-linear patterns present in the data set.

4. RESULT AND DISCUSSION

For the present study, the daily spot price (Rupees/Quintal) of agricultural commodity; mustard in Mumbai market for the period 1st January, 2009 to 14th February, 2012 are used. The data is collected form Ministry of Consumer's Affairs, Government of India. Out of 1140 total observations, 1080 have been used for model estimation and remaining 60 observations are used for validation. Summary statistics of mustard spot price is given in Table 1. To validate the stationarity of the series, two tests namely Augmented Dickey-Fuller test and Philips-Peron test are used. Results of the stationarity tests are reported in Table 2. The result indicate that spot price time series data of mustard in Mumbai is stationary.

Table 1. Summary statistics of mustard spot price

Statistic	Series	Statistic	Series
Observation	1140	Standard Deviation	320.41
Mean	2849.89	Kurtosis	1.60
Median	2900.00	Skewness	-0.75
Mode	2750.00	Coefficient of Variation (%)	11.24

Table 2. Testing for stationarity

ADF test statistic				PP test statistic			
Single mean	With trend	Probability		Single mean	With trend	Probability	
		Single mean	With trend			Single mean	With trend
-6.24	-8.44	<0.001	<0.001	-7.65	-7.81	<0.001	<0.001

The autocorrelation function (ACF) and partial autocorrelation function (PACF) for the actual price series were investigated and it has been found that though the stationarity tests validated that the series is stationary, but plot of ACF shows a slow decay towards zero indicating the possible presence of long memory (Paul *et al*. 2015). Therefore, presence of long memory is tested as discussed in Section 2.1.

In the earlier article of Paul *et al.* (2015) the classical parameter estimation methods like GPH, semiparametric methods are compared with Wavelet methods. The study was performed in both simulation as well as in real data set and found that wavelet methods of parameter estimation outperformed the conventional parameter estimation methods, based on these results, the wavelets methods used for parameter estimation in the present study. The estimate of long memory parameter d by wavelet method is found to be 0.372 (sd=0.19). The fractional differenced series with parameter (d) as 0.372 is computed.

Here we estimated different ARFIMA specifications for the data under consideration. On the basis of smallest values of AIC, BIC and absolute log-likelihood, the best ARFIMA model was chosen (Table 3). The estimate of the parameters along with z-statistics for the selected ARFIMA models are given in Table 4, which indicates evidence of long memory in the series with $0<d<0.5$. Therefore, empirical evidence shows that as the lag length increases the autocorrelations decay hyperbolically to zero.

Table 3. Log likelihood, AIC and BIC values of different ARFIMA models

Models	Log-likelihood	AIC	BIC
ARFIMA(*1,d,1*)	-5695.16	11398.3	14950.3
ARFIMA(*0,d,1*)	-5789.37	11498.0	14986.0
ARFIMA(*1,d,0*)	-5700.48	11407.0	14961.0
ARFIMA(*2,d,0*)	-5996.86	11401.7	15953.7
ARFIMA(*0,d,2*)	-5950.90	11909.8	16461.8
ARFIMA(*2,d,2*)	-6075.70	11973.2	16489.1

Table 4. Parameter estimates of ARFIMA model

Parameters	Estimates	Std. Error	z-value	Pr(>|z|)
Constant	2849.9	43.255	65.886	<0.001
AR1	0.913	0.0141	64.838	<0.001
MA1	0.132	0.039	3.324	<0.001

As we discussed in hybrid methodology in section 3 the model contains both linear and non-linear components. The linear components of the model are estimated by ARFIMA model and for nonlinear components, residuals are needed to check for their linearity assumption (Zhang 2003). Plot obtained in Fig. 1 indicates that the residuals of ARFIMA model are nonlinear and the BDS test is confirmed the presence of nonlinearity (Table 5).

Table 6. ANN parameters

Particulars	ANN parameter
Cross validation	10 fold
Optimum lag	3
Optimum hidden node	2
learning algorithm	Gradient descent
Network type	(3,2,1) Feed forward
Activation function	Linear Sigmoidal
Learning rate	0.0005
Momentum	0.004
Converge at	112 epochs
Total no. of parameters	19

Table 5. BDS test for linearity

Dimension (m)	Epsilon (ε)		Statistic	Probability	Dimension (m)	Epsilon (ε)		Statistic	Probability
2	eps(1)	160.26	65.71	<0.001	3	eps(1)	160.26	90.27	<0.001
	eps(2)	320.52	48.63	<0.001		eps(2)	320.52	54.22	<0.001
	eps(3)	480.78	42.73	<0.001		eps(3)	480.78	44.10	<0.001
	eps(4)	641.03	44.05	<0.001		eps(4)	641.03	43.7	<0.001

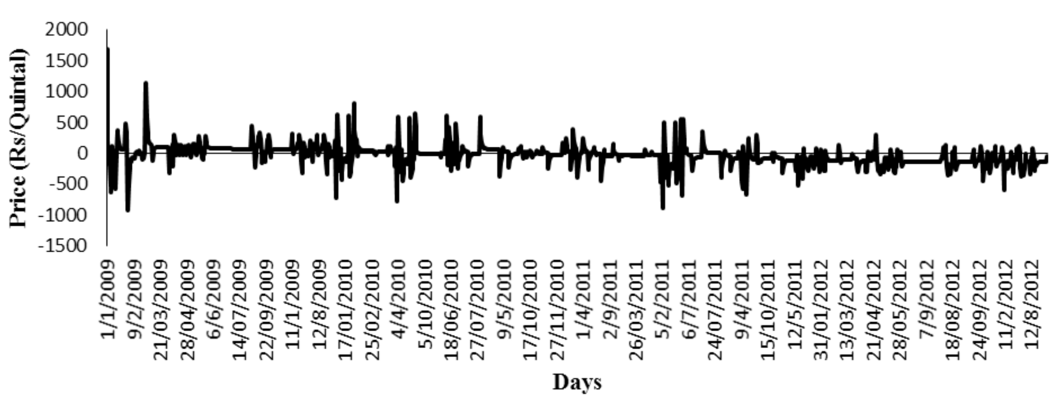

Fig. 1. Plot of residuals of ARFIMA model

According to hybrid methodology, as the residuals are found to be nonlinear, they can be estimated and predicted individually by ANN model. The optimum parameters of the ANN are given in Table 6 where, optimum number of lags are 3, number of hidden nodes are 2 and 1 is number of output. Forecast values of ARFIMA and Residuals predicted by ANN model are summed and prediction efficiency of ARFIMA and Hybrid model are compared (Table 7). Mean absolute percentage error (MAPE) of the models are computed and it is found that MAPE of the proposed methodology is less as compared to the ARFIMA model. The graphical comparison between both the models are presented in Fig. 2.

Table 7. Forecast accuracy of ARFIMA and Hybrid (ARFIMA + ANN) method

Lead Period	Actual Price	ARFIMA	% deviation	ARFIMA +ANN	% deviation	Lead Period	Actual Price	ARFIMA	% deviation	ARFIMA +ANN	% deviation
1	3000	3140.48	4.68	2985.83	0.47	31	2900	3085.94	6.41	2788.37	3.85
2	3000	3131.22	4.37	3002.29	0.08	32	2700	3061.28	13.38	2997.07	11.00
3	3000	3127.28	4.24	3000.94	0.03	33	2700	2952.69	9.36	2700.91	0.03
4	3100	3125.66	0.83	3000.88	3.20	34	2900	2905.48	0.19	2829.17	2.44
5	3100	3174.01	2.39	3097.12	0.09	35	2900	2982.94	2.86	2751.16	5.13
6	3300	3195.17	3.18	3127.89	5.22	36	2900	3016.77	4.03	2982.49	2.84
7	3300	3302.42	0.07	3166.44	4.05	37	2900	3031.61	4.54	2925.83	0.89
8	2760	3349.23	21.35	3270.34	18.49	38	2900	3038.16	4.76	2992.27	3.18
9	2760	3105.3	12.51	2681.40	2.85	39	2900	3041.11	4.87	2923.66	0.82
10	2780	2999.12	7.88	2764.47	0.56	40	2700	3042.48	12.68	2928.81	8.47
11	2780	2962.75	6.57	2795.32	0.55	41	2900	2945.24	1.56	2428.92	16.24
12	2700	2946.99	9.15	2820.10	4.45	42	2900	3000.9	3.48	2887.14	0.44
13	2780	2901.05	4.35	2722.88	2.05	43	3100	3025.24	2.41	2970.45	4.18
14	2700	2920.3	8.16	2807.98	4.00	44	3100	3133.87	1.09	2821.42	8.99
15	2700	2889.61	7.02	2770.70	2.62	45	2900	3181.29	9.7	3135.53	8.12
16	2900	2876.33	0.82	2778.79	4.18	46	3000	3104.11	3.47	2872.79	4.24
17	2900	2968.57	2.36	2963.54	2.19	47	2900	3119.54	7.57	3021.09	4.18
18	2900	3008.85	3.75	2937.09	1.28	48	2900	3077.39	6.12	2863.50	1.26
19	2700	3026.49	12.09	2927.93	8.44	49	2900	3059.12	5.49	2869.56	1.05
20	2700	2936.34	8.75	2635.00	2.41	50	2900	3051.25	5.22	2877.02	0.79
21	2810	2897.15	3.1	2762.40	1.69	51	2900	3047.91	5.1	2935.57	1.23
22	2900	2934.04	1.17	2851.72	1.66	52	2900	3046.55	5.05	2903.79	0.13
23	2900	2994.27	3.25	2815.39	2.92	53	2900	3046.04	5.04	2478.50	14.53
24	3100	3020.6	2.56	2937.24	5.25	54	2900	3045.92	5.03	3158.59	8.92
25	3100	3130.1	0.97	2945.17	4.99	55	2900	3045.95	5.03	2869.38	1.06
26	3300	3177.9	3.7	3256.80	1.31	56	2900	3046.06	5.04	2778.28	4.20
27	3300	3296.75	0.1	2840.36	13.93	57	2900	3046.2	5.04	2721.91	6.14
28	3100	3348.62	8.02	3235.89	4.38	58	2900	3046.35	5.05	2917.68	0.61
29	2900	3273.38	12.88	2983.85	2.89	59	2900	3046.5	5.05	2947.72	1.65
30	2900	3142.76	8.37	2866.33	1.16	60	3000	3046.66	1.56	2833.57	5.55
	MAPE		**5.41**	**MAPE**	**3.99**		**MAPE**		**5.41**	**MAPE**	**3.99**

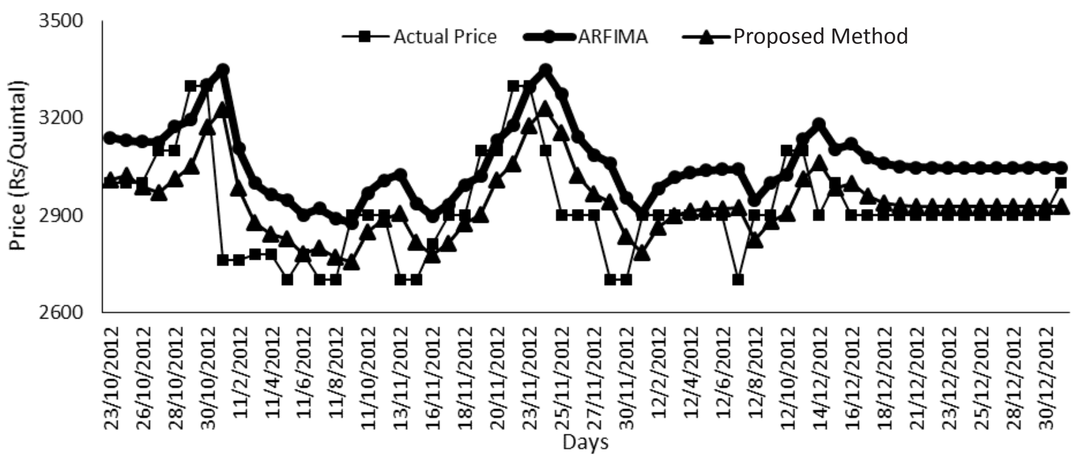

Fig. 2. Plot of ARFIMA v/s proposed methodology

5. CONCLUSION

Long memory time series has been analyzed by using ARFIMA models which are based on linear structure. ARFIMA models are not always adequate for long memory time series that have both linear and non-linear structures. In this context, the hybrid method which combines both linear and nonlinear part can be an effective way to improve forecasting performance. Based on the results obtained in this work one can infer that a hybrid model of ARFIMA and ANN increase forecasting accuracy. This is also an important result for the ANN studies in the future. This approach can be further extended by using some other machine learning techniques for varying autoregressive and moving average orders so that practical validity of the model can be well established.

ACKNOWLEDGMENTS

The authors would like to thank ICAR-IASRI New Delhi, SSIU-ISI Bangalore and Arnab Datta, Bangalore for helping in this work.

The authors would like to thank the anonymous reviewer for his/her valuable comments which are important in improving the quality of manuscript.

REFERENCES

Aarthi, R.S., Muralidharan, D. and Swaminathan, P. (2012). Double compression of test data using Huffman code. *J. Theo. Appl. Inform. Tech.*, **39(2)**, 104-113.

Asadi, S., Tavakoli, A. and Hejazi, S.R. (2012). A new hybrid for improvement of auto-regressive integrated moving average models applying particle swarm optimization. *Expert Sys. Appl.*, **39(5)**, 5332-5337.

Box, G.E.P. and Jenkins, G.M. (1970). *Time Series Analysis: Forecasting and Control.* Holden-Day, San Francisco.

Brock, W.A., Dechert, W.D., Steinman, J.A. and Lebaron, B. (1996). A test for independence based on the correlation dimension. *Eco. Rev.*, **15**, 197-235.

Chaabane, N. (2014). A hybrid ARFIMA and neural network model for electricity price prediction. *Inter. J. Elect. Power Energy Sys.*, **55**, 187-194.

De Gooijer, J.G. and Kumar, K. (1992). Some recent developments in non-linear time series modelling, testing, and forecasting. *Inter. J. Forecast.*, **8**, 135-156.

Dickey, D. and Fuller, W. (1979). Distribution of the estimators for autoregressive time series with a unit root. *J. Amer. Statist. Assoc.*, **74**, 427-431.

Faruk, D.O. (2010). A hybrid neural network and ARIMA model for water quality time series prediction. *Engg. Appl. Artif. Intell.*, **23**, 586-594.

Geweke, J. and Porter-Hudak, S. (1983). The estimation and application of long-memory time-series models. *J. Time Series Anal.*, **4**, 221-238.

Ghosh, H., Paul, R.K. and Prajneshu (2010). Wavelet frequency domain approach for statistical modeling of rainfall time-series data. *J. Statist. Theo. Prac.*, **4(4)**, 813-825.

Gil-Alana, L.A., Barros, C. and Peypoch N. (2014). Long memory and fractional integration in the housing price series of London and Paris. *Appl. Eco.*, **46(27)**, 3377-3388.

Granger, C.W.J. and Joyeux, R. (1980). An introduction to long-memory time-series models and fractional differencing. *J. Time-Series Anal.*, **4**, 221-238.

Hsua, N.J. and Tsaib, H. (2009). Semiparametric estimation for seasonal long-memory time series using generalized exponential models. *J. Statist. Plann. Inf.*, **139**, 1992-2009.

Jensen, M.J. (1999). Using wavelets to obtain a consistent ordinary least squares estimator of the long-memory parameter, *J. Forecast.*, **18**, 17-32.

Jha, G.K. and Sinha, K. (2014). Time-delay neural networks for time series prediction: an application to the monthly wholesale price of oilseeds in India. *Neural Comp. Appl.*, **24(3)**, 563-571.

Khashei, M., Bijari, M., and Raissi, A. (2009). Improvement of auto-regressive integrated moving average models using fuzzy logic and artificial neural networks (ANNs). *Neurocomputing*, **72,** 956-967.

Khashei, M., Bijari, M., and Raissi, A. (2012). Hybridization of autoregressive integrated moving average (ARIMA) with probabilistic neural networks (PNNs). *Comput. Ind. Engg.*, **63(1)**, 37-45.

Lima, L.R., and Xiao, Z. (2010). Is there long memory in financial time series. *Appl. Finance. Eco.*, **20(6)**, 487-500.

Lu, Z. and Guegan, D. (2011). Estimation of time-varying long memory parameter using wavelet method. *Comm. Statist. - Sim. Comput.*, **40(4)**, 596-613.

Mohamed, B. (2009). Comparison of non-parametric and semiparametric tests in detecting long memory. *J. Appl. Statist.*, **36(9)**, 945-972.

Nielsen, M., and Frederickson, P.H. (2005). Finite sample comparison of parametric, semiparametric, and wavelet estimators of fractional integration. *Econ. Rev.*, **24(4),** 405-443.

Paul, R.K., Prajneshu, and Ghosh, H. (2011). Wavelet methodology for estimation of trend in Indian monsoon rainfall time-series data. *Ind. J. Agric. Sci.*, **81(3),** 96-98.

Paul, R.K., Prajneshu, and Ghosh, H. (2013). Wavelet frequency domain approach for modelling and forecasting of indian monsoon rainfall time-series data. *J. Ind. Soc. Agril. Statist.*, **67(3)**, 319-327

Paul, R.K. (2014). Forecasting wholesale price of pigeon pea using long memory time-series models. *Agric. Econ. Res. Rev.*, **27(2)**, 167-176.

Paul, R.K., Samanta, S. and Gurung, B. (2015). Monte Carlo simulation for comparison of different estimators of long memory parameter: An application of ARFIMA model for forecasting commodity price. *Model Assist. Statist. Appl.*, **10(2)**, 117-128.

Paul, R.K., Gurung, B. and Paul, A.K. (2015). Modelling and forecasting of retail price of arhar dal in Karnal, Haryana. *Ind. J. Agric. Sci.*, **85(1)**, 69-72.

Pektas, A.O. and Cigizoglu, H.K. (2013). ANN hybrid model versus ARIMA and ARIMAX models of runoff coefficient. *J. Hydrology*, **500**, 21-36.

Percival, D B. and Walden, A.T. (2000). *Wavelet Methods for Time Series Analysis*. Cambridge University Press, Cambridge.

Phillips, P.C.B. and Perron, P. (1988). Testing for unit roots in time series regression. *Biometrika*, **75**, 335-346.

Ramalingam, S. (2010). A unique book on long memory time series is reviewed. *J. Statist. Comput. Sim.*, **80(4)**, 475-475.

Ray, M., Rai, A., Ramasubramanian, V. and Singh, K.N. (2016). ARIMA-WNN hybrid model for forecasting wheat yield time series data. *J. Ind. Soc. Agril. Statist*., **70(1),** 63-70**.**

Robinson, P.M. (1995). Log-periodogram regression of time-series with long-range dependence. *Ann. Statist.*, **23(3),** 1048-1072.

Shan, R., Dai, H., Zhao, H. and Liu, W. (2015). Forecasting study of Shanghai's and Shenzhen's stock markets using a hybrid forecast method. *Comm. Statist. - Sim. Comput.*, **44(4)**, 1066-1077.

Sowell, F. (1992). Maximum likelihood estimation of stationary univariate fractionally integrated time series models. *J. Econ.*, **53**, 165-188.

Tzouras, S., Anagnostopoulas, C. and McCoy, E. (2015). Financial time series modeling using the Hurst exponent. *Physica A: Statistical Mechanics and its Applications*, **425,** 50-68.

Zhang, G.P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*, **50**, 159-175.

Zhang, G., Patuwo, E.B. and Hu, M.Y. (1998). Forecasting with artificial neural networks: the state of the art. *Inter. J. Forecast*. **14**, 35-62.