Forecasting Wholesale Price of Pigeon Pea Using Long Memory Time-Series Models

Ranjit Kumar Paul

Indian Agricultural Statistics Research Institute, New Delhi-110012

Abstract

The fractional integration is a generalization of integer integration, under which time-series are usually presumed to be integrated of order zero or one. In this regard, the autoregressive fractionally integrated moving-average (ARFIMA) model along with its estimation procedure is investigated. ARFIMA model searches for a non-integer differencing parameter $d$ to difference the data to capture long memory. The model has been applied for modelling and forecasting of daily wholesale price of pigeon pea ($Cajanus cajan$) in the Amritsar and Bhatinda markets and the all-India maximum, minimum and modal prices of pigeon pea. Augmented Dickey-Fuller (ADF) test and Philips Perron (PP) test have been used for testing the stationarity of the series. Autocorrelation (ACF) and partial autocorrelation (PACF) functions have shown a slow hyperbolic decay indicating the presence of long memory. In all the five price series, long memory parameters are found to be significant. On the basis of minimum AIC values, the best model was identified for each series. To this end, evaluation of forecasting was carried out with root mean squares prediction error (RMSPE), mean absolute prediction error (MAPE) and relative mean absolute prediction error (RMAPE). The residuals of the fitted models have been used for diagnostic checking. Out-of sample forecast of wholesale prices of pigeon pea has been computed up to February, 2014. The R software package has been used for data analysis.

Key words: ADF test, ARFIMA model, long memory, PP test, pigeon pea, stationarity

JEL Classification: C32, C53

Introduction

In time series modelling, the fitted model should produce theoretical correlations close to the sample correlations calculated from the data. While, correlations of a time series are expected to vanish ultimately when observations are far apart in time, the speed of decay can be different. For example, in an Autoregressive Moving Average (ARMA)-type process, the correlations decay exponentially as the time lag increases; in other time series, the decay can occur at a much slower hyperbolic rate. The latter type of time series is said to have long memory or long-range dependence and is frequent in economic time series. Although most economic time series are non-stationary and do require differencing of some kind, it is not necessarily true that taking first differences and then using an ARMA model will be the best remedy. In Box-Jenkins analysis, it is assumed that if the series is non-stationary, the first difference will be well-behaved, as long as there are no seasonal components. In particular, it is likely that differenced series will have rapidly decaying autocorrelations and will be free of trend-like behaviour, so that it can be well described by a stationary invertible ARMA model. But, this is not always the case. The literature on ARMA models is very vast. There are several applications of these models and their extensions in agriculture; for example, Paul and Das (2010; 2013) and Paul et al. (2013a; b).
For modelling time series in the presence of long memory, the autoregressive fractionally integrated moving-average (ARFIMA) model is used. The ARFIMA model searches for a non-integer parameter, $d$, to differentiate the data to capture long memory. Regarding long memory, the useful entry points to the literature are the surveys by Robinson (1995) and Baillie et al. (1996), who have described the developments in the econometric modelling of long memory, and of Beran (1994) who has reviewed long-memory modelling in other areas. The existence of non-zero $d$ is an indication of long memory and its departure from zero measures the strength of long memory. Long memory is also called fractal structure because of non-integer $d$.

Fung et al. (1994) have investigated long memory in four daily currency futures price series. Similar studies have been conducted on stock markets (Lo, 1991; Chow et al., 1995), inflation (Scacciavillani, 1994; Hassler and Wolters, 1995), gold prices (Cheung and Lai, 1993) and foreign exchange (Booth et al., 1982). The results are mixed, but all authors agree that identification of long memory is highly significant in at least two senses: (i) the time span and strength of long memory will be an important input for investment decisions regarding investment horizons and composition of portfolios; and (ii) predictions about price movements will improve. It is also noticeable that research methodologies on long memory time series have developed very fast. During the 1980s, the classical rescaled range (R/S) analysis was the major tool. Entering the 1990s, the methods were diversified with the modified R/S analysis and the ARFIMA model as a new technique. The long memory study on agricultural futures markets is at the beginning. Helms et al. (1984) have analyzed the short series of one commodity using only the classical R/S techniques. So there is a need to investigate the long memory behaviour in the agricultural prices. The present paper has investigated the structure of long memory in daily wholesale prices of pigeon pea in different markets in India.

India is the largest producer of pigeon pea with a contribution of around 85 per cent to the world’s total production. The prices of pigeon pea in India are based on the Minimum Support Price set by the government. The pigeon pea imports are allowed in the country without any import restrictions. The present study will take advantage of the new developments in the statistical methods to analyze the time series data of wholesale price of pigeon pea in different markets of India for modelling and forecasting purposes.

**Long Memory Process**

Long memory in time-series can be defined as autocorrelation at long lags (Robinson, 1995). According to Jin and Frechette (2004), memory means that observations are not independent (each observation is affected by the events that preceded it). The autocorrelation function (acf) of a time-series $y_t$ is defined as:

$$\rho_k = \frac{\text{cov}(y_t, y_{t-k})}{\text{var}(y_t)} \quad \ldots(1)$$

for integer lag $k$. A covariance stationary time-series process is expected to have autocorrelations such that $\lim_{k \to \infty} \rho_k = 0$. Most of the well-known class of stationary and invertible time-series processes have autocorrelations that decay at an exponential rate, so that $\rho_k \approx |m|^k$, where $|m|<1$ and this property is true, for example, for the well-known stationary and invertible ARMA($p,q$) process. For long memory processes, the autocorrelations decay at an hyperbolic rate which is consistent with $\rho_k \approx Ck^{2d-1}$, as $k$ increases without limit, where $C$ is a constant and $d$ is the long memory parameter.

Based on Beran (1995, pp. 41-66), a stationary process with long memory has the following qualitative features:

- Certain persistence exists. The observations tend to stay at high levels in some periods, and at low levels in some other periods.
- During short-time periods, there seems to be periodic cycles. However, looking through the whole process, no apparent periodic cycles could be identified.
- Overall, the process looks stationary.

Quantitatively, for a stationary process, these features could be described as:

- When adding more observations, the variance of the sample mean, $\text{var}(Y)$, decays to zero at a slower rate than $n^{-1}$, which is the rate at which a white noise decays, and is asymptotically equal to a constant $g$ times $n^{c}$ for some $0 < c < 1$. 

• The correlation $r_j$ decays to zero slowly and is asymptotically equal to a constant times $j^c$ for some $0 < c < 1$.

### Testing of Long Memory

Hurst exponent (H), produced by the rescaled range (R/S) analysis is used to test the presence of long memory in a time-series. It was developed and applied to economic price analysis by Booth et al. (1982) and Helms et al. (1984). For a given time-series, the Hurst exponent measures the long-term non-periodic dependence, and indicates the average duration the dependence may last. The R/S analysis first estimates the range $R$ for a given $n$:

$$R(n) = \max_{1 \leq j \leq n} \sum_{i=1}^{n} (Y_j - \bar{Y}) - \min_{1 \leq j \leq n} \sum_{i=1}^{n} (Y_j - \bar{Y})$$  \hspace{1cm} \text{(2)}$$

where, $R(n)$ is the range of accumulated deviation of $Y(t)$ over the period of $n$ and $\bar{Y}$ is the overall mean of the time-series. Let $S(n)$ be the standard deviation of $Y$, over the period of $n$.

For a given $n$, there exists a statistic

$$Q(n) = R(n)/S(n)$$  \hspace{1cm} \text{(3)}$$

Here, $n$ is the time scale to split total observations $T$ into $\text{int}[T/n]$ segments where $\text{int}[.]$ denotes the integer part of [.]. There will be $\text{int}[T/n]$ estimates of $R(n)/S(n)$ for a given $n$. The final $R(n)/S(n)$ is the average of $\text{int}[T/n]$'s $R(n)/S(n)$. As $n$ increases, the following holds:

$$R(n)/S(n) = \alpha \cdot n^H$$

where $\alpha$ is a constant.

or

$$\log[R(n)/S(n)] = \log \alpha + H \log n$$  \hspace{1cm} \text{(4)}$$

Thus, $H$ is a 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. $H$ is always greater than 0. 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. The relationship between Hurst exponent and long memory parameter is: $H = 1 - d$.

### ARFIMA Model

Fractional integration is a generalization of integer integration, under which time-series are usually presumed to be integrated of order zero or one. For example, an autoregressive moving-average process integrated of order $d$ [denoted by ARFIMA($p, d, q$)] can be represented as Equation (5):

$$(1 - L)^d \phi (L)Y_t = \theta (L)u_t$$  \hspace{1cm} \text{(5)}$$

where, $u_t$ is an independently and identically distributed (i.i.d.) random variable with zero mean and constant variance, $L$ denotes the lag operator; and $\phi (L)$ and $\theta (L)$ denote finite polynomials in the lag operator with roots outside the unit circle. For $d = 0$, the process is stationary, and the effect of a shock $u(t)$ on $y(t + j)$ decays geometrically as $j$ increases. For $d = 1$, the process is said to have a unit root, and the effect of a shock $u(t)$ on $y(t + j)$ persists into the infinite future. In contrast, fractional integration defines the function $(1 - L)^d$ for non-integer values of the fractional differencing parameter $d$.

For $-0.5 < d < 0.5$ the process $y(t)$ is stationary and invertible. For such processes, the effect of a shock $u(t)$ on $y(t + j)$ decays as $j$ increases, but the rate of decay is much slower than for a process integrated of order zero. More precisely, the acf for zero-integrated processes decays geometrically, whereas the acf for a fractionally-integrated process decays hyperbolically, with the sign of autocorrelations being the same as of $d$. In this sense, fractional integration captures long memory dynamics more parsimoniously than non-integrated ARMA processes. In the use of ARFIMA($p,d,q$) models, correct specification of $p$ and $q$ leads to inconsistent estimation of AR and MA coefficients, but also of long memory parameter $d$, as does over-specification of both, due to a loss of identifiability.

### Estimation of Long Memory Parameter

For estimating the long memory parameter, GPH estimator proposed by Geweke and Porter-Hudak (1983) was used in the present investigation. A brief description of the method is given below:

This method is based on an approximated regression equation obtained from the logarithm of the spectral density function. The GPH estimation
procedure is a two-step procedure, which begins with the estimation of \( d \). This method is based on least squares regression in the spectral domain, exploits the sample form of the pole of the spectral density at the origin: \( f, (\lambda) \sim \lambda^{-2d}, \lambda \to 0 \). To illustrate this method, we can write the spectral density function of a stationary model \( y_t, t = 1, \ldots, T \) as

\[
f_y (\lambda) = \left[ 4 \sin^2 \left( \frac{\lambda}{2} \right) \right]^{-d} f_y (\lambda)
\]

where, \( f_y (\lambda) \) is the spectral density of \( \varepsilon_t \), assumed to be a finite and continuous function on the interval \([-\pi, \pi]\). Taking the logarithm of the spectral density function \( f_y (\lambda) \) the log-spectral density can be expressed as Equation (9):

\[
\log(f_y (\lambda)) = \log(f_y (0)) - d \log\left[ 4 \sin^2 \left( \frac{\lambda}{2} \right) \right] + \log\left( \frac{f_y (\lambda)}{f_y (0)} \right)
\]

where, \( \log\left( \frac{f_y (\lambda)}{f_y (0)} \right) \) is a constant, \( \log\left[ 4 \sin^2 \left( \frac{\lambda}{2} \right) \right] \) is the exogenous variable and \( \log\left( \frac{f_y (\lambda)}{f_y (\lambda_j)} \right) \) is a disturbance error. The GPH estimate requires two major assumptions related to asymptotic behaviour of the equation

\[
H_1: \text{for low frequencies, we suppose that } \log\left[ \frac{f_y (\lambda)}{f_y (0)} \right] \text{ is negligible.}
\]

\[
H_2: \text{the random variables } \log\left( \frac{I_y (\lambda_j)}{I_y (\lambda_j)} \right); \quad j = 1, 2, \ldots, m \text{ are asymptotically iid.}
\]

Under the hypotheses \( H_1 \) and \( H_2 \), we can write the linear regression as

\[
\log\left( \frac{I_y (\lambda_j)}{I_y (\lambda_j)} \right) = \alpha - d \log\left[ 4 \sin^2 \left( \frac{\lambda_j}{2} \right) \right] + e_j
\]

where, \( e_j \sim \text{iid } (-c, \pi^2/6) \).

Let \( y_j = - \log\left( 4 \sin^2 \left( \frac{\lambda_j}{2} \right) \right) \) the GPH estimator is the OLS estimate of the regression \( \log I_y (\lambda_j) \) on the constant \( \alpha \) and \( y_j \). The estimate of \( d \) is given by Equation (12):

\[
\hat{d}_{GPH} = \frac{\sum_{j=1}^{m} (y_j - \bar{y}) \log\left( \frac{I_y (\lambda_j)}{I_y (\lambda_j)} \right)}{\sum_{j=1}^{m} (y_j - \bar{y})^2}, \text{ where, } \bar{y} = \frac{1}{m} \sum_{j=1}^{m} y_j
\]

Robinson (1995), Hurvich et al. (1998) and Tanaka (1999) have analyzed the GPH estimate in detail. Under the assumption of normality for \( y_n \), it has been proved that the estimate is consistent and asymptotically normal.

**Illustration**

We have used the daily wholesale prices of pigeon pea (Arhar) in two markets, namely Amritsar and Bhatinda and the all India maximum, minimum and modal wholesale prices of pigeon pea for the period 1 January, 2012 to 31 December, 2013 are used. The data were collected from the website of Ministry of Consumer’s Affairs, Government of India (http://consumeraffairs.nic.in/consumer/index.php). The data for the period 1 January, 2012 to 31 October, 2013 have been used for model building and the remaining data have been used for model validation purpose. The time series plots of the prices are exhibited in Figure 1. A perusal of the plots indicates that the datasets are
stationary. In order to test for stationarity, Augmented Dickey-Fuller unit root test (Said and Dickey, 1984) and Phillips-Perron unit root test (Phillips and Perron, 1988) were conducted, and the results of these tests are given in Table 1. A perusal of Table 1 indicates that all the series are stationary. If there is an apparent trend in the dataset, then stationarity test with trend is used, otherwise stationarity test with single mean is applied.

**Structure of Autocorrelations**

For a linear time series model, typically an autoregressive integrated moving average [ARIMA (p,d,q)] process, the patterns of autocorrelations and partial autocorrelations could indicate the plausible structure of the model. At the same time, this kind of information is also important for modelling non-linear dynamics. The long lasting autocorrelations of the data suggest that the processes are non-linear with time-varying variances. The basic property of a long memory process is that the dependence between the two distant observations is still visible. For the series of daily wholesale price, autocorrelations were estimated up to 100 lags, i.e., \( j = 1, \ldots, 100 \). The autocorrelation functions of these series have been plotted in Figure 2. A perusal of Figure 2 indicates that these do not decay exponentially over time span, rather there is a hyperbolic decay of the autocorrelation functions towards zero and they show no clear periodic patterns. There is no evidence that the magnitude of autocorrelations became small as the time lag, \( j \), became larger. No seasonal and other periodic cycles were observed.

**Testing Stationarity**

**ADF Test**

The ADF test examines the null hypothesis that a time series \( y_t \) is \( I(1) \) against the alternative that it is \( I(0) \), assuming that the dynamics in the data have an ARMA structure. The ADF test is based on estimating the test regression (13):

\[
\Delta y_t = \beta' D_t + \pi y_{t-1} + \sum_{j=1}^{p} \psi_j \Delta y_{t-j} + \epsilon_t \quad \ldots (13)
\]

where, \( D_t \) is a vector of deterministic terms (constant, trend, etc.). The \( p \) lagged difference terms, \( \Delta y_{t-j} \), are
used to approximate the ARMA structure of the errors, and the value of $p$ is set so that the error $\varepsilon_t$ is serially uncorrelated. The error-term is also assumed to be homoskedastic. Under the null hypothesis, $\Delta y_t$ is $I(0)$ which implies that $\pi = 0$. The ADF t-statistic is then the usual t-statistic for testing $\pi = 0$. The ADF test has been applied in the present data sets and the results are reported in Table 1.

**Phillips-Perron Unit Root Tests**

The Phillips-Perron (PP) unit root tests differ from the ADF tests mainly in how they deal with serial correlation and heteroscedasticity in the errors. In particular, where the ADF tests use a parametric autoregression to approximate the ARMA structure of the errors in the test regression, the PP tests ignore any...
serial correlation in the test regression. The test regression for the PP tests is

$$\Delta y_t = \beta' D_t + \pi y_{t-1} + u_t$$  

...(14)

where, $u_t$ is I(0) and may be heteroskedastic. The PP tests correct any serial correlation and heteroscedasticity in the errors $u_t$ of the test regression by directly modifying the test statistics. Under the null hypothesis that $\pi = 0$, the PP statistics have the same asymptotic distributions as the ADF t-statistic. The advantage of the PP tests over the ADF tests is that the PP tests are robust to general forms of heteroscedasticity in the error-term $u_t$. Another advantage is that the user does not have to specify a lag length for the test regression. The PP test has also been applied in the present data sets and the results are reported in Table 1.

The most common method for estimating the fractional integration parameter $d$ is the ARFIMA time series method (Robinson, 2003). We have estimated different ARFIMA specifications, as described previously. Based on the smallest AIC value, the best ARFIMA model was selected. AIC values and log likelihood are reported in Table 2. The minimum AIC and log likelihood (L) are marked bold for individual market. Estimate of the parameters along with t-statistics for the selected ARFIMA models are provided in Table 3.

These estimates indicate the evidence of long memory in five price series with $0 < d < 0.5$. Positive values of the fractional differencing parameter indicate a short of long-memory known as persistence. Persistence is characterised by positive autocorrelations, and exhibit low variance at low frequencies. Note that, when $d$ parameter is positive and significant, then the series may have infinite conditional variance.

The results show that significant $d$ parameter ranges from 0.052 to 0.489 (All India Maximum Price has the highest long memory parameter). Hence, empirical evidence shows that the lag length increases the autocorrelations decay hyperbolically to zero.

**Diagnostic Checking**

The model verification is concerned with checking the residuals of the model to see if they contained any systematic pattern which still could be removed to improve the chosen ARFIMA. This has been done through examining the autocorrelations and partial autocorrelations of the residuals of various lags. For this purpose, autocorrelations of the residuals were computed and it was found that none of these autocorrelations was significantly different from zero at any reasonable level. This proved that the selected ARFIMA model was an appropriate model for forecasting the data under study.

**Validation**

One-step ahead forecasts of wholesale price along with their corresponding standard errors using naïve approach for the period 01 November, 2013 to 31 December, 2013 (total 40 data points excluding market holidays) in respect of above fitted model were computed. The attractive feature for fitted ARFIMA model is that all the observed values lie within one standard error of forecasts.

---

Table 1. Stationarity testing

<table>
<thead>
<tr>
<th>Market</th>
<th>ADF test statistic</th>
<th>PP test statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Single mean With trend</td>
<td>Probability</td>
</tr>
<tr>
<td>Amritsar</td>
<td>5.46 7.08 0.0238 0.0283</td>
<td>-3.29 -3.76 0.0165 0.0199</td>
</tr>
<tr>
<td>Bhatinda</td>
<td>6.85 9.43 0.0010 0.0010</td>
<td>-3.70 -4.27 0.0047 0.0039</td>
</tr>
<tr>
<td>All India</td>
<td>Maximum 13.11 42.71 &lt;.0001 &lt;.0001</td>
<td>-5.12 -9.23 &lt;.0001 &lt;.0001</td>
</tr>
<tr>
<td>Minimum</td>
<td>8.32 11.55 0.0402 0.0010</td>
<td>-3.53 -4.81 0.0195 0.0006</td>
</tr>
<tr>
<td>Modal</td>
<td>15.43 27.00 &lt;.0001 &lt;.0001</td>
<td>-5.55 -7.35 &lt;.0001 &lt;.0001</td>
</tr>
</tbody>
</table>
Table 2. Log likelihood and AIC values of different ARFIMA models

<table>
<thead>
<tr>
<th>Market</th>
<th>ARFIMA (1,d,1)</th>
<th>ARFIMA (1,d,0)</th>
<th>ARFIMA (0,d,1)</th>
<th>ARFIMA (2,d,1)</th>
<th>ARFIMA (2,d,0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amritsar</td>
<td>-2143</td>
<td>-2144</td>
<td>-2172</td>
<td>-2144</td>
<td>-2144</td>
</tr>
<tr>
<td>Bhatinda</td>
<td>-2228</td>
<td>-2230</td>
<td>-2237</td>
<td>-2224</td>
<td>-2230</td>
</tr>
<tr>
<td>All-India maximum</td>
<td>-2417</td>
<td>-2417</td>
<td>-2415</td>
<td>-2415</td>
<td>-2414</td>
</tr>
<tr>
<td>All-India minimum</td>
<td>4841.547</td>
<td>4839.085</td>
<td>4836.55</td>
<td>4839.021</td>
<td>4836.38</td>
</tr>
<tr>
<td>price</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All-India modal</td>
<td>4827.095</td>
<td>4833.798</td>
<td>4834.679</td>
<td>4832.306</td>
<td>4825.111</td>
</tr>
</tbody>
</table>

Table 3. Parameter estimates of ARFIMA Model

<table>
<thead>
<tr>
<th>Market</th>
<th>Parameters</th>
<th>Estimate</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amritsar</td>
<td>d</td>
<td>0.077</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>AR1</td>
<td>0.915</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Bhatinda</td>
<td>d</td>
<td>0.052</td>
<td>0.040</td>
</tr>
<tr>
<td></td>
<td>AR1</td>
<td>1.615</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td></td>
<td>AR2</td>
<td>-0.623</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td></td>
<td>MA1</td>
<td>0.821</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>All-India maximum</td>
<td>d</td>
<td>0.489</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td></td>
<td>AR1</td>
<td>-0.223</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td></td>
<td>AR2</td>
<td>-0.128</td>
<td>0.0168</td>
</tr>
<tr>
<td>All-India minimum</td>
<td>d</td>
<td>0.093</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td></td>
<td>AR1</td>
<td>1.1467</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td></td>
<td>AR2</td>
<td>-0.149</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td></td>
<td>MA1</td>
<td>0.784</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>All-India modal</td>
<td>d</td>
<td>0.477</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td></td>
<td>AR1</td>
<td>-0.157</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td>AR2</td>
<td>0.183</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>

RMSPE = \frac{1}{40} \sum_{i=1}^{40} \left( \frac{(y_{t+i} - \hat{y}_{t+i})^2}{y_{t+i}} \right) \quad \ldots(16)

RMAPE = \frac{1}{40} \sum_{i=1}^{40} \left( \frac{|y_{t+i} - \hat{y}_{t+i}|}{y_{t+i}} \right) \times 100 \quad \ldots(17)

A perusal of Table 4 reveals that in all the price series data, RMAPE is less than 5 percent, indicating the accuracy of the models used in the study.

Table 4. Validation of Models

<table>
<thead>
<tr>
<th>Market</th>
<th>MAPE (%)</th>
<th>RMSPE</th>
<th>RMAPE (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amritsar</td>
<td>195.964</td>
<td>204.773</td>
<td>3.5</td>
</tr>
<tr>
<td>Bhatinda</td>
<td>323.303</td>
<td>333.535</td>
<td>4.8</td>
</tr>
<tr>
<td>All-India maximum</td>
<td>352.963</td>
<td>366.503</td>
<td>4.7</td>
</tr>
<tr>
<td>All-India minimum</td>
<td>168.629</td>
<td>194.520</td>
<td>3.3</td>
</tr>
<tr>
<td>All-India modal</td>
<td>173.679</td>
<td>177.470</td>
<td>3.1</td>
</tr>
</tbody>
</table>

Forecasting

One step ahead out of sample forecast of wholesale price of pigeon pea for the above five markets during the period 01 January, 2014 to 28 February, 2014 have been computed. The fitted model along with the actual data points are also depicted in Figure 3 to visualize the performance of fitted models.
Conclusions

Long memory time series have been analysed by using ARFIMA models which are based on linear structure. ARFIMA \((p, d, q)\) presents attenuation of hyperbolic rate, and at the same time we can describe and analyze the short-term memory and long-term memory. Model parameter \(d\) reflects the long memory among the time series observed values, and model parameter \(p, q\) reflect short memory among the time series observed values. The long memory systems are characterized by their ability to remember events in
the long history of time series data and their ability to make decisions on the basis of such memories. The study has revealed that the ARFIMA model could be used successfully for modelling as well as forecasting of daily wholesale price of pigeon pea in different markets. The model has demonstrated a good performance in terms of explained variability and predicting power. The findings of the present study have provided direct support for the potential use of accurate forecasts in decision-making for the wholesalers, retailers, farmers as well as consumers.

Acknowledgements

The author is highly grateful to the anonymous referee for the valuable suggestions on improving the quality of this manuscript.

References


Revised received: February, 2014; Accepted May, 2014