Indian Journal of Agricultural Sciences **88** (6): 916–23, June 2018/Article

Disparity in the wages of agricultural labourers in India: An interval-valued data analysis

B S YASHAVANTH[1] and ARNAB KUMAR LAHA[2]

[1]*National Academy of Agricultural Research Management, Hyderabad 500 030*

Received:19 August 2017; Accepted: 27 March 2018

ABSTRACT

This study explores the interval-valued data analysis techniques to witness the spatial disparity in the wage rates of farm labourers in India. Farm labourers constitute more than half of the total workforce engaged in Indian agriculture. Also, farmers' expenses towards labour charges account for more than 50 per cent of the total variable cost of production for most crops.Using the time series data on the nominal farm wage rates paid at different agriculturally important states, the interval-valued series are built. The inflation-adjusted real wage rates are found and both nominal and real wage rate data are used to find the average range of the farm wage rates over the agricultural years for a decade. Using the time series analysis techniques, viz. autoregressive integrated moving average–artificial neural network (ARIMA-ANN) hybrid model and vector autoregressive moving average (VARMA) model, the interval-valued data on nominal wage rates are modelled and the best model for forecasting is identified using forecast evaluation methods. The results established the presence of spatial disparity and the forecasts indicated that this disparity is not going to narrow down in future unless some policy intervention takes place.

Key words: Interval-valued data, Spatial disparity, Time series, Wage rate

Agriculture plays a vital role in the context of Indian economy. Eventhough the share of agriculture in India's GDP has been declining, yet agriculture and its allied sectors contributes nearly 14% to India's GDP. In addition to share in GDP, agriculture is the key source of livelihood for just about half of the Indian population. Despite the share of agriculture and allied sectors in India's GDP declined from 51.9 per cent in 1951 to 14.4 per cent during 2011-12, the share in workforce remained high at 54.6 per cent declining by merely 15 percentage points during the same period.

The agricultural workforce can be divided in to two categories, viz. cultivators and agricultural labourers. A person is classified as cultivator if he or she is engaged in cultivation of land owned or held from government or held from private persons or institutions for payment in money, kind or share. A person who works on another person's land for wages in money or kind or share is regarded as an agricultural labourer. A comparison across two time periods, 2004-05 and 2011-12, indicates that while there was an increase in the size of the total workforce in the country, the size of the agricultural workforce reduced by 30.57 million people (FICCI report 2015). But, even if the agricultural workforce has reduced, it is true only with respect to the cultivators. Cultivators constituted 71.9 per cent of the total agricultural workforce in 1951 which has declined to 45.1 per cent in 2011. During the same period, the proportion of agricultural labourers has grown from 28.1 per cent to 54.9 per cent (www.iasri.res.in).

Being an important component of Indian agriculture, the labour costs have risen at a faster pace compared to other inputs which are part of the overall operational cost of cultivation. Cost of cultivation data shows that labour accounts for more than 50 per cent of the total variable cost of production for most crops. Reduction in supply of labourers along with the support by a number of government schemes, including the Mahatma Gandhi National Rural Employment Guarantee Act (MGNREGA) has led to an escalation in farm wages which is adversely impacting the profitability of the farmer. As a matter of fact, the average daily wage rate of a male worker has risen from ₹ 78 per day in 2007-08 to ₹ 232.2 per day in 2013-14; whereas for a female worker it rose from ₹ 57.9 per day to ₹ 182.5 per day during the same period. Rural wages have been growing by 17% annually on average since 2006-07 outstripping urban wages. This may be attributed to the implementation of MGNREGA due to which a temporary shortage of labour has been reported during the peak agricultural season. The Commission for Agricultural Costs and Prices (CACP) in its report on price policy for *kharif* crops of 2013-14 mentioned that labour

[1]Scientist (e mail: yashavanthbs@gmail.com), ICAR-National Academy of Agricultural Research Management, Hyderabad 500 030. [2]Professor (e mail: arnab@iima.ac.in), Indian Institute of Management, Ahmedabad 380 015

costs have risen by almost 20% per annum, during the 3 years (2011-12 to 2013-14), which have pushed the cost of production in agriculture (CACP report 2016).

Many authors have attempted to study the labour related economics in India. Kalpana (1973) found that irrigation or multiple cropping generates a positive response in the daily wage rate for agricultural labourers as well as wage income. Barbara and Nandini (2001) tried to map India's unorganized labour world which also includes agricultural labourers. Akhtar and Azeez (2012) in their study about migration have held MGNREGA responsible for the increase in agricultural labour wages. Berg *et al*. (2012) found that on an average MGNREGA boosts the real daily agricultural wage rates by 5.3%. Gulati *et al*. (2013) examined the trends in farm wages over two decades (1990s and 2000s) at all India level and across major agricultural states and tried to identify the plausible factors that may be influencing farm wages. Reddy (2013) tried to identify and analyse the factors influencing the female agricultural wages in three revenue divisions of Chittoor district in Andhra Pradesh. FICCI, in its report on labour in Indian agriculture, has identified many factors that are leading to rise in the labour wages and has suggested several policy measures to address the problems related to labour shortage. Bhattarai *et al*. (2014) have analyzed impact of MGNREGA on both agricultural and non-agricultural wage rates separately for both the genders.

Though the wage rates of agricultural labour have increased in all the states of India, there exists huge difference in wages paid in different states. While anagricultural labourer, on an average, earns ₹ 628 per day in Kerala, he earns ₹158 per day in Madhya Pradesh, as on February 2014 (Directorate of Economics and Statistics). This marks a need for studying the dispersion in the daily wage rates of agricultural labour which will help the policy makers to bring out the necessary policy changes to narrow down the disparity benefitting both farmers as well as labourers. In addition, forecasting the daily wage rates that are to be paid in the future will help the farmers to plan their activities accordingly to maximize their profit. In this study, we analyze the historical data on daily wage rates of agricultural labour and forecast the same using interval-valued time series modeling.

MATERIALS AND METHODS.

Most of the modelling and forecasting methods are based on classical single-valued time series data. Alternatively, we have used interval-valued data, i.e. chronologically collected intervals of minimum and maximum values of a variable and are generally considered in the field of symbolic data analysis (Bock and Diday 2000). Interval valued data have the advantage of taking into account variability and/or uncertainty present in the values reducing the amount of uncertainty relative to that found in single-valued data (Neto and Carvalho 2010, Xiong *et al*. 2014). Interval-valued time series (ITS) modelling and forecasting have been applied successfully in various fields like finance (Arroyo *et al*. 2011), energy markets (García-Ascanio and Maté 2010) and agriculture marketing (Lin and Gonzalez-Rivera 2016). The present study is directed at analyzing the spatio-temporal variations in the agricultural labour wages using interval-valued data analysis techniques. An attempt is also made to forecast the labour wages.

An interval valued variable is one which can be recorded in intervals of two real values, one minimum and the other maximum value it can take. When interval-valued variables are collected in an ordered sequence over time it is said to be an interval-valued time series. Consider that at each point in time (t=1,2,…,n), an interval is described as a two dimensional vector $[X_t^L, X_t^U]$ of real valued components representing the upper boundary $X_t^U$and lower boundary X_t^L such that $X_t^L \leq X_t^U$. Thus an interval valued time series is

$$[X_1^L, X_1^U], [X_2^L, X_2^U], \ldots, [X_n^L, X_n^U] \quad (1)$$

where n denotes the number of intervals of the time series. Specifically, an observed interval at time t is $\mathbf{I}_t$, represented as

$$\mathbf{I}_t = \begin{bmatrix} X_t^U \\ X_t^L \end{bmatrix} \quad (2)$$

Accordingly, an interval-valued time series is a set of intervals $\mathbf{I}_t$, each recorded at a specific time t and generally equidistant in time.

The interval values time series in (1) can be represented using two another time series: the interval mid-point series X_t^C and the half range interval series X_t^r. These are given by

$$X_t^C = \frac{X_t^U - X_t^L}{2} \text{ and } X_t^r = \frac{X_t^U - X_t^L}{2} \text{ where t=1,…,n.}$$

As mentioned earlier, the wage rates are not uniform all over the country. Hence, it is not advisable to use only one value for the entire nation as it does not take into account the variability present between the states. Let $[X_t^L, X_t^U]$, t=1,…,n be the interval time series of the daily wage rate recorded at equal time interval. Assuming that the daily wage rates of all the states at time $t$follow uniform distribution in the interval $[X_t^L, X_t^U]$, the mean and variance can be computed as below:

$$\text{Mean, } \phi_{1t}=(X_t^L + X_t^U)/2$$

$$\text{Variance, } \phi_{2t}=(X_t^U - X_t^L)^2/12$$

The random variables $\omega=(\phi_{1t}, \phi_{2t})$ are called as the internal parameters of the t^{th} interval. Assuming that the internal parameters are normally distributed, the Maximum Likelihood Estimators of the parameters can be found and they can be used to estimate the average range of daily wage rates of agricultural labour in India. Laha (2016) used this technique to find average range of temperature during the months of October-December in India.

Several ITS modelling and forecasting techniques are proposed by different authors. Some of them are popular in the classic single-valued time series modelling techniques like exponential smoothing methods (Arroyo *et al*. 2007), vector error correction model (Cheung 2007, Franco *et al*. 2006) and interval multi layer perceptron (iMLP) model

(Roque *et al*. 2007). Han *et al*. (2008) proposed an interval linear model for ITS modelling. Maia *et al*.(2008) empirically found that ARIMA-Neural Network Hybrid model performs better than the individual ARIMA and Neural Network models applied independently. Garcia-Ascanio and Mate (2010) found that vector autoregression (VAR) performs better than iMLP in forecasting electric power demand. Lin and Gonzalez-Rivera (2016) compared six different ITS models and found through empirical studies that VAR outperforms. Based on this preliminary information, we selected VARMA model, which is an extension of VAR models making inclusion of moving average terms and ARIMA-ANN hybrid models for our analysis.

An often-used methodology in handling and predicting time series is known as the Box–Jenkins method or simply ARIMA. Though ARIMA models are widely used, they are best suited only for linear models as they fail to capture the non-linear structures. A wide variety of ANN models are used in such cases where non-linear structures are to be captured. The popular Multi Layer Perceptron (MLP) networks with two layers, one hidden and one output layer connected acyclically are very often used for non-linear time series modelling. In the MLP networks, the relation between the output x_t and inputs $x_{t-1}, x_{t-2}, \ldots, x_{t-p}$ is as below:

$$X_t = \alpha_0 + \sum_{j=1}^{q} \alpha_j . g(\beta_{0j} + \sum_{i=1}^{p} \beta_{ij} x_{t-i}) + \varepsilon_t \tag{3}$$

where α's and β's are the model parameters, p is the number of input nodes, q is the number of hidden nodes and g is the transfer function. In case of ANN autoregressive models (ANNAR), the lagged variables x_{t-i} (i=1,2,…,p) are the inputs. Such an ANNAR model is represented as NNAR (p,q). Among several transfer functions, the logistic function given is most often used.

According to Zhang (2003), a time series is composed of a linear autocorrelation structure and nonlinear component as

$$x_t = L_t + N_t \tag{4}$$

where L_t and N_t denote the linear and non-linear components, respectively. Hence, a hybrid model which can capture both linear and non-linear components may perform better than the individual models. In our case of ITS, the ARIMA-ANN hybrid methodology is applied in two steps. In the first step, the ARIMA models are applied to the interval mid-point series X_t^C and the half range interval series X_t^r. The values predicted by these models for the lower and upper bounds of the interval, $\hat{L}_t^U$ and $\hat{L}_t^L$, respectively, are given by:

$$\hat{L}_t^U = \hat{X}_t^C + \hat{X}_t^r \text{ and } \hat{L}_t^L = \hat{X}_t^C + \hat{X}_t^r$$

In the next step, the residual series, e_t^c and e_t^r, obtained from the ARIMA models which are given by:

$$e_t^r = X_t^r - \hat{X}_t^r \text{ and } e_t^r = X_t^r - \hat{X}_t^r$$

are modeled using the ANN to obtain the predicted non-linear component. The non-linear component, $\hat{N}_t^U$ and, $\hat{N}_t^l$ are given by:

$$\hat{N}_t^U = \hat{e}_t^c - \hat{e}_t^r \text{ and } \hat{N}_t^L = \hat{e}_t^c - \hat{e}_t^r$$

Finally, the predicted values from the hybrid ARIMA-ANN model areobtained as

$$\hat{I}_t = \begin{bmatrix} \hat{X}_t^U \\ \hat{X}_t^L \end{bmatrix} = \begin{bmatrix} \hat{L}_t^U + \hat{N}_t^U \\ \hat{L}_t^L + \hat{N}_t^L \end{bmatrix} \tag{5}$$

The multivariate VARMA models are remarkably similar to univariate ARMA models. For two time series, $\mathbf{x}_t = [x_{1t}, x_{2t}]^T$, a bi-variate VARMA ($p,q$) model can be represented as below.

$$x_t = \mu + \Phi_1 x_{t-1} + \ldots + \Phi_p x_{t-p} + \Theta_1 \eta_{t-1} + \Theta_q \eta_{t-q} + \eta \tag{6}$$

where μ is a 2×1 vector of means, Φ_i and Θ_j (i=1,…p; j=1,…q) are 2×2 matrices of parameter estimates and η_t is a vector of residual series.

For the ITS, $\mathbf{I}_t$, two kinds of VARMA models can be fit. First, VARMA-I model using the time series of upper and lower bounds of the interval, X_t^U and X_t^L respectively, given as:

$$\hat{I}_t = \begin{bmatrix} \hat{X}_t^U \\ \hat{X}_t^L \end{bmatrix} = \mu + \Phi_1 \begin{bmatrix} X_{t-1}^U \\ X_{t-1}^L \end{bmatrix} + \ldots + \Phi_P \begin{bmatrix} X_{t-p}^U \\ X_{t-p}^L \end{bmatrix} + \Theta_1 \begin{bmatrix} \eta_{t-1}^U \\ \eta_{t-1}^L \end{bmatrix} + \ldots + \Theta_1 \begin{bmatrix} \eta_{t-P}^U \\ \eta_{t-P}^L \end{bmatrix} + \eta_t \tag{7}$$

and the other using the interval mid-point series X_t^C and the half range interval series X_t^r

$$\begin{bmatrix} X_t^c \\ X_t^r \end{bmatrix} = ¼ + ¦_1 \begin{bmatrix} X_{t-1}^c \\ X_{t-1}^r \end{bmatrix} + \ldots + ¦_p \begin{bmatrix} X_{t-p}^c \\ X_{t-p}^r \end{bmatrix} + \text{˜}_1 \begin{bmatrix} \eta_{t-1}^c \\ \eta_{t-1}^{Lr} \end{bmatrix} + \ldots + \text{˜}_q \begin{bmatrix} \eta_{t-p}^c \\ \eta_{t-p}^r \end{bmatrix} + \cdot_t \tag{8}$$

from which the ITS, $\mathbf{I}_t$ can be obtained using the arithmetic operations.

The performance of a forecasting model can be evaluated using loss functions which are based on the error terms. The forecast evaluation measures available for the classic single valued time series methods can easily be extended to the ITS scenario. The commonly used Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE) calculated separately for upper and lower bounds are used in our study. They are given by:

$$RMSE^U = \sqrt{\frac{(\widehat{X}_t^U - X_t^U)^2}{n}} \text{ and and}$$

$$MAPE^U = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{X_t^U - \widehat{X}_t^U}{X_t^U}\right| \text{, for the upper bounds and,}$$

$$\text{and } MAPE^U = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{X_t^L - \widehat{X}_t^L}{X_t^L}\right| \text{ for the lower bound}$$

Among the competing ITS models, the one which has low RMSE and MAPE values for both upper and lower bounds is considered as best suited for the data under consideration. Along with this, the Interval Average Relative Variance (ARV^I) proposed by Maia and Carvalho (2011) is used. The ARV^I statistic is given as:

$$\mathrm{ARV}^{\mathrm{I}} = \frac{\sum_{j=i}^{m}(\mathbf{I}_{j+1}-\hat{\mathbf{I}}_{j+1})'(\mathbf{I}_{j+1}-\hat{\mathbf{I}}_{j+1})}{\sum_{j=i}^{m}(\mathbf{I}_{j+1}-\bar{\mathbf{I}})'(\mathbf{I}_{j+1}-\bar{\mathbf{I}})}$$

$$= \sqrt{\frac{\sum_{j=1}^{m}(X_{j+1}^{U}-X_{j+1}^{U})^2+\sum_{j=1}^{m}(X_{j+1}^{L}-X_{j+1}^{L})^2}{\sum_{j=1}^{m}(X_{j+1}^{U}-\bar{X}^{U})^2+\sum_{j=1}^{m}(X_{j+1}^{L}-\bar{X}^{L})^2}} \quad (9)$$

The ARV^I statistic is suitable for comparing the errors in the predictions provided by a reference model with those of predictions given by the average value of the series. Therefore, lower ARV^I values lead to better forecasts. Moreover, a value of ARV^I equal to 0 indicates perfect forecasting, whereas a value of ARV^I equal to 1 indicates that the model only achieves forecasts of an average value.

The data used for this study is compiled by the Labour Bureau. The data consists of nominal daily wage rates (monthly) of agricultural labour (Man) from January, 2004 to April, 2015 in 16 different states of India. The nominal farm wages are derived by averaging wages over five types of farm operations – ploughing, sowing, weeding, transplanting and harvesting – at the state level. The states for which the data are available are Andhra Pradesh (AP), Assam (AS), Bihar (BH), Gujarat (GJ), Haryana (HA), Himachal Pradesh (HP), Karnataka (KA), Kerala (KE), Madhya Pradesh (MP), Maharashtra (MH), Orissa (OR), Punjab (PJ), Rajasthan (RJ), Tamil Nadu (TN), Uttar Pradesh (UP) and West Bengal (WB). These 16 states together comprise 93% of the agricultural labour force in the country (CACP report 2013). For these states, the monthly consumer price indices (general) for agricultural labour reported by the Labour Bureau are collected and used to obtain the inflation-adjusted real wage rates. The real wage rates are calculated by using January, 2004 as the base. Data pertaining only to male labour is considered for the study as the scope of this study is to witness the regional dispersion in the agricultural labour wage rates. In this regard, the conclusions drawn on wage rate of men labourer can be straight away extended to the women labourer. The data is available at www.indiastat.com.

RESULTS AND DISCUSSIONS

Before proceeding to ITS modeling of wage rates, it is essential to study the behavior of the wage rate time series, state wise, over the years. For this purpose, the 16 states considered in the study are sub divided into 4 groups based on their geographical location into southern, eastern, western and central and northern states as in the Table 1. The table also gives the descriptive statistics about the state wise nominal and real wage rates. A perusal of Table 1 indicates that the Kerala state recorded the highest mean nominal wage rate of ₹ 325.9 per day and real wage rate of ₹ 205.3 per day, respectively, followed by Himachal Pradesh. This is in line with what Kalpana (1973) found in 1960s which she attributed to the increased bargaining power due to unionization of the labourers in Kerala. The state of Madhya Pradesh has recorded the lowest figures in both real and nominal wage rates. Looking at the standard deviations, it is apparent that nominal wage rates have been highly varying in Kerala and stable in Gujarat. The real wage

Table 1 State-wise descriptive statistics for the average daily wage rate

Region	States	Nominal daily wage rate				Real daily wage rate			
		Mean	SD	Min	Max	Mean	SD	Min	Max
Southern	AP	132.2	68.2	50.3	249.6	76.8	18.8	48.6	107.1
	KA	114.2	66.6	52.9	253.7	67.0	15.7	52.4	102.4
	KL	325.9	143.9	176.4	652.3	205.8	28.9	152.1	275.5
	TN	172.9	104.2	76.2	439.6	106.3	30.2	75.2	184.7
Eastern	AS	111.6	49.5	61.8	237.9	69.9	10.6	56.7	102.9
	BH	103.6	54.6	53.9	229.5	62.7	15.8	47.8	102.4
	OR	102.4	48.6	50.1	208.4	61.3	12.0	44.1	83.8
	WB	119.6	58.4	55.8	242.1	72.4	13.8	52.5	105.7
Western	GJ	98.0	40.5	58.4	198.0	60.4	7.5	50.7	81.4
	MH	117.4	58.9	58.0	230.9	66.6	12.6	40.4	92.6
	RJ	147.6	72.2	69.9	320.2	80.8	14.9	57.3	120.4
Central and Northern	HA	176.0	84.8	83.8	357.1	100.0	16.3	80.3	140.1
	HP	261.9	60.9	177.5	363.3	150.5	17.3	121.5	184.2
	MP	85.7	41.6	46.6	181.9	50.2	10.4	38.8	79.2
	PJ	166.1	78.9	85.0	312.0	95.3	17.4	72.8	127.4
	UP	108.4	49.0	58.5	208.7	65.8	12.0	50.8	90.6

Table 2 State-wise overall growth rate in daily wage rates over the years 2004-2015

States	Nominal wage rate (%)	Real wage rate (%)
AP	370.3	86.0
KA	370.2	83.0
KL	214.0	26.6
TN	405.3	113.1
AS	241.4	49.8
BH	323.7	89.1
OR	301.2	62.4
WB	301.2	74.7
GJ	233.1	38.0
MH	283.2	48.4
RJ	311.9	50.7
HA	306.2	54.1
HP	103.9	43.1
MP	286.6	68.5
PJ	224.3	28.3
UP	251.9	52.7

rate has been highly varying in Tamilnadu. Tamilnadu also recorded overall growth rate of 405.3% in nominal wage rates and 113.1% in real wage rates, highest among all the states, whereas Punjab reported the lowest (Table 2).Table 3 gives the average range of daily wage rates of agricultural labour in India, calculated for each agricultural year, starting in July with the onset of monsoon. The calculations were made for all the 16 states and for 15 states excluding Kerala, since the wage rates are substantially high in Kerala affecting the all India values. Subsequently, the nominal daily wage rates from Kerala were not included for deriving the interval valued time series for modelling and forecasting nominal wage rates. It is apparent from the table that the differences in wage rates have been steadily widening over the time.

Also, in a nutshell, these results indicate that there exists a disparity in the wages paid at different regions of the country. Himanshu (2006) found that this variation is limited not just across the states, but is also found across different agro-climatic regions within a state. This spatial variation in wage rates can be attributed to regional variation in agricultural productivity (Jose 1988), level of urbanization, irrigation and demand for non-agricultural labour (Acharya 1989), agricultural productivity per worker and non-agricultural employment (Srivastava and Singh 2006). Similar spatial disparity was found in England during the time of industrial revolution (Clark 2001) and early nineteenth century (Lyle 2007).

For ITS modelling, the overall data of 136 months were divided in to two parts: the first part consisting the first 124 months was used for building the ITS model (training); and the second part consisting the last 12 months was used for evaluating the models (testing) built using the training data. The ITS data used for training the models are plotted in Fig 1. From the figure, it is perceptible how the difference in nominal wage rates has increased over the years.

According to Tsay (2010), the foundation of time series analysis is stationarity. The upper bound series X_t^U, the lower bound series X_t^L, the mid-point series X_t^c and the half range interval series X_t^r are tested for stationarity. The Augmented Dickey-Fuller unit root test is performed for this purpose. The results of the test are given in Table 4. The ADF test is performed for the original series and the

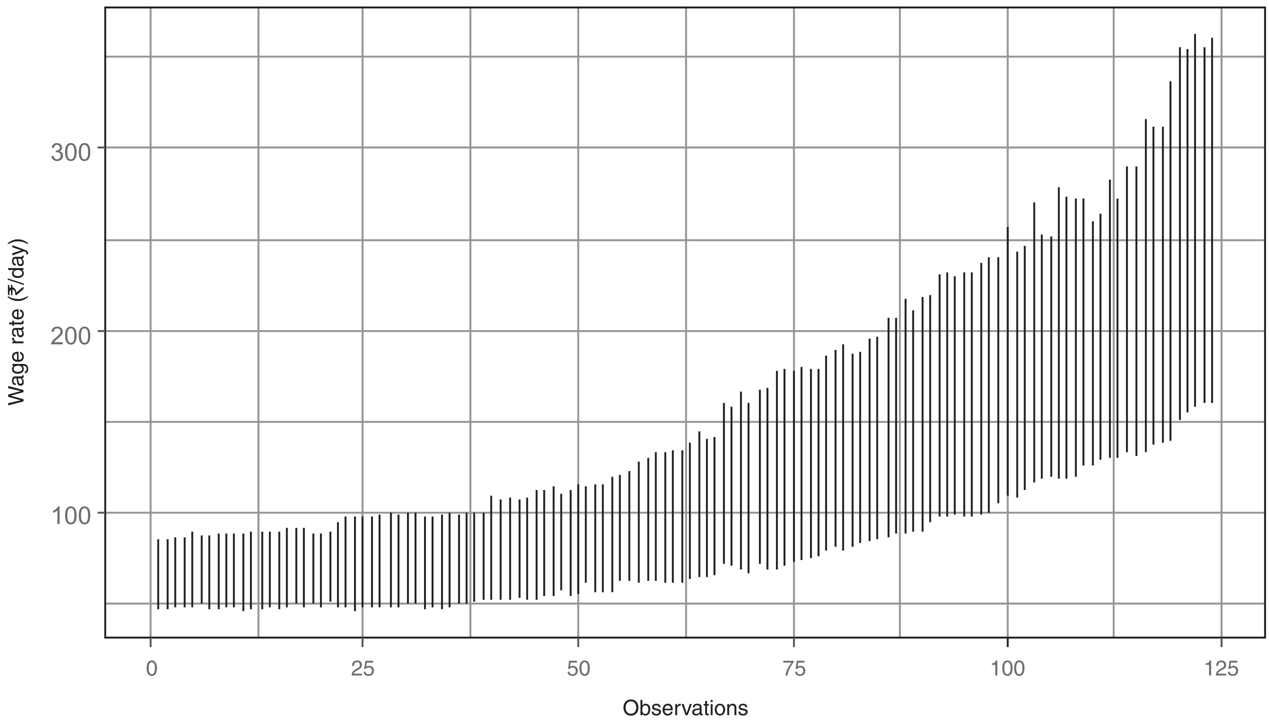

Fig 1 Interval-valued time series of daily agricultural wage rate used for model building

Table 3 Annual average range of daily agricultural wage rate

Agricultural year	Nominal daily wage rates (₹/Day)		Real daily wage rates (₹/Day)	
	All India	Excluding Kerala	All India	Excluding Kerala
2004-05	[114.7, 132.7]	[67.6, 70.2]	[110.4, 128.5]	[65.2, 66.5]
2005-06	[112.7, 133.3]	[69.1, 75.9]	[104.9, 127.1]	[64.2, 67.6]
2006-07	[118.3, 144.7]	[71.5, 80.7]	[104.6, 129.7]	[61.5, 65.3]
2007-08	[123.8, 139.8]	[79.9, 88.8]	[105.6, 112.6]	[63.3, 66.4]
2008-09	[123.8, 170.2]	[90.9, 105.4]	[93.9, 122.3]	[65.1, 69.7]
2009-10	[151.8, 197.4]	[112.7, 130.1]	[100.6, 132.6]	[65.5, 89.0]
2010-11	[192.7, 221.8]	[129.5, 155.3]	[119.5, 126.8]	[81.4, 98.3]
2011-12	[225.7, 276.7]	[156.9, 181.6]	[126.9, 151.6]	[96.8, 102.9]
2012-13	[280.9, 308.9]	[185.0, 209.5]	[143.3, 153.0]	[98.4, 109.8]
2013-14	[300.3, 407.7]	[207.9, 265.6]	[135.4, 175.4]	[102.3, 130.2]
2014-15	[371.3, 419.6]	[213.8, 276.3]	[152.3, 173.2]	[120.3, 131.5]

results indicated that the series are non-stationary signifying the necessity for differencing. The differenced series are found to be stationary at 5% level of significance. Once the stationarity is achieved, the three ITS models were fit: first, the ARIMA-ANN Hybrid using the mid-point series X_t^c and the half range interval series X_t^r; second, the VARMA-I model using the upper bound series X_t^U and the lower bound series X_t^L; and third, the VARMA-II model which uses the mid-point series X_t^c and the half range interval series X_t^r. Various candidate models of different order were fit and the best model is chosen by using Akaike Information Criteria (AIC). Accordingly, ARIMA(1,1,1)-NNAR(1,1), the VARMA (1,1) and VARMA(0,3) were selected as the suitable models for ARIMA-ANN hybrid, VARMA-I and VARMA-II models, respectively. The fitted models were evaluated for their ability to forecast using the forecast evaluation measures, the results of which are given in Table 5, for training data set and testing data set.

The results from Table 5 indicate that VARMA-II model which is fit using the mid-point series X_t^c and the half range interval series X_t^r has performed better than the other two models. The model is as below:

$$\begin{bmatrix} X_t^c \\ X_t^r \end{bmatrix} = \begin{bmatrix} 1.72 \\ 0.74 \end{bmatrix} + \begin{bmatrix} 0.11 & 0.24 \\ 0.01 & 0.32 \end{bmatrix}\begin{bmatrix} e_{t-1}^c \\ e_{t-1}^r \end{bmatrix} + \begin{bmatrix} -0.31 & 0.42 \\ -0.29 & 0.44 \end{bmatrix}$$
$$\left[\begin{bmatrix} e_{t-2}^c \\ e_{t-2}^r \end{bmatrix}\right] + \begin{bmatrix} -0.08 & -0.01 \\ -0.08 & 0.13 \end{bmatrix}\begin{bmatrix} e_{t-3}^c \\ e_{t-3}^r \end{bmatrix}$$

The fitted values along with actual values, for both lower and upper bounds are plotted in Fig 2. Subsequently, the VARMA-II model is used for forecasting the nominal daily wage rate of agricultural labour for the agricultural year 2016-17. The forecasted values for the upper and lower bounds of nominal daily wage rates along with the differences between them are reported in Table 6. It is pertinent from the results that the spatial disparity is going to widen in the forth coming years, unless some major policy interventions take place. In such situations where spatial disparity is found in wage rates, migration of labourers

Table 4 Results of Augmented Dickey-Fuller unit root test for stationarity

Series	Original Series		Differenced Series	
	ADF Statistic	p-value	ADF Statistic	p-value
Lower Bound	1.425	>0.99	-4.751	<0.01
Upper Bound	1.037	>0.99	-4.185	<0.01
Mid-point	1.642	>0.99	-3.627	0.034
Half range	-0.452	0.983	-4.962	<0.01

from low wage areas to the high wage areas would help in narrowing down the difference (Kaur *et al*. 2011). But, may be the accessibility to regular income by working in MGNREGA activities has lead to decreased migration of agricultural labourer. The resulting increase in expenses towards labour wages has forced the farmers to demand higher minimum support prices.

Agricultural labour wages form an important component of the farmers' expenses in India. Over the years, the wages of farm labourer has gone up. Many authors have held the MGNREGA responsible for this increase. Though there is an increase in farm wages all over the country, there exists

Table 5 Comparison of models using forecast evaluation measures

Data	Model	Forecast accuracy measure				
		RMSE		MAPE (%)		ARV[I] (×10-3)
		Lower bound	Upper bound	Lower bound	Upper bound	
Training	ARIMA-ANN	1.94	6.15	1.85	2.25	5.94
	VARMA-I	2.01	6.57	2.00	2.49	6.88
	VARMA-II	1.87	6.37	1.92	2.27	6.29
Testing	ARIMA-ANN	13.53	19.25	6.28	4.19	8.78
	VARMA-I	8.44	35.86	4.39	6.97	8.62
	VARMA-II	7.73	20.86	3.75	4.31	8.19

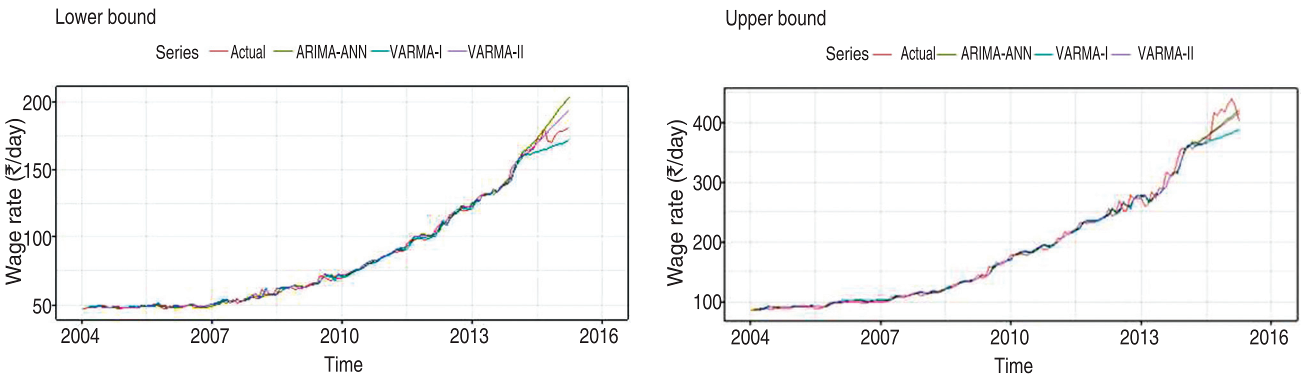

Fig 2 Plot showing actual and fitted values for the upper and lower bounds of the interval time series

Table 6 Forecasted interval bounds of average daily wage rate for the agricultural year 2016-17

Months	Forecasted values of daily wage rate (₹/day)		
	Lower bound	Upper bound	Range
Jul-16	194.44	453.82	259.38
Aug-16	195.42	456.27	260.85
Sep-16	196.40	458.73	262.32
Oct-16	197.38	461.18	263.79
Nov-16	198.37	463.63	265.26
Dec-16	199.35	466.08	266.73
Jan-17	200.33	468.53	268.21
Feb-17	201.31	470.99	269.68
Mar-17	202.29	473.44	271.15
Apr-17	203.27	475.89	272.62
May-17	204.25	478.34	274.09
Jun-17	205.24	480.79	275.56

a disparity between the states. Using interval-valued data analysis techniques, we studied this spatial dispersion in the wage rates of agricultural labour in India. The wage rates are found to be phenomenally high in Kerala and very low in Madhya Pradesh. Tamilnadu has reported the maximum growth rate in farm wages during the study period. The regional variations in agricultural productivity, irrigation, urbanization and demand for non-agricultural labour are some of the factors bringing the spatial disparity. Using interval-valued time series analysis techniques, we modelled the nominal wage rates to forecast the future values. We tried the ARIMA-ANN hybrid and VARMA models for the modelling. The VARMA model considering the mid-value and half-range series is found to be the best model based on forecast evaluation measures. Subsequently, it is used for forecasting the future values.

REFERENCES

Acharya S. 1989. Agricultural wages in India: A disaggregated analysis.*Indian Journal of Agricultural Economics* **44**(2): 121–39.

Akhtar S M J and Azeez N P A. 2012. Rural employment guarantee programme and migration. *Kurukshetra* **60**(4): 11–5.

Arroyo J, Muñoz San Roque A, Maté C and Sarabia A. 2007. Exponential smoothing methods for interval time series. (*In*) *Proceedings of the 1st European Symposium on Time Series Prediction,* pp 231–40.

Arroyo J, Espínola R Maté C. 2011. Different approaches to forecast interval time series: a comparison in finance. *Computational Econ*omics **37**(2): 169–91.

Barbara H W and Nandini G. 2001. Mapping India's world of unorganised labour. *Socialist Register* 89–118.

Berg E, Bhattacharyya S, Rajasekhar D and Manjula R. 2012. Can rural public works affect agricultural wages? Evidence from India. CSAE working paper.

Bhattarai M, Kumar R, Sandhya S and Bantilan C. 2014. Whether MGNREGS has affected agricultural wage rate in Andhra Pradesh? A panel modeling across 23 districts from 2000 to 2011. Socioeconomics discussion paper. ICRISAT, Patancheru.

Bock H H and Diday E. 2000. Analysis of Symbolic Data: Exploratory Methods for Extracting Statistical Information from Complex Data. Springer Science & Business Media, Berlin, Germany.

CACP Discussion Paper 5. 2013. Rising farm wages in India – The 'Pull' and 'Push' factors. Commission for Agricultural Costs and Prices.

Cheung Y W. 2007. An empirical model of daily highs and lows. *International Journal of Finance & Economics,* **12**(1): 1–20.

Clark G. 2001. Farm wages and living standards in the Industrial Revolution: England, 1670–1869. *Economic History Review* **3**: 477–505.

FICCI Report. 2015. Labour in Indian Agriculture: A growing challenge.

Franco M, Blanco D, Blequett W, Guglia M and Alvarado E. 2006. Cointegration methodology and error correction model used to forecast the electricity demand of the Venezuelan electric system - period 2004–2024. *IEEE Transactions on Power Systems*: 1–8.

García-Ascanio C, Maté C. 2010. Electric power demand forecasting using interval time series: a comparison between VAR and iMLP. *Energy Policy* **382**: 715–25.

Gulati A, Jain S and Satija N. 2014. Rising farm wages in India – The 'Pull' and 'Push' factors. *Journal of Land and Rural Studies* **2**(2): 261–86.

Han A, Hong Y, Lai K and Wang S. 2008. Interval time series analysis with an application to the sterling-dollar exchange rate. *Journal of Systems Science and Complexity* **21**(4): 550–65.

Himanshu. 2006. Agrarian crisis and wage labour: A regional perspective. *Indian Journal of Labour Economics* **49**(4): 835–54.

Jose A V. 1988. Agricultural Wages in India. *Economic and Political Weekly* **23**(26): 46–58.

Kalpana B. 1973. Factors affecting wage rates for agricultural labour. *Economic and Political Weekly* **8**(26): 56–64.

Kaur B, Singh J M, Garg B R, Singh J and Singh S. 2011. Causes and impact of labour migration: A case study of Punjab agriculture. *Agricultural Economics Research Review*. **24**: 459–66.

Laha A K. 2016. Statistical challenges with big data in management sciences. *(In)* pp 41-55. Pyne S, Rao B L S P and Rao S B (Eds). *Big Data Analytics*, Springer.

Lin W and Gonzalez-Rivera G. 2016. Interval-valued time series models: Estimation based on order statistics exploring the Agriculture Marketing Service data. *Computational Statistics and Data Analysis* **100**: 694–711.

Lyle M. 2007. Regional agricultural wage variationsin early nineteenth-century England. *Agricultural History Review* **55**(1): 95–106.

Maia A L S and Carvalho F A T. 2011. Holt's exponential smoothing and neural network models for forecasting interval-valued time series. *International Journal of Forecasting* **27**(3): 740–59.

Maia A L S, Carvalho F A T and Ludermir T B. 2008. Forecasting models for interval-valued time series. *Neurocomputing* **71**: 3344–52.

Neto L E A and Carvalho F A T. 2010. Constrained linear regression models for symbolic interval-valued variables. *Comput. Stat. Data Anal*. **54**(2): 333–47.

Reddy E L. 2013. Factors determining female agricultural wages: A case study in Chittoor district (Andhra Pradesh). *International Journal of Humanities and Social Science Invention* **2**(8): 30–9.

Roque A M, Mate C, Arroyo J and Sarabia A. 2007. iMLP: applying multi-layer perceptrons to interval-valued data. *Neural Processing Letters* **25**: 157–69.

Srivastava R and Singh R. 2006. Rural Wages during the 1990s: A Re-estimation.*Economic and PoliticalWeekly*, **41**(38): 4053–62.

State of Indian Agriculture.2016. Ministry of Agriculture and Farmer's Welfare, Government of India.

Tsay R S. 2010. *Analysis of Financial Time series*, Third edition. John Wiley and Sons, New Jersey.

Xiong T, Bao Y K and Hu Z Y. 2014. Multiple-output support vector regression with a firefly algorithm for interval-valued stock price index forecasting. *Knowledge Based Systems* **55**: 87–100.

Xiong T, Li C and Bao Y. 2017. Interval-valued time series forecasting using a novel hybrid Holt^{I} and MSVR model. *Economic Modelling* **60**: 11–23.

Zhang G. 2003. Time series forecasting using a hybrid ARIMA and neural networkmodel. *Neurocomputing,* **50**: 159–75.