International Journal of Science Engineering and Advance Technology (IJSEAT)
ISSN 2321-6905, Vol. 9, Issue 4, April -2021

Forecasting of Potato Prices in Telangana using FFNN and Hybrid in Time Series Models

M. Udayshankar [1] , M. Raghavender Sharma [2]
[1]Research Scholar, [2]Assistant Professor,
Department of statistics, Osmania University, Hyderabad, Telangana.
Email: udayshankar.malyala@gmail.com, drmrsstatou@gmail.com

Abstract:

This paper discussed for develop a forecasting models and predict the future prices of potatoes in Telangana. A historical potato data received from Agmart.com. The forecasted models have been determined from auto regressive integrated moving average (ARIMA) and Feed forward neural network (FFNN) and HYBRID for the monthly average prices of potato in Telangana. The performances of the forecasted models are estimated using the error measures like MAE (mean absolute error) and RMSE (root mean square error) and MAPE (Mean absolute percentage error). The result shows that, the HYBRID model performing better than the ARIMA and FFNN.

KEY WORDS: ARIMA, FFNN, HYBRID, MAPE, RMSE and MAE

1 INTRODUCTION

Potato is the major food item in the family people and small and large scale of business. One potato will have Vitamin c, B16, potassium, zinc. These vitamins are keep skin as smooth. Also, potato is a source of fiber and it helps us to control the several diseases like cholesterol and blood sugar etc. There are multiple value added products from potato like potato chips, starch, potato power etc. The world third largest potato producing is India. In Telangana, potato cultivate in Medak (4530ha and 90760 MT), Rangareddy (394 ha and 7880 MT) and Khammam (1970 ha and 39,400 MT) districts. The objective of this paper is to identify the better the idea's and provide the good methodology for estimate the future price of potato for planning farmers, business segments and government sectors.

2. Material and Methodology:

A historical data of monthly average prices of potato during January 2010 – December 2019 (containing 132 Observations). Monthly average prices of potatoes received from Agmart.com. The idea is to select the best model by building the ARIMA, FFNN and HYBRID models. In R software is used for preparing the model building and MS-Excel is used for preparation of charts and tables.

2.1. Box Jenkins methodology

In ARIMA model have multiple features like to determine the minimum number of parameters and. The Box Jenkins methodology has four steps. Which are Identification, Estimation, Diagnostic checking, Forecasting (Bogahawatte and Sivapathasundaram, 2012). Identify is the order of p, q, d are determined by checking the actions of Autocorrelation and partial autocorrelation function and ACF and PACF are useful to obtain the stationary and non stationary. Hence, the model is selected then the parameters should be estimated. The Diagnostic test is to check the selected model and predict the future prices. In the model section, multiple models were tested and finalized from the minimum error measures. Several models have been tested for the training data and identified one model based on significance of the parameters and model adequacy.

$$\phi(B)\nabla^{\mathrm{d}} Z_t = \theta(\mathrm{B})\, a_t$$

Where ∇^{d} is $(1-B)^d$ and
$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p$
$\theta(B) = (1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q)$
d , D is to to find stationary in non-seasonal component and seasonal component and
a_t is a white noise process with zero mean and variance

2.2. Neural network

Artificial Neural Network is a mathematical model, influenced functional point of biological neural network. ANN is a dynamic forecasting model and it's consists of a correlated group of artificial neurons and it act to information utilize an approximate process to determination. Hence, develop a feed forward neural network (FFNN) model to predict the future forecasting monthly average prices of potato in Telangana. In this neural network (FFNN) contains three layers networks and it includes of an input layer, hidden layer and output layer. In this paper, total number of inputs is three as lag 1, lag2, four hidden layers and one output will get in the final stage. The following table provides the information about the network model, involves the covariates variable, number of units rescaling methods, hidden layer consists of number of hidden layers and number of units in hidden Layer 4 and activation function as

Hyperbolic Tangent. Output layer consists of dependent variables and number of units and rescaling method for scale dependent and activation function is identity and Error functions using sum of squared.(Rama Krishna et al., 2013 and Haykin, 1999,), [5][20].

Table 2.1: Information about the Network

Input Layer	Covariates 1 Number of Units[a] Rescaling method of covariates	Lag1, Lag2 1Normalized
Hidden Layer	Number of hidden layers Number of units in the hidden layer 1[a] Activation function	1 4 Hyperbolic Tangent
Output Layer	Dependent variable 1 Number of units Rescaling method of scale dependent Activation function Error function	Potato prices 1 Normalized Identity Sum of squares

2.3. HYBRID Model:

HYBRID model is combination of linear and non linear models. Which is autoregressive integrated moving average method and feed forward neural network. Let the Hybrid model is

$$Z_t = w\, L_t + (1\text{-}w) N_t$$

Where Z_t is the observation at time period t and L_t represent linear model

N_t represents Non-linear models and w represents weights.

For this predicted values are determined from Seasonal ARIMA and FFNN. To find the weights (w), a simulation process was used by taking the weights range between (0, 1) with an increment of 0.01.

$$\begin{pmatrix} \boldsymbol{HYBRID} \\ \boldsymbol{Predicted} \\ \boldsymbol{Values} \end{pmatrix} = \boldsymbol{W} \begin{pmatrix} \boldsymbol{ARIMA} \\ \boldsymbol{Predicted} \\ \boldsymbol{Values} \end{pmatrix} + (\boldsymbol{1} - \boldsymbol{W}) \begin{pmatrix} \boldsymbol{FFNN} \\ \boldsymbol{Predicted} \\ \boldsymbol{Values} \end{pmatrix}$$

The error measures are computed for the newly created HYBRID models. Then choose the best HYBRID model based on the minimum error measures like RMSE, MAE and MAPE.

3. Data source:

The historical monthly average prices of potato were collected from January 2010 to December 2019 Agmart.com. The data set is separated in to training data (January 2010 to December 2017) and testing data for validating the developed model (January 2018 to December 2019). The both linear ARIMA and non linear ANN models were developed to the training data set and then validate the testing data. The data series plot shown in the following figure 3.1

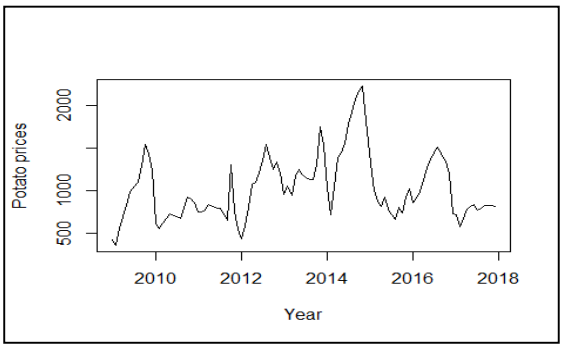


Figure 3.1: Monthly average price of potato (100 kg)

Figure 3.1 shows that, the monthly average prices of potato data are non stationary and need to be converted in non- stationary to stationary for differencing of the data. The following figure 3.2, ACF and PACF plots are presented for price of potato.

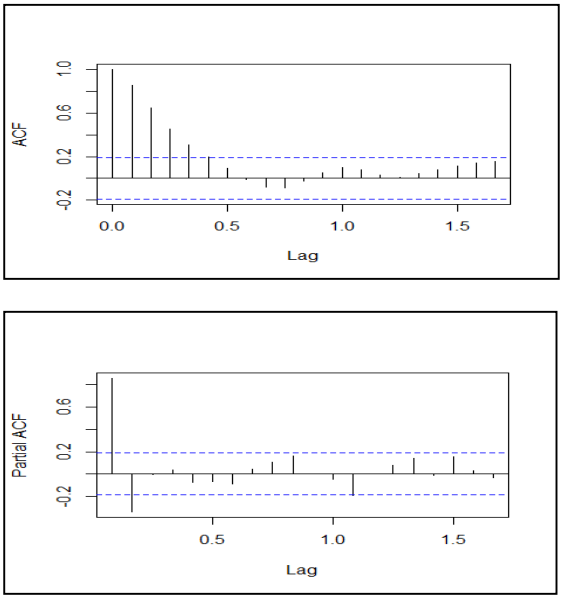


Figure 3.2: Autocorrelation and partial correlation function

In the above figures 3.2, it was observed that, the ACF is not dying at higher. This variation is shows in the plot and concluded that the time plot is non stationary. Non stationary will be converted

through applying the transformation. The below figure 3.3, is a time plot of transformed series.

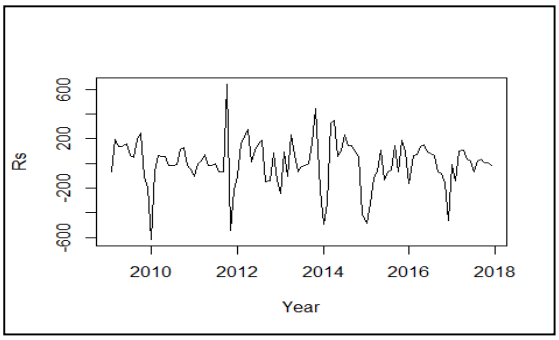
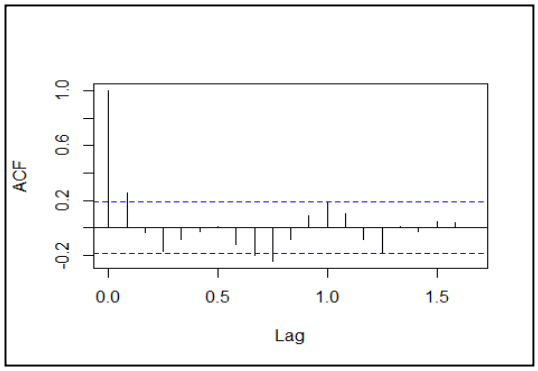


Figure 3.3: Time plot of transformed series (d=1)



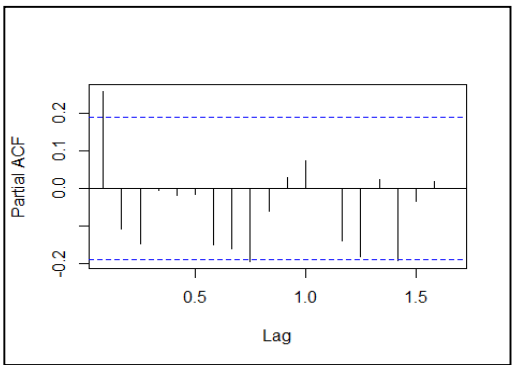


Figure 3.4: ACF, PACF plots for Transformed series prices of potato

4. Results and Discussion

4.1 ARIMA Model:

The monthly average prices of potato data was developed using Box – Jenkins methodology on the training sample. The predicted values were tested with testing sample. The training data is not a stationary from figure 3.1 and making one time difference (d=1) for non stationary data in to stationary series. Therefore the parameters are obtained by using autocorrelation and partial autocorrelation function plots. These plots show the order of non seasonal and seasonal parameters. The order of p is 2 and q value is 1 in non seasonal components and seasonal P value is 1and Q values is 1 and several orders were tested and selected the model based on the parameter significance and adequacy of the model. The parameters of the model presented in the following table.

Table 4.1 Model parameters

Parameters	Estimate	Std. Error	z value	Pr(>|z|)
ar1	-0.72	0.18	-4.01	<0.001
ar2	0.17	0.10	1.69	0.091
ma1	0.89	0.15	5.82	<0.001
sar1	0.99	0.05	19.80	<0.001
sma1	-0.94	0.18	-5.12	<0.001

The above table 3.1, substitute in the Seasonal ARIMA (2, 1, 1)(1, 0, 1), then the final model for the determining the monthly average prices of potato is

$$(1 - \emptyset_1 B - \emptyset_2 \mathrm{B}^2)(1 - B)\,(1 - B)^{12}(1 - \varphi_1 \mathrm{B}^{12})Z_t = (1 - \theta_1 \mathrm{B}\,)(1 - \Phi 1\mathrm{B})\epsilon_t$$

$$(1 + 0.72B - 0.17\mathrm{B}^2)(1 - B)(1 - B)^{12}(1 - 0.99B)Z_t = (1 - 0.89\mathrm{B}\,)(1 + 0.94\mathrm{B})\epsilon_t$$

The residulas plots are verified with help of auto correlation and partial auto correlation functions of the different appearance.

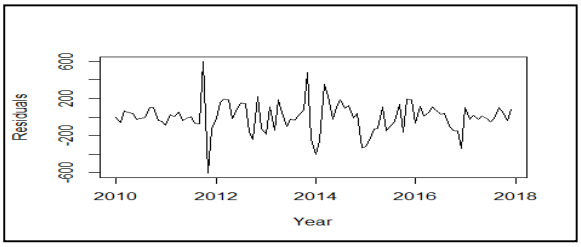
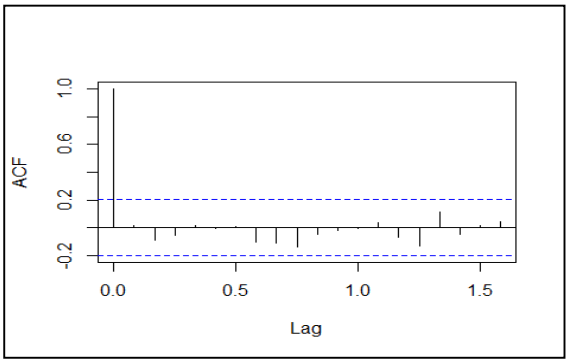


Figure 4.1: residuals series for Potato prices



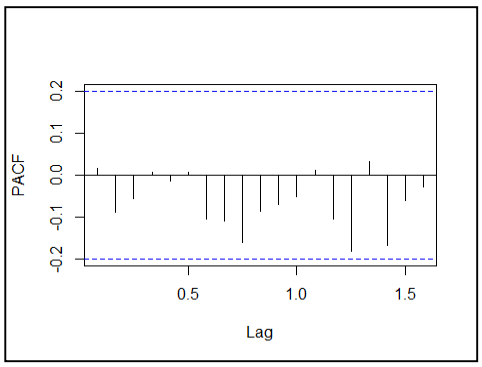


Figure 4.2: ACF and PACF residuals series for Potato prices

In the above figures, it was observed that the residuals ACF are significant difference is there from 0 to 0.2. The model adequacy is tested using the LJung box Q statistics. This Q statistics uses to verify the residuals of the data series after fitting the model. The below table consists of final model significance with Q statistic value. The hypothesis of the model is

H_0 : The fitted model is adequate

H_1: The fitted model is not adequate.

Table 4.2: The characteristics of L Jung-Box Q test statistics

Model	**Seasonal ARIMA (2, 1, 1)(1, 0, 1)**
Statistics	10.91
Degrees of freedom	14
Significance	0.69

From the above table 3.2.3, the hypothesis value is more than the 0.05 then accept H_0 and concluded that the model is adequate. Therefore the model Seasonal ARIMA (2, 1, 1)(1, 0, 1) is used for forecasting the future prices of potato. The predicted values are presented in the following table from the period January 2018 to December 2019.

Table 4.3: Forecasts of monthly average potato prices using Seasonal ARIMA (2, 1, 1)(1, 0, 1)

Month-Year	**Actual potato prices**	**Predicted potato prices**
Jan-18	821.88	724.54
Feb-18	917.93	687.78
Mar-18	1038.46	730.79
Apr-18	1430.1	814.96
May-18	1568.5	855.35
Jun-18	1563.64	857.29
Jul-18	1511.54	883.54
Aug-18	1500	899.75
Sep-18	1571.13	919.26
Oct-18	1592.94	983.20
Nov-18	1270.56	988.88
Dec-18	1107.32	875.48
Jan-19	972.89	732.39
Feb-19	933.75	694.61
Mar-19	981.93	735.83
Apr-19	1076.09	820.15
May-19	1135.33	859.38
Jun-19	1143.68	862.06
Jul-19	1109.21	887.41
Aug-19	1032.08	904.12
Sep-19	1050	922.91
Oct-19	1306.67	986.86
Nov-19	1363.39	992.01
Dec-19	1100.45	879.95

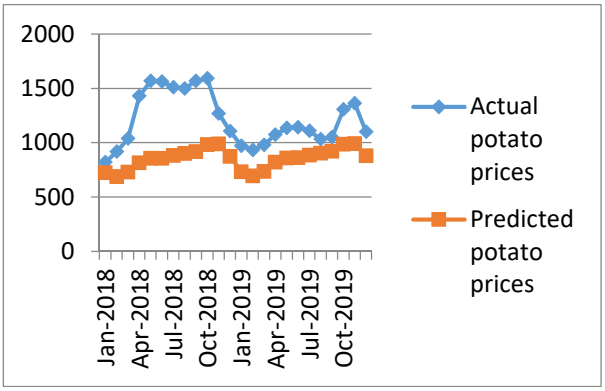


Figure 4.3: Forecasts of monthly average Potato prices

From the above figure 4.3, it was observed that the pattern for SARIMA for actual and predicted monthly average prices of potatoes are near and the error measures are listed in the below table 4.4

Table 4.4: Model performance of SEASONAL ARIMA (2, 1, 1)(1, 0, 1)

Data sets	**MAE**	**RMSE**	**MAPE**
training data	118.122	166.65	11.7799
testing data	274.033	369.995	21.8998

The Seasonal ARIMA (2, 1, 1)(1, 0, 1) model have the minimum error measures in test data as well as train data sets. From the above table, several tentative models were tested with significance and model adequacy and selected the suitable SARIMA

model for the monthly average prices of potato in Telangana. All the time the ARIMA model is not

secure for true forecast.

4.2 Feed Forward neural network model:

In Feed forward neural network is building by considering the three input layers as Lag1, Lag2 price values of monthly average price of potato and to predict the future value for forecast using FFNN model. Then the FFNN model is four layer models having 4 hidden layers and including the three input and one output layers. The hyperbolic tangent method is used for the activation function under the back propagation algorithm. The several iterative models were tested to find the optimal model based on the error measures. The FFNN model shows in the following figure 4.2.

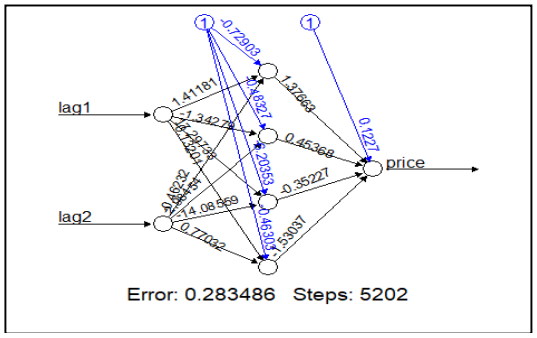


Figure 4.2: Feed forward neural network diagram.

The listed tentative model are in the below table 4.5.

Table 4.5: Possible models of FFNN model

Number of neurons in the layer			Error measures			
Input	Hidden	Output	MAPE		RMSE	MAE
2	2	1	7.29	11.79	130.06	88.95
2	3	1	7.35	11.74	131.18	89.30
2	4	1	7.26	10.23	130.67	88.35
2	5	1	7.23	11.75	129.26	88.69

From the above table 4.5, it was observed that the information on complete the number of appropriate neurons is 2 in hidden layers and the optimal network model is 2-4-1 and this model has the minimum errors MAPE, MAE and RMSE. The best model 2-4-1and parameters are shown in the following table 4.6.

Table 4.6. Model Parameters of FFNN 2-4-1

Parameter estimates		Parameter estimates	
error	2.8349 E-01	lag1.to.1layhid3	1.7297 E+01
reached.threshold	9.6051 E-03	lag2.to.1layhid3	-1.4086 E+01
steps	5.2020 E+03	Intercept.to.1layhid4	-4.6303 E-01
Intercept.to.1layhid1	-7.2903 E-01	lag1.to.1layhid4	-1.3201 E-01
lag1.to.1layhid1	1.4118 E+00	lag2.to.1layhid4	7.7032 E-01
lag2.to.1layhid1	-4.6232 E-01	Intercept.to.price	1.2270 E-01
Intercept.to.1layhid2	-4.8327 E-01	1layhid1.to.price	1.3766 E+00
lag1.to.1layhid2	-1.3428 E+00	1layhid2.to.price	4.5368 E-01
lag2.to.1layhid2	2.9845 E+00	1layhid3.to.price	-3.5227 E-01
Intercept.to.1layhid3	-6.2035 E+00	1layhid4.to.price	-1.5304 E+00

H_1 = Tanh [-0.72-1.14 $\bar{Z}_{t-1}$ +0.46 $\bar{Z}_{t-2}$]
H_2 = Tanh [-0.48+1.34 $\bar{Z}_{t-1}$ -2.98 $\bar{Z}_{t-2}$]
H_3 = Tanh [0-6.20-17.29 $\bar{Z}_{t-1}$ +14.08 $\bar{Z}_{t-2}$]
H_4 = Tanh [-0.46+0.13$\bar{Z}_{t-1}$ -0.77 $\bar{Z}_{t-2}$]

Where $\bar{Z}_{t-1} = \left(\frac{lag1-\min(lag1)}{\max(lag1)-\min(lag1)}\right)$

$\bar{Z}_{t-2} = \left(\frac{lag2-\min(lag2)}{\max(lag2)-\min(lag2)}\right)$

$\bar{Z}_{t-1}$, and $\bar{Z}_{t-2}$ are rescaled input variables.

The FFNN model equation is

O_t = 0.12+1.37 H_1 +0.45H_2 -0.35H_3 -1.53H_3

The model performance for the FFNN model is shown in the blow table

Table 4.7: FFNN Model performance

Data set	MAE	RMSE	MAPE
training data	101.15	138.80	10.70
testing data	88.35	130.67	7.26

From the above 2-4-1model performance, observed that the FFNN models gives minimum error measures in test and training models. Hence mean absolute percentage error has in test 7.26 and mean absolute error has 130.67 and root mean square has 88.35, therefore the FFNN model is better model for forecast the monthly average prices of potato in Telangana.

The predicted values are provided in the monthly average prices of potato presented in following table 4.8

Table 4.8: Forecasts using FFNN model

Month-Year	Actual Potato prices	Predicted Potato prices
Jan-18	821.88	848.27
Feb-18	917.93	858.84
Mar-18	1038.46	966.45
Apr-18	1430.10	1087.49
May-18	1568.50	1517.80
Jun-18	1563.64	1634.65
Jul-18	1511.54	1534.78
Aug-18	1500.00	1452.55
Sep-18	1571.13	1474.28
Oct-18	1592.94	1596.89
Nov-18	1270.56	1581.21
Dec-18	1107.32	1034.81
Jan-19	972.89	1046.49
Feb-19	933.75	951.06
Mar-19	981.93	946.07
Apr-19	1076.09	1012.94
May-19	1135.33	1116.52
Jun-19	1143.68	1163.97
Jul-19	1109.21	1154.40
Aug-19	1032.08	1106.23
Sep-19	1050.00	1021.52
Oct-19	1306.67	1068.57
Nov-19	1363.39	1404.36
Dec-19	1100.46	1388.53

The below figure 4.3, it shows that the FFNN models gives the better results and graph also gives good performance. FFNN forecast provide the forecast of monthly average prices of Potato in Telangana.

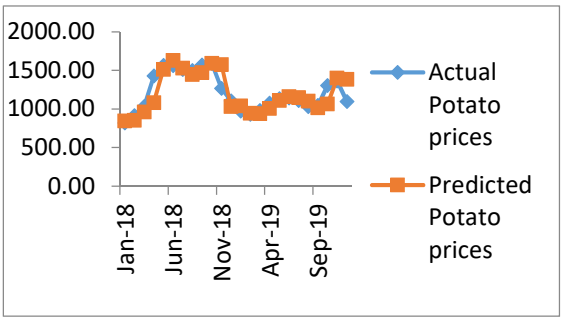


Figure 4.3: Testing sample forecasting using FFNN

4.3. Hybrid Model:

The development of HYBRID model for predicts the monthly average prices of potato in Telangana are covered. For building a HYBRID model, the linear component ARIMA model and Non linear component neural network were used. The following figure shows the root mean squared error

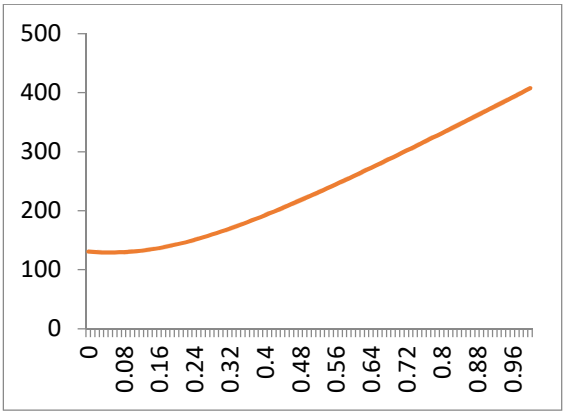


Figure 3.4.2: Weights for root mean squared error

The optimal HYBRID model is

$$\begin{pmatrix} \boldsymbol{HYBRID} \\ \boldsymbol{Predicted} \\ \boldsymbol{Values} \end{pmatrix} = \boldsymbol{W}\begin{pmatrix} \boldsymbol{ARIMA} \\ \boldsymbol{Predicted} \\ \boldsymbol{Values} \end{pmatrix} + (\mathbf{1} - \boldsymbol{W})\begin{pmatrix} \boldsymbol{FFNN} \\ \boldsymbol{Predicted} \\ \boldsymbol{Values} \end{pmatrix}$$

$$\begin{pmatrix} \boldsymbol{HYBRID} \\ \boldsymbol{Predicted} \\ \boldsymbol{Values} \end{pmatrix} = \mathbf{0.05}\begin{pmatrix} \boldsymbol{ARIMA} \\ \boldsymbol{Predicted} \\ \boldsymbol{Values} \end{pmatrix} + (\mathbf{1} - \mathbf{0.05})\begin{pmatrix} \boldsymbol{FFNN} \\ \boldsymbol{Predicted} \\ \boldsymbol{Values} \end{pmatrix}$$

The performance of model shown in the below table 4.9

Table 4.9: Hybrid model performance

Data sets	MAE	RMSE	MAPE
train data	100.79	138.25	10.17
test data	88.08	129.18	7.17

5. Comparison of ARIMA and FFNN models for potato prices in Telangana.

The forecasted prices of monthly average prices of potato in Telangana for ARIMA, FFNN and HYBRID models are in the following table.

Table 5.1: Forecasted values for ARIMA, FFNN and Hybrid Models

Month-Year	Actual prices	ARIMA predicted prices	FFNN predicted prices	HYBRID predicted prices
Jan-18	821.88	724.54	848.27	842.08
Feb-18	917.93	687.78	858.84	850.29
Mar-18	1038.46	730.79	966.45	954.67
Apr-18	1430.1	814.96	1087.49	1073.86
May-18	1568.5	855.35	1517.80	1484.68
Jun-18	1563.64	857.29	1634.65	1595.78
Jul-18	1511.54	883.54	1534.78	1502.22
Aug-18	1500	899.75	1452.55	1424.91
Sep-18	1571.13	919.26	1474.28	1446.53
Oct-18	1592.94	983.20	1596.89	1566.21
Nov-18	1270.56	988.88	1581.21	1551.59
Dec-18	1107.32	875.48	1034.81	1026.84
Jan-19	972.89	732.39	1046.49	1030.79
Feb-19	933.75	694.61	951.06	938.23
Mar-19	981.93	735.83	946.07	935.56
Apr-19	1076.09	820.15	1012.94	1003.30
May-19	1135.33	859.38	1116.52	1103.66
Jun-19	1143.68	862.06	1163.97	1148.88
Jul-19	1109.21	887.41	1154.40	1141.05
Aug-19	1032.08	904.12	1106.23	1096.13
Sep-19	1050	922.91	1021.52	1016.59
Oct-19	1306.67	986.86	1068.57	1064.48
Nov-19	1363.39	992.01	1404.36	1383.74
Dec-19	1100.45	879.95	1388.53	1363.10

The following chart shows the forecasts patterns are using ARIMA, FFNN and HYBRID models and it shows that, the compared to ARIMA, FFNN and HYBRID models are closer with high performance.

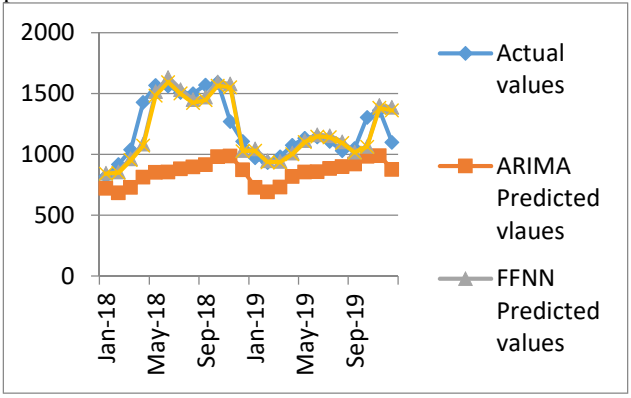


Figure 5.1: Testing sample forecasts Using ARIMA and FFNN and Hybrid models

The above graphical chart shows that, the forecasts provides the ARIMA, FFNN and HYBRID models, FFNN and HYBRID models are near and same values compared to ARIMA and they provide good performance of forecasting models. Compare to ARIMA and FFNN, HYBRID model has minimum error measure presented in the following chart 5.1

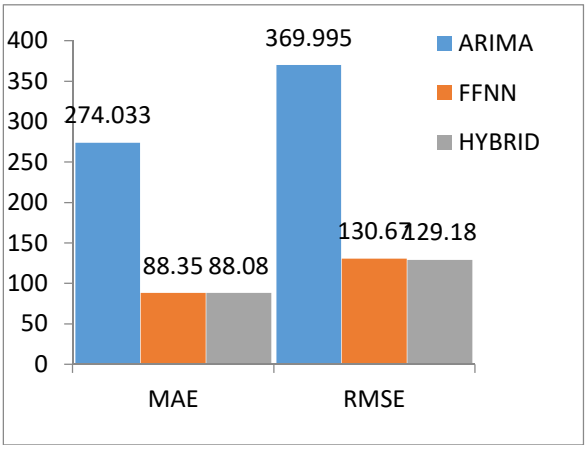


Figure 5.2: Comparison of performance for ARIMA and FFNN and HYBRID models

Table 5.1: Comparison ARIMA FFNN Model performance

	ARIMA		FFNN		HYBRID	
data sets	MAE	RMSE	MAE	RMSE	MAE	RMSE
training sample	118.122	166.65	101.15	138.8	100.79	138.25
testing sample	274.033	369.995	88.35	130.67	88.08	129.18

From the table 5.1, shows that the HYBRID model has low RMSE and MAE compare to ARIMA and FFNN. HYBRID model is appropriate than the ARIMA and FFNN model in testing sample for forecasts.

6. Conclusion:

In the above analysis, it was observed that HYBRID model is gives good forecasts result than ARIMA and FFNN models in forecasting prices of potato in Telangana. Hence, it was concluded that the HYBRID model is better model than the ARIMA and FFNN models for forecasting monthly average prices of potato in Telangana.

7. Reference:

1. Dr.M.Massarrat Ali Khan,Forecasting of Gold Prices Internal Journal Emerging Technology and Advanced Engineering. ISSN 2250-2459, ISO 9001:2008 Volume 3, Issue 3, March 2013
2. M.Basikhasteh, M.A. Movafaghpour and M.Saadatifara, 2014, "Time Series Forecasting Model for Iranian Gold Price". Indian J.Sci.Res., Vol.7, no.1, pp.1227-1232.
3. Banhi Guha and Gautam Bandyopadhyay,Gold Price Forecasting using Arima model, Journal of Advanced Management Science, Vol 4, No.2, March 2016.
4. Raghavender, M. (2009), Forecasting paddy yield in Andhra Pradesh using seasonal Time Series Model, Bulletin of Pure and applied sciences.
5. R.Rama Krishna & Naveen Kumar.B, 2013, Forecasting yield per Hectare of Rice in AP; IJMCAR, ISSN 2249-6955, Volume 3, Issue 1, March 2013, 9-14.
6. Hornik, 1993, Some New result on neural network approximation, Neural Network, 6(1993), pp. 1069-1072.
7. Adanacioglu,H.And Yercan, M., 2012,An Analysis of Tomato Prices at Wholesale level in Turkey: An Application of SARIMA model. Custose agronegocio, 8 (4): 52-75.
8. M.A.Khan, 2013, "Forecasting of Gold Prices (Box Jenkins Approach)". 'International Journal of Emerging Technology and Advanced Engineering, Vol.3, PP.662-67.
9. MAKRIDAKIS S., WHEEL WRIGHT. S.C., HYNDMAN R.J., 2003, Forecasting Methods and Applications, (John Wiley & Sons).
10. Box, G.E.P., Jenkins, G.M. and Reinsel, G.C., (1994), "Time Series Analysis Forecasting and Control", 3rd ed., Englewood cliffs, N.J.Prentice Hall.
11. Dr.M.Massarrat Ali Khan,Forecasting of Gold Prices Internal Journal Emerging Technology and Advanced Engineering. ISSN 2250-2459, ISO 9001:2008 Volume 3, Issue 3, March 2013
12. K.Murali Krishna, Dr. M. Raghavender Sharma and Dr. N.Konda Reddy,Forecasting of silver prices using Artificial Neural Networks. JARDCS, Volume 10, 06 issue 2018.
13. Pitigalaarachchi P.A.A.C, Jayasundara D.D.M, Chandrasekara N.V, 2016, Modelling and Forecasting Sri Lankan Gold Prices IJSBAR, ISSN 2307-4531.
14. . C.Gershenson, "Artificial Neural Networks for Beginners" Networks, 2003., Vol. cs.NE/0308, pp.8.
15. shahriar shafiee, Erkan Topal Elsevier,An overview of global gold Market and gold price forecasting Resources policy, 35(2010),
16. L Jung, G.M. and Box, G.E.P, (1979), " On A measure of lack of Fit in Time Series Models", Biometrika, 65297-303.
17. Adanacioglu,H.And Yercan, M., 2012,An Analysis of Tomato Prices at Wholesale level in Turkey: An Application of SARIMA model. Custose agronegocio, 8 (4): 52-75.
18. Anonymous. 2012a. Report on Fresh Fruits and Vegetables, Republic of Turkey, Ministry of Economy, 8p. [http://www.tcp.gov.tr/english/sectors/index.cfm]
19. Raghavender, M. (2010), Forecasting paddy production in Andhra Pradesh with ARIMA model, Int. J. Agricult. Stat. Sci., Vol 6 (1), pp 251-258.
20. Haykin, S.S., 1999, "Neural Networks: A Comprehensive Foundation", Upper Saddle River, N.J., Prentice Hall.