International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

An Improved Arima Model Using Ann and Svm for Forecasting Rapeseed & Mustard Production

Borsha Neog*[1], Bipin Gogoi[2], A. N. Patowary[3]

Submitted: 07/02/2024 **Revised**: 15/03/2024 **Accepted**: 21/03/2024

Abstract: Forecasting of agricultural crop production is the art of predicting production before harvest and is crucial for planning and policy making at various stages. "Rapeseed & Mustard (R & M) is the predominant oilseed crop of Assam because of its short duration. To know about futures estimates, planning is much importance to make fruitful decisions. In this context, the present study was undertaken to develop proper forecasting models for R & M of Assam. Here we have been used yearly data on production of Rapeseed & Mustard for forecasting from the year 1951 to 2018. For model building, we have used data from 1951-1998 and for model testing data from 1999 - 2018 were used for forecasting performance of the model. In this study, to analyse the past behaviour of the production of Rapeseed & Mustard to make interpretations about its future behaviour using different models Autoregressive integrated moving average (ARIMA), Artificial neural network (ANN), support vector machine (SVM) and hybrid of both ARIMA-ANN, ARIMA-SVM. For the selected crops, ARIMA (0,1,0) model was selected as a suitable model. In training, mean absolute error (MAE) for hybrid ARIMA (0,1,0)-SVM was found to be 8216.169 as compare to 8813.731 of ARIMA-ANN; 10620.825 of ARIMA (0,1,0); 10242.319 of ANN; 9831.046 of SVM. In testing, MAE for hybrid ARIMA (0,1,0)-SVM was found to be 8174.671 as compare to as compare to 9263.464 of ARIMA-ANN; 10606.565 of ARIMA (0,1,0); 10384.249 of ANN; 10139.604 of SVM. Henceforth, the performances of hybrid ARIMA-ANN and ARIMA-SVM were found to be better than that of ARIMA for both under training as well as testing data sets. So from the results we can recommend hybrid approach gives better results for forecasting of Rapeseed & Mustard production.

Keywords: *ARIMA, ANN, SVM, Forecasting, Rapeseed & Mustard.*

1. Introduction

Factual estimating model is utilized to foster exact determining by utilizing past information through distinguishing proof of patterns and examples of the information. In a monetary framework, appropriate estimate is vital in light of the fact that it would be more straightforward to plan the arrangement producers in regards to cost obsession, portion, obtainment, water system, showcasing, and capacity. This enables the government to take some initiatives such as import/export regulations, food aid programs or strategies to improve productivity and support rural development. For addressing these problems, major key instruments are statistical computing, modelling, and forecasting. Thus, the prime objective of forecasting is to give precise, scientific, and independent forecasts of crop production as early as possible.

To overcome the time series forecasting errors, combine different forecasting techniques ARIMA-SVM used [12, 14, 16]. ARIMA-ANN model is suitable for forecasting of rice yield during kharif season in West Bengal [4]. "One of the challenging issues in precision agriculture is crop yield forecasting, and numerous models have already been developed and tested.

Since climate, weather, soil, fertilizer use, and seed variety all affect crop yield, this problem necessitates the use of multiple datasets [18]. This demonstrates that crop yield prediction is not a straightforward process but rather a series of intricate steps. These days, crop yield forecast models can gauge the genuine yield sensibly, however a superior execution in yield expectation is as yet attractive" [10]. To get an outline of what has been finished on the utilization of ML in expectation of harvest, we played out a review utilizing AI procedures ANN and SVM for estimating Rapeseed and Mustard creation to settle on conclusions about its future way of behaving. It has been observed that hybrid approaches are more efficient and effective in enhancing the model's ability to forecast [13, 19].

2. Proposed Methodology

Yearly information on creation of Rapeseed and Mustard have been utilized for guaging from the year 1951 to 2018. The information from 1951-1998 were utilized for model structure and 1999 - 2018 were utilized for actually looking at the anticipating execution of the model. Programming projects SPSS, R; were utilized for displaying and determining of creation of R and M in Assam. The appropriate ARIMA model was created with the help of SPSS software. R programming bundle

[1] *Department of Agricultural Statistics, Assam Agricultural University, Jorhat-785013, Assam, India*
[2] *Department of statistics, Dibrugarh University, Dibrugarh-786004, Assam, India*
[3] *College of Fishery Science, Assam Agricultural University, Raha-782103, Assam, India*
* *Corresponding Author Email: borsha.neog@aau.ac.in*

'Figure' was utilized for demonstrating and anticipating utilizing NN and bundle 'e 1071' was utilized for displaying and guaging utilizing SVM.

2.1 Time series forecasting models:

2.1.1 The ARIMA model:

In an ARIMA model, time series variable is assumed to be a linear function of past actual values and random shocks. An ARIMA (p, d, q) model is defined by the following equation

$\phi(B)(1-B)d\ y_t=\Theta(B)\ \varepsilon_t$ (1)

Where,

$\phi(B)= 1- \phi_1 B-\phi 2B_2-\ldots-\phi pB_P$

(Autoregressive parameter)

$\Theta(B)= 1- \Theta_1 B-\Theta_2 B_2-\ldots-\Theta pB_P$

(Moving average parameter)

ε_t = White noise or Error term

d= Differencing term

B= Backshift operator i.e., $B_a Y_t=Y_{t-a}$

ARIMA strategy is completed in three phases, i.e., Recognizable proof, assessment, and demonstrative checking. The non-stationary time series must be made stationary by identifying d. A formal statistical test known as the test of unit root hypothesis or the Augmented Dickey Fuller test is used to verify stationarity. ADF test was used to check the stationarity [15]. At the phase of assessment, boundaries are assessed for the ARIMA model probably picked at the early ID stage. Boundaries assessment for ARIMA model is by and large finished through iterative least squares strategy. The ampleness of chosen model is tried at the phase of indicative checking. At this stage, testing is finished to check whether the assessed model is genuinely sufficient i.e., whether the blunder terms are background noise implies mistake terms are uncorrelated with mean zero and consistent change. Ljung-Box test is applied to the original series or residuals after fitting a model for this purpose. Box et al. provide an excellent explanation of the Ljung-Box test. [5]. Assuming the model is viewed as insufficient, the three phases are rehashed until palatable ARIMA model is chosen for the time series viable.

2.1.2 Ljung-Box test:

Greta M. Ljung, a statistician, is the name of the Ljung-Box test. Ljung and E.P. George A statistical test called the Box is used to see if a time series has autocorrelation. It's also known as the Box-Pierce test. The test finds out if errors are i.i.d. (also known as white noise). The Ljung-Box test's null hypothesis is

H_0: The alternative hypothesis is that the residuals are distributed independently.

H_1: The residuals have serial correlation and are not distributed independently.

The test statistic for the Ljung-Box test is as follows:

$Q = n(n+2)\ \Sigma p_k^2 / (n-k)$

where: n = sample size, p_k = sample autocorrelation at lag k

The test statistic Q follows a chi-square distribution with h degrees of freedom; that is,

$Q \sim \chi 2(h)$.

We reject the null hypothesis and say that the residuals of the model are not independently distributed" if $Q > \chi 2_{1-\alpha,\ h}$

2.1.3 Artificial Neural Network (ANN) model:

ANN(s) requires no earlier presumption of the information producing process, rather it is generally relied upon qualities of the information known as information driven approach. The most common time series forecasting and modeling model is the single hidden layer feed forward network. Multilayer ANNs are the network of simple processing units with three layers that create this model. The main layer is input layer, the center layer is the secret layer and the last layer is yield layer.

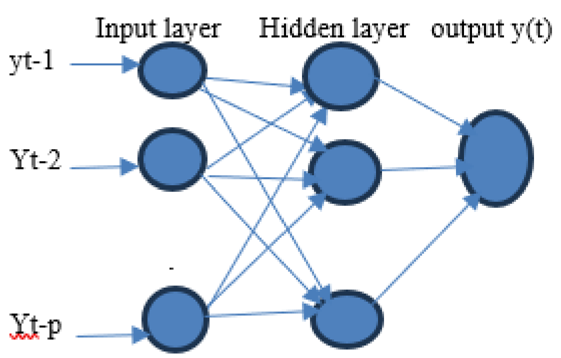

Figure 1: Neural Network architecture

"The relationship between the output (y_t) and the inputs ($y_{t-1}, y_{t-2}, \ldots, y_{t-p}$) can be mathematically represented as follows:

$Yt=f(\sum_{j=0}^{q} w_j g\ (\sum_{i==0}^{q} w_{ij}\ y_{t-i}))$

Where $w_j(j=0,1,2,\ldots,q)$ and $w_{ij}(i=0,1,2,\ldots,p; j=0,1,2,\ldots,q)$ are the model parameters often called the connection weights; p is the number of input nodes and q is the number of hidden nodes, g and f denote the activation function at hidden and output layer respectively".

2.1.4. Support Vector Machine:

Support vector machine proposed is a nonlinear calculation utilized in regulated learning system for information characterization, design acknowledgment and relapse examination [17]. The model was constructed in two stages: testing and training. during the testing and training. In the preparation step, the biggest piece of the dataset has been utilized for the assessment of the capability. "In the testing step, the speculation capacity of the model has been assessed by really taking a look at the model presentation in the little subset.

It has been utilized in a large number of uses, for example, in information mining, characterization, relapse, and time series determining [8,9,20]. The capacity of SVM is to tackle nonlinear relapse assessment issues and it makes SVM effective in time series determining.

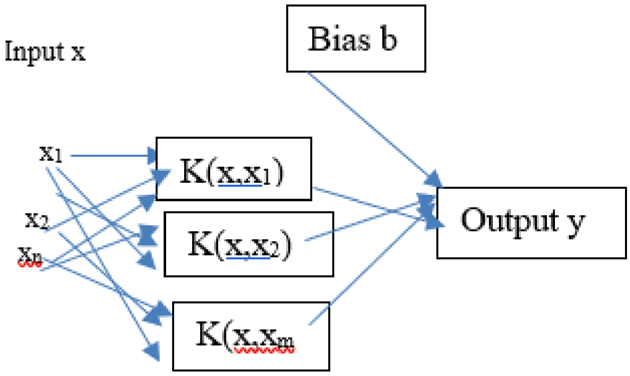

Figure 2: SVM Architecture"

Evaluation Criteria:

The most well-known mistake capability in brain networks is the number of squared blunders. Asymmetric least squares, least fourth powers, least absolute deviations, and percentage differences are additional error functions provided by various software.

2.1.4. Hybrid approach:

This approach follows the Zhang's (2003) cross breed approach, appropriately the connection among direct and nonlinear parts can composed as follow

$Y_t = L_t + N_t$

"The fundamental system of this approach is to display the direct and nonlinear parts independently by various model. The philosophy comprises of three stages. First and foremost, ARIMA model is applied to the information series to fit the direct part. Let the forecast series given by ARIMA model signified as $\widehat{L_t}$. In the subsequent step, rather than foreseeing the direct part, the residuals meant as et which are nonlinear in nature are anticipated. The residuals can be acquired by taking away the anticipated worth $\widehat{L_t}$. from genuine worth of the thought about time series y_t."

$e_t = y_t - \widehat{L_t}$

An ANN and SVM model is now used to predict the residuals. Let the forecast series given by ANN/SVM model indicated as $\widehat{N_t}$. At last, the anticipated straight and nonlinear parts are joined to produce total expectation.

$\widehat{y_t} = \widehat{L_t} + \widehat{N_t}$

Ljung-Box test is used to test for non-linearity in this study.

The graphical representation of proposed approach is expressed in the figure 3 & 4"

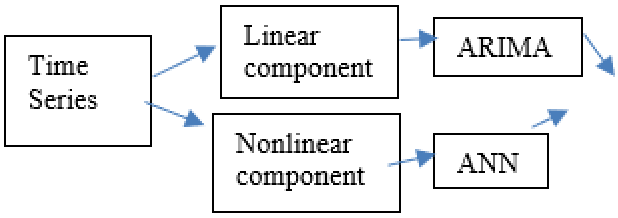

Figure 3: Schematic representation of ARIMA-ANN hybrid methodology

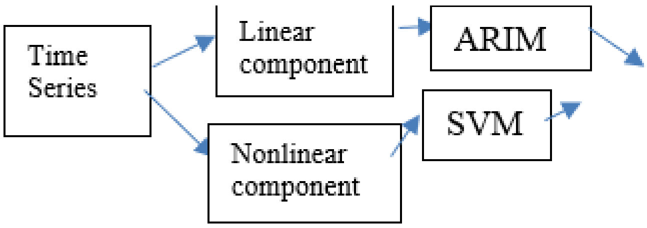

Figure 4: Schematic representation of ARIMA-SVM hybrid methodology

Forecasting Performance:

"Determining Execution of the model has been changed by figuring mean outright mistake (MAE). The model with least upsides of MAE for preparing and testing informational collection is liked for determining reason. The MAE is determined by

$MAE = \frac{1}{n}\sum_{t=1}^{n}|y_t - \hat{y}_t|$

"Where n is the total number of forecast values. Y_t is the actual value at period t and $\hat{y}_t$ is the corresponding forecast value".

3. Results and Discussion:

Table 1: Summary Statistics of Rapeseed & Mustard Production time series

Statistic	Rapeseed & Mustard Production	Statistic	Rapeseed & Mustard Production
Observation	68	Maximum	199501.00
Mean	106432.22	Standard deviation	48727.05
Median	123218.50	Skewness	0.08
Range	159270.00	Kurtosis	-1.41
Minimum	40231.00	Coefficient of variation (%)	45.78

"Here we have used combined traditional model ARIMA and soft computing techniques ANN and SVM on production of R & M crop from the year 1951 to 2018. For model building, 1951 to 1998 taken as a training set and for validation, 1999 to 2018 taken as a testing set. Sequence charts for production of R & M is shown in Figure 5".

Figure 5: Sequence charts for production of Total rice

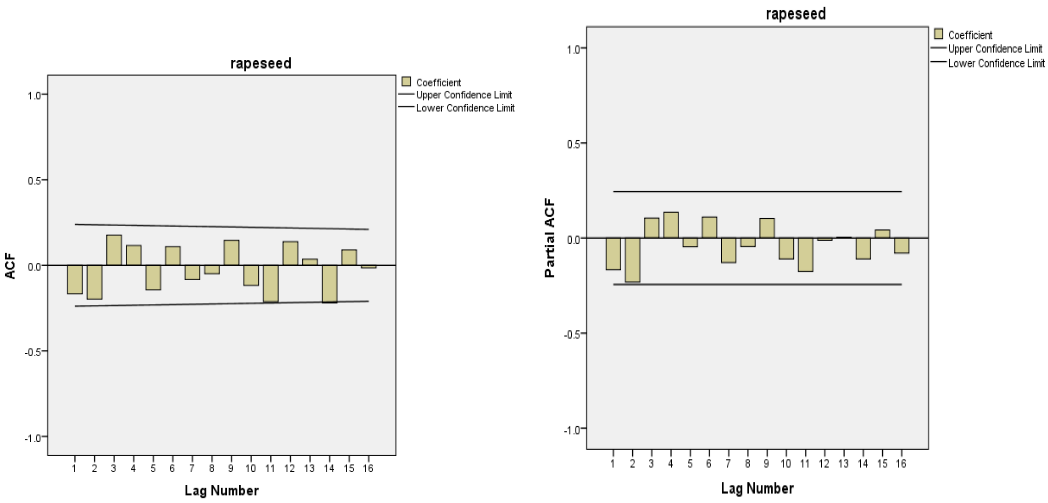

Figure 6: ACF & PACF of production of Rapeseed & Mustard

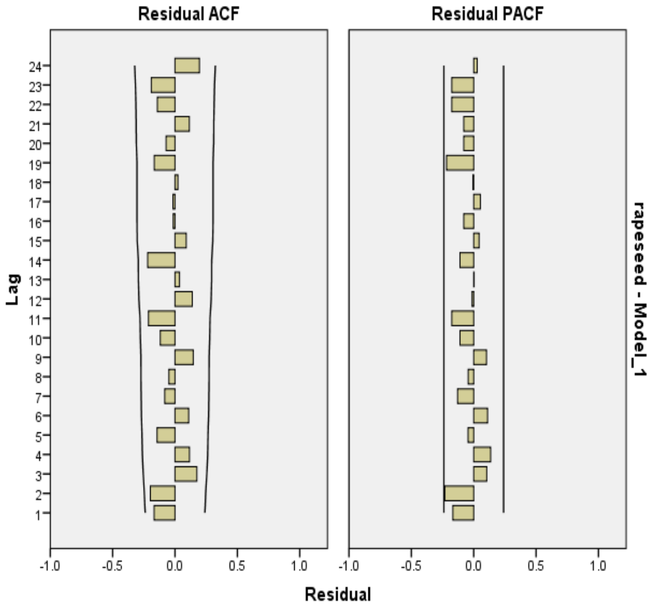

Figure 7: Residual ACF & Residual PACF of production of Rapeseed & Mustard

"The production of rapeseed and mustard from 1951 to 2018 was used in the Auto Regressive Integrated Moving Average (ARIMA) model for linear time series models to forecast Assam. In view of the base upsides of proportions of decency of fit, similar to Root Mean Square Blunder (RMSE), Mean outright Rate Blunder (MAPE), Mean Outright Blunder (MAE) given in Table 2 (BIC= 19.117), background noise utilizing Ljung-Box Q test for residuals detailed in Table 3 (p value= 0.144) and meaning of the boundary gauges given in Table 4, we chose the model ARIMA (0,1,0) as a best fitted for the creation of R and M and a similar model was found via auto.arima choice from R programming. By utilizing the direct model ARIMA (0,1,0), we determined creation esteem by 2025 which are given in Table 5 and graphically introduced in the accompanying figure 8".

"Table 2: Goodness of fit Statistics of Rapeseed & Mustard

Fit Statistic	ARIMA (0,1,1)	ARIMA (1,1,0)	ARIMA (1,1,1)	ARIMA (0,1,0)
Stationary R-squared	0.043	0.028	0.063	0.000
R-squared	0.924	0.922	0.925	0.920
RMSE	13537.528	13640.467	13494.498	13729.669
MAPE	11.034	11.010	10.680	10.935
MAE	10572.726	10614.127	10122.912	10554.118
MaxAPE	56.546	56.897	59.595	59.514
MaxAE	36792.624	38469.651	37410.472	41649.030
Normalized BIC	19.152	19.167	19.208	19.117

Table 3: Test for white noise of Rapeseed & Mustard

Model	Ljung-Box Q		
	Statistics	DF	Sig.
ARIMA (0,1,1)	17.962	17	0.391
ARIMA (1,1,0)	20.667	17	0.242
ARIMA (1,1,1)	21.619	16	0.156
ARIMA (0,1,0)	24.340	18	0.144

Table 4: Parameter Estimates of Rapeseed & Mustard

Models	Parameter Estimate	SE	t	Sig.
ARIMA (0,1,1)	0.236	0.121	1.948	0.056
ARIMA (1,1,0)	-0.166	0.122	-1.354	0.180
ARIMA (1,1,1)				
AR Lag1	0.852	0.141	6.026	0.000
MA Lag1	1.000	45.535	0.022	0.983
ARIMA (0,1,0) Constant	1982.030	1677.346	1.182	0.242

Table 5: Forecast of production of Rapeseed & Mustard

Year	Forecast	LCL	UCL
2019	185647.03	158234.86	213059.20
2020	187629.06	148862.40	226395.72
2021	189611.09	142131.82	237090.36
2022	191593.12	136768.78	246417.46
2023	193575.15	132279.68	254870.62
2024	195557.18	128411.35	262703.01
2025	197539.21	125013.43	270064.99

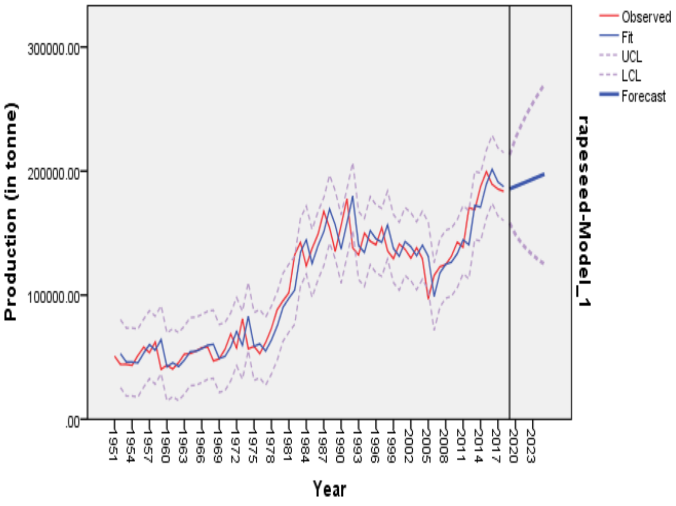

Figure 8: Forecast of production of Rapeseed & Mustard by 2025 using ARIMA (0,1,0)"

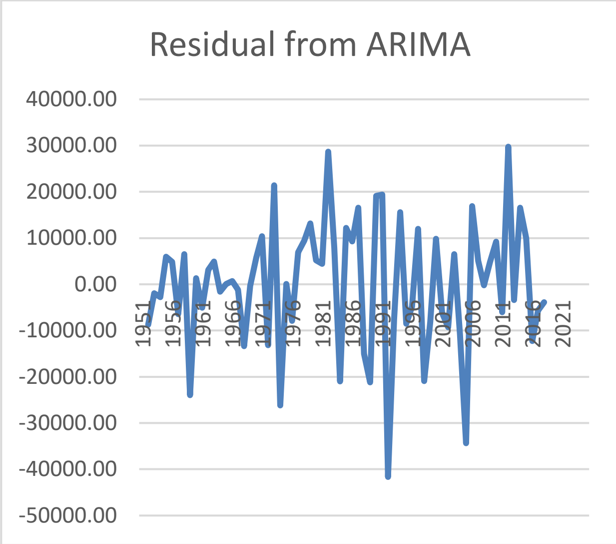

Figure 9: Residual graphs for production of Rapeseed & Mustard using ARIMA (0,1,0)

Table 6: Tests of Normality for residuals of Rapeseed & Mustard

Kolmogorov-Smirnov			Shapiro-Wilk		
Statistic	Df	P value	Statistic	Df	P value
0.074	67	0.200	0.980	67	0.343

From the above tests and diagrams it is affirmed that the residuals acquired through ARIMA (0,1,0) are nonlinear as the p esteem is non-huge. Consequently, we can apply non-direct delicate registering models i.e., fake brain organization (ANN) and support vector machine (SVM) the two methodologies have applied for displaying and estimating of residuals acquired through the chose model of ARIMA (0,1,0).

For better exactness of estimating on creation of R and M was gotten through half breed approach i.e., ARIMA (0,1,0)- ANN. The results of our experiments with various neural networks with varying hidden nodes on residuals and time delays are presented in Table 7.

Table 7: "MAE for Neural Network models for production of Rapeseed & Mustard"

Model parameters	**MAE for Training**	**MAE for Testing**
1:2s:1l	10797.968	10602.707
1:4s:1l	10809.118	10497.728
1:6s:1l	10807.493	10400.956
1:8s:1l	10811.171	10281.363
1:10s:1l	10816.912	10295.901
2:2s:1l	10438.479	10630.814
2:4s:1l	10310.212	10523.965
2:6s:1l	10254.689	10520.685
2:8s:1l	**10242.319**	**10384.249**
2:10s:1l	10250.279	10471.139
3:2s:1l	10411.137	10498.484

3:4s:1l	10435.519	10497.111
3:6s:1l	10363.016	10500.895
3:8s:1l	10318.797	10427.039
3:10s:1l	10291.017	10431.597
4:2s:1l	10469.024	10266.383
4:4s:1l	10402.223	10308.716
4:6s:1l	10389.078	10397.944
4:8s:1l	10385.676	10380.596
4:10s:1l	10417.733	10223.206
5:2s:1l	10568.002	10338.615
5:4s:1l	10516.380	10376.837
5:6s:1l	10537.215	10219.865
5:8s:1l	10533.248	10310.541
5:10s:1l	10508.193	10210.114
6:2s:1l	10722.834	10244.382
6:4s:1l	10679.276	10366.429
6:6s:1l	10698.845	10094.726
6:8s:1l	10706.902	10419.423
6:10s:1l	10681.590	10315.543

The model 2:8s:1l was found to be the best based on minimum mean absolute error (MAE) values of 10242.319 for training and 10384.249 for testing. From this best chosen model, we have determined the assessed upsides of residuals and fitted upsides of creation of R and M got by ARIMA (0,1,0) then, at that point, figure worth of creation was acquired through cross breed approach i.e., ARIMA (0,1,0)- ANN. The hybrid ARIMA-ANN goodness of fit measure was found to be 8991.752, compared to 10554.118 for the ARIMA (0,1,0). Once more, residuals acquired through ARIMA (0,1,0) were applied on the non-direct methodology i.e., support vector machine involving spiral premise capability as piece. The MAE for hybrid ARIMA-SVM was estimated and the forecast values for production that were obtained through ARIMA (zero, one, zero) were corrected with the residuals from SVM. MAE for cross breed ARIMA-SVM was viewed as 8912.512 as contrast with 10554.118 of ARIMA (0,1,0) and 8991.752 of mixture ARIMA-ANN. Consequently, the hybrid model's performance was found to be superior to ARIMA (0,1,0) alone.

For estimate the forecast value of production of R & M through hybrid approach along with forecast values of ARIMA (0,1,0) presented in the given table.

Table 8: “Experimental Results of forecast of Production of Rapeseed & Mustard”

Year	Actual values of Production	Forecast Production by ARIMA (0,1,0)	Forecast Production by Hybrid Approach using ANN	Forecast Production by Hybrid Approach using SVM
1951	50869.00			
1952	44139.00	52851.03		
1953	44175.00	46121.03		
1954	43368.00	46157.03		
1955	51293.00	45350.03		
1956	58142.00	53275.03		

1957	53667.00	60124.03	52703.63	50090.49
1958	62192.00	55649.03	63531.64	61346.02
1959	40231.00	64174.03	38468.83	35718.58
1960	43557.00	42213.03	50592.93	52147.84
1961	40488.00	45539.03	43705.09	46148.27
1962	45570.00	42470.03	50556.80	52895.72
1963	52458.00	47552.03	54806.17	56558.46
1964	52882.00	54440.03	53064.98	51828.37
1965	54936.00	54864.03	55819.74	53563.40
1966	57625.00	56918.03	57871.98	55383.28
1967	58415.00	59607.03	58093.73	55419.74
1968	47011.00	60397.03	46823.15	44275.93
1969	48627.00	48993.03	52397.35	51334.03
1970	56369.00	50609.03	58175.61	57290.97
1971	68733.00	58351.03	67966.70	65738.00
1972	57544.00	70715.03	54702.67	52373.75
1973	80938.00	59526.03	82521.31	82467.47
1974	56738.00	82920.03	52314.46	51398.94
1975	58779.00	58720.03	62285.72	67791.25
1976	52899.00	60761.03	52486.47	49857.41
1977	61836.00	54881.03	63512.26	61481.37
1978	73224.00	63818.03	71446.01	68700.39
1979	88403.00	75206.03	85640.59	84484.48
1980	95613.00	90385.03	93350.59	97029.78
1981	102000.00	97595.03	102162.53	111253.11
1982	132628.00	103982.03	133915.24	144458.88
1983	142542.00	134610.03	142965.72	145617.86
1984	123540.00	144524.03	125730.38	122411.24
1985	137737.00	125522.03	144748.68	144594.41
1986	148997.00	139719.03	151261.54	23418.88
1987	167523.00	150979.03	168451.03	163287.93
1988	154477.00	169505.03	149397.26	144276.42
1989	135243.00	156459.03	136714.26	130949.89
1990	156312.00	137225.03	162539.06	156285.90
1991	177672.00	158294.03	176027.87	169855.28
1992	138005.00	179654.03	128493.51	126028.09
1993	132429.00	139987.03	140212.63	129719.12
1994	150009.00	134411.03	154991.77	153248.30
1995	143463.00	151991.03	143602.27	150687.88
1996	140607.00	145445.03	144278.26	138243.02
1997	154572.00	142589.03	158358.87	153662.23
1998	135631.00	156554.03	135021.34	129118.85
1999	129425.00	137613.03	135571.12	129253.80
2000	141231.00	131407.03	146460.62	146208.93
2001	137056.00	143213.03	139222.17	136707.00
2002	129784.00	139038.03	134178.24	130558.22
2003	138296.00	131766.03	143640.17	142917.38
2004	129395.00	140278.03	132255.11	130532.34
2005	96992.00	131377.03	102554.09	101635.47
2006	115874.00	98974.03	122987.84	137606.23
2007	122897.00	117856.03	124015.68	131976.20
2008	124688.00	124879.03	128154.45	134048.74
2009	131493.00	126670.03	135742.04	139969.90

2010	142661.00	133475.03	146190.89	147600.43
2011	138647.00	144643.03	140690.04	137499.19
2012	170382.00	140629.03	174563.31	170340.27
2013	168977.00	172364.03	162012.16	154487.13
2014	187522.00	170959.03	182961.63	182172.90
2015	199501.00	189504.03	185128.04	187382.24
2016	189233.00	201483.03	167987.78	172170.53
2017	185564.00	191215.03	170995.71	175075.11
2018	183665.00	187546.03	171028.72	176309.46
2019		185647.03	173999.57	179913.00
2020		187629.06	169532.68	176569.00
2021		189611.09	166868.51	173523.00
2022		191593.12	165231.64	170578.00
2023		193575.15	164191.42	167574.00
2024		195557.18	163518.06	164442.00
2025		197539.21	163076.91	161164.00

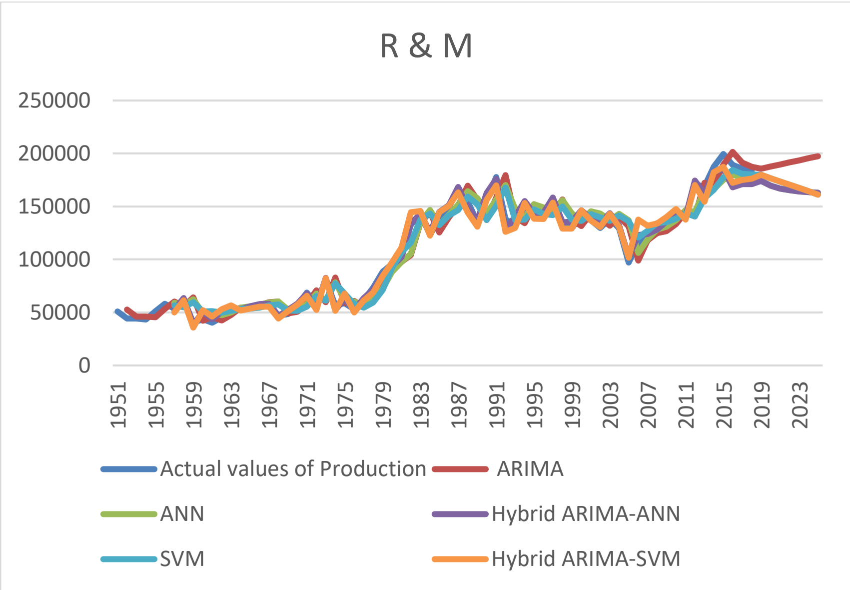

Figure 10: Comparison graphs for production of Rapeseed & Mustard using different models.

Table 9: MAE of different models for production of Rapeseed & Mustard

Data	ARIMA	ANN	SVM	ARIMA-ANN	ARIMA-SVM
Training	10620.825	10242.319	9831.046	8813.731	8216.169
Testing	10606.565	10384.249	10139.604	9263.464	8174.671

"For different models ARIMA (0,1,0), ANN (02:8s:1l), SVM, ARIMA-ANN and ARIMA-SVM, the values of MAE under training set are found to be 10620.825, 10242.319, 9831.046, 8813.731 & 8216.169 respectively, whereas the values of MAE under testing set are found to be 10606.565, 10384.249, 10139.604, 9263.464 & 8174.671 respectively. Based on these results, the model ARIMA-SVM is the suitable model for forecasting of production of R & M because of the minimum value of MAE both under training and testing set".

4.Conclusion:

"In this study, we have applied different hybrid models and compared their performances with individual traditional and soft computing models for forecasting of R & M in Assam. ARIMA (0,1,0) model was selected as suitable model for Rapeseed & Mustard based on measures of goodness of fit. For training data set MAE for hybrid ARIMA (0,1,0)-SVM was found to be 8216.169 as compare to 8813.731 of ARIMA-ANN; 10620.825 of ARIMA (0,1,0); 10242.319 of ANN;

9831.046 of SVM. For testing data set, MAE for hybrid ARIMA (0,1,0)-SVM was found to be 8174.671 as compare to as compare to 9263.464 of ARIMA-ANN; 10606.565 of ARIMA (0,1,0); 10384.249 of ANN; 10139.604 of SVM. Based on the accuracy measures, it is evident that performance of hybrid models gives better results compared with individual models. It is also suggested to explore the performance of different hybrid models such as ARIMA-MARS, ARIMA-NARX, ARIMA-NLSVRX etc. for agricultural crop forecasting".

References:

[1] Alexandridis, A.K. ; Zapranis, A.D. (2013). Wavelet neural networks: A practical guide. *Neural Networks*, 42, 1-27. http://dx.doi.org/10.1016/j.neunet.2013.01.008

[2] Alam, M.W.; Sinha, K.; Ranjan, R.K.; Ray, M; Rathod, S.; Singh, K.N. (2018). *Development of Hybrid time series models using Machine learning techniques for forecasting crop yield with covariates.* Indian Agricultural Statistics Research Institute, New Delhi.

[3] Banakar, A.; Azeem, M.F. (2008). Artificial wavelet network and its application to neuro-fuzzy models. *Appl. Soft. Computing,* 8 (4) , 1463-1485. https://doi.org/10.1016/j.asoc.2007.10.020

[4] Banik, A; G. Raju; S. Shukla (2021). Rice yield forecasting in West Bengal using hybrid model. Data Science and security, *Proceedings of IDSCS 2021*. PP. 222-231

[5] Box, G. E. P.; Pierce, D. A. (1970). Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Models. *Journal of the American StatisticalAssociation.* 65 (332):1509-1526. doi:10.1080/01621459.1970.10481180.

[6] Box, G.E.P.; Jenkins, G.M.; Reinsel, G.C. (1994), *Time Series Analysis: Forecasting and Control (3rd ed.)*. Holden-Day, San Francisco.

[7] Broock, W.; Scheinkman, J.A.; Dechert, W.D.; LeBaron, B. (1996). A test for independence based on the correlation dimension. *Econ. Rev.*, 15, 197-235. https://doi.org/10.1080/07474939608800353

[8] Cao, L.J.; E.H. Tay, 2003. Support vector with adaptive parameters in financial time series forecasting. *IEEE Trans. Neural Network*, 14:1506-1518. DOI: 10.1109/TNN.2003.820556

[9] Flake, G.W.; S. Lawrence, 2002. *Efficient SVM regression training with SMO. Machine learning,* 46:271-290

[10] Filippi, P., Jones, E.J., Wimalathunge, N.S., Somarathna, P.D.S.N., Pozza, L.E., Ugbaje, S.U., Bishop, T.F.A., 2019a. An approach to forecast grain crop yield using multilayered, multi-farm data sets and machine learning. *Precis. Agric*. 1–15. https://doi. org/10.1007/s11119-018-09628-4.

[11] G. M. Ljung, G. E. P. Box (1978). On a Measure of a Lack of Fit in Time Series Models. *Biometrika*. 65 (2): 297–303. doi:10.1093/biomet/65.2.297

[12] Hervai, S.; Osborn, DR.; Birchenhall, CR. (2004). Linear versus neural network forecast for European industrial production series. *Int J Forecast* 20:435-446. https://doi.org/10.1016/S0169-2070(03)00062-1

[13] Jha, G.K.; Sinha, K. (2014). Time-delay neural networks for time series prediction: an application to the monthly wholesale price of oilseeds in India. *Neural Comput. Appl*., 24(3). http://dx.doi.org/10.1007/s00521-012-1264-z

[14] Kohzadi, N.; Boyd, MS.; Kermanshahi, B.; Kaastra, I. (1995). A comparison of artificial neural network and time series models for forecasting commodity prices. *Neurocomputing* 10:169-181. https://doi.org/10.1016/0925-2312(95)00020-8

[15] Makridakis, S.; Wheelwright, S.C.; Hyndman, R.J. (1998). *Forecasting: Methods and Applications (3rd ed.)*. Wiley, Chichester.

[16] Neog, B.; Gogoi, B.; Barua, S.; Mahanta, S.; Sarmah, K. (2023). Development of hybrid time series models for forecasting autumn rice production in Assam. *The Pharma Innovation,12(2): 3376-3384. ISSN (E): 2277-7695*

[17] Thoranin, Sujjaviriyasup.; Komkrit, Pitiruek. (2013). Hybrid ARIMA-Support Vector Machine Model for Agricultural Production Planning. *Applied Mathematical Sciences*, 7(57), 2833 – 2840. https://doi.org/10.12988/ams.2013.13251.

[18] Vapnik, V.N. (1995). The nature of Statistical learning Theory. 1st edn., Springer-Verlog, New York, ISBN: 0-387-94559-8. http://dx.doi.org/10.1007/978-1-4757-2440-0

[19] Xu, X., Gao, P., Zhu, X., Guo, W., Ding, J., Li, C., Wu, X. (2019). Design of an integrated climatic assessment indicator (ICAI) for wheat production: a case study in Jiangsu Province, *China. Ecol. Ind.* 101, 943–953. https://doi.org/10.1016/j.ecolind.2019. 01.0591

[20] Zhang, G. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing,* 50, 159-175. https://doi.org/10.1016/S0925-2312(01)00702-0

[21] Zhao, C. Y.; H.X. Zhang; M.C.; Liu, Z.; D. Hu; B.T. Fan. (2006). Application of support vector machine (SVM) for prediction toxic activity of different data sets. *Toxicology,* 217: 105-119. https://doi.org/10.1016/j.tox.2005.08.019