Predicting rice yield for Bangladesh by exploiting weather conditions

Conference Paper · December 2017

DOI: 10.1109/ICTC.2017.8191047

4 authors:

Md. Akter Hossain
East Delta University
6 PUBLICATIONS 20 CITATIONS

Mohammed nazim uddin
East Delta University
36 PUBLICATIONS 220 CITATIONS

Mohammad Arif Hossain
New Jersey Institute of Technology
31 PUBLICATIONS 266 CITATIONS

Yeong Min Jang
Kookmin University
297 PUBLICATIONS 2,813 CITATIONS

Predicting Rice Yield for Bangladesh by Exploiting Weather Conditions

Md. Akter Hossain[1], Mohammed Nazim Uddin[1], Mohammad Arif Hossain[1], and Yeong Min Jang[2]
[1]School of Science, Engineering & Technology, East Delta University, Chittagong, Bangladesh.
[2]Department of Electronics Engineering, Kookmin University, Korea.
Email: mdakterhossain293@gmail.com, nazim@eastdelta.edu.bd, arif.h@eastdelta.edu.bd, yjang@kookmin.ac.kr

***Abstract*—Due to climate change and its impact on agriculture, accurate estimate of future crops yield is very critical for low-lying countries like Bangladesh. Rice is the most consumed crop for Bangladesh with an annual production of more than 40 million tons per year. Despite being gifted with multiple seasons suitable for cultivating variety of rice, its productivity in Bangladesh has been changing following climatic variationsin this region over the last decades. In this paper, we present an approach, called WPSRY (Weather-based Prediction System for Rice Yield), for forecasting rice yield in different regions of Bangladesh. The proposed approach (WPSRY) firstly builds a model to predict the weather parameters applying Neural Networks (NN), and then estimates the rice yields applying Support Vector Regression (SVR) that uses as inputs predicted weather from NN as well as current agricultural data. Simulation demonstrates that WPSRY approach achieves promising prediction accuracy.**

Keywords—rice yield prediction; artificial neural network; support vector machine; means absolute error; mean relative error.

I. INTRODUCTION

More than half of 15 million hectares of total land in Bangladesh is being used for cultivation. Rice is one of the main crops that growwell in those vast areas of land and supplies 60-70% of calories to large population in this region [1]. Although the six different seasons favor the cultivation of different categories of rice throughout the year, Bangladesh is highly affected by global climate change [2] [3] and regularly experiences natural calamities such as floods due to heavy monsoon rainfall, droughts, and tropical cyclones, etc. Thus monitoring the growth and estimating the rice yield is very critical for different parties including government, commodity traders and farmers in order to plan, harvest, storage. Accurate prediction is also critical for ensuring food security.

There are three major crop varieties, e.g., Aus, Aman and Boro. Those categories crops are grown according to different time schedules of year. The cultivation seasons March-June, July-October, and November to December for Aus, Aman and Boro, respectively. The weather of Bangladesh is hot and humid during March-June, moist and rainy during July-October and cold and dry during November to December. Each of these varieties of rice shows vulnerability to somewhat different environmental stresses. Hench reliable prediction for rice yields is very challenging.

Predicted meteorological data for future can provide significance source of information to the forecasting models and helps to achieve reliable prediction for rice yield. Unfortunately weather stations network in Bangladesh is not too strong to supply weather predictions for agricultural areas located outside main cities. However, it is possible to find patterns and build models that can predict the future weather from historical data monitored for past years.

In this study, we present a new approach for rice yield forecasting in Bangladesh based on machine learning algorithms. The proposed approach employs predicted weather and previously recorded information for rice yield as the data source and works in two stages. Firstly, it obtains the education guess (prediction) of weather parameters for each separate agricultural region in Bangladesh. To develop prediction models for weather data it applies NN. Secondly, it builds another prediction for rice yield forecasting that uses as inputs, the outputs from NN based weather prediction model, and also past yield data. For yield prediction, it considers SVR as the prediction algorithm. Both NN and SVR are widely cited machine learning algorithms and successfully applied in many areas of engineering, science and agriculture.

This paper has been organized as follow. Section II provides a brief review of the previous works. Section III describes the proposed approach for rice yield prediction and Section IV presents the experimental setup. Section V describes the evaluation methods and presents the results. Finally, Section VI concludes the paper.

II. RELATED WORK

Forecasting crops yield has been an active area of research and a great variety of methods have been investigated. This section briefly reviews the prominent methods in the relevant literature.

Sarker et al. [4] applied widely cited regression technique to a set of rice yield data from Bangladesh. They used temperature and rainfall as inputs in their proposed approach and concluded that both rainfall and temperature have statistically significant impact on the yield of rice in Bangladesh.

Basak [5] employed the Decision Support System for Agro Technology Transfer (DSSAT) model to investigate the relationship of rice yield with the elements present in the environment. Based on a simulation results, it was shown that that temperature has a negative impact on the crops yield.

Shanmuganathanet al. [6] studied the relationship daily temperature and grapevine yield in New Zealand. They applied NNs to cluster the attributes that had similar values, and then chi-square test used to test the statistical significance among the related elements. They found that temperature variance along with precipitation, humidity and wind-speed has an important step in determining quality and the wine yield.

Jaikla et al. [7] developed a new approach for rice yield prediction using the SVR. The prediction method used for this study was divided into 3 phases - soil nitrogen prediction, rice stem weight prediction and rice grain weight prediction. They also compared the obtained results with that from commercial software (named as DSSAT4) that implements Crop Simulation Model (CSM-Rice simulation model).

The potential of several data mining techniques for rice yield prediction was investigated by Dey et al. [8]. They compared the performance of Multiple Linear Regression (MLR), Adaptive Boosting (AdaBoost), SVR and the Modified Nonlinear Regression (MNR). The parameters of these algorithms were optimized on training data and then were used along with the independent variables were used to make prediction for test data. As the evaluation measures they considered Mean Squared Error (MSE), Root Means Squared Error (RMSE), Mean Absolute Error (MAE) and R-square.

The above review confirms that weather data has significant impact on the rice yield and employing them as the inputs to the forecasting model may result significant improvement in prediction accuracy [e.g. 4, 6]. In this study, we further extend weather based prediction approach for rice yield but in a different way. Contrast to the previous studies that used current weather data to predict future rice yield, we use predicted weather data for future for forecasting rice yield.

III. WEATHER-BASED PREDICTION SYSTEM FOR RICE YIELD (WPSRY)

In this model, WPSRY (Weather-based Prediction System for Rice Yield), we have been predicting the rice yields on the basis of different predicted variables of weather. In order to do so, the first and foremost thing is acquiring each parameter values of Weather statistics and the Agriculture data monthly mean records of each district of 2012-2015 from the Bangladesh government records. For some of the districts particular year's climatic parameters or production data was not available, hence calculated those values using machine learning Weka tools [9]. Now, preparing the datasets for applying them into multilayer perceptron NN; we splitting our original datasets into three subsets, where 70% data split

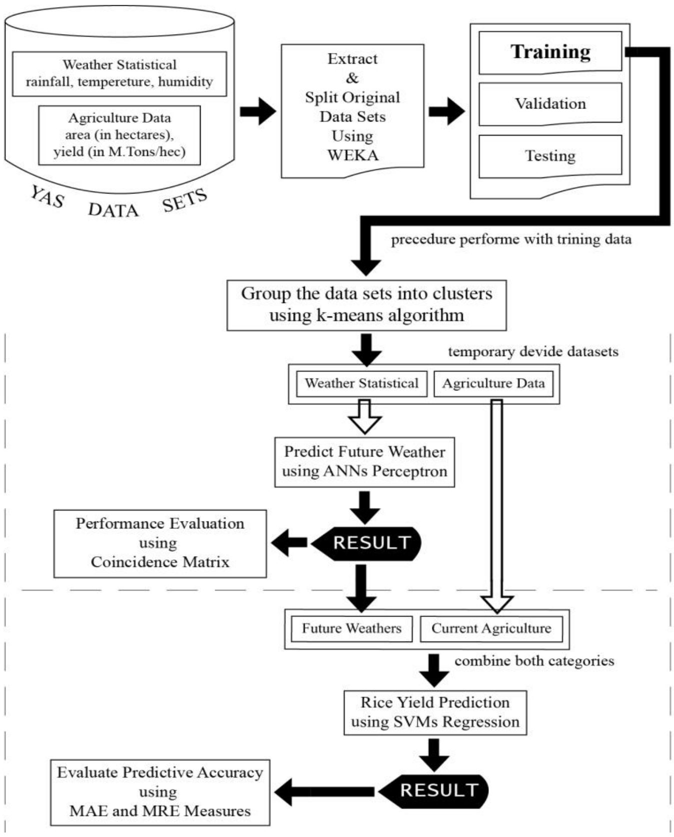

Fig. 1. Architecture of System Model (WPSRY)

as training and 15% as for validation and other 15% for testing data sets which construct with 11 categories of 500 elements. The splitting process has done by using Weka tools. Then, cluster datasets parameter values into groups using K-means algorithm. Since, our original data sets consist of 500 elements of 11 categories including agriculture data sets. So, for accomplish first approach we need all of elements of weather statistical with 9 categories after temporary omitting agriculture data sets. Another Important note, there are three different states also considered with weather attributes: average, minimum, and maximum, as shown in Fig. 2. At last, datasets were saved in ***.mat*** format for applying multilayer perceptron technique to forecast weather by using MatLab. Then, the coincidence matrix has used for evaluate the results of weather statistical. Afterward, combine the predicted parameter values of weather with current agriculture data to carry outthe next operation. For our study, SVM Regression has used to reach the desire target of yields. Here, SVRs implementations have done by using Python programming language [10]. At last, by using following measures: Means Absolute Error (MAE) and Mean Relative Error (MRE); we have analyzed the define results.

A. *Artificial Neural Network*

A neural network contains a set of inter-connected entities, called units or nodes. Each entity is designed to imitative its biological copy, the neuron. ANN technique is based upon the model of biological nervous system. The learning process has done by setup the network input nodes and assigning weights (W_{ij}) & bias (θ). Next in recall, we have to compute the net function with (1), then, outputs have measure in fixed format of 0's and 1's, as follow with (2).

$$net = \sum_i W_{ij} X_i - \theta \quad (1)$$

$$Y_j = \begin{bmatrix} 1 & net > 0 \\ & if \\ 0 & net \leq 0 \end{bmatrix} \quad (2)$$

Now, If $(T_j - Y_j) \neq 0$; than update weights and bias parameters (3) (4), where, T_j denoted for target and Y_j denoted for output, and σ denoted the learning rate. At last, by using (5), replace the updated values with old values, here, Δ define the update parameters.

$$\Delta W_{ij} = \sigma(T_j - Y_j) X_i \quad (3)$$

$$\Delta\theta_j = -\sigma(T_j - Y_j) \quad (4)$$

$$\begin{gathered} W_{ij} = W_{ij} + \Delta W_{ij} \\ \theta_j = \theta_j + \Delta\theta_j \end{gathered} \quad (5)$$

The above process with repeated until every input pattern is satisfied with according to (6).

$$(T_j - Y_j) == 0 \quad (6)$$

Figure 2, shows the build model of multilayer perceptron neural network, it has 9 input neurons, O hidden neurons and 21 output neurons corresponding to the 21 regions. In study [11], the authors successfully performed an operation by divided the regions of Bangladesh into 4 parts according to 21 different districts. Every district considered as class and clusters them with respective class for each region. Hence, we also consider the same 21 regions for weather statistical and also for crop yields. Here, O ensembles have build of NNs and then select the best one. Each ensemble E_k combines the predictions of q neuron network with the same number of hidden neurons but based on different initialization of weights.

B. *Support Vector Machine Regression*

SVMs construct a hyper-plane or hyper-plane sets in a high-dimensional space, which can be used for classification, regression, or any other tasks. SVR depends only on a subset of the training data, because the cost function for building the model ignores any training data close to the model prediction [12, 13]. Training the SVR means concluded according to (7) (8).

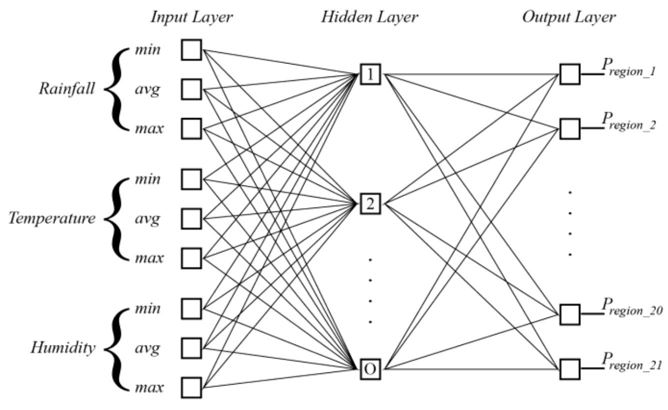

Fig. 2. Multilayer Perceptron Neural Network

$$\text{Minimize: } \frac{1}{2}\|w\|^2 \quad (7)$$

$$\text{Subject to: } \begin{aligned} y_i - (w, x_i) - b \leq \varepsilon \\ (w, x_i) + b - y_i \leq \varepsilon \end{aligned} \quad (8)$$

Where, x_i is training sample with target value, y_i. The inner product plus intercept $(w, x_i) + b$ is the prediction for that sample, and ε is a free parameter that serves as a threshold: all predictions have to be within a ε range of the true predictions. Slack variables are usually added into the above to allow for errors and to allow approximation in the case the above problem is infeasible.

IV. EXPERIMENTAL DESIGN

A. *Problem Statement*

Taking measure of the weather conditions becomes an imperative point if one is to predict the yield of the crop. Given data sets –

- A time series of previous weather data for selected location monthly measure: $W = [W_1, W_2, W_3, \dots\dots, W_m]$, where W_i is the weather data for month i. The dimensional vectors of weather of the maximum, minimum and average monthly temperature (T), humidity (H) and rainfall (R), $W_i = [T^i_{max}, T^i_{min}, T^i_{avg}, H^i_{max}, H^i_{min}, H^i_{avg}, R^i_{max}, R^i_{min}, R^i_{avg}]$.
- Agriculture data sets in Area (in hectares), A; and Yields (in M.Tons/hectare), Y.

B. *Datasets Collection*

All the datasets used in this paper have sourced from the publicly available records of the Government for the year 2006 to 2015. For this study, we collected and merged our required data sets from the "Yearbook of Agricultural Statistics-2015", which is a statistical report generated by the Bangladesh Bureau of Statistics (BBS) [14]. It is a comprehensive report that consists of several statistical attributes including weather statistical that also sourced from Bangladesh Meteorological Department (BMD). Also some other weather data collected from BMD's weather forecasting reports. In this way data for our study of the last 4 years was collected – 2012, 2013, 2014, and 2015.

C. *Used Tools*

A fitting application known as *nftool*, which we used to make sure that how fits our datasets are. Process performed by train a network where *Levenberg-Marquardt* learning algorithm was selected for calculate datasets by running into network. This algorithm typically takes more memory but less time. Training automatically stops when generalization stops improving, as indicated by an increase in the means square error of the validation samples. Also *Weka* tool have used for splitting dataset, calculate missing data and evaluate results.

D. Method Used For Comparison

The method used for comparison (M-1) is very similar to the recently proposed method by data mining [8]. They classify the datasets into two formats, independent and dependent. Yield data considered as dependent data where other weather and area datasets was denoted as independent. Then usage of Multiple Linear Regression (MLR), Adaptive Boosting (AdaBoost), Support Vector Machine Regression (SVMR) and the Modified Nonlinear Regression (MNR), equation is implemented on the training set to find out parameter values. However, there are several important differences between this method and our method.

Firstly, we apply k-means cluster algorithm to group the data based on their weather conditions and select the optimal clusters based on the data. However, they do not use any cluster algorithm; just manually assign the weather based on the weather report. Secondly, as inputs to the prediction model for each region we use predicted weather profile, while they use current weather data. Finally, the weather profile in our case is defined as 9-dimensional vector using information from 3 sources (rainfall, temperature and humidity) while they just used only 3 sources (rainfall, temperature and humidity).

V. Performance Evaluation

A. Evaluation Methods

A coincidence matrix have generated for evaluate predicted weather parameter values. The diagonal values represent the correct decisions made. The values outside this diagonal indicate the errors. True positive rate of a classifier is estimated by dividing the correctly classified positives by the total positive count. The false positive rate of the classifier is estimated by dividing the incorrectly classified negatives by the total negatives. The overall accuracy of a classifier is estimated by dividing the total correctly classified positives and negatives by the total number of samples. Other performance measures, such as recall, specificity and F-measure are also used for calculating other aggregated performance measures. These values are defined in Fig. 3.

		True Class	
		Positive	Negative
Predicted Class	Positive	True Positive Count (TP)	False Positive Count (FP)
	Negative	False Negative Count (FN)	True Negative Count (TN)

Fig. 3. Coincidence Matrix

From the values of Fig. 3, True Positive Rate, specificity, accuracy, F1 Score and MCC were calculated.

$$True\ Positive\ Rate = \frac{TP}{TP+FN} \tag{9}$$

$$Precision = \frac{TP}{TP+FP} \tag{10}$$

Correctly classified instances are known as Sensitivity or True Positive Rate and incorrectly classified instances are known as Specificity or Precision and can be calculated using (9) and (10) described above. Accuracy is defined as the overall success rate of the classifier and computed by (11).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{11}$$

The F1 score measures (as well known as F-measure) exactness utilizing the insights Precision and Recall. Accuracy is the proportion of true positives (TP) to all predicted positives (TP + FP). Recall is the proportion of true positives to every single actual positive (TP + FN).

$$Precision = \frac{TP}{TP+FP} \tag{12}$$

$$Recall = \frac{TP}{TP+FN} \tag{13}$$

$$F-measure = \frac{2}{\frac{1}{Presition}+\frac{1}{Recall}} \tag{14}$$

The Matthews Correlation Coefficient (MCC) is a measure of the nature of paired (two-class) arrangements. The MCC can be calculated by using the (15).

$$MCC = \frac{((TP\times TN)-(FP\times FN))}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \tag{15}$$

The study datasets has executed using Knowledge Flow in Weka tool. The Weka components connected them together in order to form a knowledge flow for processing and analyzing current datasets. A previous study also has done this with coincident matrix successfully [15]. The components selected in Knowledge flow are discussed below:

DataSources (tab): *ArffLoader* load the ARFF dataset file for the current study.

Evaluation (tab): *ClassAssigner* assigns column as class for current study dataset, training or testing. *ClassValuePicker* chooses class value to be considered as positive class. *CrossValidationFoldMaker* splits the current data set, training set or test set into folds. *Classifier Performance Evaluator* evaluates the performance of batch trained/tested classifiers of the current study.

Classifiers (tab): *Multilayer Perceptron* classifier to be used for the current study.

Following two measures uses to evaluate the predictive accuracy of rice yields, Means Absolute Error (MAE) and Mean Relative Error (MRE). These are the two most popular measures for evaluating the accuracy of any prediction. MAE and MRE are defined as follows with (16) & (17).

$$MAE = \frac{1}{M}\frac{1}{R_T}\sum_{m=1}^{M}\sum_{r=1}^{R_T}|p_m^r - \hat{p}_m^r| \tag{16}$$

$$MRE = \frac{1}{M}\frac{1}{R_T}\sum_{m=1}^{M}\sum_{r=1}^{R_T}\left|\frac{p_m^r-\hat{p}_m^r}{R}\right| \times 100\% \qquad (17)$$

Where, p_d^r and $\hat{p}_d^r$ are the actual and predicted yield outputs for month m at region r, respectively; M is the number of instances (months) in the testing data; R_T is the total number of predicted yield outputs for a month, and R is the range of the yield output.

B. Results Analysis

This part presents the results obtained after running the Multilayer Perceptron technique on weather dataset of Bangladesh. Weka was used to construct the algorithms. Different parameter set for multilayer perceptron was as follows:

GUI=true;	learningRate=0.3;
autoBuild=true;	momentum=0.2;
debug=true;	seed=0;
decay=false;	nominalToBinaryFilter=true;
hiddenLayers=a;	normalizeAttributes=true;
reset=True;	normalizeNumericClass=true;

The algorithm achieved the accuracy of 93.64%, sensitivity of 89.36% and specificity of 91.72%. The model was also evaluated and following parameters were computed which resulted in mean absolute error of 0.0423, relative absolute error of 11.6%, root mean squared error of 0.1327 and root relative squared error of 28.9%. Further it gives F1 score of 0.91 and MCC of 0.93.

The results of Aman rice is present with Fig. 4. Our method got MRE=5.82% after spending 11 min training time, where method M-1 reached to MRE=7.76%.

Figure 5, shows the results for Aus rice. Method M-1 got MRE=9.87% while our method got MRE=8.79% with 13 min training time.

Finally for boro rice, in Fig. 6 presents the results for our approach. This time, we arrived with 9.5 min training time to minimum MRE=6.28%, and the method M-1 reached MRE=7.10%.

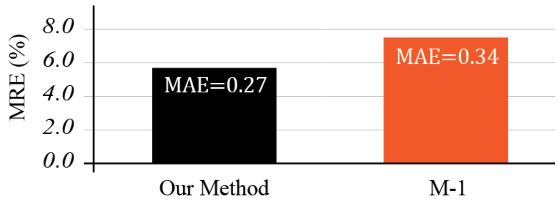

Fig. 4. Results for Amen rice yields

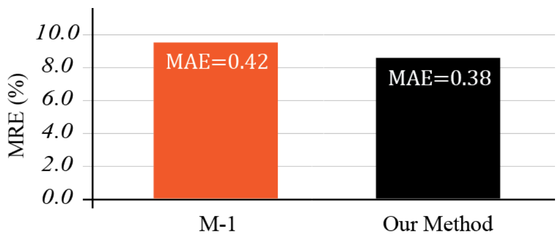

Fig. 5. Results for Aus rice yields

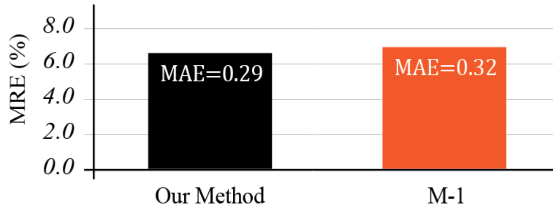

Fig. 6. Results for Boro rice yields

The predicted values for the 2016 seasons are cross-validated with the actual values from the same seasons. Table I show the predicted results for Aman for total 21 different regions of year 2016. Here, actual yield values also define with residual parameters.

Similarly, Table II shows both actual and predicted results of Aus crop for total 21 regions of year 2016.

Finally, Table III depicts results for Boro crop yield along with actual yields.

TABLE I. RESULTS for ACTUAL v/s PREDICTED YIELD for AMAN

Year	Region (total 21)	Actual Yield	Predicted Yield	Residual
2016	Barisal	2.406	2.400	-0.006
2016	Bhola	2.465	2.460	-0.005
2016	Bogra	2.374	2.264	-0.11
2016	Chandpur	1.982	1.918	-0.064
2016	Chittagong	1.568	1.589	0.021
2016	Comilla	3.081	3.109	0.028
2016	Cox's Bazar	1.254	1.112	-0.142
2016	Dhaka	2.224	2.241	0.017
2016	Dinajpur	2.983	3.027	0.044
2016	Faridpur	2.180	2.019	-0.161
2016	Feni	3.112	3.119	0.007
2016	Jessore	3.765	3.760	-0.005
2016	Khulna	1.641	1.633	-0.008
2016	Madaripur	1.141	0.998	-0.143
2016	Mymemsingh	1.318	1.220	-0.098
2016	Patuakhali	2.028	2.098	0.07
2016	Rajshahi	2.097	1.999	-0.098
2016	Satkhira	2.160	2.060	-0.1
2016	Sylhet	2.655	2.801	0.146
2016	Tangail	1.967	0.882	-1.085
2016	Rangpur	1.774	1.611	-0.163

TABLE II. RESULTS for ACTUAL v/s PREDICTED YIELD for AUS

Year	Region (total 21)	Actual Yield	Predicted Yield	Residual
2016	Barisal	2.965	2.859	-0.106
2016	Bhola	2.584	2.512	-0.072
2016	Bogra	2.417	2.404	-0.013

2016	Chandpur	2.885	2.889	0.004
2016	Chittagong	2.160	2.082	-0.078
2016	Comilla	2.659	2.637	-0.022
2016	Cox's Bazar	2.959	3.050	0.091
2016	Dhaka	2.881	2.717	-0.164
2016	Dinajpur	2.452	2.414	-0.038
2016	Faridpur	2.861	2.327	-0.534
2016	Feni	2.782	2.705	-0.077
2016	Jessore	1.127	1.107	-0.02
2016	Khulna	2.405	1.410	-0.995
2016	Madaripur	3.405	3.400	-0.005
2016	Mymemsingh	3.154	2.200	-0.954
2016	Patuakhali	3.112	3.111	-0.001
2016	Rajshahi	2.654	2.655	0.001
2016	Satkhira	2.425	2.415	-0.01
2016	Sylhet	2.909	2.899	-0.01
2016	Tangail	2.009	2.011	0.002
2016	Rangpur	3.001	3.006	0.005

TABLE III. RESULTS for ACTUAL v/s PREDICTED YIELD for BORO

Year	Region (total 21)	Actual Yield	Predicted Yield	Residual
2016	Barisal	2.965	2.859	-0.106
2016	Bhola	2.584	2.512	-0.072
2016	Bogra	2.417	2.404	-0.013
2016	Chandpur	2.885	2.889	0.004
2016	Chittagong	2.160	2.082	-0.078
2016	Comilla	2.659	2.637	-0.022
2016	Cox's Bazar	2.959	3.050	0.091
2016	Dhaka	2.881	2.717	-0.164
2016	Dinajpur	2.452	2.414	-0.038
2016	Faridpur	2.861	2.327	-0.534
2016	Feni	2.782	2.705	-0.077
2016	Jessore	1.127	1.107	-0.02
2016	Khulna	2.405	1.410	-0.995
2016	Madaripur	3.405	3.400	-0.005
2016	Mymemsingh	3.154	2.200	-0.954
2016	Patuakhali	3.112	3.111	-0.001
2016	Rajshahi	2.654	2.655	0.001
2016	Satkhira	2.425	2.415	-0.01
2016	Sylhet	2.909	2.899	-0.01
2016	Tangail	2.009	2.011	0.002
2016	Rangpur	3.001	3.006	0.005

VI. CONCLUSION

The study shows that, a successful interaction among rice yield and the weather statistical of different region within certain error threshold. This will enable people concerned in the field of agricultural to take first the educated guess of weather variables; so as to minimize their loss in the coming years. This will enable to have a better way to think about more accurate results.

REFERENCE

[1] A. Rahman, Leonid Roytman, Nir Y. Krakauer, Mohammad Nizamuddin, and Mitch Goldberg, Use of Vegetation Health Data for Estimation of Aus Rice Yield in Bangladesh, Sensors, vol 9(4), pp. 2968-2975, 2009.

[2] R. Mahmood, Air temperature variations and rice productivity in Bangladesh: a comparative study of the performance of the YIELD and the CERES-Rice models, Ecological Modelling, vol 106(2-3), pp. 201-212, 1998.

[3] Anwar, A. “Climate change impacts and adaptation assessment in Bangladesh”. Clim. Res. 1999, 12, pp. 109–116.

[4] Md. A. R. Sarker et al., “Exploring the relationship between climate change and rice yield in Bangladesh,” Elsevier Agricultural Systems vol. 112, pp. 11-16, October, 2012.

[5] J. K. Basak. “Climate Change Impacts on Rice Production in Bangladesh: Results from a Model”, URL: *http://www.unnayan.org/documents/Climatechange/climchange_impacts_rice_production.pdf/*.

[6] S. S. et al., "Data Mining Techniques for Modelling the Influence of Daily Extreme Weather Conditions on Grapevine, Wine Quality and Perennial Crop Yield," *IEEE Second International Conference on Computational Intelligence, Communication Systems and Networks* (CICSyN) , pp. 90-95, 2010.

[7] R. J., S. A. and A. J., “Rice Yield Prediction using a Support Vector Regression method”, Proceedings of ECTI-CON 2008, IEEE, pp. 29-32, 2008.

[8] U. K. Dey, A. H. Masud and Md. N. Uddin, “Rice Yield Prediction Model Using Data Mining”, *IEEE International Conference on Electrical, Computer and Communication Engineering (ECCE)*, pp. 321-326, 2017.

[9] Weka 3: Data Mining Software in Java, Machine Learning Group at the University of Waikato, Official Web: *http://www.cs.waikato.ac.nz/ml/Weka/index.html/, accessed on 5th June 2017*.

[10] Understanding Support Vector Machine via Examples, URL: *https://sadanand-singh.github.io/posts/svmpython/*.

[11] Md. M. Rahman and N. Haq, R. M Rahman, “Application of Data Mining Tools for Rice Yield Prediction on Clustered Regions of Bangladesh”, *17th International Conference on Computer and Information Technology (ICCIT)*, pp. 8-13, 2014.

[12] Smola, Alex J.; Schölkopf, Bernhard (2004). “A tutorial on support vector regression”, Statistics and Computing. 14 (3): 199–222, URL: *http://eprints.pascal-network.org/archive/00000856/01/fulltext.pdf/*.

[13] Smola, Alex J.; Scholkopf, Bernhard. “A tutorial on support vector regression”, URL: *http://www.svms.org/regression/SmSc98.pdf/*.

[14] Bangladesh Bureau of Statistics. Bangladesh Bureau of Statistics, Yearbook of Agricultural Statistics of Bangladesh. Government of the People’s Republic of Bangladesh, 2015, URL: *http://www.bbs.gov.bd/, last assessed on 13th April, 2017*.

[15] N. Gandhi, O. Petkar and L. J. Armstrong, “Rice Crop Yield Prediction Using Artificial Neural Networks”, *IEEE International Conference on Technological Innovations in ICT For Agriculture and Rural Development (TIAR 2016)*, pp. 105-110, 2016.