Quest Journals
Journal of Architecture and Civil Engineering
Volume 6 ~ Issue 7 (2021) pp: 05-09
www.questjournals.org

Research Paper

Arma Based Crop Yield Prediction Using Temperature and Rainfall Parameter Classifications

Ananthi.V[1], Meenaroshini.D[2], Bhavatharani.K[3], Dr.Venkateshguru[4]
[1]*UG-cse KSR college of engineering*
[2]*UG-cse KSR college of engineering*
[3]*UG-cse KSR college of engineering*
[4]*Assistant Professor – cse KSR college of engineering*

ABSTRACT
The main aim of this project is to provide a methodology for crop yield prediction based on the historical climatic and production data. Crop yield prediction based on the previous years of temperature and rainfall can help farmers take necessary steps to improve crop yield in the coming season.Crops are sensitive to various weather phenomena such as temperature and rainfall. Therefore, it becomes crucial to include these features when predicting the yield of a crop.In this work, ARMA (Auto Regressive Moving Average) method is used to forecast crop yield. Past ten years of data set is taken for temperature and rainfall for our country. Yield prediction is then carried out using a Fuzzy logic algorithm to better judge the crop yield.In addition, this project classifies the data set records using KNN to predict the model for future test record data sets. The project is designed using R Studio. The coding language used is R 3.4.4.
KEYWORDS: *DataSet, SVM Algorithm, Temperature, Rainfall, Crop yield prediction.*

Received 20 June, 2021; Revised: 03 July, 2021; Accepted 05 July, 2021 © The author(s) 2021. Published with open access at www.questjournals.org

I. INTRODUCTION

Data mining is still in its infancy, companies in a wide range of industries - including retail, finance, health care, manufacturing transportation, and aerospace - are already using data mining tools and techniques to take advantage of historical data. Data mining or knowledge discovery is the computer assisted process of digging through and analyzing enormous sets of data and then extracting the meaning of the data. Tools of data mining predict behaviors and future trends, allowing businesses to make proactive,knowledge-driven decisions. They can answer business questions that traditionally were too time consuming to resolve. They scour databases for hidden patterns, finding predictive information that experts may miss because it lies outside their expectations.

Agriculture is one of the most important economic sectors in India. It plays an important role in rural development and sustainability. The level of agriculture may decrease due to factors like unpredicted rainfall, climate change, use of excessive pesticides etc. The main aim of this study is to provide a methodology for crop yield production based on the historical climatic and production data. Crop yield prediction based on the previous years of temperature and rainfall can help farmers take necessary steps to improve crop yield in the coming season.Understanding crop yield can help ensure food security and reduce impacts of climate change. The project tried to develop a method such that the crop yield can be predicted beforehand using only temperature and rainfall of previous years. Accurate rainfall prediction is a difficult task because rainfall depends on various features such as cloud cover, evapotranspiration, and many other climatic factors but we wanted to extract useful information about crop yield using only two features i.e. temperature and rainfall. The proposed method uses ragAuto Regressive Moving Avee and **Seasonal** ARMA models to predict temperature. Since we want to predict temperature from past values, these models suit our needs. A time-series created from the dataset is fed into the model to predict temperature. Similarly, ARMA and ARMA with exogenous variables (ARMAX) models are used to predict rainfall. The ARMAX model is used in case of rainfall so that other factors such as cloud cover, temperature and evapotranspiration can also be taken into account. We used a fuzzy logic system to predict yield. The fuzzy model takes in the predicted values from the model with least errors and gives the yield for that season.

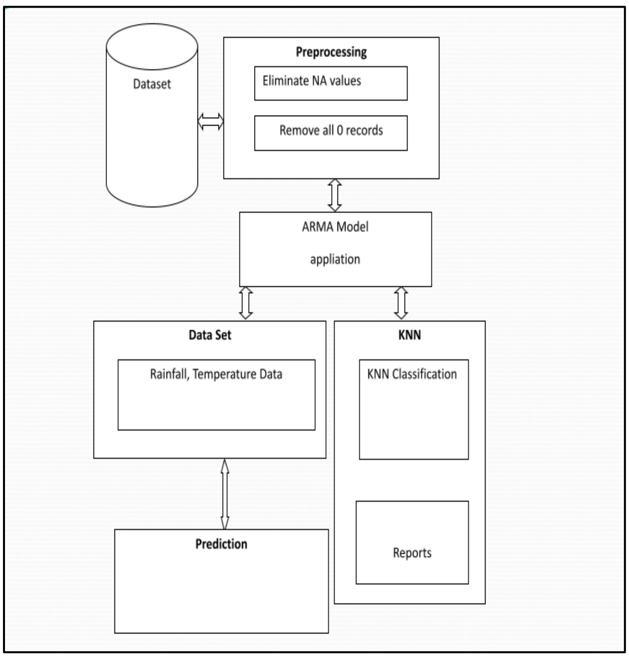
II. RELATED WORKS

In this paper [1] the authors stated that most of greenhouse growers desire a determined amount of yields in order to accurately meet market requirements. The purpose of this paper is to explore the dynamics of neural networks in forecasting crop (tomato) yield using environmental variables; here they aimed at giving accurate yield amount. They used the Adaptive Neuro-Fuzzy Inference System (ANFIS). The input to ANFIS is several parameters derived from the crop growth model (temperature, Co2, vapor pressure deficit (VPD), yield, and radiation).

ANFIS has only one output node, the yield. One of the difficult issues in predicting yield is that remote sensing data do not go long back in time. Therefore any predicting effort is forced to use a very restricted number of past years in order to construct a model to forecast future values. The system is trained by leaving one year out and using all the other data. They then evaluate the deviation of their estimate compared to the yield of the year that is left out. The procedure is applied to all the years and the average forecasting accuracy is given.ANFIS is considered as a class of adaptive networks that perform as a framework for adaptive fuzzy inference [6] systems. Generally, it is a multilayer feed forward adaptive network where each node realizes a particular node function of its corresponding inputs and the nodes in ANFIS include adaptive and fixed ones, and ANFIS is characterized with the parameter set that is the union of the parameter sets associated with all adaptive nodes. The use of a neuro-fuzzy system for crop yield [7] estimate has some interesting characteristics. All the variables that are input into the system are associated with varying degrees of accuracy. Some ambiguity comes from measurement error and generality. Using fuzzy sets instead of the actual values as inputs, we aim at shifting to the semantics of the data rather than its measure [6]. It is well known that with neuro–fuzzy modeling there is the alternative to use a fuzzy set as the output. In this case, yield would be expressed for example as low, normal of high with each of those three borders corresponding to a fuzzy set.They did not imply however that seeking a crisp (non-fuzzy) value is a more exact approach than seeking a trend expressed in fuzzy sets (low, medium, etc.). But although the accuracy of prediction is probably the same in both expressions of desired output, people are more used to and feel more confident in looking at number rather than a membership function.

III. PROPOSED METHODOLOGY

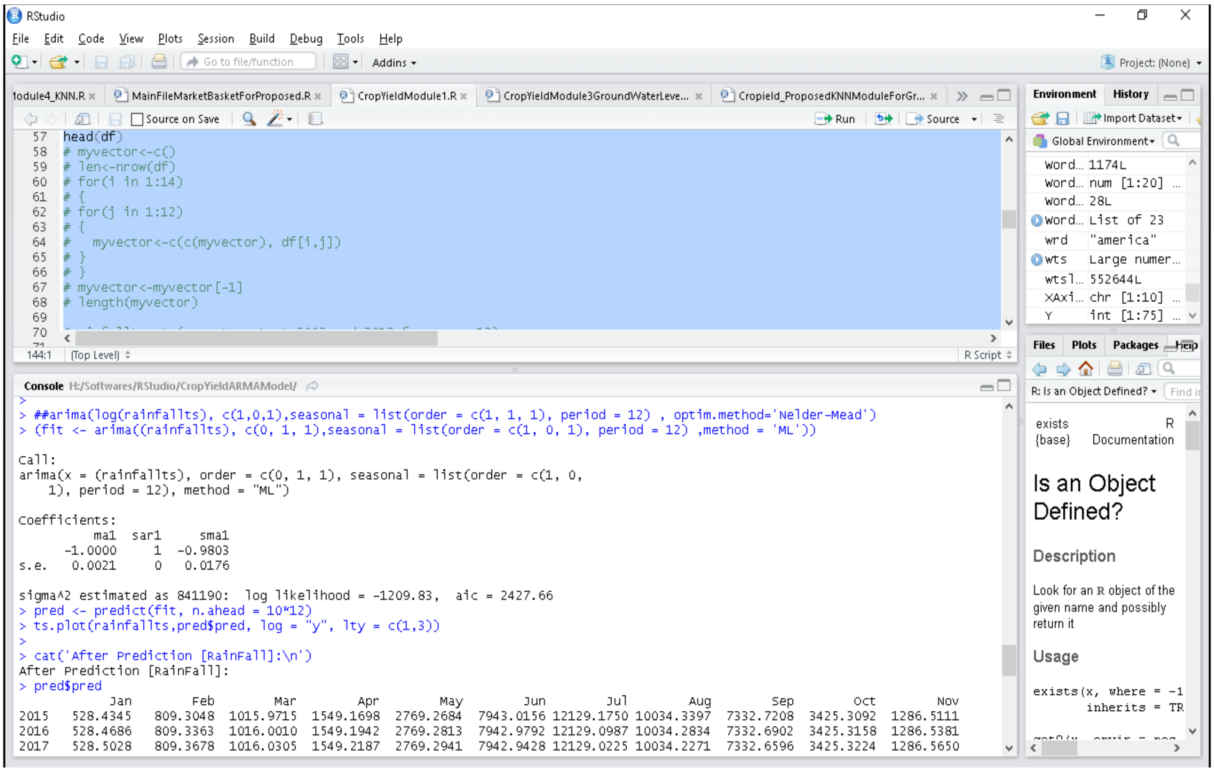
In proposed system, all the existing methodology is carried out. Like existing system, here also, the data set is taken from Indian meteorological sites saved in excel files as 'csv' files. It contains temperature and rainfall data. These data are taken for text pre-processing first and then converted into time series data set format. Then ARMA model is used to predict the future years' levels and then a fuzzy logic is applied to classify the crop yield in future years. In addition, KNN is applied for classification and so it is found to be suitable especially if the data set is having more number of records is contains outlier data.The KNN model yields better accuracy only if the test record is exactly matched with any of the training data records.KNN is applied so could be preferred when the data set grows larger.KNN is applied so could be preferred when the outlier data is more.



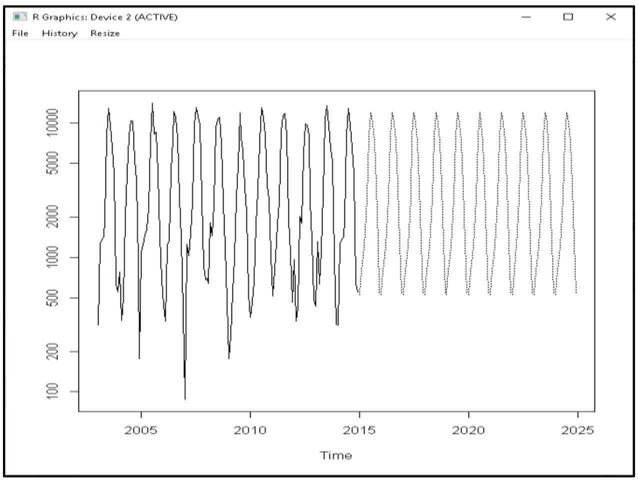
It deals with all the analysis that takes up in developing the project. Each structure has to be thought of in the developing of the project, as it has to serve the end user in a user-friendly manner. One must know the type of information to be gathered and the system analysis consist of collecting, Organizing and evaluating facts about a system and its environment. The main objective of the system analysis is to predict the crop yield for future years. The prediction needs to be analyzed well. The details are processed through coding themselves. It will be controlled by the programs alone.The Proposed system accessing process to solves problems what occurred in existing system. The current day-to-day operations of the organization can be fit into this system. Mountainly operational feasibility should include on analysis of how the proposed system will affects the organizational structures and procedures.The cost and benefit analysis may be concluded that computerized system is favorable in today's fast moving world.The assessment of technical feasibility must be based on an outline design of the system requirements in terms of input, output, files, programs and procedure. The project aims to predict the crop yield for future years. The current system aims to overcome the problems of the existing system. The current system is to reduce the technical skill requirements so that more number of users can access the application.

IV. RESULT

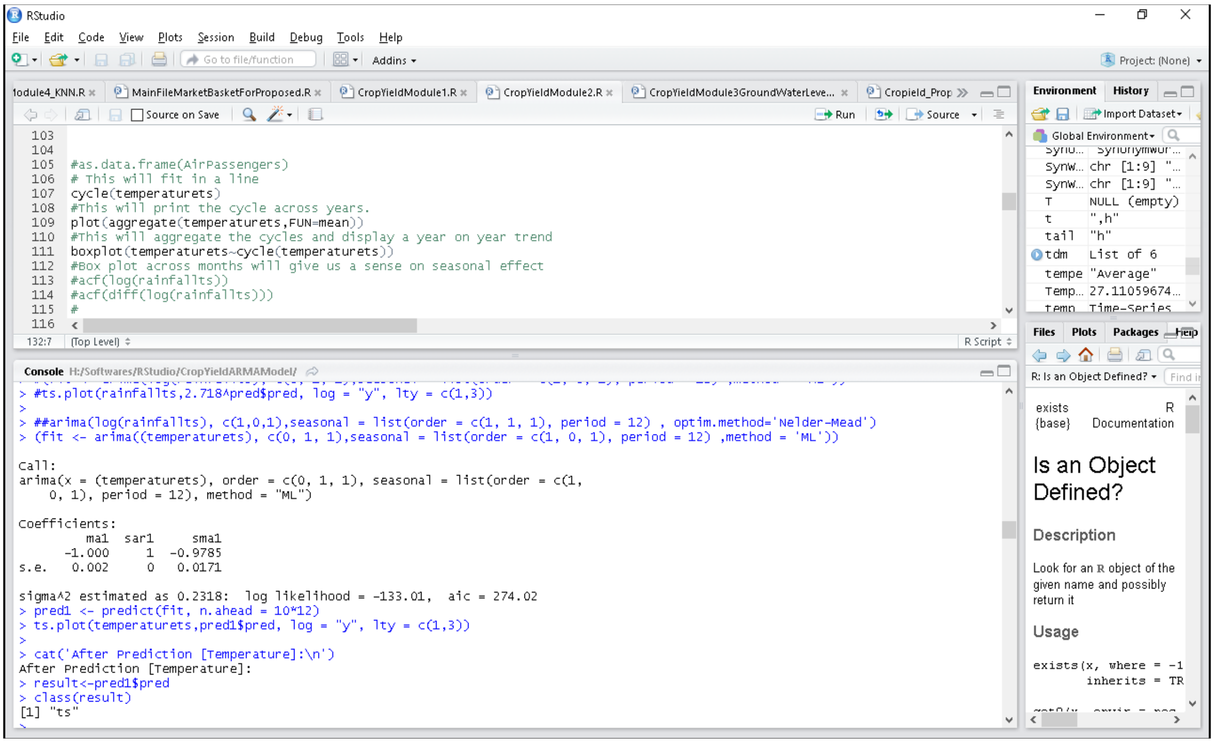
4.1. RAINFALL PREDICTION



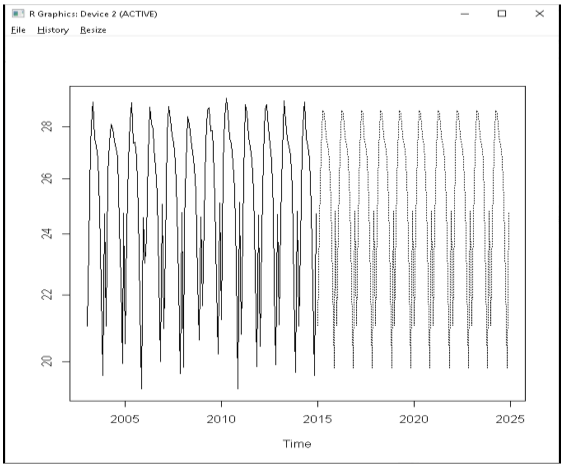
OUTPUT



4.2. TEMPERATURE PREDICTION



OUTPUT



V. CONCLUSION

According to the results, temperature is best predicted by the ARIMA model and the accuracy of predictions made for rainfall by ARMA model is also good. Rainfall, which is an important factor for the prediction of crop yield is difficult to estimate precisely. Climate factors may change due to other remaining variables which may influence the prediction of rainfall.Also, the proposed work makes use of fuzzy logic to estimate crop yield which works on a set range rather than discrete values, therefore, the error in predicted rainfall data does not cause problems as long as the difference between actual and estimated values is not drastic. The model can successfully predict crop yield for a given year when the rainfall and temperature values for the previous years is known. It will be helpful in analyzing the past data and so as to predict the future levels. In future, logistic regression can be applied to further classify the data.The main aim of this project is to provide a methodology for crop yield production based on the historical climatic and production data.Crop yield prediction based on the previous years of temperature and rainfall can help farmers take necessary steps to improve crop yield in the coming season. Understanding crop yield can help ensure food security and reduce impacts of climate change.Crops are sensitive to various weather phenomena such as temperature and rainfall. Therefore, it becomes crucial to include these features when predicting the yield of a crop. Weather forecasting is a complicated process. In this work, ARMA (Auto Regressive Moving Average) method is used to forecast crop yield. Past ten years of data set is taken for temperature and rainfall for our country. Yield prediction is then carried out using a Fuzzy logic algorithm to better judge the crop yield.In addition, this project data set records are used in KNN approach to predict the model for future test record data sets. The project is designed using R Studio. The coding language used is R 3.4.4.

REFERENCES

[1]. Sandip Rakshit, "**R for Beginners**"Edition 1, 2017, McGraw-Hill Education, ISBN: 9789352604555, 9352604555.

[2]. Mike. McGrath, **"R for Data Analysis in easy steps: R Programming Essentials",** Kindle Edition. **Sold by:** Amazon Asia-Pacific Holdings Private Limited.

[3]. Mark Gardner, **"Beginning R: The Statistical Programming Language",** Kindle Edition.

[4]. P. Kumar "Crop yield forecasting by adaptive neuro fuzzy inference system." Mathematical Theory and Modeling 1.3, pp. 1-7, 2011.

[5]. I. Kaushik, and S.M. Singh. "Seasonal ARIMA model for forecasting of monthly rainfall and temperature." Journal of Environmental Research and Development 3, no. 2, pp. 506-514, 2008.

[6]. S. Veenadhari, B. Misra, and C. D. Singh. "Machine learning approach for forecasting crop yield based on climatic parameters." 2014 International Conference on Computer Communication and Informatics, pp. 1-5. IEEE, 2014.

[7]. S. Prabakaran, P.N. Kumar, and P.S.M. Tarun. "Rainfall prediction using modified linear regression", 2006.

[8]. S.K. Mohapatra, A. Upadhyay, and C. Gola. "Rainfall prediction based on 100 years of meteorological data." 2017 International Conference on Computing and Communication Technologies for Smart Nation (IC3TSN), pp. 162-166. IEEE, 2017.

[9]. L.A.Zadeh, "Fuzzy sets," Information and Control, pp. 338-353, 1965.

[10]. Ehret, D. L., B. D. Hill, D. A. Raworth, and B. Estergaard. Artificial neural network modeling to predict cuticle cracking in green house peppers and tomatoes. Computers and Electronics in Agric. 61(2): 108-116. 2008.

[11]. Pucheta, J., D. Patiño, R. Fullana, C..Neurodynamic programming-based optimal control for crop growth in precision agriculture. In Proc. 16th IFAC World Congress. Available at: www.nt.ntnu.no/users/skoge/prost/ proceedings/ ifac2005/ Papers/ Paper5047.html. Schugurensky, and B.Kuchen.2005.

[12]. Kiran Mai,C., Murali Krishna, I.V.,and A.VenugopalReddy, 2006.Data Mining of Geospatial Database for Agriculture Related Application. Proceedings of Map India. New Delhi. ((http://www.gisdevelopment.net/proceedings/mapindia/2006/agriculture/mi06agri_12 4.htm).

[13]. Jorquera H, Perez R, Cipriano A, Acuna G(2001). Short term forecasting of air pollution episodes. In. Zannetti P (eds) Environmental Modeling 4. WITPress, UK.

[14]. Rajagopalan B. Lall U (1999) A k- nearest- neighbor daily precipitation and other weather variables. WatResResearch35(10) :3089 – 3101.

[15]. Tripathi S, Srinivas VV, Nanjudiah RS (2006).Down scaling of precipitation for climate change scenarios: a support vector machine approach, J. Hydrology 330-337.

[16]. Zelu Zia (2009). An Expert System Based on Spatial Data Mining used Decision Tree for Agriculture Land Grading. Second International Conference on Intelligent Computation Technology and Automation.Oct10-11,China.

[17]. Jun Wu, Anastasiya Olesnikova, Chi- Hwa Song, Won Don Lee (2009).The Development and Application of Decision Tree for Agriculture Data. IITSI :16-20.

[18]. Dhanya,C.T. and D. Nagesh Kumar, 2009. Data mining for evolution of association rules for droughts and floods in India using climate inputs. J. of Geo. Phy.Res.114:1-15.